ATRIUM

Od pixelů ke znalostem

Automatizace popisu archeologických fotografií pomocí počítačového vidění

Petr Pajdla, Ronald Harasim, David Novák, David Spáčil, Olga Lečbychová

Archivy, knihovny, muzea v digitálním světě (AKM) 2025 (26. – 27. 11. 2025)

ATRIUM

Advancing Frontier Research in the Arts and Humanities

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

- Horizon EU project
- Focus on workflows and demonstrators

Transnational Access – Travel Grants

Individual access & Summer schols

Current call for applicants is open until **31 December, 2025**!

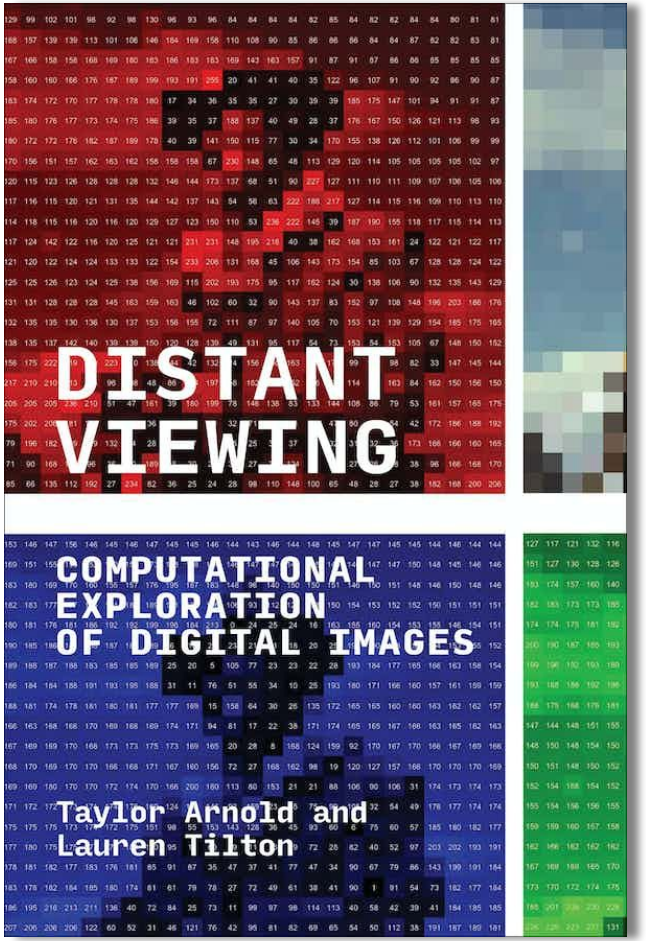
http://www.atrium-research.eu

See poster for more information https://doi.org/10.5281/zenodo.17434302

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:
  - **ca. 140k** documents, ie. **~1.5m** pages
  - **ca. 500k** of various photographs
- **Improve/enhance** metadata in the Archaeological map of the Czech Republic and ARIADNE services.

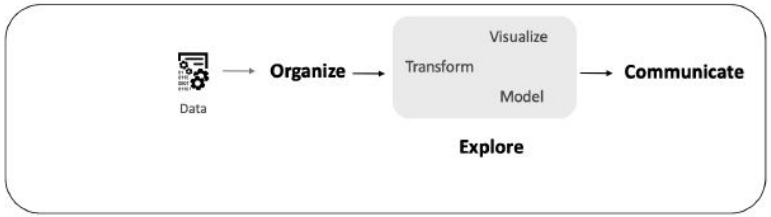
(Images: Archive of the Institute of Archaeology, Czech Acad. Sci, Brno)



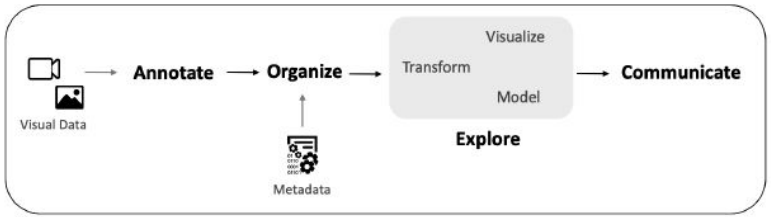
Distant viewing

*"The process of **distant viewing** applies **computer-vision algorithms** to **automatically interpret** a layer of **meaning within images** through the creation of **structured annotations**."*

Arnold & Tilton 2023: Distant Viewing. Computational Exploration of Digital Images. p. 25



Data Science Pipeline (Structured Data)

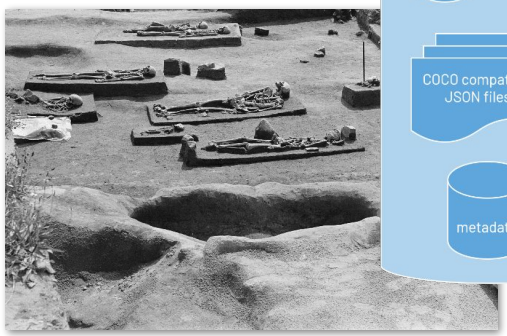


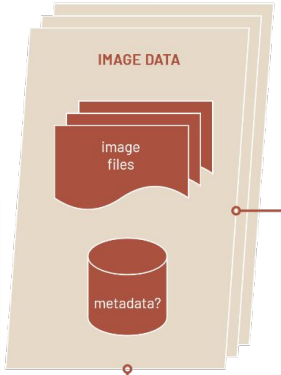
Pipeline for Distant Viewing

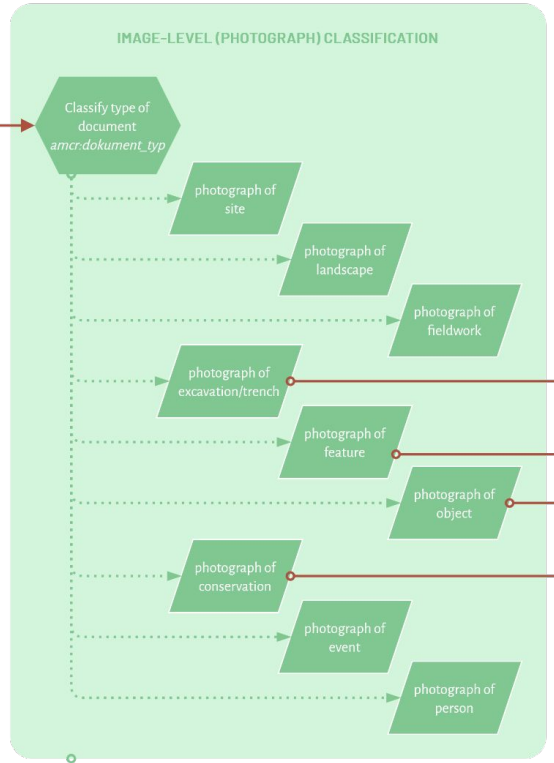
Fig 2.1, p 35

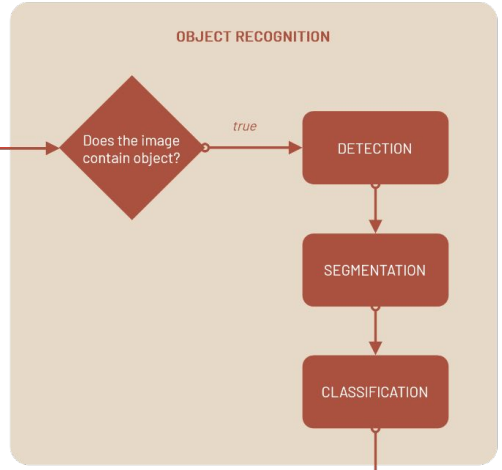
ATRIUM

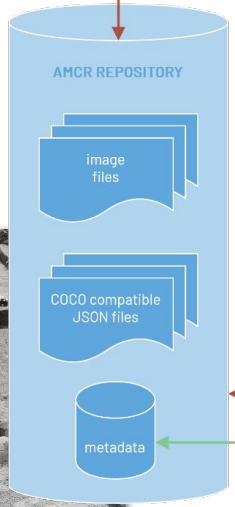
Workflow











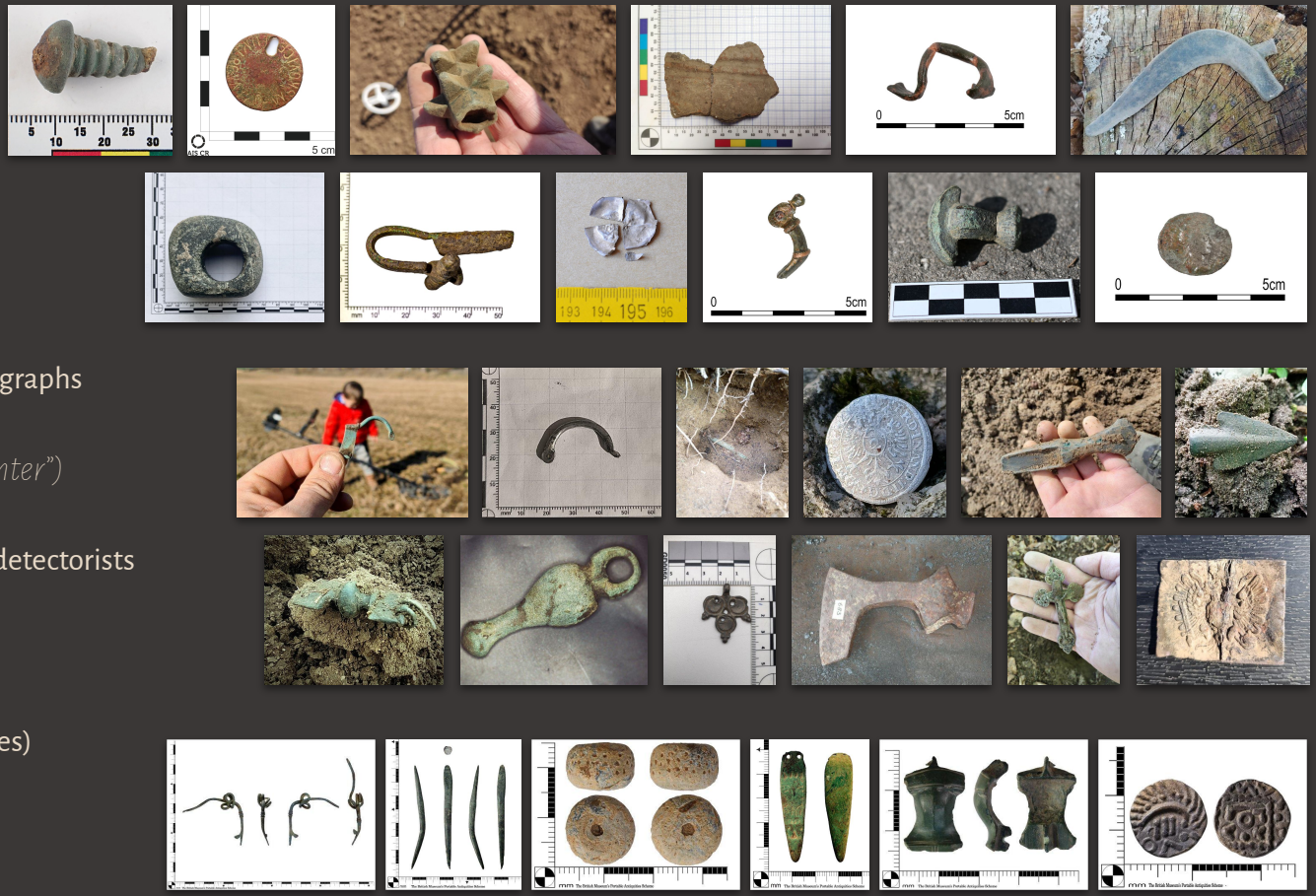
(Image DOIs: 10.60585/M-FT-110736000, 10.60585/M-FT-660173300, 10.60585/M-FT-900144845)

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)
- *Montelius dataset*
  - includes drawings and images scanned from published resources



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

ATRIUM

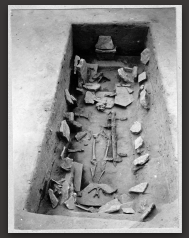
Datasets /2

*Photographs of **fieldwork***

- *ARUP digital photographs of fieldwork collection*
  - **>60k** photographs

- *ARUP & ARUB archival photographs*
  - **~300k** photographs
  - various scenes, fieldwork, people, finds, sites etc.
  - a lot of legacy b/w photographs

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





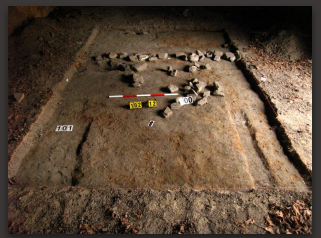












ATRIUM

Automatic Image Annotation Workflow

Copy to clipboard

Steps to fine-tune pre-trained image recognition model for domain-specific applications

This workflow outlines a process for fine-tuning a pre-trained image recognition model to enhance its ability to recognize specific object categories that are underrepresented or entirely absent in its original training dataset. The primary goal is to create a lightweight machine learning (ML) model capable of annotating images using terms from domain-specific controlled vocabularies. This facilitates more accurate and consistent image annotation in specialized contexts.

The workflow serves the dual purpose of improving the model's performance on domain-specific data and streamlining the image annotation process. By iteratively combining manual annotation, automated annotation using the fine-tuned model, and model re-training, the workflow supports efficient creation of

Details

ACCESS

License Creative Commons Attribution 4.0 International MIT License

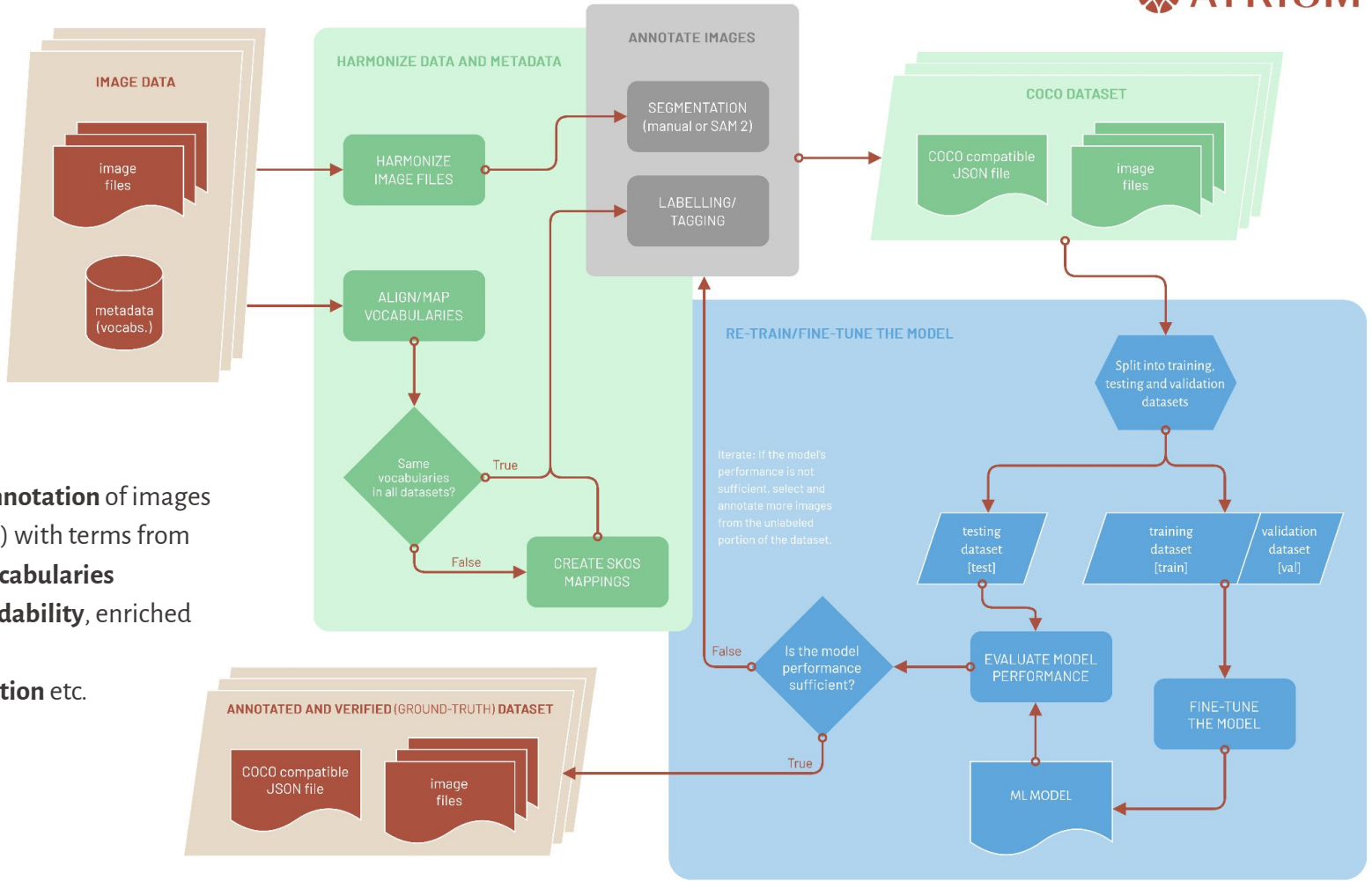
CATEGORISATION

Activity Annotating Segmenting Tagging Editing Machine Learning Pattern Recognition Storing

https://marketplace.sshopencloud.eu/workflow/G6ck4w

ATRIUM

Training workflow



Goals:

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

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
```

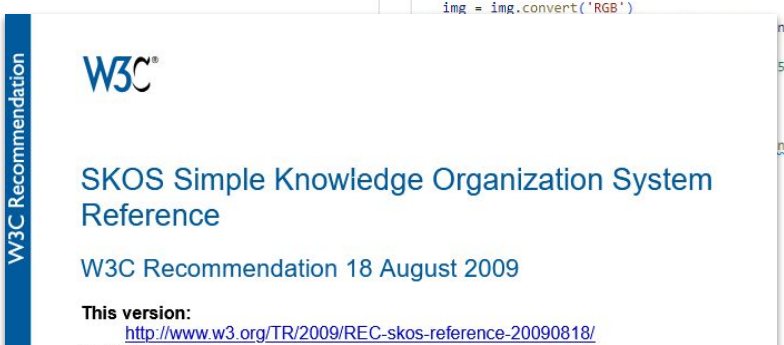
```python
def convert_image_mode(img: Image.Image) -> Image.Image:
    """Convert image to appropriate mode for JPEG saving."""
    original_mode = img.mode

    # 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'))
```

Alignment of controlled vocabulary terms

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

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



ATRIUM

Annotation & segmentation

- CVAT (*Computer Vision Annotation Tool*) https://www.cvat.ai/
- Make Sense AI (https://www.makesense.ai/)

- Segment Anything Model (**SAM2**) Semantic segmentation https://ai.meta.com/sam2/

- Result: **COCO** compatible **JSON**

- Annotation of **AMCR-PAS dataset**
  - **9 762** images, **9 469** annotated (**~97 %**)
  - **10 749** objects (annotations)
  - speed: **~50** objects per hour

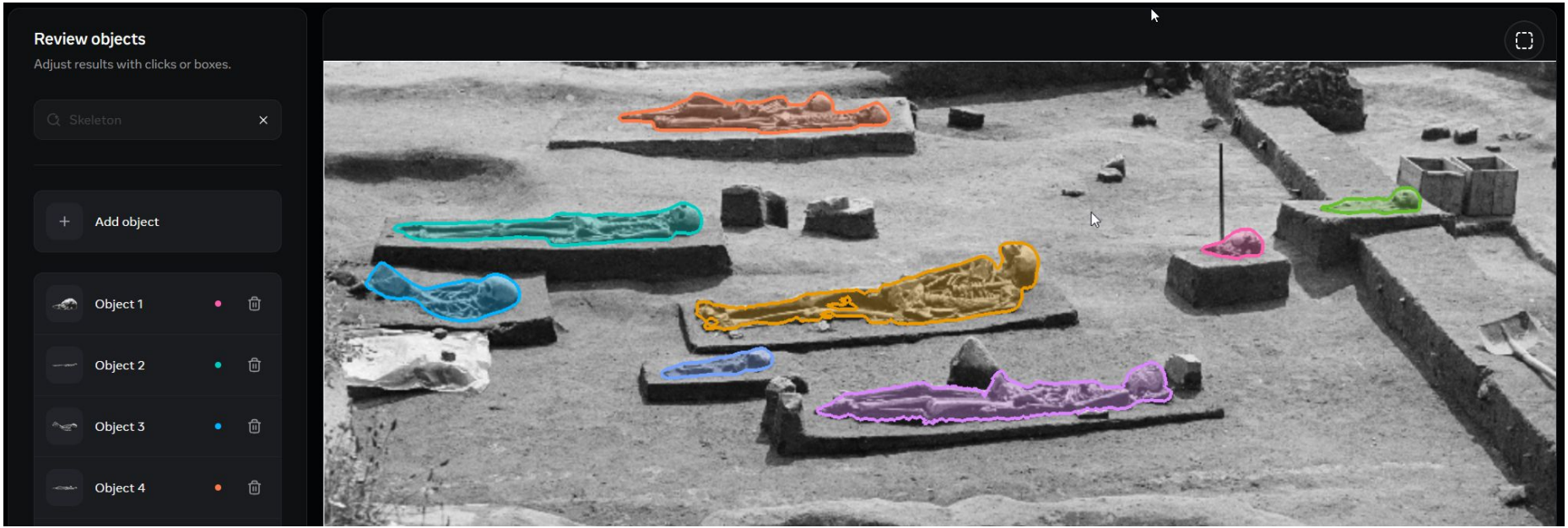

```
"annotations": [
    {
        "id": 1,
        "image_id": 1,
        "category_id": 1,
        "segmentation": [
            [
                555.0,
                682.0,
                548.0,
                683.0,
                529.0,
                697.0,
                518.0,
                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

Sidenote

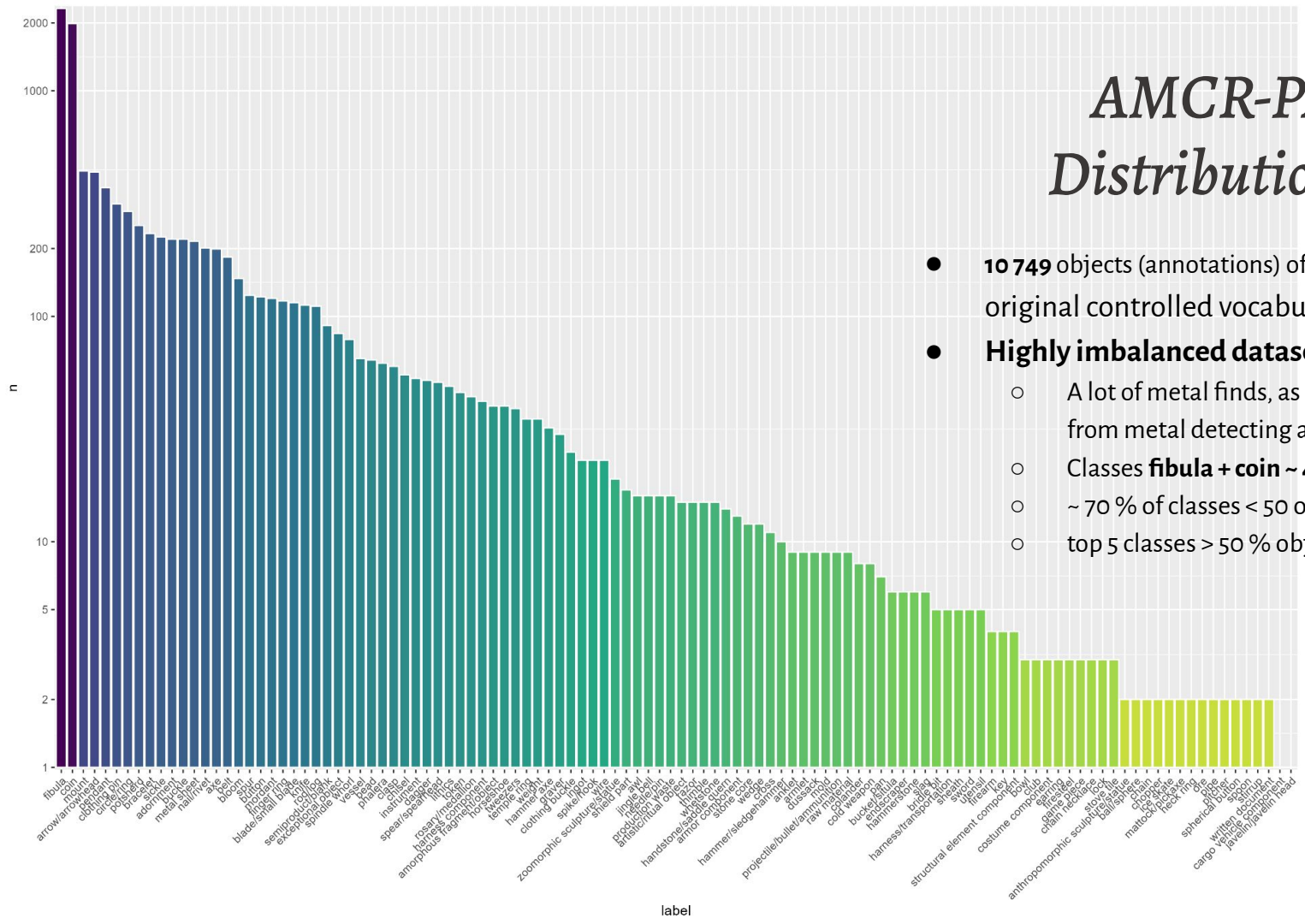
SAM3 (*Segment Anything Model*) by Meta (November 2025, https://ai.meta.com/sam3/) → text prompt



https://aidemos.meta.com/segment-anything

ATRIUM

AMCR-PAS dataset: Distribution of classes



- **10 749** objects (annotations) of **112 classes**, original controlled vocabulary has *207 classes*...
- **Highly imbalanced dataset**
  - A lot of metal finds, as the dataset records finds from metal detecting activities.
  - Classes **fibula + coin ~ 40 % of the dataset**
  - ~ 70 % of classes < 50 objects
  - top 5 classes > 50 % objects

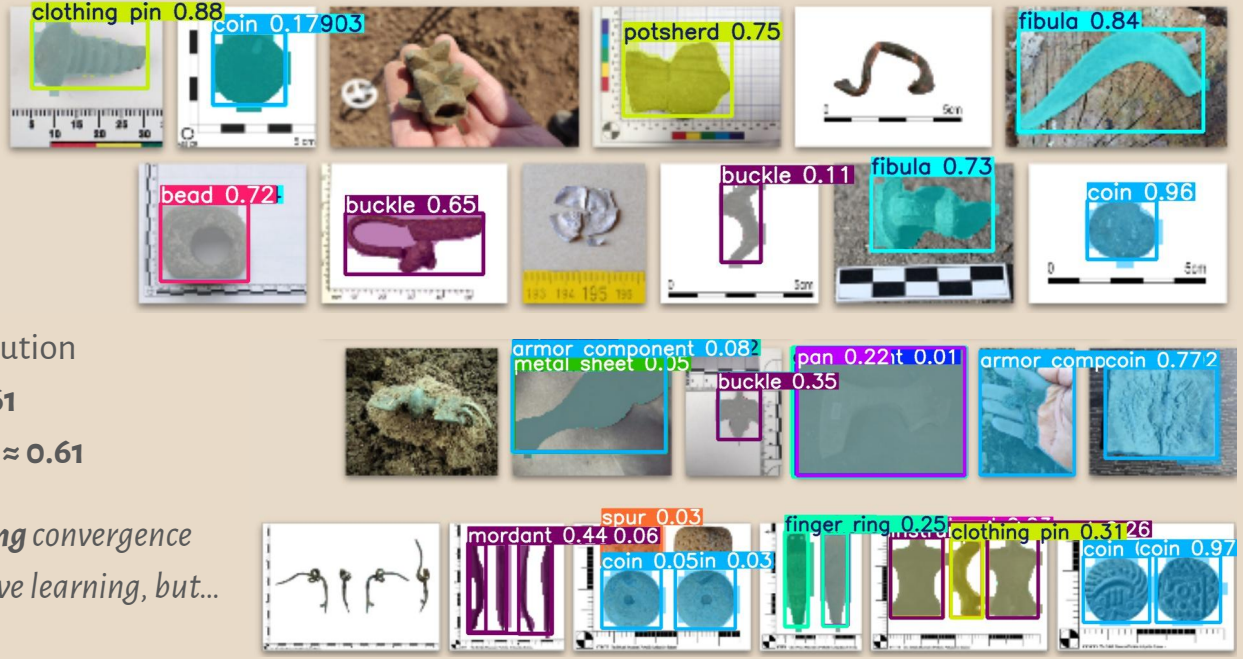
ATRIUM

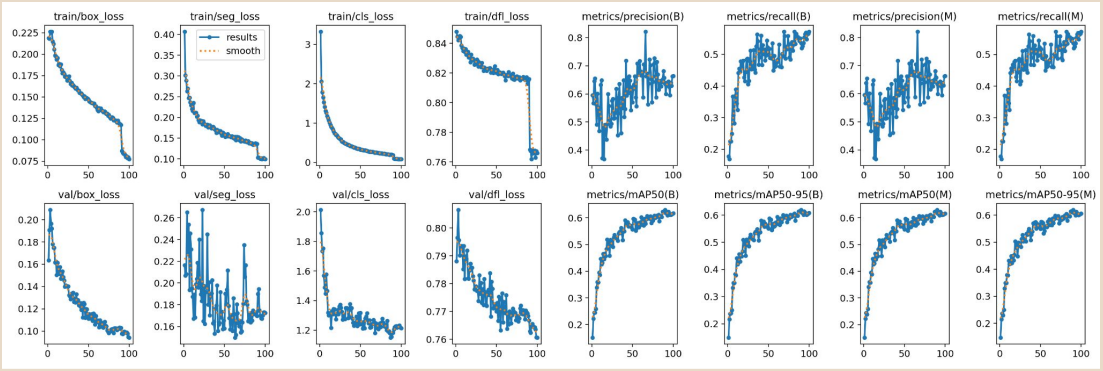
Training

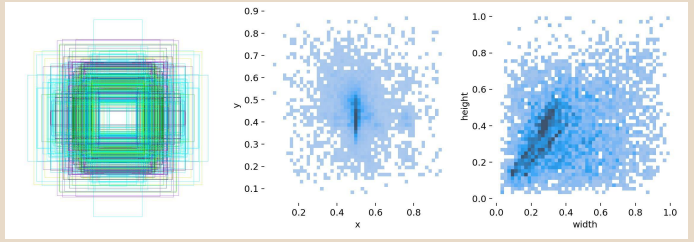
YOLO v11 seg., 100 epochs, PyTorch

- Best results with **340x340px** resolution
- Detection (BBox) **mAP50–95 ≈ 0.61**
- Segmentation (mask) **mAP50–95 ≈ 0.61**

*Pilot training sessions demonstrate **promising** convergence with steady loss reduction, indicating effective learning, but... the dataset is highly **imbalanced**!*



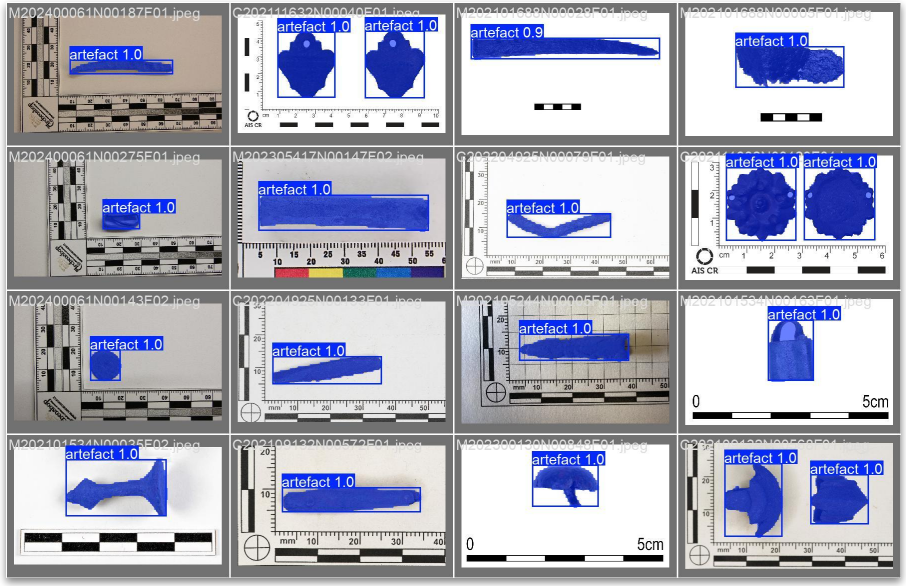




ATRIUM

Next steps...

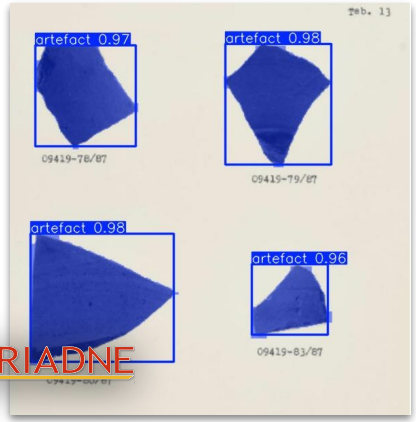
- **Continue annotation** of the datasets
  - ML models help with further annotations
  - *Ground-truth* datasets **grow**
- Deal with **imbalance** in the datasets
  - Data augmentation strategies/synthetic data
  - Class-aware sampling
- Cope with **variability** of angles and viewpoints
  - A fibula viewed from a *side* and from *top* looks very differently...
- *proceed, iterate...*

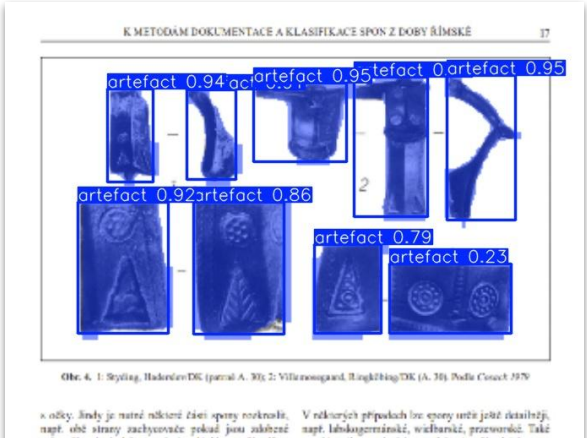


Planned results:

- *Workflow* – already published ✓

  https://marketplace.sshopencloud.eu/workflow/G6ck4w
- *Data processing 'pipeline' in* Archaeological map of the Czech Republic
- *Datasets* – published at the end of the project?
- *Code* – utilities for dataset curation etc., *soon?*
- *Enriched metadata in* Archaeological map of the Czech Republic *and* ARIADNE





ATRIUM

Archaeological Information System of the Czech Republic

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

Ronald Harasim
Czech Academy of Sciences, Institute of Archaeology, Brno

David Novák
Czech Academy of Sciences, Institute of Archaeology, Prague

David Spáčil
Czech Academy of Sciences, Institute of Archaeology, Brno

Olga Lečbychová
Czech Academy of Sciences, Institute of Archaeology, Brno

Thank you for listening!

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

Funded by the European Union

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

ARUB Institute of Archaeology Czech Acad Sci, Brno

INSTITUTE OF ARCHAEOLOGY OF THE CAS PRAGUE