



## Navigating the challenges of aggregation

Aleksandra Nowak  
[aleksandra.nowak@man.poznan.pl](mailto:aleksandra.nowak@man.poznan.pl)  
Poznań Supercomputing and Networking Center

Everywhere

Search collections of Polish libraries, archives, museums...



Advanced

On-line collections of Polish cultural and scientific institutions. We have more than 4.8 million objects!

***Zapraszamy na portal tematyczny FBC Czasopisma!***

Wszystkich zainteresowanych zbiorami prasy z kolekcji polskich instytucji kultury i nauki zapraszamy do nowego serwisu tematycznego FBC Czasopisma, dostępnego pod adresem <http://czasopisma.fbc.net.pl/>. Serwis ten oferuje dostęp do kilku tysięcy tytułów czasopism zebranych ze zbiorów wszystkich instytucji dostarczających dane do FBC. Dane w portalu FBC Czasopisma można przeszukiwać na podstawie tytułu i słów kluczowych związanych z tematyką

*FBC recommends*



Thaddeus Kościuszko



Music magazines



Sacral sculpture



Polish women novelists



Fashion during times of the Polish People's Republic



Physical exercises



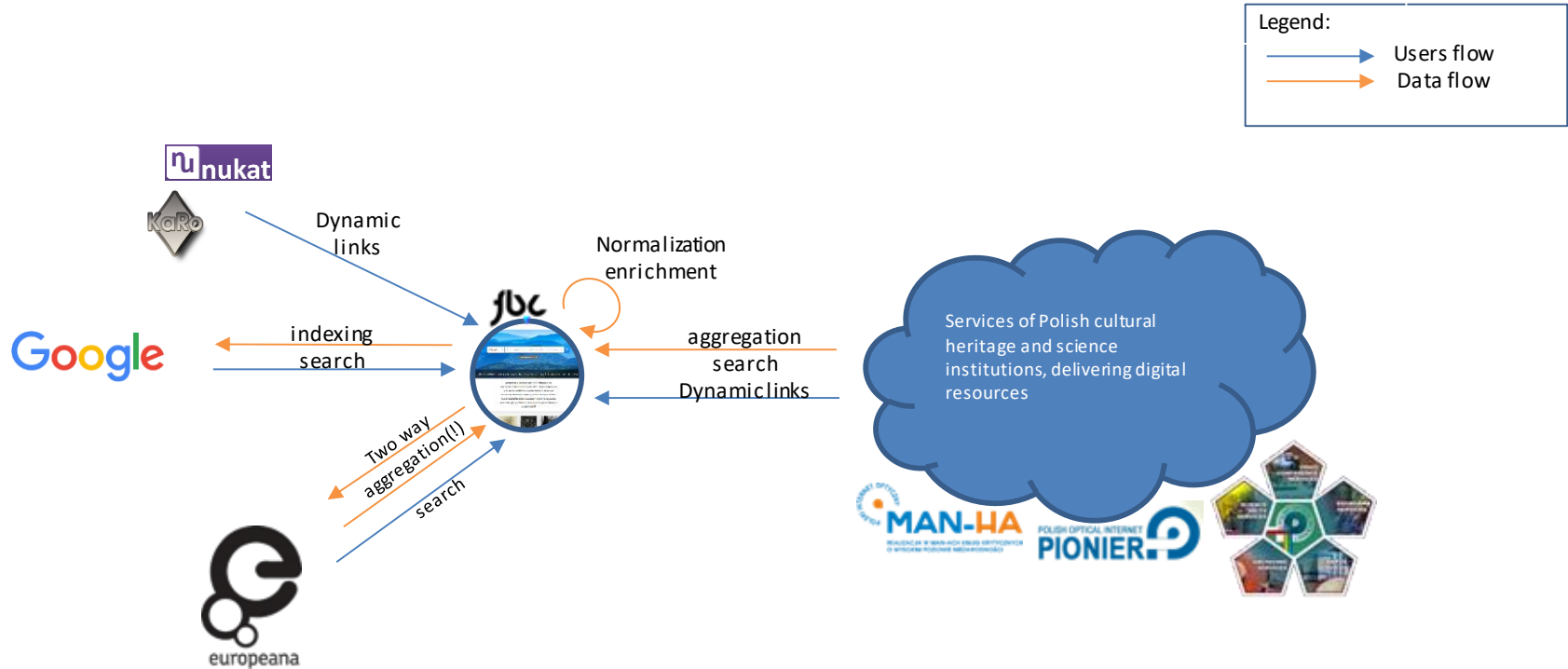
Maps of Europe



Advertising posters

## FBC – Digital Libraries Federation

platform for digital resources of Polish cultural heritage and science



## FBC – Digital Libraries Federation

### statistics

---

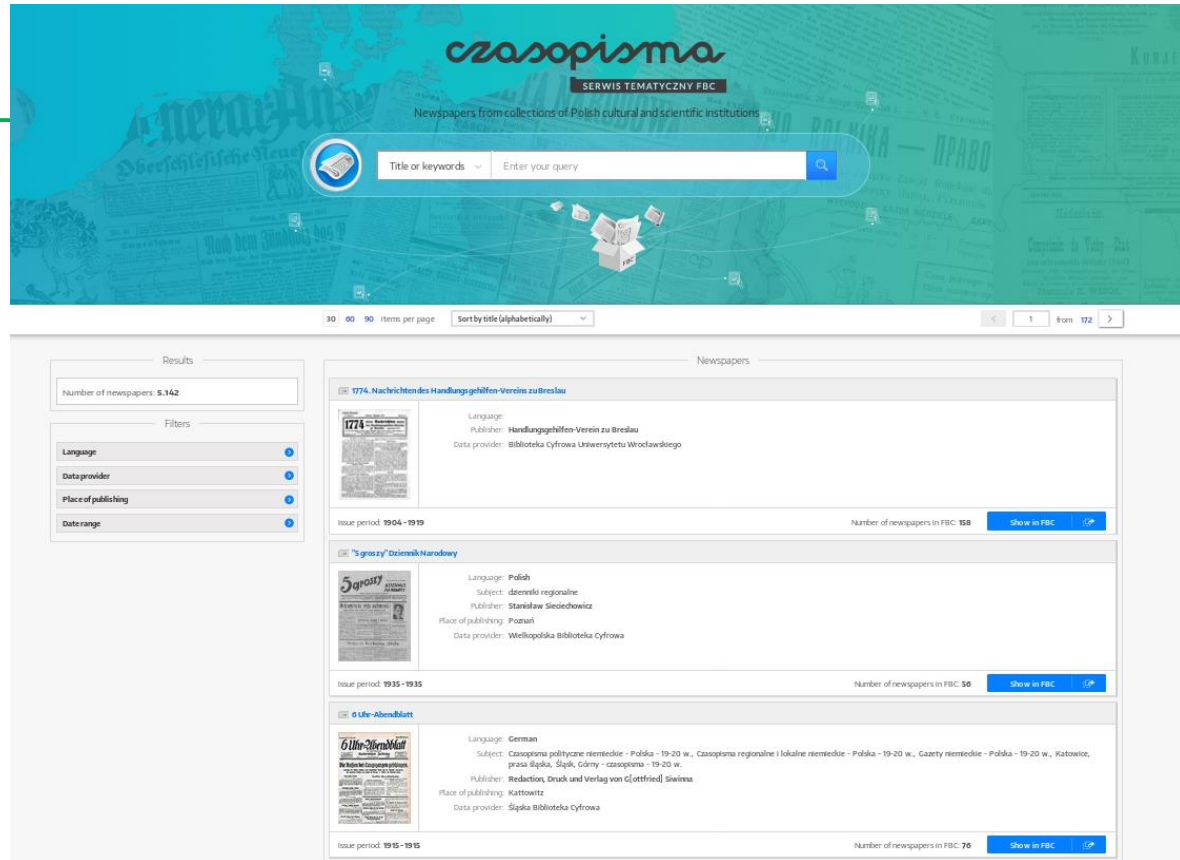
- Started in 2007
- 4,825,781 objects at the moment (3,658,491 in free access)
- 132 data providers: digital libraries, digital museums, institutional repositories
- Traffic:
  - 1mln sessions, 650k users per year



## FBC Journals

<http://czasopisma.fbc.net.pl/>

- First thematic portal of FBC
- Dedicated tool to search for and view journals and newspapers available in FBC



The screenshot displays the 'czasopisma' website, a thematic portal for FBC. The header is teal with the 'czasopisma' logo and the text 'SERWIS TEMATYCZNY FBC Newspapers from collections of Polish cultural and scientific institutions'. A search bar with a magnifying glass icon is located below the header. The main content area is white and shows a list of newspaper entries. Each entry includes a thumbnail image of the newspaper cover, its title, language, publisher, place of publishing, and data provider. The first entry is '1774. Nachrichten des Handlungs-gehilfen-Vereins zu Breslau', the second is '5 groszy Dziennik Narodowy', and the third is '6 Uhr-Abendblatt'. The interface also includes a sidebar with filters for language, data provider, place of publishing, and date range, and a results section at the top left showing the total number of newspapers found (5,142).

## Challenges of aggregation – lessons learned

### data & aggregation

---

#### ***Find the balance between standards and data quality***

- Dublin Core is not for everybody
  - Institutions were "forced" to use Dublin Core, and they mapped their data to it, in a lot of cases data meaning was lost
  - It was also too limited, and we couldn't make some functionalities work with it
  - A tailored data format was created on the Dublin Core basis (and introduced in many data sources)

#### ***Consider cooperation instead of doing it all by yourself***

- After Polish Ministry of Culture started to fund digitisation, a lot of small institutions created their own services instead of working with bigger institutions (e.g. regional digital library)
- A lot of them cannot really maintain these services anymore after the funding is finished

## Challenges of aggregation - lessons learned

### data quality

---

#### *Define/agree on standards at the very beginning*

- Think if you can standardize any of the fields in the metadata (e.g. language, rights, type).
  - Use one of well-known standards (e.g. ISO codes for language, <http://rightsstatements.org/> for rights)
  - You can work it out together with your data providers
- Without it data providers will do it by themselves and you will have to handle multiple different self-established standards at the aggregator level

# Challenges of aggregation - lessons learned

copyrights

---

## ***Validate the data!***

- Problems in our metadata:
  - no rights information available
  - intellectual property rights mistaken for physical ownership of the object
- Validate and:
  - Give feedback
  - Do not accept records with the information missing or when it's invalid\*

\*it might be too restrictive solution



## Challenges of aggregation - lessons learned

### institutional buy-in

---

#### ***Increase the awareness about your aggregator and Europeana***

- We organized several workshops about FBC, Europeana, digital humanities, first individually, then in cooperation with National Audiovisual Institute (FINA) under "Tu Europeana" program funded by the Ministry of Culture
- what worked "better": at some point Ministry of Culture changed how digitization grants were evaluated, there were bonus points if data was in the end available in Europeana.
  - That fired back a little bit as only few institutions did it well, the rest left it for the very end of the project, some have not delivered their data



# Poznań Supercomputing and Networking Center

affiliated to the Institute of Bioorganic Chemistry of the Polish Academy of Sciences,

ul. Jana Pawła II 10, 61-139 Poznań, POLAND,

Office: phone center: (+48 61) 858-20-01, fax: (+48 61) 850-25-01,

e-mail: [office@man.poznan.pl](mailto:office@man.poznan.pl), <http://www.psnc.pl>