

artresearch.net

62nd joint meeting of the CIDOC CRM SIG

23–26 March 2026

Artem Kozlov (artem@rem.sh)

Pharos is an international association that currently brings together twelve art **historical photo archives** based in Europe and North America. It is committed to creating a freely available digital research platform that provides interoperable access to millions of photo archive images and associated scholarly documentation.

1.67M

WORKS

2.3M

PHOTOS

1.37M

ARTWORKS WITH
IMAGES

82.1% of all works





2.14M

PHOTOS WITH
IMAGES

93.4% of all photos

Data integration status across Pharos partner institutions

| PARTNER | STATUS |
|--|--------|
|  Bibliotheca Hertziana | Live |
|  Bildarchiv Foto Marburg | Live |
|  Fondazione Zeri | Live |
|  Frick Photoarchive | Live |
|  KHI Florence | Live |
|  Paul Mellon Centre | Live |
|  RKD Netherlands | Live |
|  Warburg Institute | Live |

| PARTNER | STATUS |
|---|---------------|
|  Dumbarton Oaks | On dev |
|  I Tatti, Harvard | New data soon |
|  National Gallery of Art | To be done |
|  Yale Center for British Art | To be done |
| | |
| | |
| | |
| | |

Core components powering artresearch.net

- 1. **ResearchSpace**

- 2. **QLever**

- High-performance RDF/SPARQL database

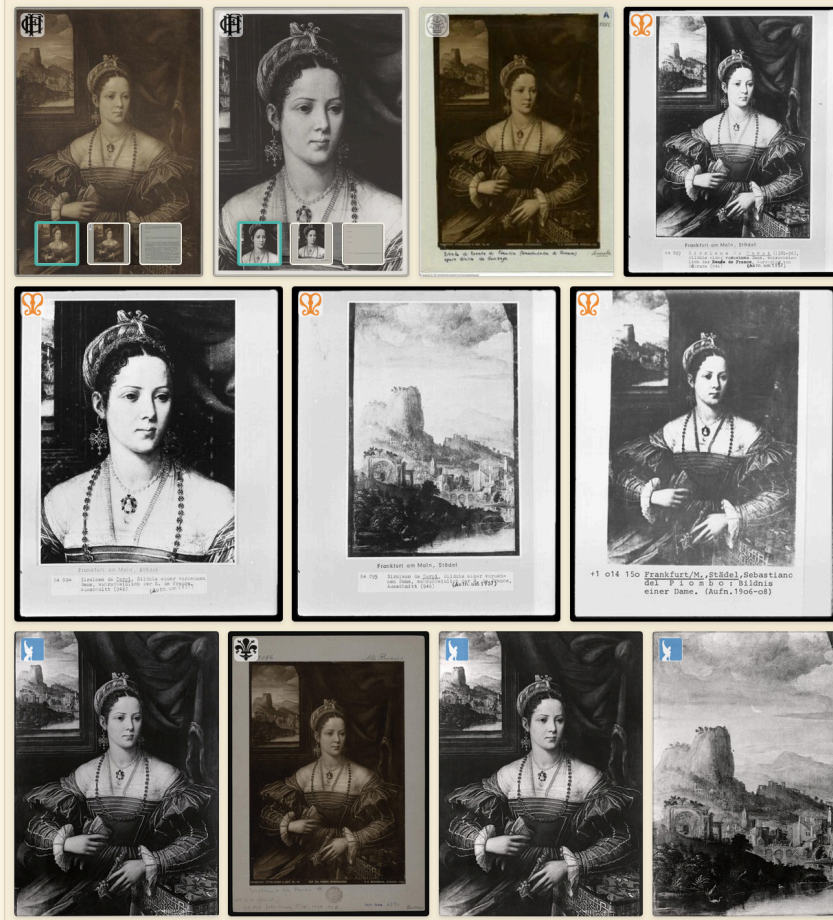
- 3. **USearch — Vector Index**

- Qwen3-8b — embedding model for metadata search
- google/siglip2-so400m-patch16-naflex — embedding model for image search

System is alive only when it is used

Photos Similar Works

Visual Clusters



Portrait of a Lady.

3 more titles from 5 institutions

CREATOR / ATTRIBUTION

GIROLAMO DA CARPI

- **Girolamo da Carpi**
- **Girolamo da Carpi, attributed to**

CAMPAÑA, PEDRO DE

- **Campaña, Pedro de**
- **Campaña, Pedro de, attributed to**
- **Campaña, Pedro de, ?**

ALTERNATIVE ATTRIBUTION

PARMIGIANINO

- **Parmigianino** earlier attribution
- **Parmigianino, ?**

SEBASTIANO DEL PIOMBO

- **Sebastiano del Piombo**
- **Sebastiano del Piombo, attributed to**

SODOMA · **Sodoma**

Reconciling artist names across institutions via E13 Attribute Assignment



1. SPARQL Endpoint

- Publicly available at artresearch.net/sparql

2. Data Model Documentation

- WIP documentation available but not properly published yet
- gist.github.com/aindlq/e37944f24b6ea349ad61bd326170826b

1. **Data is not static**
2. **Original schema and data model also changes**
3. **It is not enough to look into representative example, one needs to analyze the whole dataset**
4. **Over time people change their mind about what is important and what is not**

```
-- a2750 - Teilbauwerk-Name (Name of the Part of the Building)
-- Names the specific part of the building (e.g., "Sala del Mascherone").
when (or_ [exists [x|a2740|], exists [x|a2750|]]) (
  P46 ---> (E22, relativeUriT "/place/part/{name}_{type}" [("name", [x|a2750/text()|]), ("type", [x|a
  P2 --> (E55, P.built_work_part),
  [x|a2750|] @> appellation_0_1 P.preferred_name [x|normalize-space(text())|],
  [x|a2740|] @> e55_type "builtwork_part_type",

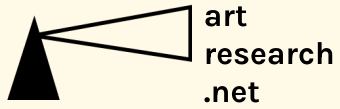
  -- main work is part of this specific builtwork part
  P46 --> (E22, typedNamedUri work),

  P156 ---> (E53, relativeUri "/place") ==> [

    -- a2780 - T-B-Stelle (Placement of the Building Part)
    -- Pinpoints the exact location within the named part (defined by a2740/a2750)
    -- where the object is found (e.g., "Ostwand," "Kuppelpendentifs").
    [x|a2780|] @> (
      P89i ---> (E53, relativeUriT "/sub-place/{name}" [("name", [x|normalize-space(text())|])]) ==
      appellation_0_1 P.preferred_name [x|normalize-space(text())|],

      -- location of the work
      P55i --> (E22, typedNamedUri work)
    ]
  ]
)
```

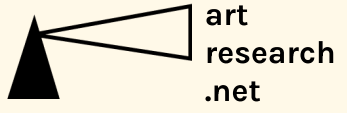
1. **Inspired by George Bruseker's notation**
2. **It is a Haskell code**
3. **Compile time checks on XML schema level and CIDOC-CRM properties**
4. **Very fast — mapping of 100k XML records in 10 seconds**
5. **Works great with AI coding assistants (Claude Code, Antigravity)**



Pharos welcomes new partner institutions — particularly **photographic archives** — interested in providing interoperable access to their collections through a shared CIDOC-CRM-based research infrastructure.

For partnership inquiries and further information:

info@artresearch.net



NLnet Grant

Vector based similarity search index for QLever database

→ nlnet.nl/project/QLever-similarity

"... the tendency to get carried away by the ordinary words, and to let the words suggest issues that are not really there. One typical example is the use of the word 'knowledge' in recent computer work. The very word knowledge, being a noun in the grammatical sense, suggests that knowledge is something, some substance, had by people. Then it follows immediately that this something may be represented in our computers, and we have '**knowledge representations**'. The lure of words has been known by many thoughtful people, including Otto Jespersen, the linguist, and William James and Gilbert Ryle, the philosophers. The matter comes up in many of the present writings... I avoid the word knowledge, talking instead of **human knowing, an activity.**"

— Peter Naur, *Computing: A Human Activity* (1992)