What is the problem with the Copyright Clearance Octopus? (identifying orphan works at NSL)

Szabolcs Dancs
acting director of National Library Services

25 April, 2019
National Library of Estonia, Tallinn
National Library Objectives

• provide access to information, i.e.
  – to a variety of document types (incl. books, journals, maps, pictures)
    • of various legal status:
      – public domain
      – under copyright
        » out-of-commerce
        » in-commerce works,

» orphan works

„For the purposes of establishing whether a work or phonogram is an orphan work, the organisations ... shall ensure that a diligent search is carried out (Article 3,1.).”

impl. in Hungary by Government Decree 138/2014 (IV. 30)
Sources of Diligent Search in 2012/28/EU

(a) legal deposit, library catalogues and authority files maintained by libraries and other institutions;

(b) the publishers' and authors' associations in the respective country;

(c) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;

(d) the databases of the relevant collecting societies, in particular reproduction rights organisations;

(e) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works);
Sources of Diligent Search in 2012/28/EU

(a) legal deposit, library catalogues and authority files maintained by libraries and other institutions;
(b) the publishers' and authors' associations in the respective country;
(c) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;
(d) the databases of the relevant collecting societies, in particular reproduction rights organisations;
(e) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works).
Sources of Diligent Search in Gov. Decr. 138/2014 (IV. 30)

a) the database of the Hungarian National Union Catalogue (MOKKA);
b) legal deposit, Hungarian National Bibliography;
c) library catalogues and authority files maintained by libraries and other institutions;
d) the publishers' and authors' associations;
e) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;
f) the databases of the relevant collecting societies;
g) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works);
h) database of the National Electronic Document Delivery System.
Sources of Diligent Search in Gov. Decr. 138/2014 (IV. 30)

a) the database of the Hungarian National Union Catalogue;
b) legal deposit, Hungarian National Bibliography;
c) library catalogues and authority files maintained by libraries and other institutions;
d) the publishers' and authors' associations;
e) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;
f) the databases of the relevant collecting societies;
g) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works);
h) database of the National Electronic Document Delivery System.
Copyright „filtering” – a semi-automated method

Authority data and bibliographic data

- Public domain
- Under copyright

Sources listed in the Government Decree 138/2014.

- Orphan work
- Not orphan work

License management

- To be provided by the permission of the copyright holder
- Not to be provided
Cooperation with European Union Intellectual Property Office (EUIPO)

sources listed in the Government Decree 138/2014.

EUIPO Orphan Work Database

orphan work
 Uploaded (in Excel) by NSL, forwarded by HIPO

registered orphan work

feedback

digital library service

accessed orphan work
8 sources = 8 tentacles of the Copyright Clearance octopus

- a) the database of the Hungarian National Union Catalogue (MOKKA);
- b) legal deposit, Hungarian National Bibliography;
- c) library catalogues and authority files maintained by libraries and other institutions;
- d) the publishers' and authors' associations;
- e) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;
- f) the databases of the relevant collecting societies;
- g) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works);
- h) database of the National Electronic Document Delivery System.
Diligent search in an automated way

Aim:
• automated queries of relevant databases
• automated data exchange with EUIPO database

Problems:
• lack of relevant databases (those of listed in the directive)
• lack of relevant databases with standard interface
• no use of standards (ONIX)
• lack of relevant metadata in library databases (requirement of autopsy; in the future: RDA? GDPR?)
Current process of copyright clearance

• based on a 12 pages **policy** on identifying and using orphan works (implemented after consulting with the Hungarian Intellectual Property Office) and a 11 pages **training paper** (addressed new employees responsible for copyright clearance)

• the policy includes
  – a list of relevant sources of diligent search (richer than in the Gov. Dec.)
  – detailed regulation of the use end registration of orphan works
A moral and financial responsibility

a) the database of the Hungarian National Union Catalogue (MOKKA);
b) legal deposit, Hungarian National Bibliography;
c) library catalogues and authority files maintained by libraries and other institutions;
d) the publishers' and authors' associations;
e) existing databases and registries, WATCH (Writers, Artists and their Copyright Holders), the ISBN (International Standard Book Number) and databases listing books in print;
f) the databases of the relevant collecting societies;
g) sources that integrate multiple databases and registries, including VIAF (Virtual International Authority Files) and ARROW (Accessible Registries of Rights Information and Orphan Works);
h) database of the National Electronic Document Delivery System

Should a diligent search be considered to be carried out with good faith if it didn’t take into account databases that were online unaccessible and/or officially unknown?
Personal factor and ...

• 1 librarian engaged in diligent search (in part-time); currently on maternity leave; skills:
  – information and librarian studies (Master)
  – deep interest in copyright issues (thesis on CC)
  – aware of the problem of accessibility of digital assets in libraries and the benefits of orphan work regime
• 1 public employee temporarily engaged in diligent search (in part-time); skills:
  – historian (Master)
  – painstaking
  – just getting familiar with copyright, orphan work, etc.
Colleague with librarian background:
- quite fast work (as it is possible), focusing on information sources defined in our orphan work policy (based on the laws)
- more effective in the sense of exploiting possibility to facilitate access offered by orphan work regime

Social employee with historian background:
- not fast work with a high guarantee of diligence and good faith
Access-module under development

- part of the new national library platform
- based on ELDORADO
- cooperation with Books In Print company (data exchange in ONIX; opusnet.hu)
- using copyright statements (https://rightsstatements.org/en/) to be recorded in the field '506, in MARC 21; plus terms intended to use:
  - Public Domain – for documents published more than 140 years before the digitisation (to be automatically added to the records);
  - in copyright – freely usable (e.g. for preservation purposes); in copyright – freely usable (e.g. via access through dedicated terminal);
  - CC = Creative Commons (plus CC statements), etc.
library catalogues and authority files maintained by libraries and other institutions
• authority databases maintained by Petőfi Literary Museum
• voluntary register of works maintained by the Hungarian Intellectual Property Office
• Hungarian National Scientific Bibliography (a register for works written by Hungarian scientists)

the publishers' and authors' associations
• one database with online availability (Hungarian Writers' Association)

the databases of the relevant collecting societies and journalist associations
• 3 collecting societies
• 5 journalist associations

other internet resources
• online biographical cyclopaedia (e.g. Hungarian Catholic Lexicon; Lexicon of Hungarian Jews)
Use Case 1 – Monographs (I)

One document – more works (of different type/various creators)

- WORK 1: main text
- WORK 2: translation
- WORK 3: illustration (incl. cover)
- WORK 4: editing series? (no significant contribution)
Use Case 1 – Monographs (II)

One document – more works (of different type/various creators)

WORK 1
Stein Aurél (1862-1943)

WORK 2
Halász Gyula (1881-1947)
photos: Stein Aurél (1862-1943)
cover: ? → HUNGART (coll. soc.)
portrey of S. A.: ? → HUNGART

WORK 3:1-3

WORK 4
Lóczy Lajos (1849-1920)
Article written ...

- by editorial employee → publisher as right holder
- by contributor → permission by author is required

Problem: in order to decide there is a need to access archives of contracts (if there exist)?

*Sloboda* (1945-1948) – journal of Antifascist Front of Slavs living in Hungary

- no confirmed information about successor organisation (probably: Democratic Association of Slovaks Living in Hungary (1948-), later: Association of Slovaks Living in Hungary (1990-))
Problem of cooperation

interoperability: cooperation on the level of IT systems
• lack of "metadata-awareness": role of metadata, using controlled vocabularies, namespaces (national and other), dataflow optimization are getting more and more relevant
• general problem: nonexploitation of standards

cooperation of stakeholders
• social and market background
• public sector: reluctance, mistrust, lack of win-win approach
• private sector: obsolete business models (long tail, role of niches); lack of funding models inspiring creation of new business models
DSM Directive – new perspectives


- a new regime enabling easier access to works defined as *out-of-commerce*
- **text and data mining exception** serves as an opportunity to establish projects such as *Corpus* in BnF (supporting researchers in data mining of copyrighted materials [e.g. legal deposit])
Thanks for attention!