ATRIUM

Archeologický ústav AV ČR, Brno

Workflow pro automatické rozpoznávání archeologických nálezů pomocí umělé inteligence (AI)

Workflow for Automated Archaeological Artifact Recognition using AI

Petr Pajdla & Ronald Harasim

Motivation and objectives

- **Manual processing** of archaeological archives is a tedious, time consuming, and **challenging task** prone to human errors and inconsistencies…
- This makes it a prime candidate for **automation**.
- Archives of **ARUP** & **ARUB** hold around:
    - **~140k** documents, ie. **~1.5m** pages
    - **500k** of various photographs
- **Improve/enhance** metadata in the Archaeological map of the Czech Republic and ARIADNE services.

(Imgaes: Archive of the Institute of Archaeology, Czech Acad. Sci, Brno)

ATRIUM

Advancing Frontier Research in the
Arts and Humanities

- Horizon EU project
- Focus on workflows and demonstrators

*“**Facilitating access** to digital **research infrastructures** and advancing frontier **knowledge in the arts and humanities** — across **disciplines, languages** and **media.**”*

Transnational Access – Travel Grants

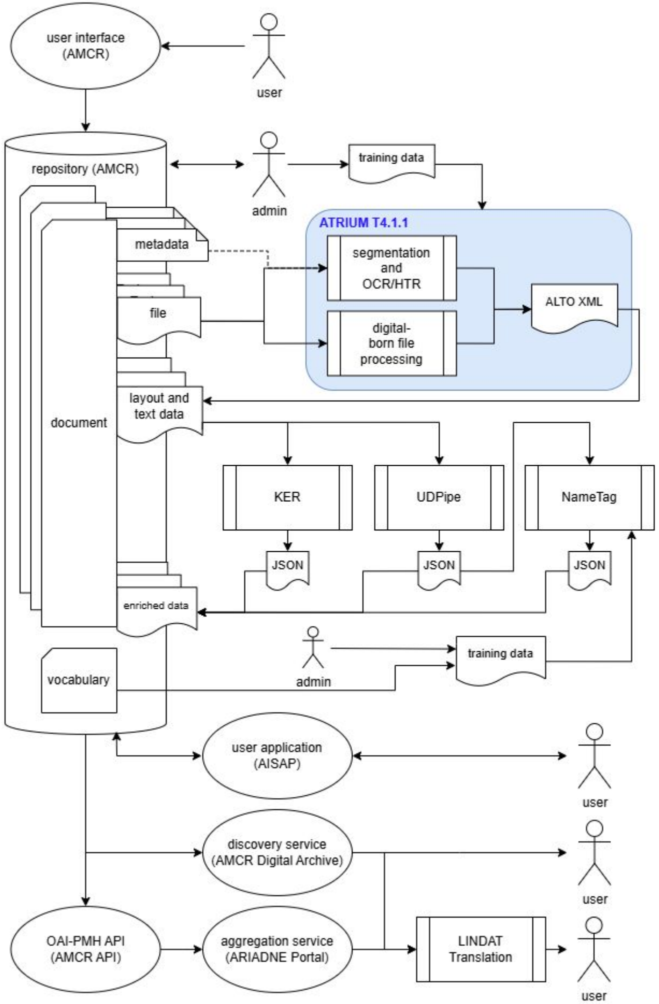
- Current call for applicants is open until 31 May, 2025!
  http://atrium-research.eu

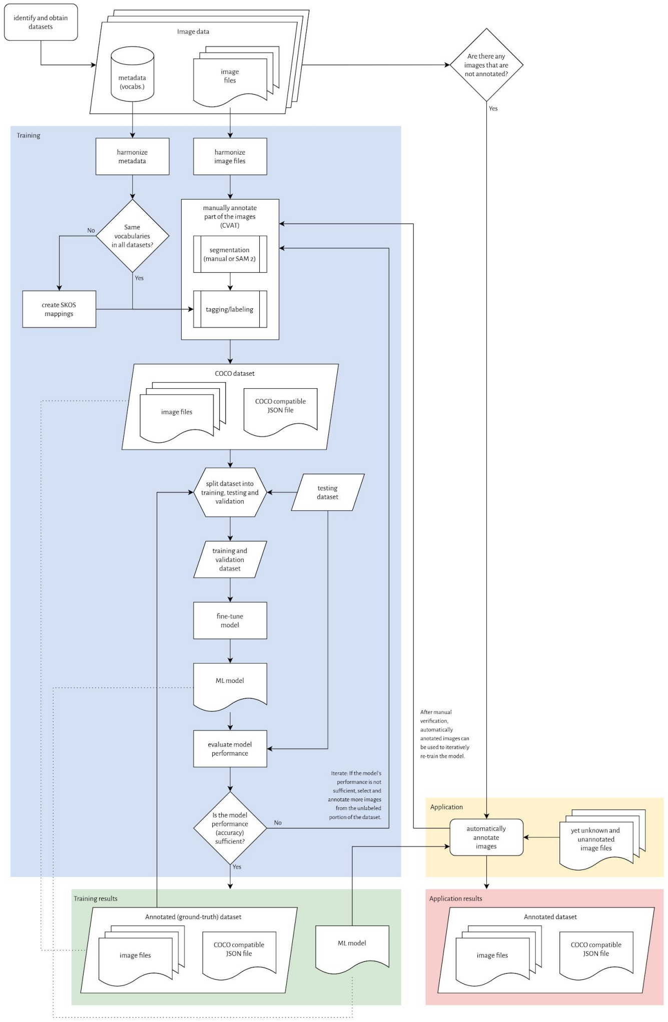
Workflows overview

- **Text-based** workflow

Model for page classification https://github.com/ufal/atrium-page-classification
(details in the poster by Dana Křivánková and Kate Lutsai)

- **Image-based** workflow

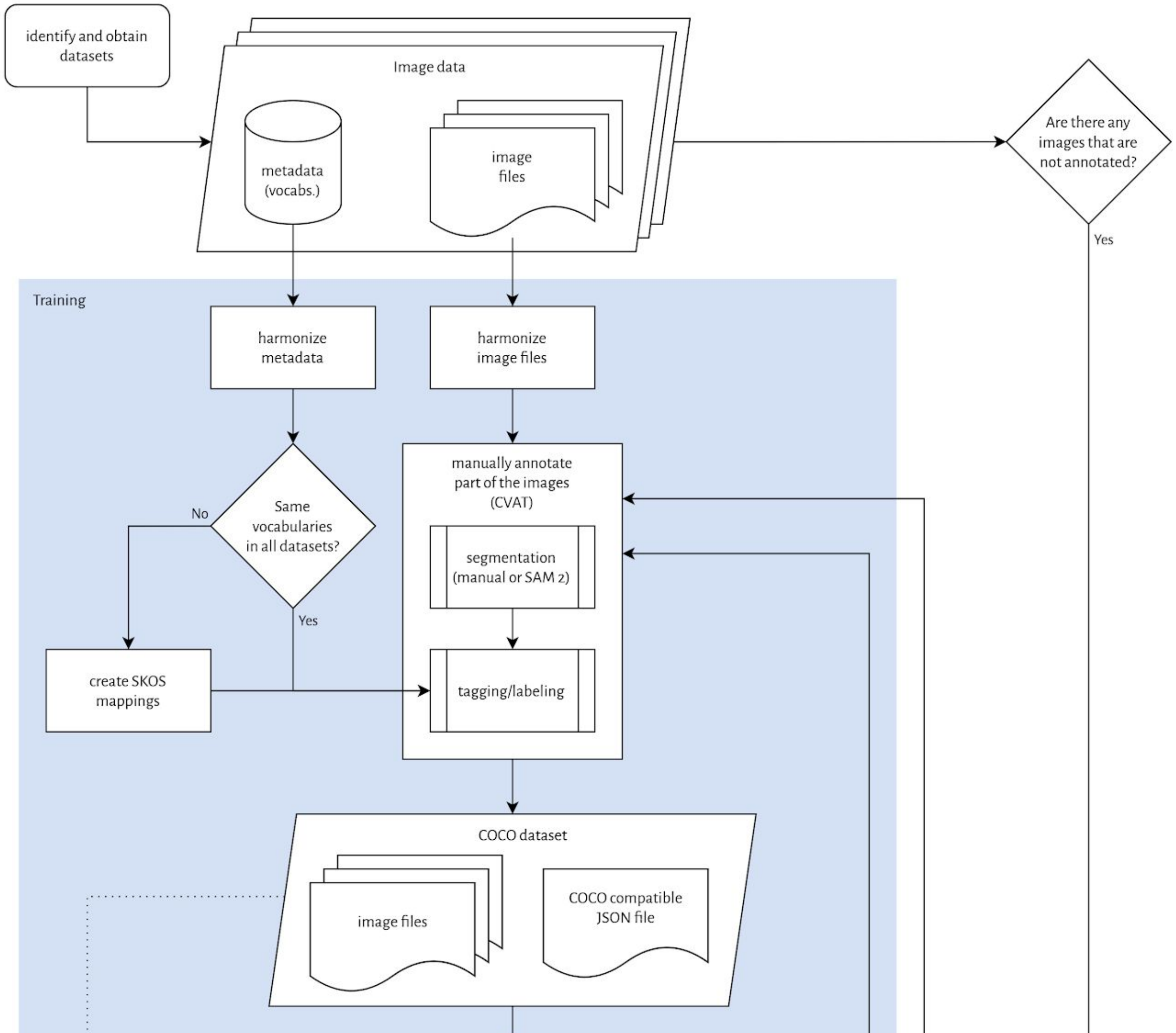




Automatic image annotation workflow

Goals:

- **Automatic annotation** of images (photographs) with terms from controlled vocabularies
- Enhanced **findability**, enriched **metadata**
- Object **extraction** etc.



(1) Consider the goal and use-case

(Image DOI: 10.60585/M-FT-110736000)

"This archaeological excavation reveals an ancient burial with human skeletal remains surrounded by stone artifacts and pottery fragments. Most notably, a prominent iron sword lies centrally alongside the body, suggesting a warrior or high-status burial. A stone headrest or marker is visible at the far end of the rectangular pit."

(Claude 3.7 Sonnet)

- ***Viewing* using LLMs**
  - Description in natural language followed by vocabulary matching
  - Optionally output in machine-readable format (eg. JSON)
  - *Expensive* in terms of computing resources and/or price
- ***Traditional* computer vision models**
  - Versatile and lightweight models
  - *"Garbage in, garbage out"*, does not hallucinate
  - *Expensive* in terms of time/workforce invested in the dataset creation

ATRIUM

(2) Gather and harmonize image data

Image data *come in all shapes and sizes…*

- Transform to a common bitmap format (JPEG)
  - TIFFs might be challenging (transparency, multi-page files etc.)
- 3-channel input expected (RGB)
- Further optimization depends on the model selection…

```python
def process_tiff_image(img: Image.Image) -> Image.Image:
    """Process TIFF images, handling multiple frames and bit depths."""
    # Count frames
    try:
        n_frames = 0
        while True:
            img.seek(n_frames)
            n_frames += 1
    except EOFError:
        pass

    # Return single-frame TIFF as is
    if n_frames <= 1:
        img.seek(0)
        return img

    # Find best frame (with most content)
    best_frame, max_std = 0, 0
    for i in range(n_frames):
        img.seek(i)
        std_dev = sum(abs(px - 128) for px in img.convert('RGB').convert('L').getdata()) / (img.width *
        if std_dev > max_std:
            max_std, best_frame = std_dev, i

    img.seek(best_frame)
    return img
```

(3) Align controlled vocabulary terms

- skos:mappingRelation
  - skos:closeMatch
  - skos:**exactMatch**
  - skos:broadMatch
  - skos:narrowMatch
  - skos:relatedMatch

https://marketplace.sshopencloud.eu/workflow/yLX6Bd

```python
def convert_image_mode(img: Image.Image) -> Image.Image:
    """Convert image to appropriate mode for JPEG saving."""
    original_mode = img.mode

    # Convert based on mode
    if original_mode in ('I;16', 'I'):
        img = img.point(lambda i: i * (255.0 / 65535.0)).convert('L')
    elif original_mode == 'F':
        img = img.convert('RGB')
    nal_mode == 'P' and 'transparency' in img.info):
    55, 255, 255))
    nel('A'))
```

W3C Recommendation

W3C®

SKOS Simple Knowledge Organization System Reference

W3C Recommendation 18 August 2009

This version:

http://www.w3.org/TR/2009/REC-skos-reference-20090818/

ATRIUM

Datasets /1

*Photographs of **finds***

- *AMCR-PAS*
  - **>7.5k** finds, **~10k** photographs
- *“Lovec pokladů”* (*“Treasure hunter”*)
  - **>300k** finds
  - photographs by metal detectorists
  - quality varies greatly
- *Portable Antiquities Scheme*
  - **>750k** finds (with images)
  - high quality of images
- *Montelius dataset*
  - includes drawings and images scanned from published resources





(Images: AMCR-PAS, Lovec pokladů, Portable Antiquities Scheme)

ATRIUM

Datasets /2

*Photographs of **fieldwork***

- *ARUP digital photographs of fieldwork collection*
  - **>60k** photographs
- *ARUP & ARUB archival photographs*
  - **~300k** photographs
  - various scenes, fieldwork, people, finds, sites etc.
  - mostly b/w photographs

(Images: Digital Archive of the Archaeological Map of the Czech Republic)















ATRIUM

(4) *Annotate images*

- Various **annotation platforms**...
  - CVAT (Computer Vision Annotation Tool)
    https://www.cvat.ai/
  - VIA (VGG Image Annotator)
    https://www.robots.ox.ac.uk/~vgg/software/via/
- BBoxes vs **segmentation**
- Segment Anything Model (**SAM2**)
  - Semantic segmentation
  - https://ai.meta.com/sam2/
- **COCO** compatible **JSON**

- Annotation of **AMCR-PAS dataset**
  - **9764** images, **7303** annotated (**~65 %**)
  - **7426** objects (annotations)
  - speed: **~50** objects per hour


```
"annotations": [
    {
        "id": 1,
        "image_id": 1,
        "category_id": 1,
        "segmentation": [
            [
                555.0,
                682.0,
                548.0,
                683.0,
                529.0,
                697.0,
                518.0,
                701.0,
                492.0,
                717.0,
                429.0,
                766.0,
                382.0,
                857.0,
                361.0,
                910.0,
                361.0,
                931.0,
                354.0,
                949.0,
                342.0,
                1026.0,
                350.0,
                1038.0,
                383.0,
                1057.0,
                416.0,
                1059.0,
                431.0,
                1065.0,
                438.0,
                1063.0,
                462.0,
                1063.0,
                512.0,
                1046.0,
                553.0,
                1046.0,
                586.0,
                1041.0,
```

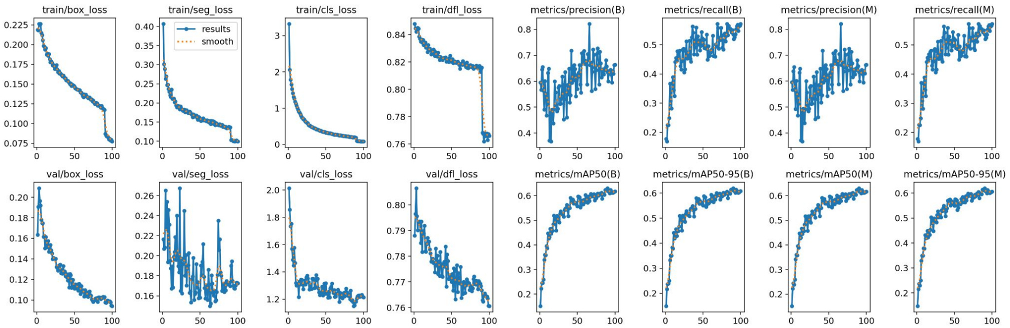

(Image DOIs
10.60585/M-FT-110736000 and
10.71928/M-202300087-N00394)

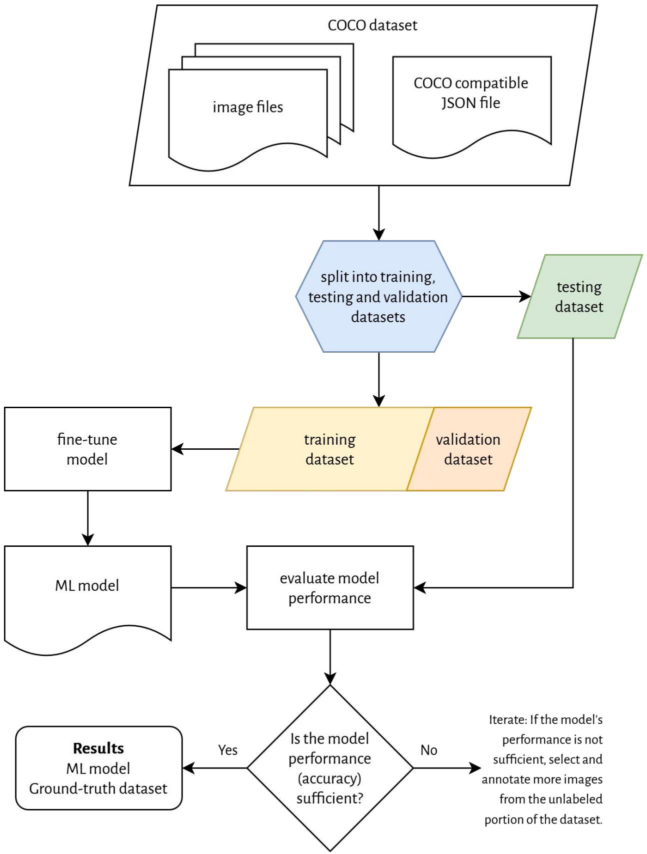
ATRIUM

(5) Re-train/fine-tune the model

- Split into **training**, **validation** and **testing** datasets
  - **Imbalanced** dataset – proportional representation of classes...

- **YOLO v11 seg.**, 100 epochs, PyTorch
  - Best results with **340x340px** resolution
  - Detection (BBox) mAP50–95 ≈ 0.61
  - Segmentation (mask) mAP50–95 ≈ 0.61



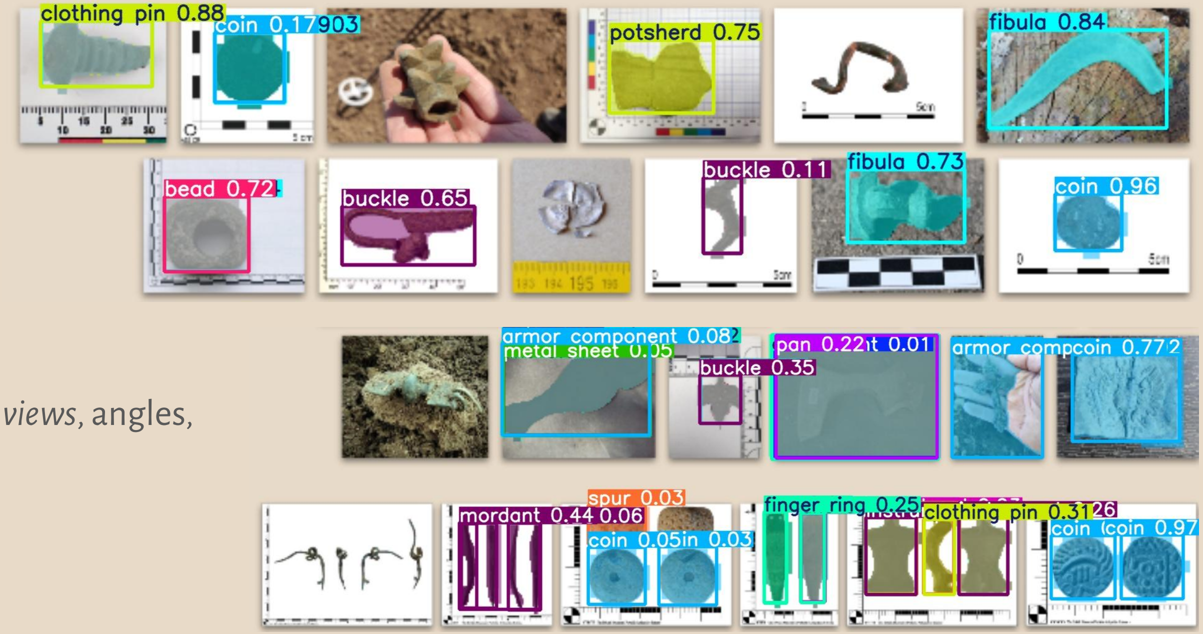


ATRIUM

(6a) Apply the ML model

Apply the model, verify the results, iterate...

- First *good* version of the model helps with further annotations...
- *Ground-truth* dataset grows

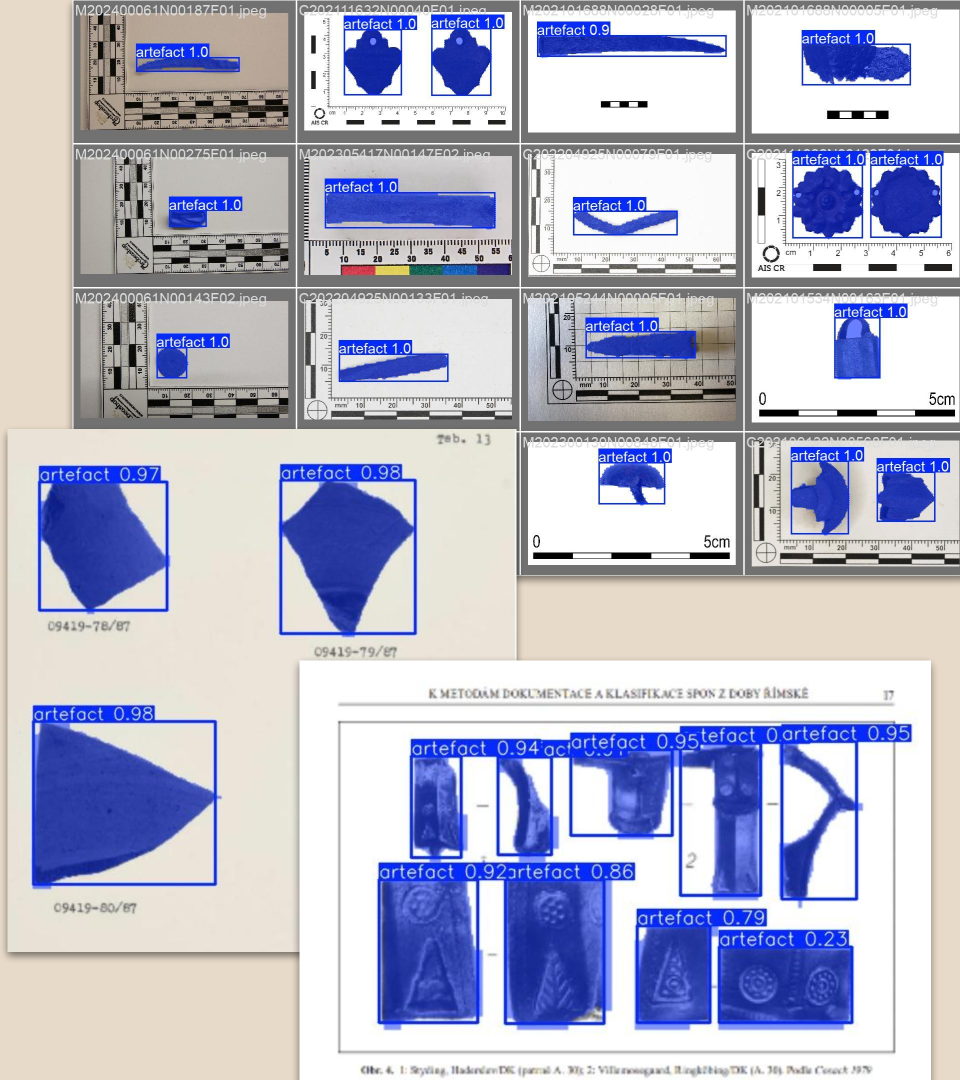


TODOs:

- Cope with **imbalanced dataset**
- Cope with **variability** of different *views*, angles, sizes etc.
- Enlarge *ground-truth* dataset...

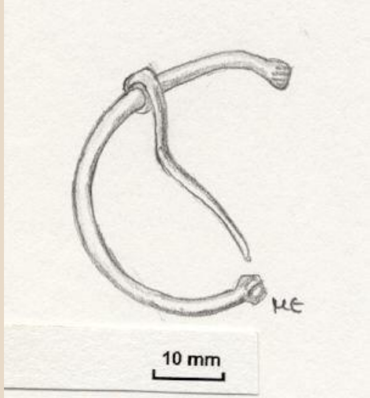
(6b) Artifact Database

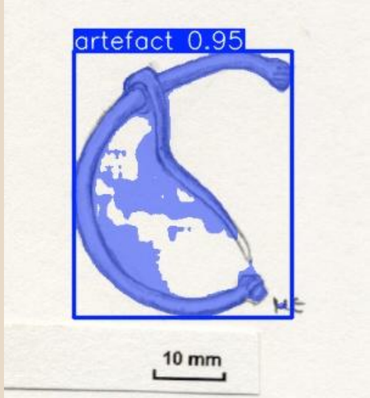
1. **Input:** Photographs containing artifacts.
2. **Detection & Segmentation:** A pre-trained model automatically finds and precisely outlines (segments) artifacts in the photographs.
3. **Artifact Extraction:** Identified artifacts are "cut out" from the photographs as separate images.
4. **Vectorization (Visual Transformer):** Each extracted artifact is processed by a Visual Transformer, which creates a unique semantic vector (a numerical "fingerprint" of the artifact).
5. **Database Storage:** All vectors representing the artifacts are stored in a specialized vector database.



(6b) Searching for similar Artifacts

1. **Input (Query):** A new image of an artifact for which we want to find similar ones.
2. **Query Processing:**
   a. The input image undergoes the same **detection, segmentation, and extraction** process as during database creation.
   b. The extracted artifact from the query is then converted into a **vector** using the same Visual Transformer.
3. **Database Comparison:** The query artifact's vector is compared against all vectors in the artifact database.
4. **Output:** Display of artifacts from the database whose vectors are mathematically most similar to the query vector.





ATRIUM

Sidenote: LLM Fine-Tuning & Deployment

Models used:

- Qwen 2.5 VL
- Gemma 3 27B

Our Local Setup:

- **Hardware:** In-house NVIDIA RTX 4090 GPUs (24GB VRAM)
- **Fine-tuning Engine: Unsloth** library (for efficient training)
- **Serving Framework: Ollama** (for local model utilization)

Focus: Tailoring advanced LLMs for our specific applications using our own infrastructure.

unsloth

MAiA

Managing Artificial Intelligence in Archaeology

*'MAIA is here to **investigate** Artificial Intelligence (**AI**) **applications** within **archaeology**.'*

- https://www.maiacost.eu/
- COST Action CA23141
- > 200 members...
- At least 3-4(?) **surveys** planned...
- Training **school** on **data management for AI applications** at **ADS** (York, 2nd week of September?)

Working groups:

1. State of the Art: AI and Archaeology
2. Digital Comparative Collections and AI Training Data for Archaeology
3. AI and Archaeological Research
4. Dissemination and Communication

ATRIUM

ATRIUM

Thank you for listening!

For more information, please visit https://www.atrium-research.eu/

Funded by
the European Union

Petr Pajdla
pajdla@arub.cz
Czech Academy of Sciences, Institute of Archaeology, Brno

Ronald Harasim
Czech Academy of Sciences, Institute of Archaeology, Brno
harasim@arub.cz

...and special thanks to the ARUB & ARUP annotation team!
Kristína Rašlová, Zuzana Kopáčová, Eva Buchtová, Tomáš Chlup
David Spáčil, David Novák, Olga Lečbychová