From print to digital: First steps in collecting digital music publications in UK legal deposit libraries

Amelie Roper
Cambridge University Library, University of Cambridge, UK

Abstract
As a result of the 2013 Non-Print Legal Deposit Regulations, the United Kingdom’s legal deposit libraries acquired two large collections of digital music publications in PDF format: 43,165 from Music Sales and 13,167 from Faber Music. These constitute their back catalogues for the period 2013 to 2018. This article considers the genres of music that were collected, the relationship between printed and digital output and intended users. The increasing prevalence of digital publication in the United Kingdom’s music publishing industry was evident, with 99% of the content collected from these firms in this period in digital format. The ‘near duplication’ of content through the production of variant editions contributed greatly to the volume of output. Although 71% of content was popular music, other musical genres were also represented. Most of the publications were intended for performance rather than academic study. Despite the tendency to reprint extracts from printed publications digitally, there is currently little potential to switch UK music publishers from print to digital deposit. To truly reflect the nation’s cultural heritage, future collecting will need to embrace the breadth of the United Kingdom’s digital music publishing industry. Legal deposit libraries will need to collect publications with accompanying audiovisual content, those in multiple parts, proprietary music notation file formats and interactive content delivered via apps. This will require workflows that can accommodate publishers producing a handful of publications and those publishing at scale. The sustainability of collecting relies on close cooperation between libraries, composers, publishers, distributors, aggregators and music notation software providers.

Keywords
digital music publications, Faber Music, legal deposit, Music Sales

Corresponding author:
Amelie Roper, Cambridge University Library, West Road, Cambridge CB3 9DR, UK.
Email: alr39@cam.ac.uk
In the 1520s, the English lawyer, publisher and printer John Rastell made a discovery that had a decisive impact on the history of music publishing. By pioneering the technology of printing music by single impression using movable type, he revolutionised the industry, making it possible to print notation more quickly and economically (King, 1971, 1975). Fast forward five centuries and the use of the handpress and movable type has long been abandoned in favour of digital production methods. From the 1980s, this involved the use of notation software, with the resultant publications still issued in printed format. In the decades that followed, the advent of the internet and the availability of a greater range of open-source notation software programs resulted in the use of digital methods for both the production of music publications and their distribution to customers. Although music continues to be sold in print, digital publications are now a steady growth area of the industry, with research undertaken in 2018 indicating that 53% of UK music publishers were issuing content in digital format, an increase from 37% in 2014 (Mazzucco, 2018: 14; Patterson, 2014: 10).

**Non-print legal deposit**

The circulation of music publications in digital format is a development whose significance rivals Rastell’s earlier innovations, and which has mirrored the increasing availability of books, journals and maps electronically. Recognising the proliferation of digital publications and the need to offer them the same protection as print, the Legal Deposit Libraries Act was passed in 2003, paving the way for a consultation period during which procedures and processes were formulated to bring digital publications under the umbrella of legal deposit. Rather than direct legislation, this was achieved by means of cross-sectoral self-regulation in the form of the Legal Deposit Libraries (Non-Print Works) Regulations 2013. These are an instrument of secondary legislation that complement the primary legislation of the 2003 Act. Extending the long-standing arrangements that are in place for printed publications, they promote comprehensive collecting by giving the United Kingdom’s six legal deposit libraries the right to collect and archive published digital material, and make it available to users. The aim is to form a national collection of non-print publications, and the content as a whole has come to be referred to as non-print legal deposit.

The Regulations mitigate against the development of a ‘digital black hole’ that could result if digital content, especially born-digital content (content with no printed counterpart), is not collected (Smith and Cooke, 2017: 176). UK publishers are required to deposit the files and associated metadata for their digital publications with the British Library. This content is then replicated at four nodes: the two British Library sites (Boston Spa and St Pancras), the National Library of Wales at Aberystwyth and the National Library of Scotland at Edinburgh. These four institutions access this content directly from their local nodes, whereas the other legal deposit libraries – Cambridge University Library, the Bodleian Library and Trinity College Dublin – obtain access via a secure network. Publications are made available to users on dedicated terminals in these institutions’ reading rooms via a two-step process: content must be searched for via the catalogue and then viewed via Ericom’s AccessNow software. The automatic sharing
of content between institutions means that, unlike printed publications, publishers need only deposit their digital output once. Since the Regulations came into force on 6 April 2013, the legal deposit libraries have collected over 550,000 e-books in EPUB and PDF formats, 6.7 million e-journal articles in PDF format and 700 TB of archived websites.

A phased approach to collecting was adopted, starting with e-books, e-journals and digital maps. Following the establishment of a foundation with these formats, the first steps in collecting digital music publications were taken in April 2017 by means of a 2-year pilot project. This was led by the British Library with input from the legal deposit libraries’ Sheet Music Task Group, a cross-institution body that was already in existence to promote a cooperative approach to the collection of legal deposit music publications. The distinctive nature of digital music publications in terms of their format and metadata attracted comparison with the collection of digital maps. However, while there are broad similarities, there was one important difference. Born-digital mapping had been received from the Ordnance Survey as part of a voluntary deposit arrangement from 1998, so the implementation of non-print legal deposit legislation required evolution of systems already in development (Fleet and Hatfield, 2017: 189). By contrast, there had been no significant collecting of digital music publications prior to the pilot project. The handful that had been acquired had been stored securely but were not accessible to users, and no end-to-end workflows existed for ingest either at scale or in small numbers. There was therefore a limited foundation on which to build, a situation compounded by the fact that no attempts at systematic collecting had been undertaken internationally.

In light of the groundbreaking nature of collecting non-print legal deposit music publications, a preparatory study was undertaken prior to the pilot project by the British Library’s newly appointed Digital Music Curator (Patterson, 2014). As well as surveying the United Kingdom’s digital music publishing landscape, this provided a deeper understanding of the ways in which digital music publications were distinct from e-books. Key differences include the absence of ISBNs from most publications, the use of proprietary music notation file formats as well as PDF for delivery, the presence of added functionality such as playback and transposition and the issuing of publications comprising multiple instrumental and/or vocal parts. These might be delivered as a single file or separately. If delivered separately, the system for collecting would need to recognise the relationship between the parts and present this in a meaningful way to users (Chesser and Ridgewell, 2019: 91).

The findings of the 2014 study paved the way for the pilot project whose principal aims were to design workflows to ingest and provide access to digital music publications and to begin collecting content from two publishers: Music Sales and Faber Music. This involved collecting their back files of digital content published between 2013 and 2018, which comprised over 56,000 publications in PDF format – 43,365 from Music Sales and 13,167 from Faber Music. On the basis of research undertaken by John Rink (2020), it is possible to distinguish between ‘plain’ publications available in PDF format, publications that have been enhanced for performers, and dynamic editions of a scholarly nature. All of the content collected during the pilot project falls squarely into this first category. However, recognising that this was only a part of the United Kingdom’s digital music publication output, an additional strand of the project
undertook research into the publishing landscape, in order to update the findings of the 2014 study (Mazzucco, 2018).

Faber Music and Music Sales were chosen due to their status as major industry players whose volume of publications would allow workflows to be tested at scale. Founded in 1965 as a sister company to the book publisher Faber and Faber, Faber Music is one of the leading independent British publishers of classical, contemporary, media and popular music. Alongside works by Ralph Vaughan Williams, Malcolm Arnold, Gustav Holst and Benjamin Britten, their catalogue features a distinguished list of internationally recognised composers including George Benjamin, Thomas Adès, Oliver Knussen, Jonathan Harvey and Anders Hillborg. Faber Music also handles Carl Davis’s film scores, Andrew Lloyd Webber’s musical *Cats* and the choral and concert works of Howard Goodall. In addition, they publish tutors and educational publications, including those by Fanny Waterman, Paul Harris and Pam Wedgwood.

At the time of planning the project, Music Sales was also a well-established British music publishing and distribution firm. The Music Sales Group owned over 60 international publishing houses and 30 imprints, many of them iconic names in their field, including the UK publishers Chester Music and Novello. Since then, Music Sales has sold its physical and online divisions to Milwaukee-based publisher Hal Leonard and rebranded itself Wise Music Group. As an American publisher, Hal Leonard publications are not eligible for collection under legal deposit. However, Wise Music Group’s London offices house the teams for all of its publishing imprints as well as promotional staff for Chester Music, Novello and Campbell Connelly. Their website also describes the London offices as ‘the control centre for all new media and internet activities’. These administrative arrangements pose challenging questions with respect to what is eligible for collection under legal deposit regulations, both in print and digitally.

For the purposes of this study, this knotty issue has been put to one side in order to concentrate on evaluating the 56,532 publications that were acquired during the pilot project. Behind this statistic lie a range of specialist activities including creating a workflow for collection, preservation and access, liaising with Faber Music and Music Sales to obtain the content and metadata, and transforming the publisher metadata into MARC format. This article does not consider these activities in any detail, instead focusing on the publications themselves. It considers the genres of music that were collected, the relationship between printed and digital output and how the publications are intended to be used. This is achieved by analysis of the back catalogues at scale together with close-readings of individual publications. This provides a deeper understanding of the digital music publishing industry that will help to inform future collecting activity.

**Defining digital music publications**

Before analysing these collections in greater depth, it is useful to take a step back to consider what is meant by the term ‘digital music publication’. Donald Krummel (2001) has commented that a music publisher ‘issues musical editions that consist primarily of musical notation, whether for performance or study’. This points towards the prevalence
of musical notation as a defining characteristic of music publications. In the digital world, this allows us to distinguish digital music publications from e-books about music and digital audio. The latter is particularly important because content that is composed solely of digital sound and/or moving images is not eligible for collection under the Regulations. This mirrors the situation with audiovisual content more generally, which has never been subject to legal deposit, but which has been collected under voluntary arrangements instead. Government guidance published in 2013 clarified that audiovisual content that supports a digital publication, such as practice tracks that accompany notated music, may be collected (Great Britain, Department for Culture, Media and Support, 2013). The collection of ‘incidental’ audiovisual content poses particular challenges to legal deposit libraries, since, in a similar manner to a set of separately published instrumental parts, it needs to be linked to the primary publication and delivered to users alongside the notation.

It is also important to consider what constitutes ‘publication’. George Sturm (2000: 628) has described music publishing as ‘the art of bringing a musical product to a public’. This aligns with the framework for non-print legal deposit, which does not indicate what constitutes a publisher or describe the features of a digital publication, instead defining published items by the act of making content publically available. Public availability is especially pertinent in the world of digital publishing, since it points towards a crucial distinction between digital production and digital publication. In terms of the Regulations, something is only published digitally when it is made available to the public in a digital format. Thus, the existence of a digital file for a music publication does not constitute digital publication unless that file goes on to be made available to the public digitally through a publisher or distributor’s website. By extension, while printed music publications are underpinned by digital files, these files themselves are not digital publications, because digital is not the final publication format.

This distinction is vital in enabling legal deposit libraries and publishers alike to understand what is eligible for deposit. It also has ramifications for the breadth of material that may be collected. Recent research showed that at least 350 music publishers are currently active in the United Kingdom (Mazzucco, 2018: 4). However, many composers make their publications available on publically available websites. Under the terms of the Regulations, this constitutes publishing, thereby vastly expanding the amount of content that is in scope for collection. The ability of composers to self-publish is facilitated by services provided by notation software companies. Avid’s Sibelius software, for example, has a cloud sharing option which enables composers to share their scores online. A similar service is provided by Noteflight Marketplace, a branch of the notation software community Noteflight. This goes hand-in-hand with the movement of music notation software to the cloud. As Kimmy Szeto (2018: 99) has observed, this blurs the distinction between editing, storing and publishing, since every saved version of a composition can be stored online automatically, which, if permission is set to publically accessible, equates to instantaneous publication. There are also dedicated self-publishing sites, including Score Exchange and Composers Edition, and the retailer Sheetmusicplus.com (Sheet music plus, 2020) has opened its website to self-publishing
composers. This points towards an overlap between the acts of publication and distribution that is especially prevalent in the world of digital publishing.

In addition, many UK companies provide access to digital sheet music publications without considering themselves to be publishers. For example, 8notes.com is run and owned by Red Balloon Technology, a music and technology company based in St Albans. Established in 2001, it currently provides access to over 11,000 digital publications. Advertising enables them to provide a large percentage of the content free of charge, alongside costed add-on services including play-along tracks, parts and transpositions. Meanwhile, the London-based Great Scores provides access to over 84,000 digital publications, with the website inviting users to ‘browse sheet music then print on your printer’. This points towards the role of digital music publishing in contributing to ‘print to play’ services (Rink, 2020). This is currently a pervasive feature of the industry which is driving the ways in which musicians engage with digital content. Moreover, while paper-based rehearsals and performances are slowly giving way to those using screen readers, wireless pedals, cloud storage and annotation tools, this is far from widespread, and many musicians still prefer to use physical copies. ‘Print to play’ is not a natural fit with the Regulations. These only allow a portion of a publication to be printed out and used for non-commercial research or private study, largely mirroring fair dealing copying. Just as with printed publications, the ability to obtain a partial copy of a digitally published musical work is likely to be of limited value to users.

Internationally, two well-established providers of digital music publications are the International Music Score Library Project/Petrucci Music Library (IMSLP) and the Choral Public Domain Library (CPDL). These mainly offer scanned published scores in the public domain at no cost. Formed in 2006, IMSLP has come to be known as the ‘Wikipedia of Music’ since it harnesses the power of crowdsourcing by allowing individuals to upload their own digital scores. As of July 2020, it contained over half a million scores. While they may not be the latest scholarly editions, nor may it be easy for users to understand exactly which edition they are accessing, IMSLP is nevertheless extremely popular with performing musicians, particularly amateurs. Subscription-based suppliers also exist that aggregate content from publishers. These include Alexander Street’s Music Online: Classical Scores Library, a source for scores of the classical canon as well as lesser-known contemporary works, and nKoda, whose website describes their content as ‘the largest digital library of sheet music ever created on subscription’. In addition, internationally, a growing number of music publishers offer digital access to their publications at a smaller scale, often through custom-made apps or e-readers. These include Bärenreiter, Henle, Eulenberg, Schott and Edition Peters.

Drawing together these strands, in the context of non-print legal deposit, a digital music publication can be defined as content composed primarily of notated music which is publically available, either freely or with a charge. A variety of file formats may be used, including proprietary music notation file formats and PDF. The term ‘digital music publication’ is preferable to ‘digital sheet music’, ‘electronic score’, ‘digital score’, ‘digitised score’ and numerous variants thereon. These terms have specific meanings or historical associations that are misleading in a non-print legal deposit context. Although ‘digital sheet music’ is the term used in the British Library catalogue, by implying sheets
of paper, it conflates the digital and physical formats. While in practice, many digital music publications are currently designed to be printed, this is not always the case, nor will this practice necessarily persist in the future. In addition, for some audiences, ‘sheet music’ implies popular music (Riley and Dalmau, 2007: 132–133). While the digital format is vital for the distribution of popular music, as we shall see, it is also used for other genres. The terms ‘electronic score’ and ‘digital score’ imply the alignment of parts in score format. This does not apply to all digital music publications, with some comprising scores and parts or instrumental tutors. Meanwhile ‘digitised music’ suggests the production of a digital edition from an analogue source by scanning or photography. This production method is used for some digital music publications and has been particularly important in enabling the publication of an increasing amount of content from manuscripts. However, the focus on the means of production rather than the nature of the output means it is only representative of a subset of the industry.

**Overview of collecting**

The starting point for analysis was a report generated by the British Library containing MARC records for the digital publications collected from Faber Music and Music Sales during the pilot project. This constituted the descriptive metadata for 56,532 publications, in practice roughly equivalent to their digital back catalogues for the period 2013 to 2018. To enable comparison between digital and printed publications, MARC records were also obtained for the material deposited in print by these two publishers for the equivalent period. This resulted in a dataset comprising metadata for a total of 57,202 publications (see Figure 1). At the broadest level, this showed that 99% of publications collected from Faber Music and Music Sales under legal deposit legislation between 2013 and 2018 were digital. The number of printed items published by Music Sales and Faber Music was almost identical. However, Music Sales’ digital publications constituted 76% of the total published output overall, reflecting their significance in the industry more broadly.

|          | Print          | Digital        |
|----------|----------------|----------------|
| Faber Music: | 336            | 13,167         |
| Music Sales: | 334            | 43,365         |
| **Total** | **670**        | **56,532**     |
| **Grand total (all print and digital):** | **57,202**     |                |

**Figure 1.** Overview of publications collected from Faber Music and Music Sales, 2013–2018.

Taken as a whole, the files and metadata had a number of noteworthy features. While no audiovisual material was included, a small number of e-books in PDF format were deposited. This is symptomatic of a broader issue that arises when publishers produce
both digital music publications and e-books: there is currently no automated way of distinguishing between them at the point of ingest. Prior to the pilot project, this problem had already been encountered in the context of Cambridge Scholars Publishing. This publisher was known to the library community as a book rather than a music publisher and has been depositing its output digitally. Alongside e-books, unexpectedly, this has included around 40 digital music publications. These were ingested as e-books, necessitating later updating of their catalogue records. This updating does not arise out of any desire to ‘finesse’ the metadata – although it is worth noting that the metadata requirements for books and music are distinct. At the most basic level, the material type needs to be updated from ‘e-book’ to ‘digital sheet music’ in order to support faceted searching by format. As collecting expands, this is an area that the legal deposit libraries will need to continue to monitor.

In addition, most of the digital music publications did not include dates, meaning that dates of publication in the catalogue records had to be inferred. Moreover, assigning a publication start date of 2013 to the digital back catalogues is somewhat arbitrary and simply reflects the year in which the Regulations came into force. In practice, some of the content might have been published earlier, and the dataset is more accurately viewed as a snapshot of what was available digitally in 2018, the year in which the back files were deposited at the British Library. This meant that it was not possible to compare the printed and digital data by year. It is also worth noting that the metadata that each of the publishers supplied was very different. This made the process of mapping it to MARC more complex, and, when it came to analysis, meant that it was sometimes necessary to use different fields in each publisher’s dataset to answer the same question.

**Genre**

A key question for both users and library professionals is what genres of music are being published digitally. The dataset enables this question to be addressed from two complementary angles: musical style (classical, popular, jazz, film, folk and educational) and performing forces (solo vocal, solo instrumental, choral, chamber and orchestral). Using the subject terms in the metadata for both the printed and digital publications, each publication was classified as either popular, classical, jazz, film, folk or educational. For printed publications, educational works dominated, comprising 37% of output, closely followed by classical (33%). Popular music accounted for 17% of the printed catalogue, film music and jazz were both 6% and folk music was 1%. The picture for digital publications was very different, with popular music accounting for 71% of output. Film was 7%, jazz 6% and folk and educational were 3% each. Overall, this suggests that the digital format is primarily being used for the publication of popular music, whereas print is more widely used for educational and classical publications. At the same time, it should be noted that the sheer volume of digital publications means that they make a significant contribution to all musical styles.

A surprising finding was that classical music accounted for 10% of digital output. This was higher than expected and can be explained in part by the appearance of popular classical repertoire in multiple editions. Handel’s ‘Water music’, for example, appears
no fewer than 13 times in the digital publications deposited by Music Sales. To further understand this trend, the composers and artists principally responsible for digital output were analysed using data in the MARC 100 field (main entry, personal name) for the Faber Music publications and the 720 field (added entry, uncontrolled name) for the Music Sales publications. Sixty-seven different classical composers were found in the Faber digital catalogue and 205 in the Music Sales digital catalogue. Of these, the most popular across both catalogues were J.S. Bach, Beethoven and Mozart. There were also a small number of contemporary classical composers. Peter Maxwell Davies, Thea Musgrave, Giles Swayne and Judith Weir, for example, were all represented in the Music Sales catalogue, as were composers from the younger generation, such as Helen Grime and Philip Venables.

To further refine our understanding of genre, where possible in the context of the metadata supplied, every publication was classified as choral, solo vocal, solo instrumental, chamber music or orchestral/band (the latter category also including opera and oratorio). Classes were assigned on the basis of data in the MARC 250 field (edition statement) for the digital publications and the 655 field (index term for genre/form) for the printed publications. Solo vocal works dominated the digital publications, with 61% of publications from the Faber Music and Music Sales digital catalogues involving solo voice and piano and/or guitar, compared with 24% in the printed catalogues. This aligns with the dominance of popular music in digital output. Conversely, chamber and orchestral music formed only 1% of the digital deposit, compared to 51% of print. In terms of solo instrumental works, music for piano dominates both formats but was particularly prevalent digitally, accounting for 15% of total output. Both formats included only a small number of choral works (8% in the digital catalogue and 2% in the printed catalogue).

**Overlap between print and digital publishing**

Another important question for both legal deposit libraries and publishers is the extent to which material published digitally and in print overlaps. The Regulations state that any content that is published identically in printed and digital formats need only be collected once, in the best format for preservation. The savings to publishers are attractive, and, as a result, many book publishers have switched to digital deposit. Several music publishers have also discussed switching to digital deposit with the British Library. However, to date, this has not been appropriate, since, while the content was produced digitally, it was only made available to the public in print. This confusion between digital production and digital publication was reinforced by the publisher survey undertaken during the pilot project, with one firm commenting that ‘everything we have published and ever will publish exists in a digital format’ (Mazzucco, 2018: 22). Furthermore, even when music publishers are identified that are eligible to switch from print to digital deposit, the long-term preservation needs of any publications in proprietary notation file formats will need careful consideration.

The content acquired during the pilot project provides an ideal test case for exploring the overlap between print and digital publishing. Analysis of the MARC 245 fields (title
and statement of responsibility) showed that only 30 printed publications exhibited a close match with the digital catalogues. This suggests that the overlap between content being published in print and digitally is less than 0.1\%. Moreover, given that this finding is based on the metadata rather than comparing the physical printed publications to their potential digital matches, the degree of overlap could be smaller or even non-existent. In addition, even if the publications were identical, due to the continuing presence of unique print and unique digital publications, any switch to digital deposit would be partial, resulting in a ‘mixed deposit’ model. This would be likely to be administratively more burdensome to both libraries and publishers than maintaining the status quo. This is an important finding because it indicates that, at present, there is very little potential to switch music publishers from print to digital deposit. This is reinforced by the survey undertaken as part of the pilot project, which indicated that, while 53\% of the industry is now publishing content digitally, only 14\% of that content is exclusively digital (Mazzucco, 2018: 4).

Faber Music’s *Mastering the piano with Lang Lang* was excluded when calculating the degree of overlap but is nevertheless a useful case study. It appears in five volumes in print but has also been published separately as an app produced in partnership with Tido Music. This incorporates interactive music, audio playback and coaching videos. This is one of a growing number of examples of printed music publications that are also produced in enhanced versions digitally. For example, Bärenreiter, TouchPress and Deutsche Grammophon have collaborated to produce an app version of Beethoven’s Ninth Symphony that allows users to switch between four recordings of the work, including one with video, that are synched with different versions of the score. Due to its complexity, music delivered via bespoke apps was not collected during the music pilot project. However, it was considered more generally as part of a parallel project on emerging formats (Smith and Cooke, 2017).

In addition, although there was very little identical replication of material published in print and digitally, closer analysis indicates a marked trend to republish extracts from printed publications digitally. Forty-four per cent of Faber Music’s digital publications have an ‘also published in’ note in their catalogue records, pointing customers to printed volumes containing that particular item as part of a compilation. An example is ‘Another day of sun’ from *La La Land*. The Faber Music deposit included three digital editions of this work: easy guitar (British Library DRT ELD.DS.DSM.6119), piano/vocal (DRT ELD.DS.DSM.6117) and easy piano (DRT ELD.DS.DSM.12471). In addition, Faber Music published two printed compilations containing this song: an easy piano version entitled *La La Land: music from the motion picture soundtrack* (British Library G.1513.h.(2.)) and a vocal score, *La La Land: music from the motion picture soundtrack* (G.1216.u.(4.)). Although this trend had already been identified in the 2014 study, the content collected during the pilot project indicates just how prevalent it is.

Sometimes digital reissuing does not involve new typesetting, instead reproducing content directly from printed publications. For example, Faber Music’s digital catalogue contains two editions of Gustav Holst’s ‘I vow to thee my country’ (DRT ELD.DS.DSM.7708 and DRT ELD.DS.DSM.8111). These bear the notes ‘Also published in: Classic FM: Play The Classics’ and ‘Also published in: It’s Never Too Late
To Play Classics. The digital versions contain page numbers that reflect their printed sources. Significantly, the typesetting in both digital editions is exactly the same, including layout, dynamics, fingering and other markings. This indicates a reuse of content that originated in printed publications, which is then carried over to the digital format. This suggests that, just like early digital editions of texts, some digital music publications can be considered ‘spill-overs from the print medium’ (Gabler, 2010: 48).

Duplicates and near duplicates
The appearance of Holst’s ‘I vow to thee my country’ twice in Faber Music’s digital catalogue is symptomatic of a wider trend. Analysis of the MARC 245 fields indicates that as much as 31% of the content appears twice or more in the Faber Music digital catalogue, and 56% in the Music Sales catalogue. Although it is most common for

![Diagram](image-url)

**Figure 2.** (a) Duplication of Faber Music digital content based on close match on MARC 245 field. (b) Duplication of Music Sales digital content based on close match on MARC 245 field.
publications to appear twice, as Figures 2(a) and (b) show, many publications appear much more frequently than this. Although some content appears multiple times in the printed catalogues, this happens much less frequently, with 48 of the Faber Music and 35 of the Music Sales printed publications exhibiting very close matches on the 245 field. Moreover, closer inspection of the printed “duplicates” indicates that these were usually different publications rather than identical reprints, often as a result of the publication of the same repertoire in different arrangements or score formats.

Faber Music’s most frequently occurring digital publications were ‘Over the rainbow’, ‘The bells of St Mary’s’ and ‘Rudolph the red nosed reindeer’ (all appearing 12 times), while for Music Sales it was ‘Silent night’ (38 times), ‘Mamma Mia’ (34 times) and ‘Jingle bells’ (31 times). As indicated by Figures 3(a) and (b), some of the top 10 digital publications are common to both publishers. The tendency of both firms to print the same works is borne out by closer analysis. The second movement from Dvořák’s Ninth Symphony opus 95 in E minor (the ‘New World’ Symphony) has been popularised though its use in the Hovis ‘Boy on the bike’ advertisement. First aired in 1973, this was relaunched in 2019 and has since been voted the most iconic UK advert of all time. It was collected no fewer than 21 times in the pilot project, with 12 publications issued by Music Sales and 9 by Faber Music. These comprise seven versions for piano, three for clarinet solo, one for clarinet and piano, two for alto saxophone, one for tenor saxophone, three for flute solo, one for melody line and chords, one for solo violin, one for violin and piano and one for cello and piano. This points towards the existence of repertoire that is in sufficiently high demand for digital publication by multiple firms to be economically viable. By contrast, the overlap in content published in print is more limited. Thus, while identical repertoire is published by multiple firms in print and digitally, it is much more prevalent in the world of digital publications. This can be accounted for by the cheaper production and distribution methods and the ease with which short, popular works can be made available as stand-alone digital publications.

Closer analysis of the digital publication of multiple works with the same title indicates that only a small proportion are actually duplicates – that is, identical reprints of another publication. The two occurrences of Erik Satie’s Gnossienne number 1 in the Faber Music catalogue (DRT ELD.DS.DSM.5618 and DRT ELD.DS.DSM.8344) are useful examples in this respect. The typesetting of the musical notation is identical, but, since the header title and footer are different, they are not strictly duplicates, but rather ‘near duplicates’ – that is publications that appear to be identical on the basis of the data in the 245 field, but where closer inspection of other metadata fields and/or the publications themselves show that they are in fact different. The widespread occurrence of near duplicates in both publishers’ digital back catalogues is an important feature for legal deposit libraries to be aware of, since they have the potential to confuse users, and also point towards the need for the presence of sufficiently detailed metadata to allow them to be distinguished.

Near duplicates occur for a number of reasons. One of the most prevalent is the presence of the same repertoire in multiple arrangements for different instruments. For example, Faber Music published five digital editions of Queen’s ‘I want it all’, of which three are for piano, voice and guitar (DRT ELD.DS.DSM.936, DRT ELD.DS.DSM.940
### (a)

| Title                                              | Number of Faber Music publications | Number of Music Sales publications |
|----------------------------------------------------|------------------------------------|------------------------------------|
| 1 Over the rainbow [Harold Arlen; E.Y. Harburg]     | 12                                 | 20                                 |
| 2 The bells of St Mary’s [A. Emmett Adams; Douglas Furber] | 12                                 | 0                                  |
| 3 Rudolph the red nosed reindeer [Johnny Marks]     | 10                                 | 1                                  |
| 4 Smells like teen spirit [Nirvana]                 | 10                                 | 3                                  |
| 5 Summer nights (from ‘Grease’) [Warren Casey; Jim Jacobs] | 10                                 | 0                                  |
| 6 Blueberry Hill [Vincent Rose; Larry Stock]        | 9                                  | 10                                 |
| 7 My baby just cares for me [Walter Donaldson; Gus Kahn] | 9                                  | 4                                  |
| 8 You’re beautiful [James Blunt]                    | 9                                  | 5                                  |
| 9 I’ve got you under my skin [Cole Porter]          | 8                                  | 2                                  |
| 10 Let’s face the music and dance [Irving Berlin]    | 8                                  | 2                                  |

### (b)

| Title                                              | Number of Music Sales publications | Number of Faber Music publications |
|----------------------------------------------------|------------------------------------|------------------------------------|
| 1 Silent night [Franz Gruber]                       | 38                                 | 4                                  |
| 2 Mamma Mia [ABBA]                                 | 34                                 | 0                                  |
| 3 Jingle bells [J. Pierpoint]                       | 31                                 | 2                                  |
| 4 Hark! The herald angels sing [Felix Mendelssohn; Charles Wesley] | 29                                 | 2                                  |
| 5 Chiquitata [ABBA]                                | 28                                 | 0                                  |
| 6 Stand by me [Ben E. King, Jerry Leiber and Mike Stoller] | 28                                 | 2                                  |
| 7 Away in a manger [William J. Kirkpatrick]        | 27                                 | 2                                  |
| 8 Dancing queen [ABBA]                             | 27                                 | 0                                  |
| 9 I have a dream [ABBA]                            | 27                                 | 0                                  |
| 10 Walking in the air (theme from ‘The Snowman’) [Howard Blake] | 27                                 | 2                                  |

**Figure 3.** (a) Faber Music most popular digital publications. (b) Music Sales most popular digital publications.
and DRT ELD.DS.DSM.941), one is for guitar tab (DRT ELD.DS.DSM.938) and one is for voice/piano reduction (DRT ELD.DS.DSM.7896). Distinguishing between these requires the user to scrutinise the metadata closely, and even then, it is often necessary to open the files to check for differences. This is likely to be frustrating for users. Although technically ambitious, it is worth noting that this could be avoided if there was an option to preview publications from the catalogue before viewing them in full via Ericom.

The presence of the same works in different keys can also result in near duplicates, and again is often only apparent to the user if the metadata is scrutinised. This is sometimes to allow a work to be performed by a greater range of instruments or singers. When key signatures with multiple sharps or flats are eliminated, it can also make works more accessible to beginners. This is also achieved by producing new editions with simplified rhythms. For example, Felix Mendelssohn’s ‘O for the wings of a dove’ appears twice in the Faber Music catalogue. One version is scored for piano and eliminates the triplets from the original (DRT ELD.DS.DSM.7457). The other is arranged for piano and guitar and is faithful to the original rhythm (DRT ELD.DS.DSM.7932).

Another reason for the appearance of near duplicates is the separate publication of instrumental parts. Sometimes, the parts are described in the 245 field, a feature that will assist users in distinguishing between similar-looking publications. This is the case with Faber Music’s *Flute basics repertoire unit 1* by Sally Adams, where the 245 field indicates that the publication at DRT ELD.DS.DSM.9228 is the flute part, while that at DRT ELD.DS.DSM.9229 is the piano part. However, it is more common for the 245 fields to be identical. Brahms’ ‘Hungarian dance’ number 4, arranged by Mary Cohen, for example, appears twice with an identical 245 field in the Faber Music deposit. It is only by looking closely at the edition statement and medium of performance fields that it becomes apparent that the publication at DRT ELD.DS.DSM.4718 is the score and part, while DRT ELD.DS.DSM.4717 is just the violin part.

The collection of music in multiple parts had been identified as a potential challenge before the pilot project started. In reality, the material collected was not as complex as anticipated, since in every case, there was a one-to-one rather than a many-to-one relationship between parts and metadata. This meant that it was not necessary to link multiple digital files to one catalogue record. However, some degree of cross-referencing between catalogue records describing separate parts from the same work is a refinement that would be beneficial to users. In addition, wider analysis indicates that the one-to-one relationship between metadata and files is not representative of the industry as whole, and, in addition, that associated sound files are likely to be collected. These are issues that will need to be addressed in the future.

Another factor that contributes to the presence of near duplicates is the trend to publish extracts from larger works separately. Those from compilations tend to have distinctive titles, and therefore do not usually look like duplicates. However, there are others where movements from a stand-alone work are published separately, often with very similar 245 fields. Examples include Music Sales’ separate publication of each of Elgar’s *Five piano improvisations* and Lennox Berkeley’s *Five short piano pieces*. Similarly, concerto movements are published separately, such as Vivaldi’s Concerto for lute, two violins and basso continuo in D major, RV 93. This includes the famous ‘Largo’
that has been used widely in popular culture, including in the television series ‘The Simpsons’ and the 2019 film of Louisa May Alcott’s book *Little women*. It is possible to purchase each of the three movements from this concerto separately from Music Sales. This is part of a wider trend to publish popular extracts from larger classical works. For example, while Music Sales do not publish a complete digital vocal score of Handel’s *Messiah*, it is possible to purchase digital extracts including ‘Evry valley’, ‘I know that my redeemer liveth’ and the ‘Hallelujah chorus’.

**Use**

The material collected during the pilot project also provides an insight into the target audience for digital music publications. This is important for legal deposit libraries since it underpins a deeper understanding of how the publications are intended to be used, which in turn can guide technical, service delivery and metadata requirements. The data on length of publications found in the physical description field (MARC 300) are particularly relevant in identifying potential user groups. This was ‘normalised’ to produce a total figure for pagination for each publication. For example, a publication described as ‘1 score (16 pages)’ was normalised to a length of 16 pages, while a publication described as ‘1 score (16 pages) + 1 part (7 pages)’ was normalised to 23 pages. Any publications that were recorded as ‘unpaginated’ or which lacked pagination for the parts were recorded as ‘pagination not identified’. The normalised paginations indicated that the vast majority of digital publications had a length of 9 pages or less, while most printed publications had a length of 20 pages or more (see Figure 4). This was consistent across both publishers. The fact that digital publications are shorter suggests that they are unlikely to include substantial prefatory or editorial matter, therefore pointing towards their intended use for amateur performance rather than professional performance or academic study.

To further aid understanding of the intended users of digital music publications, the edition statement (MARC 250 field) was analysed to identify terminology relating to musical format (full score, vocal score, instrumental parts, guitar tab etc.). Due to the wide variation in terminology, it was not possible to map the whole dataset to a common set of format terms. Instead, two different sets of format terms were defined, one for the printed publications and one for the digital publications. For printed publications, they were score (no further specification), full score, vocal score, score and part(s), playing score, set of parts and study score. For digital publications, they were guitar tab, lyrics only, instrumental part, drum chart and score. Any format terminology in the 250 field was then aligned with one of these categories. The results are shown in Figures 5(a) and (b).

Unfortunately, only 15% of the printed publications and 9% of the digital publications contained terminology in the 250 field that could be mapped to a format category. Therefore, the findings in Figure 5 are not conclusive. Nevertheless, they are indicative of some broad trends. In terms of the printed publications, the formats of vocal score, score and part, playing score and set of parts comprise 20% of the output and could definitely be used for performance. It is therefore safe to conclude that at
Figure 4. Length of publications.
least 20% of the content was intended for performance. In terms of the digital publications, guitar tab, lyrics only, instrumental part and drum chart are all highly suggestive of performance. These comprise 99% of the content that could be analysed. Considered alongside their shorter length and the absence of critical material, this suggests that, in terms of the material deposited, the digital publications were overwhelmingly intended for use by amateur performers, while the target audience of the printed publications was more mixed.

The emphasis on performers as a target audience is mirrored by the digital music publishing industry more broadly. Faber Music’s *Mastering the piano with Lang Lang* app blends the provision of repertoire for practice with aids to learning. Meanwhile, the partnership between Edition Peters and London-based Tido Music brings piano music to life by drawing together notated music, recordings, practice tools, masterclasses and written commentary. This content is aimed at more advanced musicians, a trend also apparent in industry innovations outside the United Kingdom. The German music publisher Henle’s app, for example, combines authoritative editions with features such as the ability to modify score layout, switch between different fingerings and annotate.

The collection of digital music publications under legal deposit legislation must balance long-term preservation, future use and current user needs. While preservation and future use have been partially addressed through the act of collecting, much less attention has been given to the usefulness of the content to current users.

| Format term                  | % of publications |
|------------------------------|-------------------|
| Score (no further specification) | 44                |
| Full score                   | 34                |
| Vocal score                  | 10                |
| Score and part(s)            | 7                 |
| Playing score                | 3                 |
| Set of parts                 | 1                 |
| Study score                  | 1                 |

| Format term                  | % of publications |
|------------------------------|-------------------|
| Guitar tab                   | 95                |
| Lyrics only                  | 2                 |
| Instrumental part            | 1                 |
| Drum chart                   | 1                 |
| Score                        | 1                 |

Figure 5. (a) Format of printed publications (both publishers). (b) Format of digital publications (both publishers).
Research has shown that digital music users want to access material on their handheld devices and that they are likely to engage with it on a very personal level through annotation (Hooper, 2013: 573). The ability of the content collected during the pilot project to meet current users’ needs is curtailed by printing and downloading restrictions and the necessity of accessing content on library terminals. This renders it more useful for theoretical study rather than practical use. Moreover, even this type of use would be difficult in reading rooms, since Ericom’s AccessNow software does not currently provide adequately for the need to navigate large, complex scores, link parts, play associated audio or display more than one publication at once to facilitate comparison.

To a certain extent, these issues reflect usability challenges of non-print legal deposit content more generally. The Digital Library Futures project (2017–2019) looked at the use of non-print legal deposit content in academic legal deposit libraries (Gooding et al., 2019). It concluded that the impact of the Regulations had been positive in terms of the provision of access to digital collections in unprecedented breadth and depth, and in increasingly varied formats. However, the imposition of access restrictions that mirror print legal deposit collections were found to be at odds with users’ increasing reliance on personal devices, specialist software and remote access, therefore failing to meet their information-seeking behaviour. When it comes to digital music publications, the specific needs of music users make the disconnect between users’ needs and access arrangements even more pronounced.

Conclusions

By analysing the 56,532 digital music publications collected from Music Sales and Faber Music during the pilot project and contextualising them with the printed content deposited by the same publishers, the increasingly prevalent role of digital publication in the United Kingdom’s music publishing industry is evident. Ninety-nine per cent of the content collected from these firms during the period 2013 to 2018 was in digital format. All of the publications were flat PDFs, in every instance there was a one-to-one relationship between metadata and the corresponding digital file, and no incidental audio-visual content was deposited. Although 71% of content was popular music, given the sheer volume of publications, the digital format is also a significant medium for the distribution of music in other genres, including classical music. While the classical digital publications were dominated by famous, canonical repertoire, it was also used for the publication of works by contemporary composers.

The separate publication of extracts from printed works digitally, together with the ‘near duplication’ of a large proportion of the digital content through the production of variant editions, contributed greatly to the volume of output. However, despite the tendency to reprint extracts from printed publications digitally, the content collected during the pilot project illustrated that there is currently little potential to switch UK music publishers from print to digital deposit. This is because very few publications appear in both formats in identical editions. It was also apparent that, while the bulk of the digital
music publications collected were intended for performance rather than academic study, current access arrangements fall very short of the needs of performing musicians.

These observations can be used to guide future collecting activity. The purpose of legal deposit is to ensure that the UK’s cultural heritage is acquired, preserved and accessible for future generations (Gibby and Brazier, 2012: 362). In order to achieve this, future collecting must embrace the breadth of digital music being published. This will require legal deposit libraries to grapple with issues including the collection of publications with accompanying audiovisual content, those issued in multiple parts where there is a many-to-one relationship between catalogue record and digital files, content in proprietary music notation file formats and interactive content delivered via apps. Future collecting may also include digital critical editions – that is, multidimensional combinations of digitised and enriched source materials, including score facsimiles, video and audio recordings and encodings (Wiering, 2009: 24). This suggests the need to continue to work cooperatively with collecting activity in the field of emerging formats.

Furthermore, to reflect fully the breadth of the UK music publishing landscape, it will be necessary to devise workflows that can accommodate music that is self-published and publishers who produce only a handful of digital music publications each year. Extension of the existing publisher submission portal, or development of a specific one for digital music publications, will be essential in achieving this British Library (2015). This will enable firms to upload publications themselves with relevant descriptive metadata, thereby avoiding the need to produce bespoke workflows forevery publisher. It is vital that such a portal is able to flag the format of digital music publications at the point of ingest, in order to distinguish them from e-books. This will facilitate faceted format searching in the catalogue and pave the way for automated metadata updates.

In order to ensure that content collected is useful to current as well as future users, it will also be important to address the gaps between the needs of users accessing digital music publications and the capabilities of Ericom’s AccessNow software. Part of the solution will be technical, but equally important will be a reframing of the Regulations to ensure that they align with the opportunities offered by digital humanities, rather than being based on the arrangements for printed publications. This will require government lobbying and constructive negotiations with publishers to see that the needs of all parties are met. In addition, ensuring that the delivery of non-print legal deposit content meets users’ needs goes hand-in-hand with metadata enhancements, particularly with reference to distinguishing between publications with similar titles.

Moreover, it is noteworthy that a key part of the pilot project was the production of detailed mappings of the publisher metadata to MARC format. This is a very labour intensive process that will be unsustainable with current staffing levels as new publishers are added, and existing publishers make changes to the structure of their metadata. This problem is compounded by highly variable metadata standards amongst UK music publishers. Distributors like Amazon are not widely used, and identifiers such as ISBNs and ISMNs are not universally applied to content. There is no equivalent to the book industry standard ONIX for music publishing, and more broadly, the concept of metadata standards is unfamiliar to many firms.
This points towards the need for legal deposit libraries to work even more closely with those involved in the music publishing industry. As well as addressing metadata requirements, this will involve ascertaining which content is in scope for deposit and continuing to clarify the criteria that need to be fulfilled in order for publishers to switch from print to digital deposit. Given their importance in the digital music publishing industry, it will be particularly important to increase dialogue with publishers of popular music and composers who are self-publishing. In addition, in order to create efficient workflows that are robust enough to cope with the sheer volume of material that is in scope for collection, it will be necessary to cooperate with aggregators such as nKoda as well as distributors and software providers. The United Kingdom’s digital music publishing industry is a complex and fragmented one, but also one that makes unique and important contributions to the nation’s cultural heritage. Building these relationships will enable legal deposit libraries to undertake collecting that reflects the breadth of production and that is also sustainable.

Declaration of conflicting interests
The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding
The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The author received financial support for the research, authorship and publication of this article in the form of a British Library Coleridge Research Fellowship.

ORCID iD
Amelie Roper https://orcid.org/0000-0003-3239-2208

Note
1. The dataset that supports the findings of this study is available in the British Library’s Shared Research Repository at https://doi.org/10.23636/1187 under a CC0 1.0 Universal Public Domain licence. The data are also available on the British Library’s online catalogue: http://catalogue.bl.uk.

References
8notes.com: Free Sheet Music, Riffs, Lessons and Tools for Musicians Who Play. Available at: https://www.8notes.com (accessed 17 July 2020).

Beethoven’s 9th Symphony as an App from Deutsche Grammophon and Touch Press in the Bärenreiter Edition. Available at: https://www.baerenreiter.com/en/catalogue/digital-media/app-beethovens-ninth/ (accessed 17 July 2020).

Boorman S, Selfridge-Field E and Krummel D (2001) Printing and Publishing of Music, Grove Music Online. Oxford: Oxford University Press.

British Library (2015) Publisher submission portal. Available at: https://www.youtube.com/watch?v=Er3cbIRYrwY&feature=youtu.be (accessed 17 July 2020).
Chesser R and Ridgewell R (2019) Non-print legal deposit and music in the UK: a progress report. *Brio* 56(2): 87–95.

Choral Public Domain Library. Available at: http://www2.cpdl.org/ (accessed 17 July 2020).

Composers Edition. Available at: https://composersedition.com/ (accessed 17 July 2020).

Fleet C and Hatfield P (2017) Map collecting in the digital age: implementing non-print legal deposit in the UK Legal Deposit Libraries. *Alexandria: The Journal of National and International Library and Information Issues* 27(3): 188–197.

Gabler H (2010) Theorizing the digital scholarly edition. *Literature Compass* 7(2): 43–56.

Gibby R and Brazier C (2012) Observations on the development of non-print legal deposit in the UK. *Library Review* 61(5): 362–377.

Gooding P, Terras M and Berube L (2019) Towards user-centric evaluation of UK non-print legal deposit: a Digital Library Futures White Paper (Glasgow, Edinburgh and Norwich).

Great Britain, Department for Culture, Media and Sport (2013) Guidance on the Legal Deposit Libraries (Non-Print Works) Regulations. Available at: https://www.gov.uk/government/publications/guidance-on-the-legal-deposit-libraries-non-print-works-regulations-2013 (accessed 17 July 2020).

Great Britain, Legal Deposit Libraries Act (2003) Available at: http://www.legislation.gov.uk/ukpga/2003/28/contents (accessed 17 July 2020).

Great Britain, Legal Deposit Libraries (Non-Print Works) Regulations (2013) (SI 2013/777). Available at: http://www.legislation.gov.uk/ukdsi/2013/9780111533703 (accessed 17 July 2020).

Great Scores: Browse and Play Sheet Music Then Print on Your Printer. Available at: https://www.greatscores.com/ (accessed 17 July 2020).

Hooper L (2013) Contemplating e-scores: open ruminations on the e-score, the patron, the library, and the publisher. In: *Proceedings of the Charleston Library Conference*, pp. 571–575. West Lafayette, IN: Purdue University. Available at: http://dx.doi.org/10.5703/1288284315326 (accessed 17 July 2020).

IMSLP/Petrucci Music Library: Sharing the World’s Public Domain Music. Available at: https://imslp.org/ (accessed 17 July 2020).

King A (1971) The significance of John Rastell in early music printing. *The Library*, 5th Series, 26: 197–214.

King A (1975) An English broadside of the 1520s. In: Sternfeld F and Fortune N (eds) *Essays on Opera and English Music in Honour of Sir Jack Westrup*. Oxford: Basil Blackwell, pp. 19–24.

Mazzucco E (2018) *Digital Sheet Music Publishing in the United Kingdom: Implementation of Non Print Legal Deposit Regulations* [internal report]. London: British Library.

Mazzucco E (2018) *Music Online: Classical Scores Library (Alexander Street)*. Available at: https://alexanderstreet.com/products/music-online-classical-scores-library (accessed 17 July 2020).

nKoda: *Digital Sheet Music Library*. Available at: https://www.nkoda.com/ (accessed 17 July 2020).

Patterson A (2014) *Acquiring, Ingesting and Providing Access to Digital Sheet Music at the British Library* [internal report]. London: British Library.

Riley J and Dalmau M (2007) The IN Harmony Project: developing a flexible metadata model for the description and discovery of Sheet Music. *The Electronic Library* 25(2): 132–147.

Rink J (2020) Digital editions and the creative work of the performer. *Nineteenth-Century Music Review*. DOI:10.1017/S1479409819000673 (accessed 17 July 2020).
Ryan H and Sampson W (2018) *The No-Nonsense Guide to Born-digital Content*. London: Facet Publishing.

*Score Exchange*. Available at: http://www.scoreexchange.com (accessed 17 July 2020).

*Sheet Music Plus: World’s Largest Sheet Music Selection*. Available at: https://www.sheetmusicplus.com/ (accessed 17 July 2020).

Smith C and Cooke I (2017) Emerging formats: complex digital media and its impact on the UK Legal Deposit Libraries. *Alexandria: The Journal of National and International Library and Information Issues* 27(3): 175–187.

Sturm G (2000) Music publishing. *Notes* 56(3): 628–634.

Szeto K (2018) Keeping score, digitally. *Music Reference Service Quarterly* 21(2): 98–100.

Wiering F (2009) Digital critical editions of music: a multidimensional model. In: Crawford T and Gibson L (eds) *Modern Methods for Musicology: Prospects, Proposal and Realities*. London: Ashgate, pp. 23–46.

**Author biography**

Amelie Roper joined Cambridge University Library as Research Facilitator in November 2019. Prior to this, she was Digital Music Curator, and then Research Development Manager, at the British Library. Her research interests focus on music printing and publishing from the 16th century to the present day.