Project QVIRE

Investigating the structure of medieval manuscripts with the help of AI

Evina Stein
Martin Tamajka

May 12, 2023

Evina Stein

Academic limbo, Beirut
(Manuscript studies)

Martin Tamajka

Lead Research Engineer,
KInIT
(Artificial Intelligence)

Juraj Šedivý
Comenius University

Marcel Veselý
KInIT

Peter Griger
KInIT

Marián Šimko
KInIT

QVIRE project

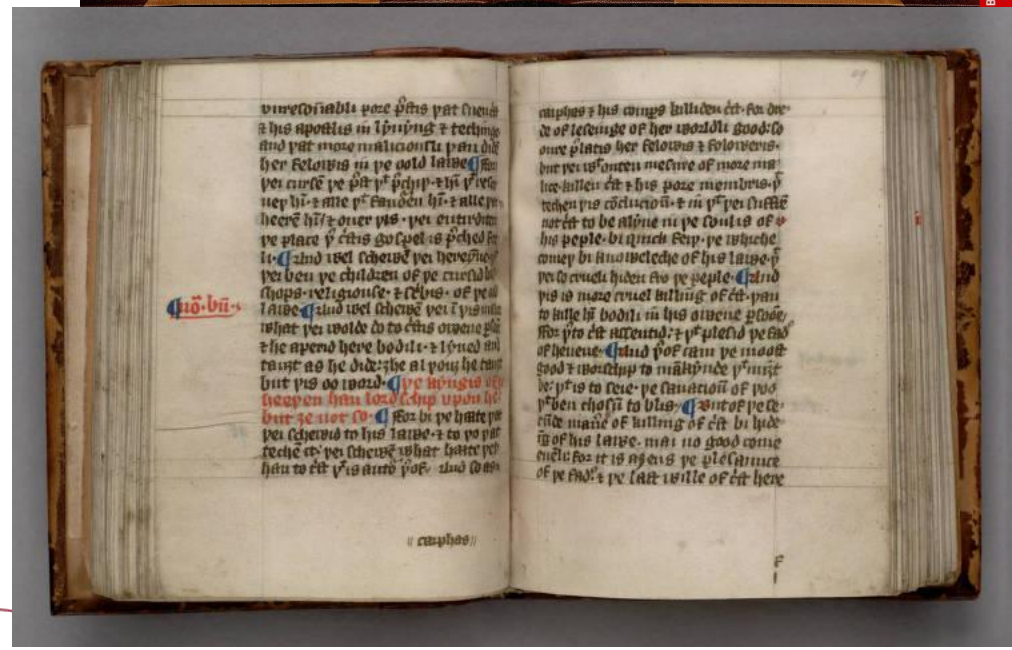
- November 2022 – March 2023 (**5 months**) at Kempelen Institute for Intelligent Technologies (KInIT) in Bratislava
- **10,000€** from the Grant of the Prime minister of the Slovak Republic for Innovation (2022)
- Objective: develop an AI model and prototype software for harvesting of **assembly cues** from digital manuscript facsimiles

What are assembly cues?

Assembly cues in the Latin West

At the beginning/end of a gathering (quire)

- **Quire marks:** mark the position of a quire in a sequence (e.g., by Roman numerals, letters, dots, etc.)
- **Catchwords**: a word copied at the end of a quire that corresponds to the first word of the following quire
- **Leaf signatures**: indicate the position of (bi)folia within a quire



What do we know about assembly cues so far?

- We rely on ‘inherited factoids’ about the medieval usage of assembly cues
- Traditional focus on manuscript content rather than the materiality of the codex meant that assembly practices were never seen as a valuable historical phenomenon or data source in their own right
- No systematic examination of assembly cues (both due to the traditional scholarly perspective and the difficulty of data collection)
- Only three limited case studies: E.A. Lowe (1928), Jean Vezin (1967), and Elena Rodríguez Díaz (1999)

> Other methods of marking the sequence of quires are to use the letters in a name,[25] or a number of points,[26] or even more exotic systems.[27] A further aid for the bookbinder are the catchwords (Reklamanten) written under the last line or word of a gathering, indicating the first word or syllable with which the first page in the next gathering began. This usage is found already in Spain in the tenth century, in France and Italy in the eleventh, and is later found everywhere.[28]

- B. Bischoff, *Latin Palaeography* (1990), p. 23

Novel (digitally-enabled) perspective

Assembly cues an example of a data source that can be fully harnessed only within a new, digitally-enabled research, **because rather than deep qualitative research on small scale, their study calls for 'shallow' quantitative data collection on a large scale**

- Large-scale digitization
- AI-assisted (semi-)automated data collection
- quantitative analysis

The potential of assembly cue research

1. **As diagnostic features**
   - Traditional research indicates that particular assembly cue types are characteristic of certain periods, regions, and contexts of production (e.g., prevalence of alphabetical quire marks in the EMA -> Roman numerals in the HMA -> catchwords in the LMA)
2. **Providing important insights into Western manuscript cultures**
   - What prompted the switch from quire marks to catchwords, where, and when?
   - Why catchwords spread and eventually replaced quire marks in the Latin West?

Additional benefits of (semi-)automated cue detection

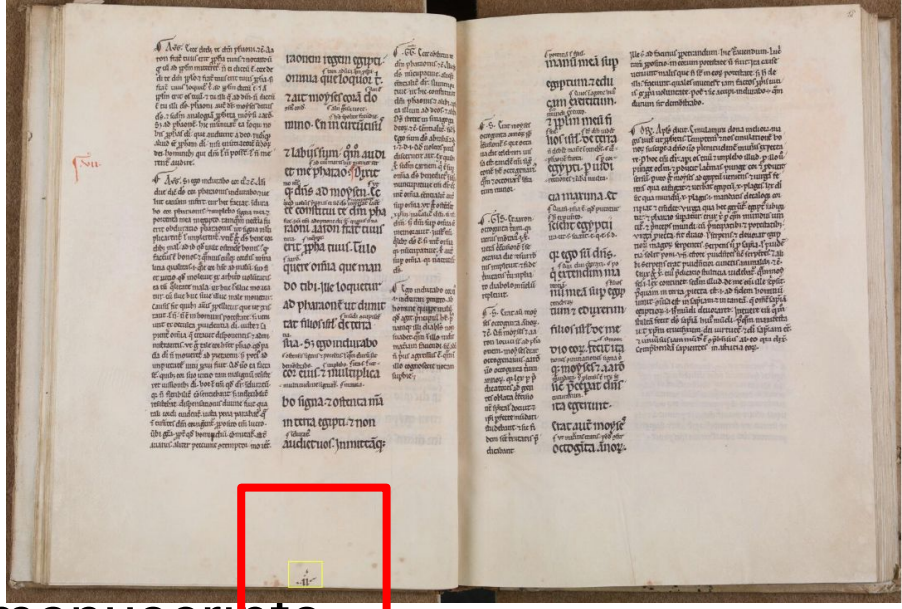
Insights into the material structure of manuscripts from digital images

- Since assembly cues code the beginning/end of quires, they allow us to gain a provisional view of the material structure of manuscripts without the need to examine them in person
- (Semi-)automated detection of physically irregular manuscripts (e.g., composite manuscripts, damaged manuscripts, manuscripts with added quires)
- (Semi-)automated codicological formulas? Or something different?

membr., de quibus vid. infra). I (2–3, typis impressa) + II (4–7) + 1 fol. (8) + (XIV + I + 1 fol.) {(9–11/14–17/19–39) + (12–13) + 18} + (XXII + I) {(40–51/54–85) + (52–53)} + 2 fol. (86–87) + XIII (88–113) + IX (114–131) + III (132–137) + 2 fol. (138–139) + 2 IV (140–155) + 2 I (156–159) + 2 fol. (160–161) + I (162–163). F. 41, 56, 59, 60 marg. inf. desectus est.

An example of a codicological formula of Leiden, Universiteitsbibliotheek, Voss. Lat. F 37

The (modest) project goals



- Create an AI model (neural network)
  - For detection of assembly cues
  - in a selected corpus (Saint-Omer) of digitized manuscripts (== images) with IIIF manifests
- Create software
  - That uses this AI model
  - to analyze historical manuscripts
  - and make it available to the community

Selecting a suitable corpus

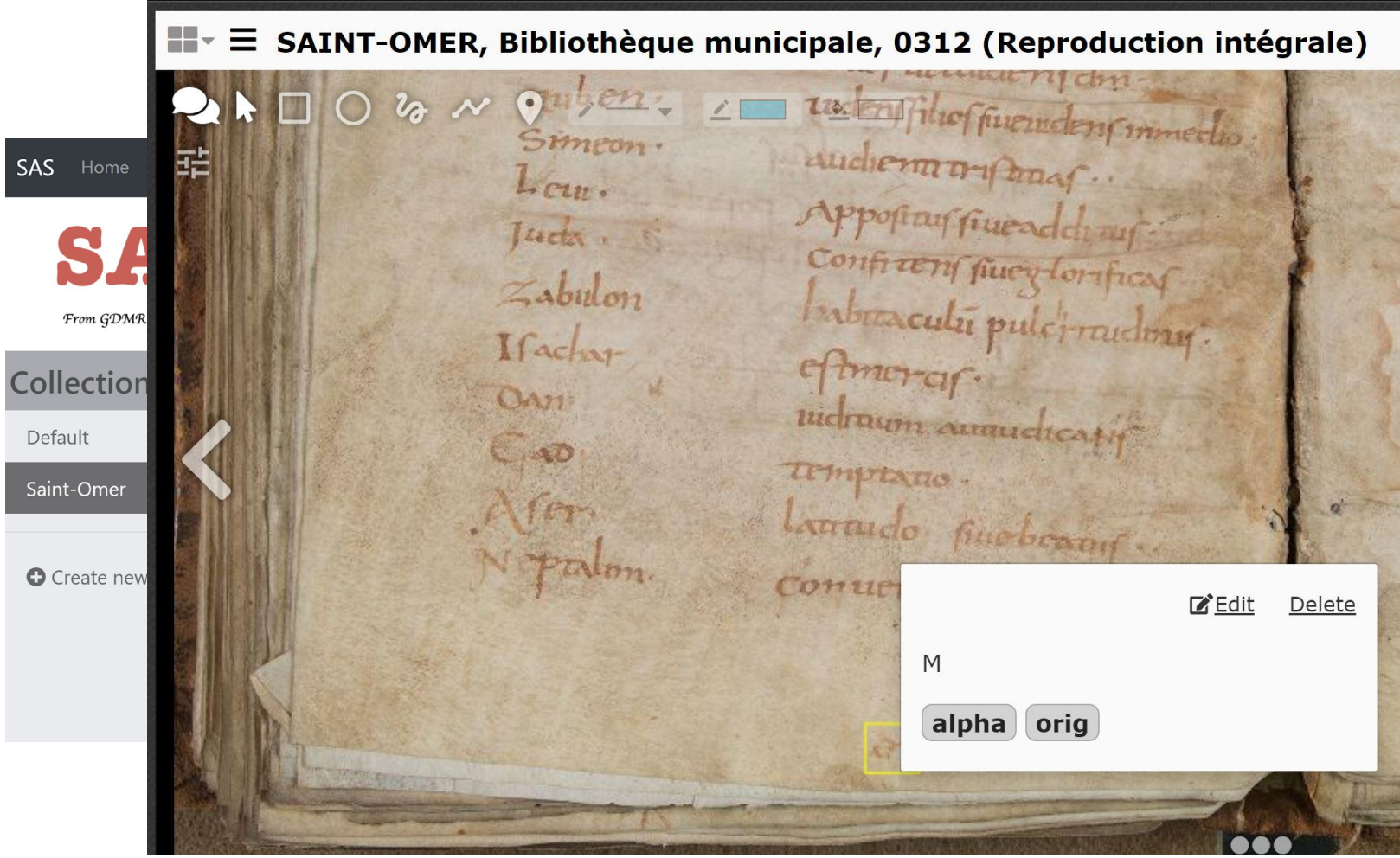
Medieval manuscripts from the municipal library of Saint-Omer (701 mss.)

- IIIF manifests readily available via BVMM
- manuscripts from all historical periods of the Middle Ages (chronological spread)
- Manuscripts mostly produced regionally (geographically anchored)

We would like to thank Dominique Stutzman (IRHT/CNRS) for consultation and assistance with data selection

Training data

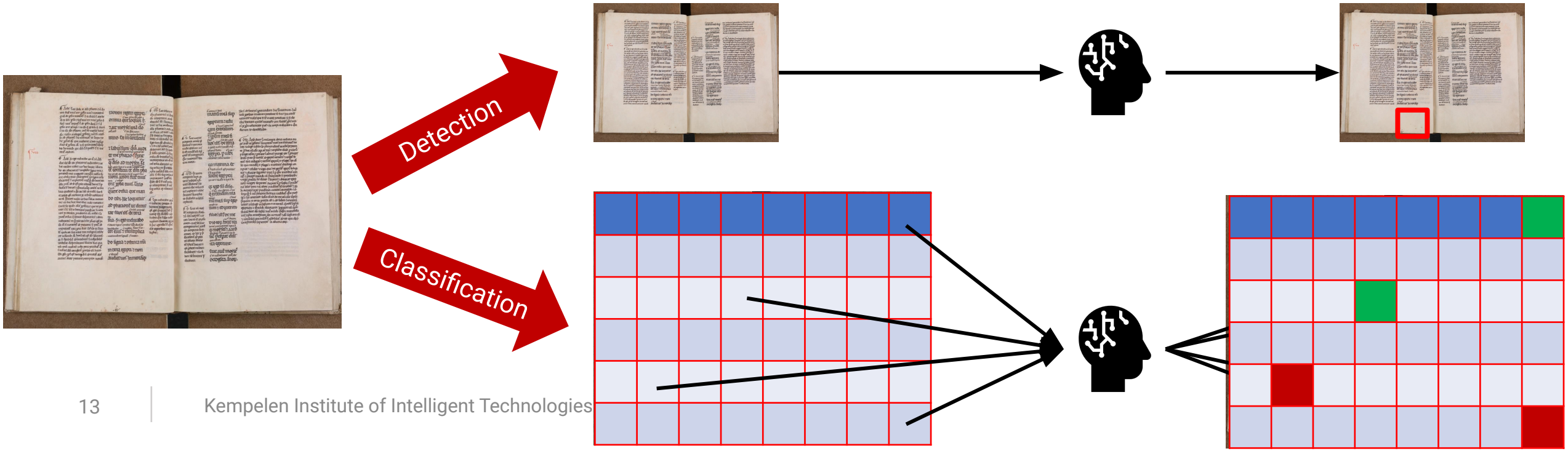
- 67 manuscripts from Saint-Omer containing assembly cues (~10 % of the corpus)
- Workflow managed with the Simple Annotation Server (SAS)
- IIIF manifests annotated in Mirador 2
  - identified by framing
  - Labelled using controlled vocabulary
  - transcribed quire marks and leaf signatures



We would like to thank Rastislav Luz (Slovak National Archive, Bratislava) for his work generating part of the training data

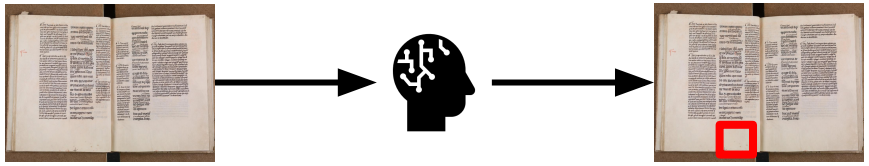
Basic information

- We wanted to train a specific model – Deep neural network
- We experimented with two approaches:



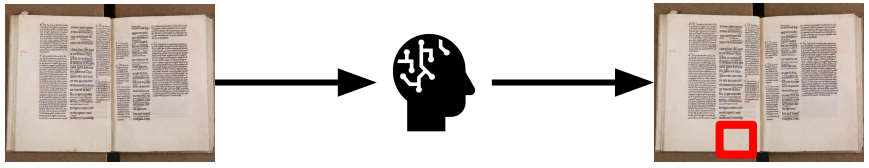
Approach I: Detection

- Original size of manuscript images tend to be around 5000x3700 (18.5 MP)
- Images are down-sampled to 1066x800 (~1MP)
  - To speed up training and to make this data compatible with existing detection models
- **Tech detail:** We used **Faster-RCNN** neural network
  - Plus, we made certain preprocessing



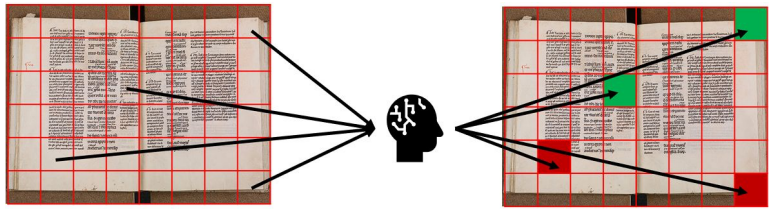
Approach I: Detection
Problems of this approach

- Down-sampling the image might be too drastic, relevant information can be lost The **quire marks can be tiny**.
- Very limited amount of training data
  - 59 manuscripts ~11.000 pages
  - ~1700 pages with quire marks

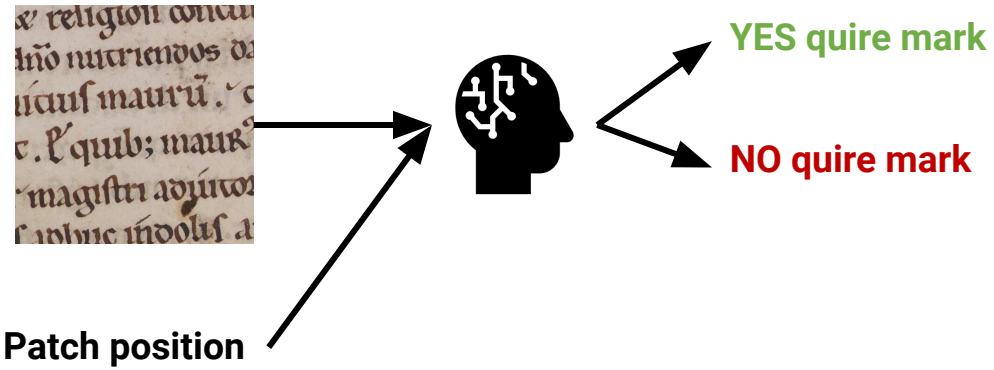


Approach II: Classification

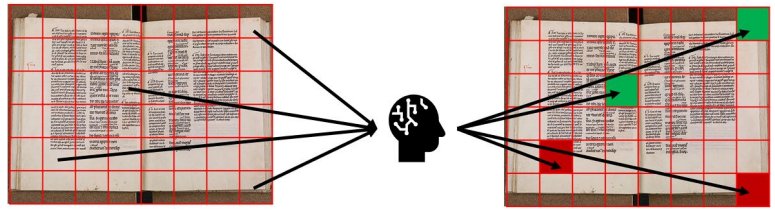
- Split up each image (page) into overlapping patches of size 512x512
  - Then, down-sample to 256x256
- Then, classify each patch separately
- **Tech detail**: We used **Resnet** neural network architecture



Approach II: Classification Problems of this approach



- All patches are classified separately
  - No global information
  - Where is the patch located in the page? Quire marks are usually close to edges.
  - We used a trick Together with patch (small image), we added **information about its position** within page as an input to the AI model



Results and final approach selection

We evaluated both models in terms of following criteria:

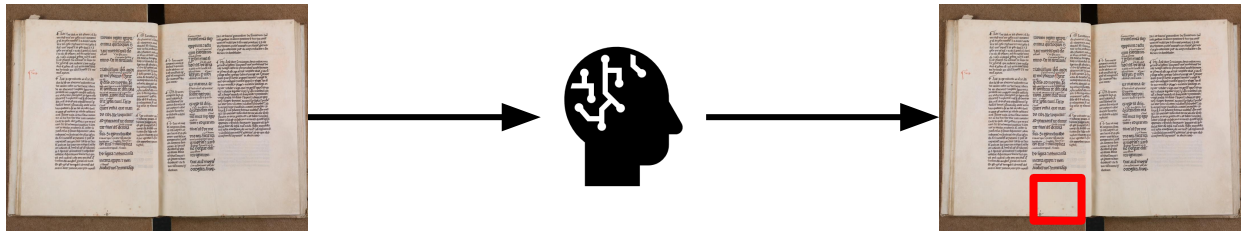
1. How many of the real quire marks did the model detect (we want it **high**)
   - Ideally, we want to detect all the quire marks in the manuscripts
2. How many “false” quire marks did the model detect (we want it **low**)
   - Ideally, we want the model to detect only real quire marks – no “hallucinations”, no false positives

In terms of statistics, we measured Precision and Recall of the models

Results and final approach selection

Winner: **Detection-based approach** (best ratio between Precision and Recall)

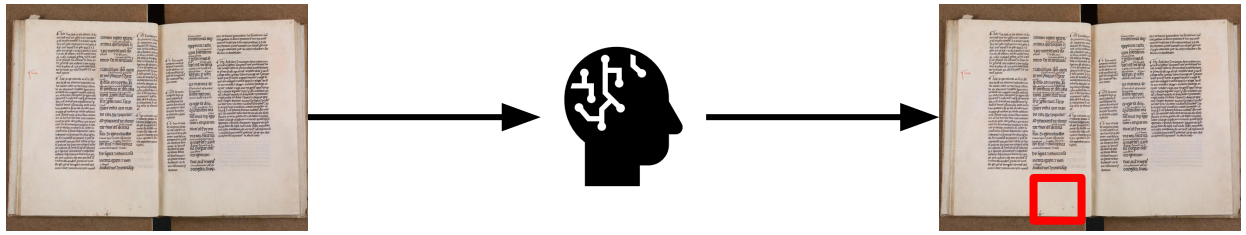
- Recall: 0.47
  - From 100 quire marks, the model detected 47 and missed 53
- Precision: 0.47
  - From 100 quire marks detected by the model, 47 were really quire marks. The rest were false positives == “hallucinations”.



Results and final approach selection

Winner: **Detection-based approach** (best ration between Precision and Recall)

- These numbers might seem low, but:
  - Very low amount of data
  - High variability of data
- We consider this result promising in terms of real-world use
  - Even if only 50% of quire marks from a single manuscript are detected, we can use these to make some assumptions (especially with big corpora)
  - Plus, with increasing number of annotated data the model will become better

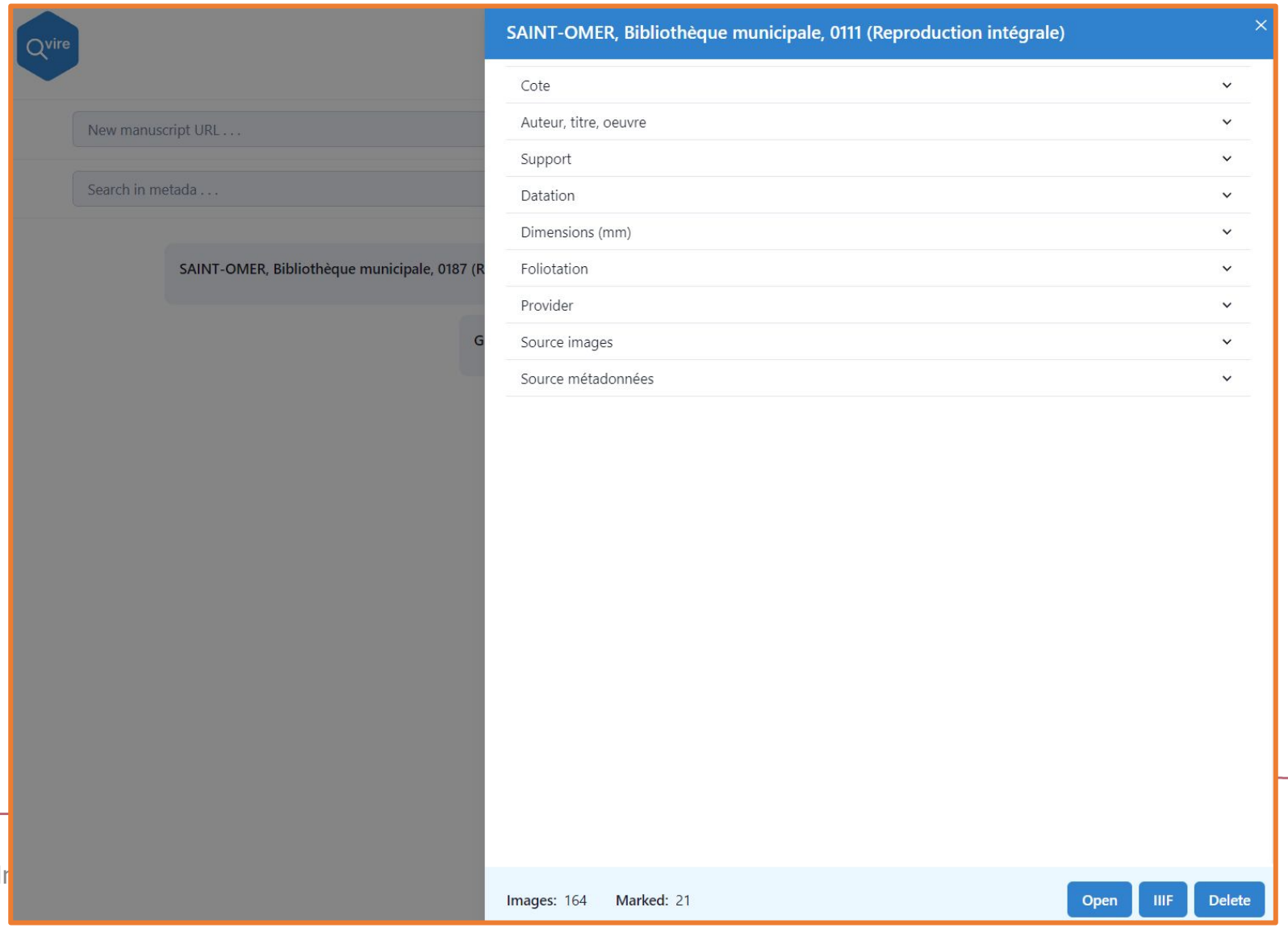


Tool prototype

http://108.143.37.22/

Tool prototype

http://108.143.37.22/



Tool prototype

http://108.143.37.22/

Future plans

- QVIRE conceived as a pilot to test feasibility of AI-assisted approach to precision data harvesting related to assembly cues
- We would like to apply for further funding to expand and improve the model
- Goal: to study assembly cues in different manuscript cultures (e.g., Greek, Hebrew, and Arabic manuscripts)

Thank you for your attention!

Project website:
https://kinit.sk/research/qvire-cultural-ai/

Here, you will find:

- Code
- Model and Data with annotations

Generated by Stable Diffusion 2.1

KINIT

Kempelen Institute
of Intelligent Technologies

Sky Park Offices
Bottova 7939/2A
811 09 Bratislava – Staré Mesto
Slovensko

www.kinit.sk