J-Ark Project

Pavel Kats, Coordinator, Jewish Heritage Network

Jewish Heritage Network

Co-financed by the Connecting Europe Facility of the European Union

European Commission

About JHN

- Global **non-profit** digital initiative for Jewish cultural heritage content
- Registered as a Foundation in the Netherlands
- Content network of more than 30 Jewish **museums, archives, libraries, and research centers** from Europe (mainly), US, and Israel
- Digital **platform of services** for cultural heritage content aggregation, online publication and promotion
- Accredited partner of Europeana, the European digital platform for culture
- Funded by the European Commission and other funders via several projects

europeana aggregator
TRUSTED PARTNER

J-Ark: Project Vision

Use **CEF Building Blocks** of the European Commission to deliver a long-term preservation platform for Jewish heritage archives.

- **Building Blocks**
  - eArchiving - for long-term preservation according to European standards
  - Europeana - for accessing, aggregating, mapping and distributing heritage collections
  - eTranslation - for content services (machine translation and anonymisation)
- **Target Audience: Heritage Archives**
  - Small and medium-size archives guarded by museums, communities, NGOs etc who are not affiliated with a major (e.g. national-level) long-term preservation backbone
- **Ecosystemic Approach**
  - Integrating components from different domains / building blocks to create an end-to-end service for an important market segment (also with an eye on future approaches - Data Spaces)

Project

- **Duration**

September 2019 - February 2023 (CEF-TC-2020-2)

- **Budget**

€393,011

- **Partners**
  - Pangeanic (Spain)
  - JHN (Netherlands)
  - KEEP SOLUTIONS (Portugal)
  - Europeana (Netherlands)
  - Brama Grodzka - Teatr NN (Poland)
  - Jewish Community of Lithuania

Co-financed by the Connecting Europe Facility of the European Union

European Commission

Product and Service

End-to-end solution covering the full spectrum of needs for small- and mid-size heritage collections:

- **Share / Ingest**

Access digital content and metadata through a variety of ways

- **Publish / Distribute**

Publish via APIs and popular heritage platforms (Europeana, national, thematic platforms)

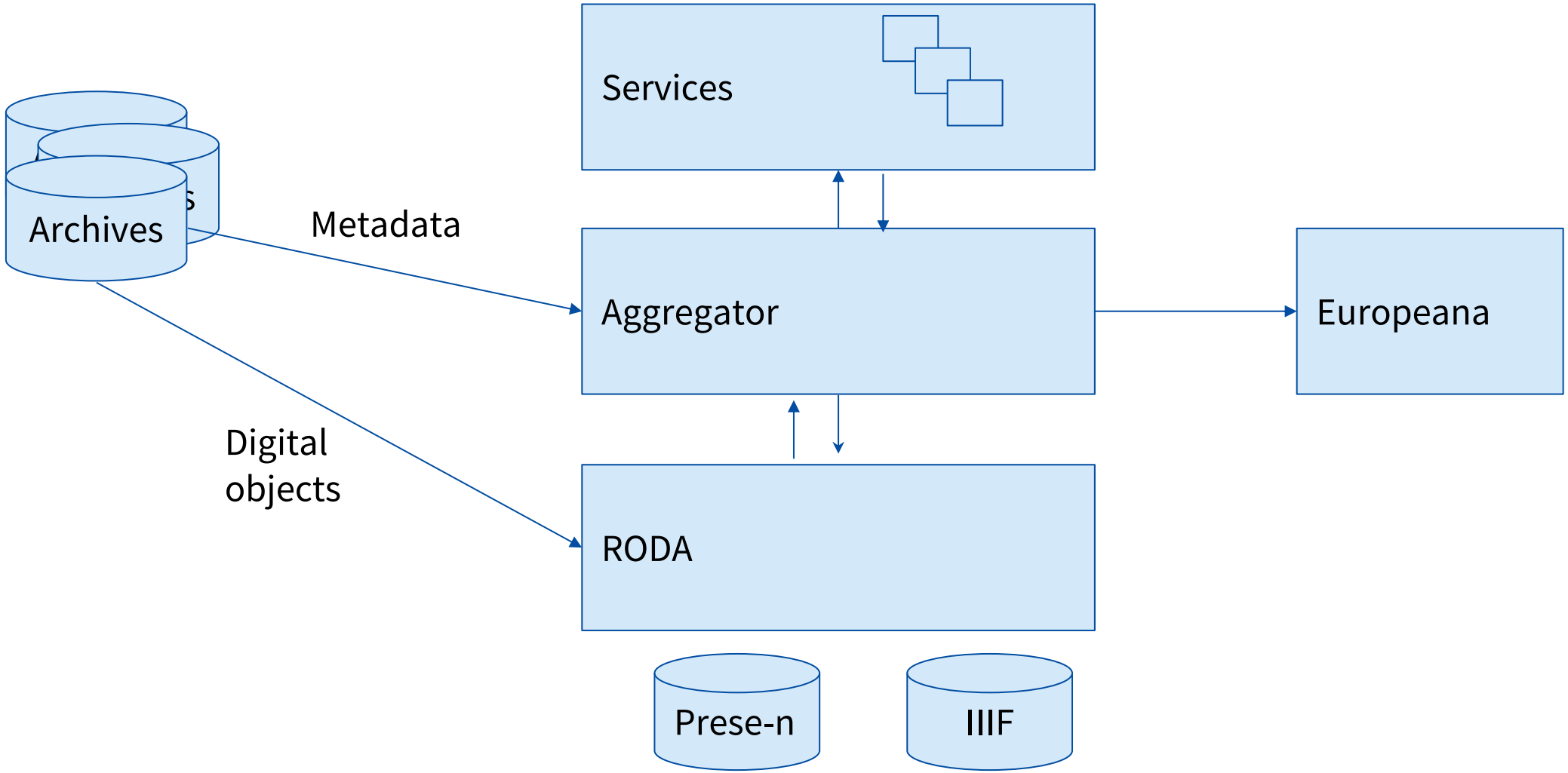
- **Process and Preserve**

Suite of content services and long-term preservation following the E-ARK standards of EC

Specific emphases:

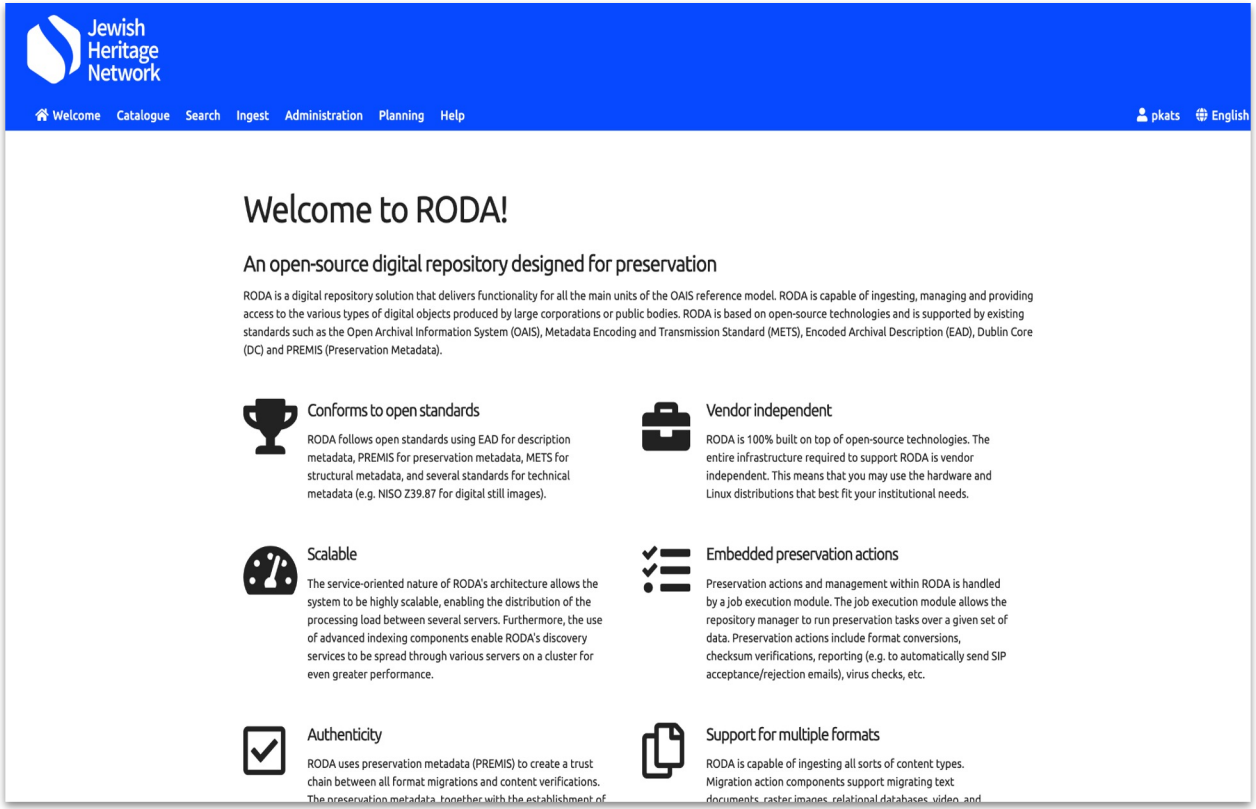
- Full SaaS solution provided on the Cloud
- Ease of use and rich training materials
- Fully multilingual

Components / Architecture



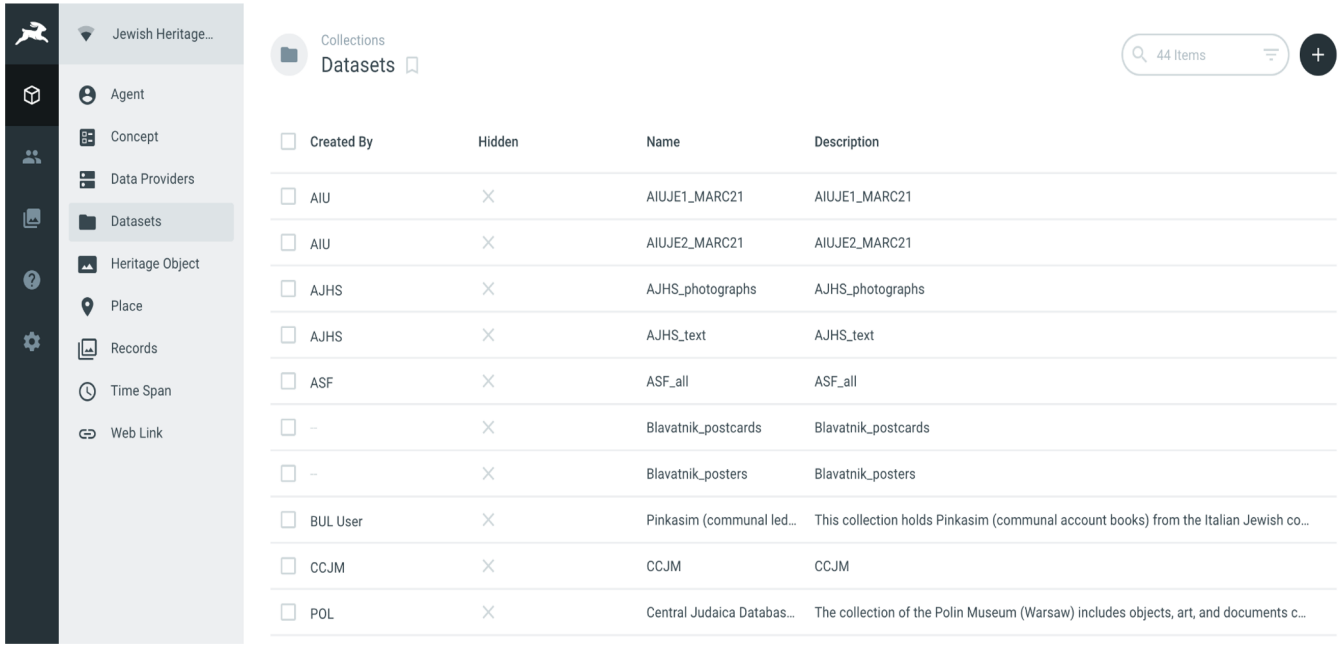
Preservation Service

- RODA 4.2.0 by KEEPS
- Supports E-ARK IP 2.0.4
- SFTP service for large transfers
- IIIF Support
    - image standard protocol service (using Cantaloupe)
    - IIIF presentation standard service
    - Universal viewer
- Monitoring stack (Prometheus + Grafana)
- EDM support (added as internal md format + presentation)



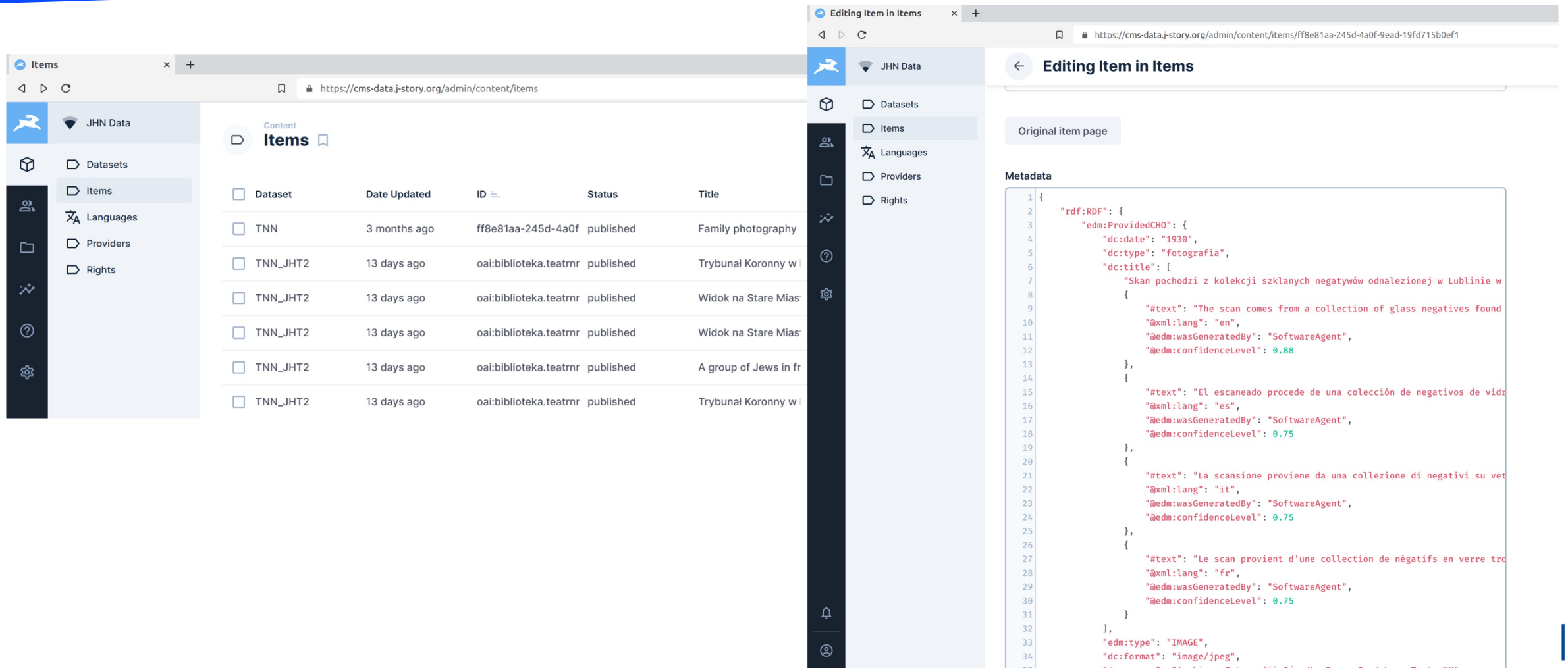
Metadata Aggregator

- Judaica Europeana aggregator by Jewish Heritage Network
- Supports different heritage formats (MARC21, DC, EDM) and custom
- Maps to EDM
- Built-in data processing flows: semantic enrichment, translation
- Based on open-source Headless CMS framework



European Commission

Record View (and EDM)



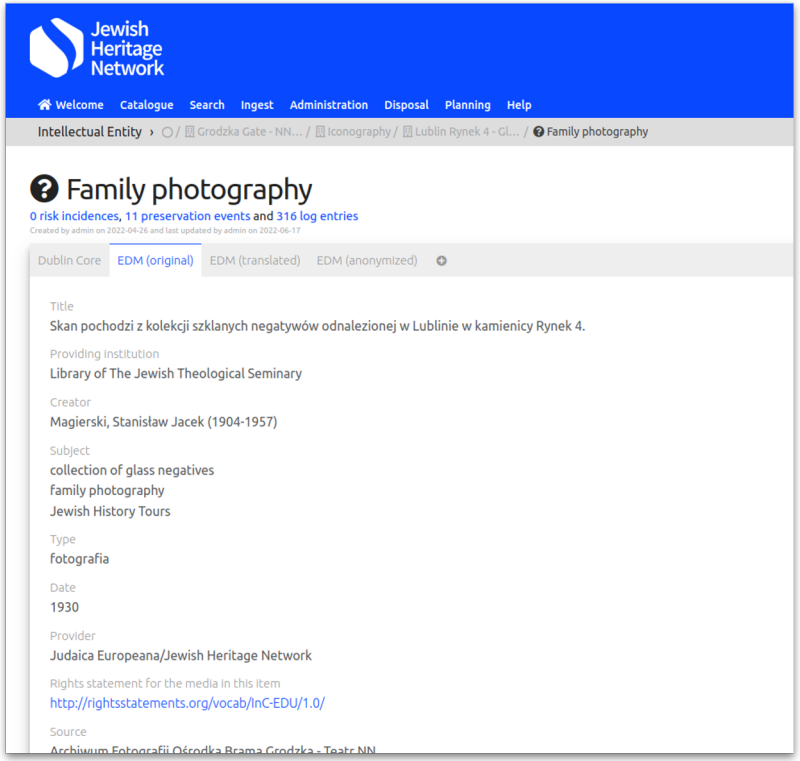
AI Content Services

- Adaptive near-human quality Machine Translation supporting all European languages and many others
- Anonymise, pseudo-anonymise, de-identify, names and other personal data to make content and communications fully compliant with GDPR and all privacy legislation.
- Developed by Pangeanic via various European projects (IDAAPTRA, Maps, Europeana Translate)

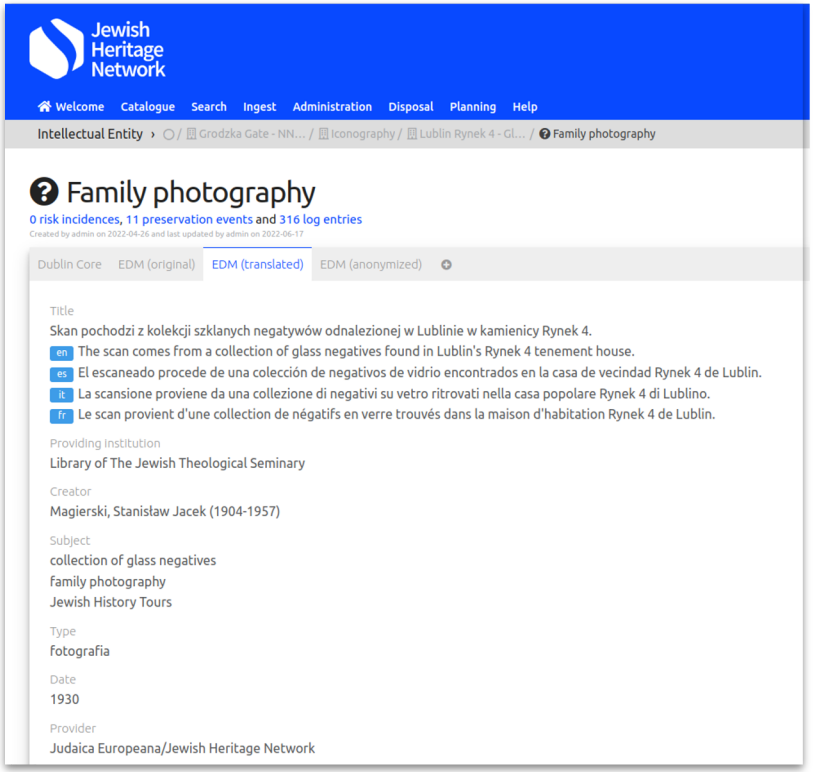
European Commission

Services: Anon. and Translation

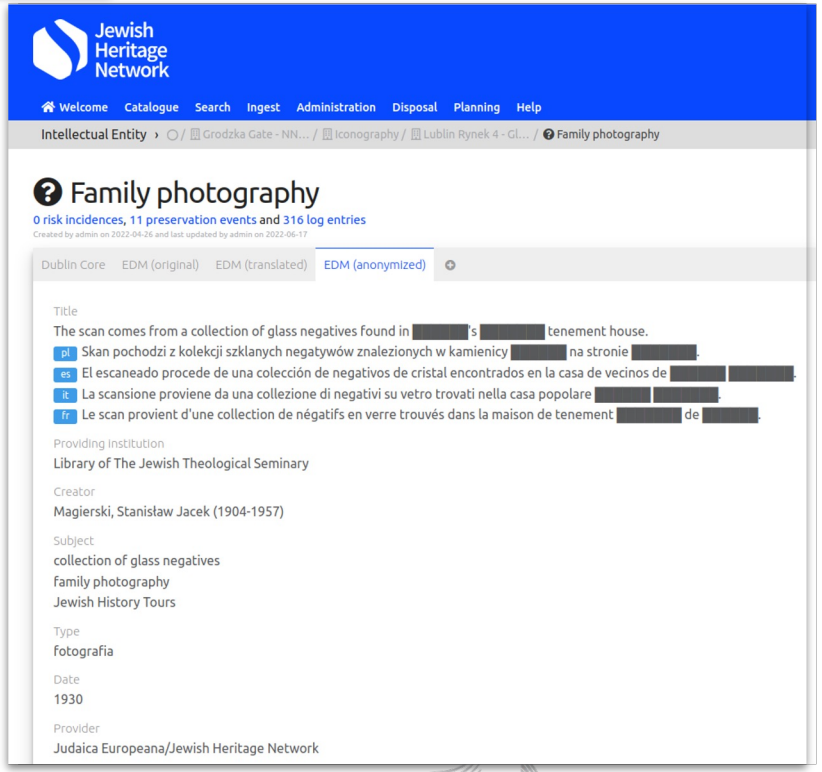
Describe



Translate



Anonymise



European Commission

EDM transformation

Describe

```
▼<edm:ProvidedCHO rdf:about="https://biblioteka.teatrnn.pl/dlibra/publication/edition/11263/content">
   <dc:creator>Magierski, Stanisław Jacek (1904-1957)</dc:creator>
   <dc:date>1930</dc:date>
   <dc:format>image/jpeg</dc:format>
   <dc:identifier>oai:biblioteka.teatrnn.pl:102244</dc:identifier>
   <dc:language>en</dc:language>
   <dc:relation>oai:biblioteka.teatrnn.pl:publication:107619</dc:relation>
   <dc:source>Archiwum Fotografii Ośrodka Brama Grodzka - Teatr NN</dc:source>
   <dc:subject>collection of glass negatives</dc:subject>
   <dc:subject>family photography</dc:subject>
   <dc:subject>Jewish History Tours</dc:subject>
   <dc:title>Skan pochodzi z kolekcji szklanych negatywów odnalezionej w Lublinie w kamienicy Rynek 4.</dc:title>
   <dc:type>fotografia</dc:type>
   <dcterms:isPartOf xml:lang="en">Jewish History Tours</dcterms:isPartOf>
   <edm:type>IMAGE</edm:type>
 </edm:ProvidedCHO>
```


European Commission

EDM transformation

Describe → Translate

```
<edm:ProvidedCHO rdf:about="https://biblioteka.teatrnn.pl/dlibra/publication/edition/11263/content">
  <dc:creator>Magierski, Stanisław Jacek (1904-1957)</dc:creator>
  <dc:date>1930</dc:date>
  <dc:format>image/jpeg</dc:format>
  <dc:identifier>oai:biblioteka.teatrnn.pl:102244</dc:identifier>
  <dc:language>en</dc:language>
  <dc:relation>oai:biblioteka.teatrnn.pl:publication:107619</dc:relation>
  <dc:source>Archiwum Fotografii Ośrodka Brama Grodzka - Teatr NN</dc:source>
  <dc:subject>collection of glass negatives</dc:subject>
  <dc:subject>family photography</dc:subject>
  <dc:subject>Jewish History Tours</dc:subject>
  <dc:title>Skan pochodzi z kolekcji szklanych negatywów odnalezionej w Lublinie w kamienicy Rynek 4.</dc:title>
  <dc:title xml:lang="en" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.88">The scan comes from a collection of glass negatives found in Lublin's Rynek 4 tenement house.</dc:title>
  <dc:title xml:lang="es" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.75">El escaneado procede de una colección de negativos de vidrio encontrados en la casa de vecindad Rynek 4 de Lublin.</dc:title>
  <dc:title xml:lang="it" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.75">La scansione proviene da una collezione di negativi su vetro ritrovati nella casa popolare Rynek 4 di Lublino.</dc:title>
  <dc:title xml:lang="fr" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.75">Le scan provient d'une collection de négatifs en verre trouvés dans la maison d'habitation Rynek 4 de Lublin.</dc:title>
  <dc:type>fotografia</dc:type>
  <dcterms:isPartOf xml:lang="en">Jewish History Tours</dcterms:isPartOf>
  <edm:type>IMAGE</edm:type>
</edm:ProvidedCHO>
```


European Commission

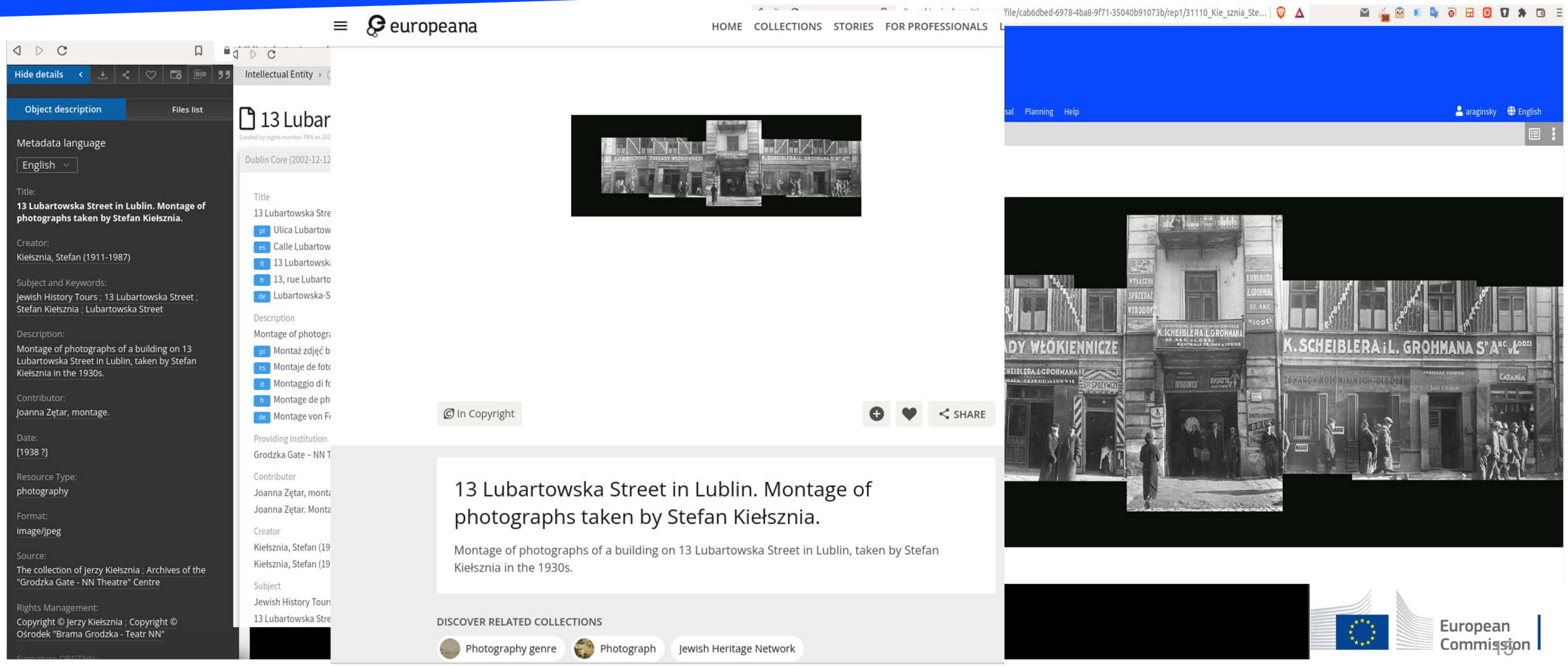
EDM transformation

Describe → Translate → Anonymise

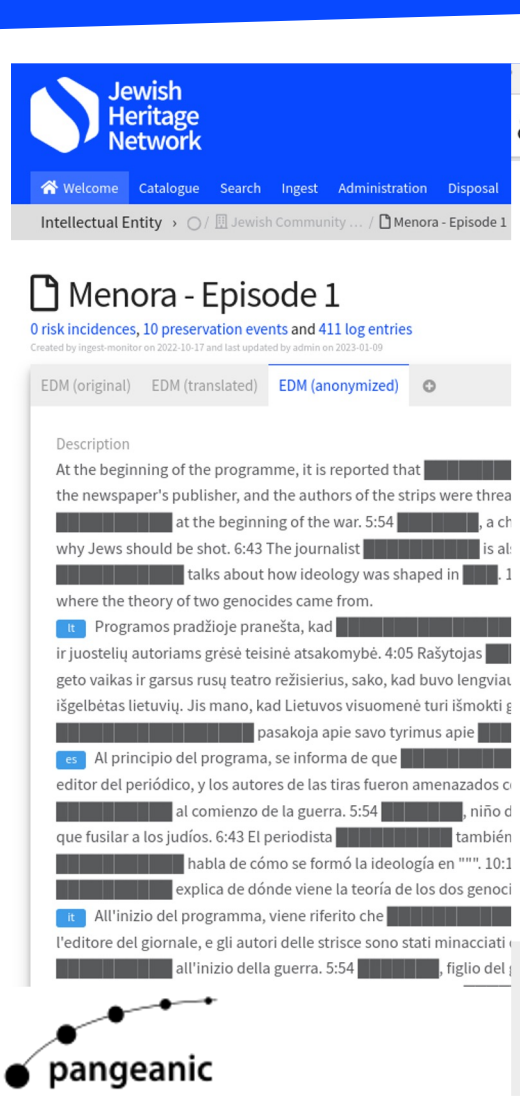
```
▼<edm:ProvidedCHO rdf:about="https://biblioteka.teatrnn.pl/dlibra/publication/edition/11263/content">
   <dc:creator>Magierski, Stanisław Jacek (1904-1957)</dc:creator>
   <dc:date>1930</dc:date>
   <dc:format>image/jpeg</dc:format>
   <dc:identifier>oai:biblioteka.teatrnn.pl:102244</dc:identifier>
   <dc:language>en</dc:language>
   <dc:relation>oai:biblioteka.teatrnn.pl:publication:107619</dc:relation>
   <dc:source>Archiwum Fotografii Ośrodka Brama Grodzka - Teatr NN</dc:source>
   <dc:subject>collection of glass negatives</dc:subject>
   <dc:subject>family photography</dc:subject>
   <dc:subject>Jewish History Tours</dc:subject>
   <dc:title xml:lang="en">The scan comes from a collection of glass negatives found in [redacted]'s [redacted] tenement house.</dc:title>
   <dc:title xml:lang="pl" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.88">Skan pochodzi z kolekcji szklanych negatywów znalezionych w kamienicy [redacted] na stronie [redacted].</dc:title>
   <dc:title xml:lang="es" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.91">El escaneado procede de una colección de negativos de cristal encontrados en la casa de vecinos de [redacted] [redacted].</dc:title>
   <dc:title xml:lang="it" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.94">La scansione proviene da una collezione di negativi su vetro trovati nella casa popolare [redacted] [redacted].</dc:title>
   <dc:title xml:lang="fr" edm:wasGeneratedBy="SoftwareAgent" edm:confidenceLevel="0.89">Le scan provient d'une collection de négatifs en verre trouvés dans la maison de tenement [redacted] de [redacted].</dc:title>
   <dc:type>fotografia</dc:type>
   <dcterms:isPartOf xml:lang="en">Jewish History Tours</dcterms:isPartOf>
   <edm:type>IMAGE</edm:type>
 </edm:ProvidedCHO>
```

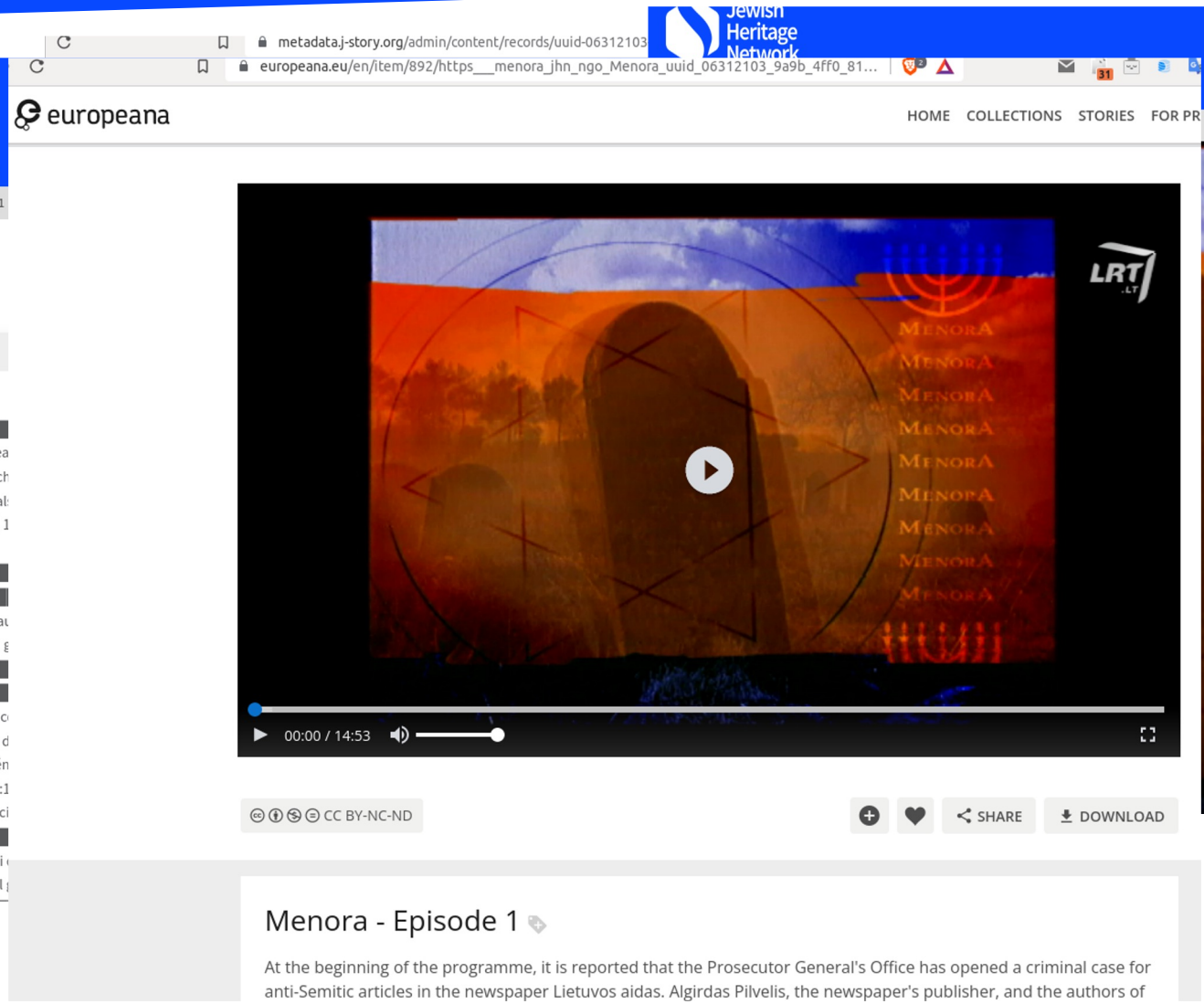

European Commission

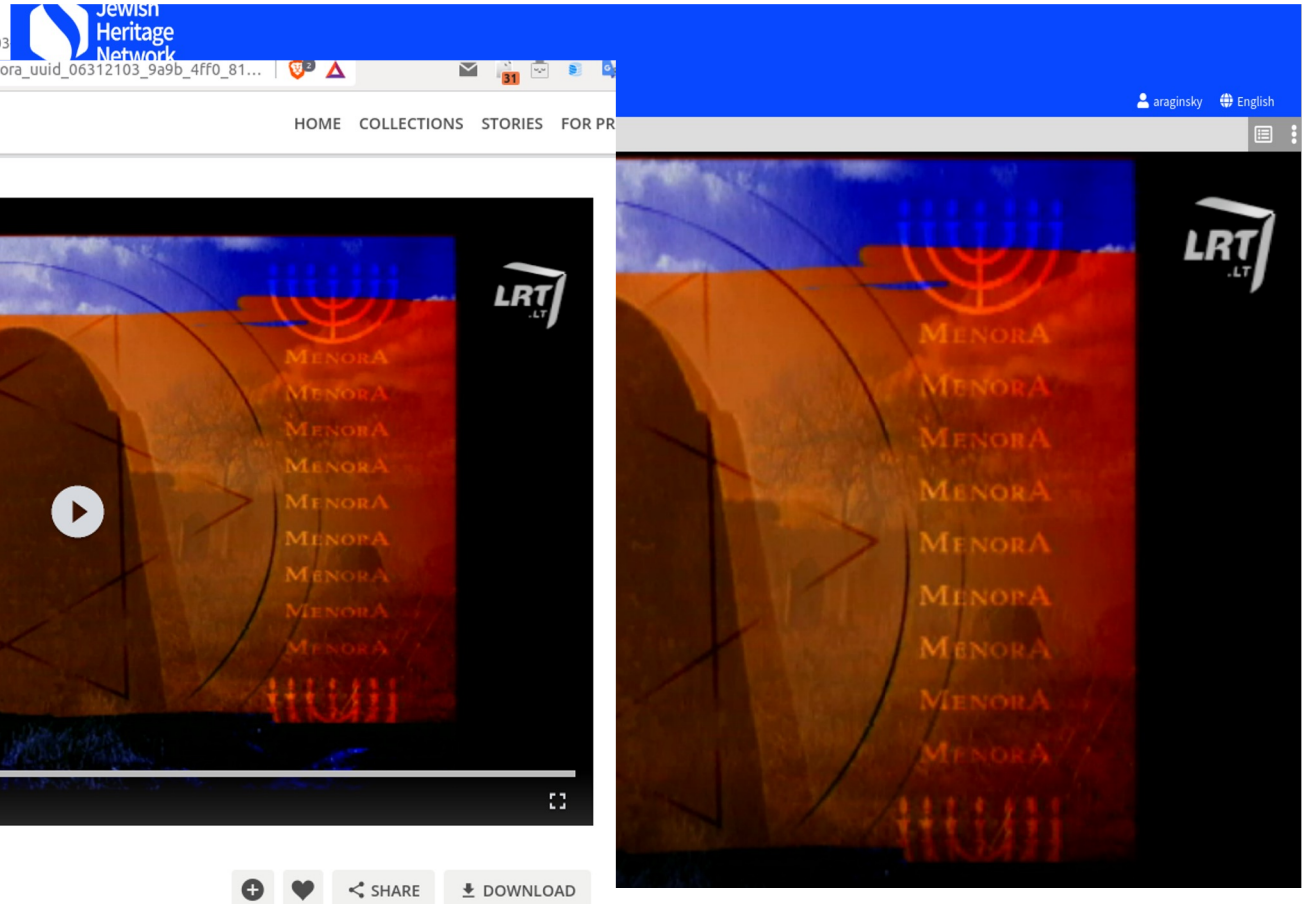
Content Partner 1 (Lublin)



Content Partner 2 (Lithuania)



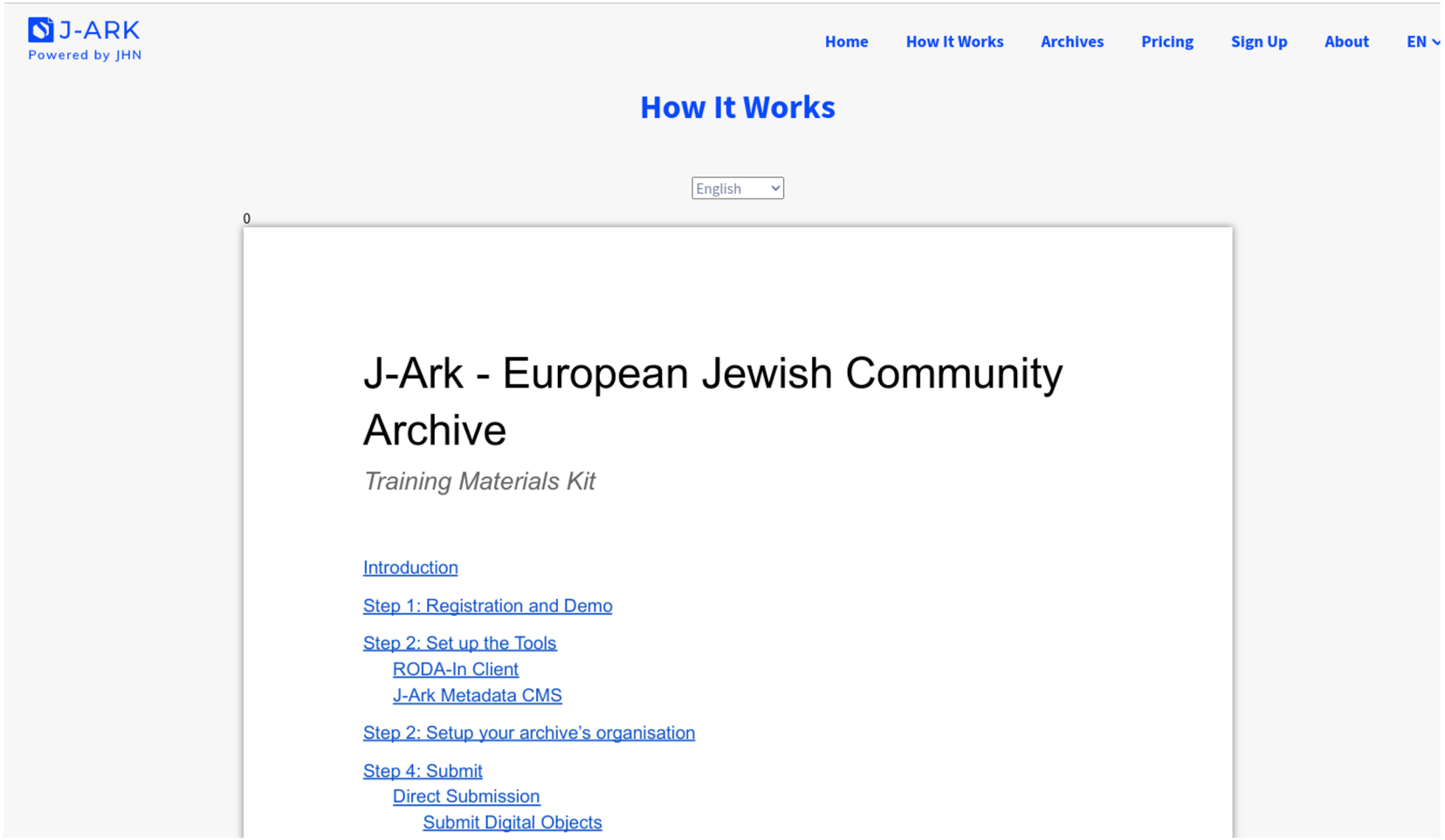




pangeanic

European Commission

Training & Onboarding Materials



Summary

- **Uniqueness of the Project**
  - Connection between three European DSIs / building blocks
- **Service Vision**
  - One solution for small heritage organisations to handle collections
- **Future Thinking**
  - Testbed for integration and data exchange between blocks / domains, crucial. Interesting case for the emerging European Data Space(s)?

Final Event

Preservation of Digital Heritage in the data space

Colleagues from the J-ARK project, European Commission and Europeana Foundation explore the future of digital preservation within the context of the common European data spaces.

pangeanic

PROGRAMME

Setting the Scene (9-10am)

To open the event we'll set the scene by hearing updates about the project from its own team, learning about the vision of the European Commission on eArchiving and common European data space for cultural heritage, and about the role of the Europeana Initiative.

Moderated by Pavel Kats (Jewish Heritage Network), with presentations from

- J-Ark project: vision, components use-cases and sustainability, Alexander Raginsky (Pangeanic)
- The eArchiving initiative, Arpad Welker (European Commission)
- The common European data space for cultural heritage, Fulgencio San Martin (European Commission)
- Common European data space for cultural heritage: update and vision, the role of the Europeana

Innovation (10.05-11.25am)

We'll continue with four sessions on various innovation initiatives in the long-term preservation and archiving domains.

Moderated by Luis Faris, KEEP Systems, with presentations followed by Q&A from

- eArchiving - what it is and how does it work, Anja Paulic (The Archives of the Republic of Slovenia)
- Information hand in hand with sustainable long-term digital preservation, Karin Bredenberg (Sydarkivera)
- Standardising the exchange of Digital Preservation know-how using Preservation Action Registries, Jack O'Sullivan (Preservica)
- Can The Internet Back up itself? SOLID has the answer, Alex Tourski (Post-Platforms) and Dr. Esther de Loof (SolidLabs)

The future (11.30-12.15pm)

The event will end with a round table discussion about the future of digital preservation and big questions our community should address.

Moderated by Pavel Kats (Jewish Heritage Network)

With Luis Faria (KEEP Systems), Kerstin Arnold (Archive Foundation Europe), Manuel Herranz (Pangeanic)

European Commission

Thank you