iPBL

Supplementing Literary Bibliography with Internet Sources. Collaborative Cataloguing

Expanding the database

0,8M bibliographical records

62K authors

122K works

19K cultural events

25K theatrical & radio plays

2K films

The ongoing R&D project iPBL aimed at expanding the source catalogue of Polish Literary Bibliography (PBL) and supplementing it with information about online materials such as electronic literature, literary blogs, websites devoted to the Polish culture, and social media profiles. Literary bibliographies focus mostly on printed and digitised materials, thus adding online sources poses a two-fold challenge:

1. Methodological challenge - how to select and describe data on the internet culture and reconcile it with the records on print culture within the same database
2. Technological challenge - how to obtain metadata of electronic texts and secure their accessibility for future researchers.

The revolution we propose is to make available, for any interested party, an even more detailed database, which is highly accessible, free and compatible with international standards.

The scope of the database will cover:

- → nationwide cultural websites and services
- → online magazines
- → blogosphere
- → literary works published on the Internet

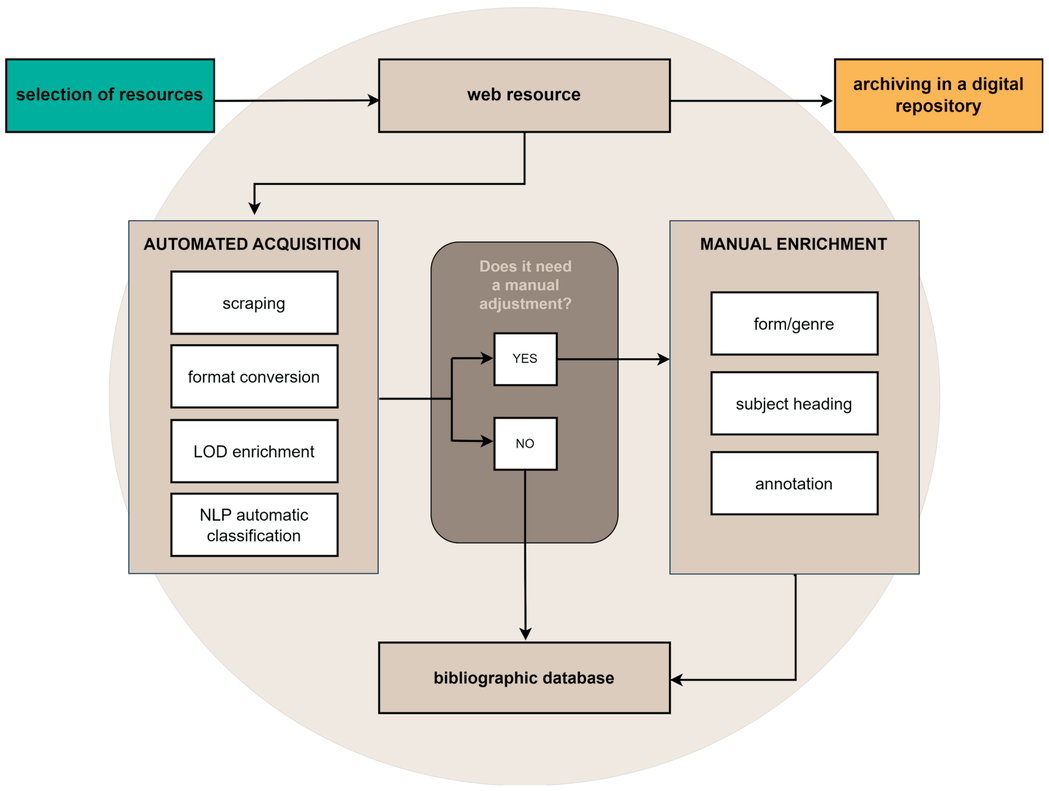
The selective and analytical bibliographic database will include:

150K records **2K** Internet sources **40K** records with autopsy descriptions **4K** data sources

iPBL supplementing literary bibliography with internet sources

Creating the bibliography of internet sources challenges traditional documentation practices, as the workflow for collecting, curating, and sharing data that will consider, from the very beginning, the human and non-human stakeholders. The project's workflow consists of the following operations:

1. Sources selection, based on the analysis of the available sources: the National Library of Poland's Internet Resources Database and Bibliographic Guide, the ISSN portal, and the Internet Archive.
2. Automated acquisition. The information will be obtained due to dedicated scraping algorithms and then converted into the internationally approved bibliographic formats (MARC21, Bibtex, RIS); the resulting bibliographic resource will be enriched with Linked Open Data structures (references to international controlled vocabularies and authority databases, e.g. including VIAF, LCSH) and automatically classified using NLP.
3. Optional manual adjustment. If the results of the automated process are unsatisfactory, the records will be subject to manual adjustment for the genre, subject heading and content annotation.
4. Archiving the source in a digital repository. The sources will be automatically submitted to the Web Archive to ensure their future availability.
   Publishing metadata record in the database. Final record will be incorporated into the PBL database.



Research potential

New internet data will enable research into Polish digital culture, and even more important, save it from oblivion. There are known cases of complete disappearance of cultural content, 20% of cited web resources become inaccessible, which makes it impossible to study. It is necessary to record data to allow any future research at all and the broaden PBL database will be the point where any academic collaboration may begin.
Digital culture is a part of the Polish cultural heritage. This heritage is largely made up of documents created in a digital form and functioning primarily on the Internet: on websites, portals, online magazines, blogosphere, and fora, as well as on websites or social profiles of institutions and people of culture. Recording data from online sources is thus a natural step in the PBL database evolution.

DIGITAL HUMANITIES CENTRE

This work was supported by the Polish Ministry of Science and Higher Education through the National Programme for the Development of the Humanities (grant number: NPRH/DN/SP/495736/2021/10, 2023-2025).

Ministerstwo Edukacji i Nauki

Any questions?
Department of Current Bibliography
pracownia.poznan@ibl.waw.pl
Polska. Bibliografia.Literacka

Visit our website:
pbl.ibl.waw.pl