ISNI and author identification in the UK publishing supply chain: a case study

Andrew MacEwan and Emma Rogoz
The British Library
ISNI Webinar for Finnish publishers
11th November 2020

The British Library and ISNI

- Member of ISO 27729 International Standard Name Identifier Committee to draft the standard
- Founding Member of ISNI with the Bibliothèque nationale de France, co-representing the Conference of European National Librarians (CENL)
- Provide QA services to ISNI-IA for build and maintenance of the ISNI database
- Foundational work with data from national library authority files
  - Virtual International Authority File (VIAF)

www.bl.uk
The British Library as ISNI Registration Agency

- Strategic shift from name disambiguation in the British National Bibliography (BNB) and in our catalogues
- Towards
  - Bridging data silos
  - Identity management for the supply chain
  - Exploiting the potential of the BNB and Cataloguing-in-Publication (CIP) workflows
  - Developing an ISNI Registration Service

The British Library as ISNI Registration Agency

- Principles
  - Embed ISNI in all cataloguing workflows
  - Automate processes as far as possible
  - Engage with the supply chain
The British Library as ISNI Registration Agency

• Challenges

• Embedding ISNI in the LC/NACO file
  – Collaboration with Library of Congress and PCC network

• Embedding ISNI in legacy bibliographic data
  – Developing matching algorithms and processes

• Building a Registration Service
  – Online User Portal

• Engagement with the supply chain
  – UK Publishers Interest Group

BL ISNI engagement activities

• Publisher Licensing Services hosted meeting on ISNI for Publishers & Imprints January 2018

• BIC Breakfast March 2018: International Identification for Authors, Publishers & Imprints
  – Included a presentation by the Bibliothèque nationale de France on the integration of ISNI in their CIP workflows
BL ISNI engagement activities

- BL hosted a series of follow up meetings with a Publisher Interest Group to explore options for:
  - Using ISNI to identify Imprints and Publishers across the supply chain
  - Explore ways for sharing existing ISNIs for authors in publisher and trade supply chain metadata
  - EDItEUR has provided valued support in cultivating publisher contacts
  - Bibliographic Data Services also a key publishers’ liaison as the BL’s CIP provider

ISNI Publisher Interest Group

Bibliographic Data Services  Harper Collins
BIC                  ISNI International Agency
British Library       Nielsen
Cambridge University Press  Pan Macmillan
EDItEUR                      Penguin/Random House
Hachette UK            Publisher Licensing Services
International ISBN Agency  Bloomsbury
And more joining…
ISNI Publisher Interest Group

- Use cases for ISNI for publishers/imprints
- Options for supplying author’s ISNIs for products
- Matching ISNIs to in-house publisher IDs & NACO IDs
- Formats for exchanging data
- How we might introduce ISNIs into the supply chain

Embedding ISNIs in the British National Bibliography

- BNB ISBNs matched against VIAF
- NACO names strings in VIAF correspond to name strings in BNB
- VIAF record cross referenced to pull the appropriate ISNI into BNB
Using the BNB to map publisher data via VIAF to ISNI

Graph Matching

- Publisher provides a name string and associated product(s)
- Equivalent ISBNs matched against BNB
- ISBNs matched against VIAF
- VIAF name strings compared to the original name to find matches (green line: similar name strings; red line: name strings differ)
- VIAF record cross referenced to the appropriate ISNI

Publisher ID to ISNI

Example Match Results

<table>
<thead>
<tr>
<th>ISBN</th>
<th>Title</th>
<th>HC/Reverse NACO</th>
<th>ISNI</th>
<th>HUK Proprietary ONIX ID</th>
<th>Person ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>9780297608134</td>
<td>The Beatles Lyrics</td>
<td>Hunter Davies</td>
<td>0000000108900319</td>
<td>31803</td>
<td>31803</td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Beatles lyrics : the unseen story behind the Hunter Davies 1936-</td>
<td>J W Ironmonger</td>
<td>000000397945825</td>
<td>34210</td>
<td>34210</td>
</tr>
<tr>
<td>9780297608226</td>
<td>Not forgetting the whale</td>
<td>John Ironmonger</td>
<td>000000397945825</td>
<td>33323</td>
<td>33323</td>
</tr>
<tr>
<td>9780297608227</td>
<td>The tipping point</td>
<td>J G Jurado</td>
<td>000000109195287</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608493</td>
<td>Woman of the dead : a thriller</td>
<td>Anthea Bell</td>
<td>000000007848681</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608493</td>
<td>Woman of the Dead : a thriller</td>
<td>Bernhard Aicher 1972-</td>
<td>0000001022790646</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608561</td>
<td>My life with Wagner</td>
<td>Anthea Bell</td>
<td>000000007848681</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608561</td>
<td>My life with Wagner</td>
<td>Christian Thielemann 1959-</td>
<td>00000011700249K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608721</td>
<td>White Mountain</td>
<td>Robert Twigger 1964-</td>
<td>000000114614538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608721</td>
<td>White mountain</td>
<td>Robert Twigger 1964-</td>
<td>000000114614538</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608752</td>
<td>Summer Requiem : a book of poems</td>
<td>Vikram Seth</td>
<td>0000000193972921</td>
<td>31988</td>
<td>31988</td>
</tr>
<tr>
<td>9780297608752</td>
<td>Summer Requiem : a book of poems</td>
<td>Vikram Seth</td>
<td>0000000193972921</td>
<td>31988</td>
<td>31988</td>
</tr>
<tr>
<td>9780297608895</td>
<td>The Second Half</td>
<td>Roddy Doyle</td>
<td>0000000108749980</td>
<td>56005</td>
<td>56005</td>
</tr>
<tr>
<td>9780297608895</td>
<td>The second half</td>
<td>Roy Keane</td>
<td>0000000045702389</td>
<td>132040</td>
<td>132040</td>
</tr>
<tr>
<td>9780297609346</td>
<td>Chicken &amp; egg</td>
<td>Dave Myers 1967-</td>
<td>0000000108321958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297609346</td>
<td>The Hairy Bikers’ Chicken &amp; Egg</td>
<td>Hairy Bikers</td>
<td>0000000108321958</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297609346</td>
<td>Chicken &amp; egg</td>
<td>Si King</td>
<td>0000000383954661</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ISBN</td>
<td>Title</td>
<td>HC/Reverse NACO</td>
<td>Original NACO</td>
<td>ISNI</td>
<td>Publisher ID to ISNI Matching Results</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------------------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>--------------</td>
<td>---------------------------------------</td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Beatles lyrics; the unseen</td>
<td>0000000011 -</td>
<td>0000000122</td>
<td>0000000122770055</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Not Forgetting the whale</td>
<td>9780297859147</td>
<td>9780297856269</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Tipping Point</td>
<td>0000000021 -</td>
<td>0000000117</td>
<td>000000011700249X</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Woman of the dead; a thriller</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>0000000108749380</td>
<td>0000000000000000014</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>000000011700249X</td>
<td>000000011700249X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Beatles lyrics: the unseen</td>
<td>00000000122770055</td>
<td>0000000122770055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Not Forgetting the whale</td>
<td>9780297859147</td>
<td>9780297856269</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Tipping Point</td>
<td>0000000021 -</td>
<td>0000000117</td>
<td>000000011700249X</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Woman of the dead; a thriller</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>0000000108749380</td>
<td>0000000108749380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>000000011700249X</td>
<td>000000011700249X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Beatles lyrics; the unseen</td>
<td>00000000122770055</td>
<td>0000000122770055</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Not Forgetting the whale</td>
<td>9780297859147</td>
<td>9780297856269</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>The Tipping Point</td>
<td>0000000021 -</td>
<td>0000000117</td>
<td>000000011700249X</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>Woman of the dead; a thriller</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td>9780297857426</td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>0000000108749380</td>
<td>0000000108749380</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9780297608134</td>
<td>My life with Wagner</td>
<td>000000011700249X</td>
<td>000000011700249X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Publisher Name & ISNI Matching Results**

<table>
<thead>
<tr>
<th>Publisher</th>
<th>Number of names</th>
<th>Number of matches</th>
<th>Success rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>1,201</td>
<td>954</td>
<td>79%</td>
</tr>
<tr>
<td>Bloomsbury</td>
<td>43,558</td>
<td>28,420</td>
<td>65%</td>
</tr>
<tr>
<td>BurleighDodd</td>
<td>681</td>
<td>35</td>
<td>5%</td>
</tr>
<tr>
<td>Channel/View</td>
<td>1,392</td>
<td>1,146</td>
<td>82%</td>
</tr>
<tr>
<td>Canongate</td>
<td>521</td>
<td>363</td>
<td>70%</td>
</tr>
<tr>
<td>CUP</td>
<td>21,298</td>
<td>16,292</td>
<td>76%</td>
</tr>
<tr>
<td>DK</td>
<td>2,103</td>
<td>1,409</td>
<td>67%</td>
</tr>
<tr>
<td>Hachette</td>
<td>10,406</td>
<td>7,820</td>
<td>72%</td>
</tr>
<tr>
<td>HarperCollins</td>
<td>1,117</td>
<td>708</td>
<td>63%</td>
</tr>
<tr>
<td>King</td>
<td>1,117</td>
<td>708</td>
<td>63%</td>
</tr>
<tr>
<td>Liverpool University Press</td>
<td>1,498</td>
<td>1,064</td>
<td>71%</td>
</tr>
<tr>
<td>PanMacmillan</td>
<td>1,642</td>
<td>1,332</td>
<td>81%</td>
</tr>
<tr>
<td>Penguin</td>
<td>14,297</td>
<td>9,638</td>
<td>67%</td>
</tr>
<tr>
<td>Pluto</td>
<td>1,189</td>
<td>1,189</td>
<td>100%</td>
</tr>
<tr>
<td>RandomHouse</td>
<td>24,060</td>
<td>16,127</td>
<td>67%</td>
</tr>
<tr>
<td>TaylorFrancis</td>
<td>107,871</td>
<td>68,878</td>
<td>64%</td>
</tr>
</tbody>
</table>

Total: 247,091 matches out of 163,400 with a success rate of 66%.
### Name de-duplication

**Resolving within and between publishers’ data files**

<table>
<thead>
<tr>
<th>Publisher</th>
<th>Number of de-duplicated IDs (across all publishers)</th>
<th>Number of de-duplicated IDs (within publisher)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atlantic</td>
<td>0</td>
<td>12</td>
</tr>
<tr>
<td>Bloomsbury</td>
<td>5409</td>
<td>619</td>
</tr>
<tr>
<td>BurleighDodds</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ChannelView</td>
<td>139</td>
<td>0</td>
</tr>
<tr>
<td>Canongate</td>
<td>3363</td>
<td>2225</td>
</tr>
<tr>
<td>CUP</td>
<td>398</td>
<td>62</td>
</tr>
<tr>
<td>Chatto &amp; Windus</td>
<td>1040</td>
<td>746</td>
</tr>
<tr>
<td>HarperCollins</td>
<td>2350</td>
<td>119</td>
</tr>
<tr>
<td>Kogan Page</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Liverpool University Press</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>PanMacmillan</td>
<td>524</td>
<td>33</td>
</tr>
<tr>
<td>Penguin</td>
<td>3449</td>
<td>846</td>
</tr>
<tr>
<td>Pluto</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>RandomHouse</td>
<td>4386</td>
<td>1788</td>
</tr>
<tr>
<td>TaylorFrancis</td>
<td>8650</td>
<td>9916</td>
</tr>
<tr>
<td>Total</td>
<td>30608</td>
<td>16374</td>
</tr>
</tbody>
</table>

**Publisher name and ISNI Matching results**

**Name de-duplication**

- **isni:000000006655121X**
  - PanMacmillan: 8348
  - Penguin: 22163
  - RandomHouse: 04069
  - Harcourt: 1

- **isni:00000000665959589**
  - Penguin: 21702
  - RandomHouse: 9495
  - Dk: 29990

- **isni:0000000066588178**
  - Penguin: 33288
  - RandomHouse: 14857

- **isni:0000000066553901**
  - HarperCollins: 15192
  - Penguin: 31521
  - RandomHouse: 30805

- **isni:0000000066578383**
  - Penguin: 29896
  - Penguin: 29160

- **isni:000000006655121X**
  - Penguin: 32908
  - Penguin: 29896

- **isni:0000000066579589**
  - Penguin: 35086
  - Penguin: 29896

- **isni:000000006657182**
  - Penguin: 35086
  - Penguin: 35086

- **isni:0000000066579589**
  - Penguin: 35086
  - Penguin: 35086

- **isni:0000000066573301**
  - Penguin: 35086
  - Penguin: 35086
Publisher ID to ISNI
Ongoing work

- Some work on improving of match rates
- Bringing NACO up to date in ISNI and re-run publisher files.
- Some failed matches due to sparse name forms and lack of ISBNs
- Embedding ISNIs in the supply chain (ONIX data)

ISNI in ONIX
Example from Harper Collins

```xml
<TitleElement>
    <TitleElementLevel>01</TitleElementLevel>
    <NoPrefix/>
    <TitleWithoutPrefix textcase="02">Boy Giant</TitleWithoutPrefix>
    <Subtitle>Son of Gulliver</Subtitle>
</TitleElement>

<Contributor>
    <SequenceNumber>1</SequenceNumber>
    <ContributorRole>A12</ContributorRole>
    <NameIdentifier>
        <NameIDType>01</NameIDType>
        <IDTypeName>HCP UK Author ID</IDTypeName>
        <IDValue>4121</IDValue>
    </NameIdentifier>
    <NameIdentifier>
        <NameIDType>16</NameIDType>
        <IDValue>00000012147035X</IDValue>
    </NameIdentifier>
    <PersonName>Michael Foreman</PersonName>
    <PersonNameInverted>Foreman, Michael</PersonNameInverted>
    <NamesBeforeKey>Michael</NamesBeforeKey>
    <KeyNames>Foreman</KeyNames>
</Contributor>

<Contributor>
    <SequenceNumber>2</SequenceNumber>
    <ContributorRole>A12</ContributorRole>
    <NameIdentifier>
        <NameIDType>01</NameIDType>
        <IDTypeName>HCP UK Author ID</IDTypeName>
        <IDValue>1897</IDValue>
    </NameIdentifier>
    <NameIdentifier>
        <NameIDType>16</NameIDType>
        <IDValue>0000000121251907</IDValue>
    </NameIdentifier>
    <PersonName>Michael Morpurgo</PersonName>
    <PersonNameInverted>Morpurgo, Michael</PersonNameInverted>
    <NamesBeforeKey>Michael</NamesBeforeKey>
    <KeyNames>Morpurgo</KeyNames>
</Contributor>
```
Publisher ID to ISNI

Next Steps

- Finalise ISNI CIP workflows with BDS for late 2020 go live
- Expansion to other publishers
- Develop workflows for new author registrations
- BL to launch ISNI portal self-registration service – Late 2020
- Options for filling in gaps
  - Priced batch services
  - CIP feedback loop
  - BL ISNI Service Portal

BL ISNI Registration Portal

- Only ISNI Registration Agencies and ISNI Members can register new ISNIs
  - Using ISNI XML via Atom Pub API
  - Submit CSV files
  - Online web interface for individual submissions
- All requests are processed through ISNI matching algorithms to prevent duplication
- Rich metadata is required to assign new unique ISNIs and assure disambiguation
BL ISNI Registration Portal

- As a Registration Agency the BL can mediate for batch or individual requests
- Processing individual requests is currently a manual process
  - Requests mostly via email
- The BL is developing an online service to allow non-members to request ISNIs directly through a mediated interaction with the ISNI database

BL ISNI Registration Portal

- Three main functions:
  - Search for an existing ISNI
  - Add data to an existing ISNI (no changes or deletions)
  - Register a new ISNI
- User experience – online, real time turnaround
- Nominal fee for ISNI registration
- QA support by BL Quality Team

[ISNI Portal]
Questions?