

# MapReader

## FOSS4G Lancaster

Living with Machines  
7 Oct 2023

Katie McDonough  
Lancaster University

### Our Partners

The  
Alan Turing  
Institute



UEA University of  
East Anglia

UNIVERSITY OF  
EXETER

Queen Mary  
University of London

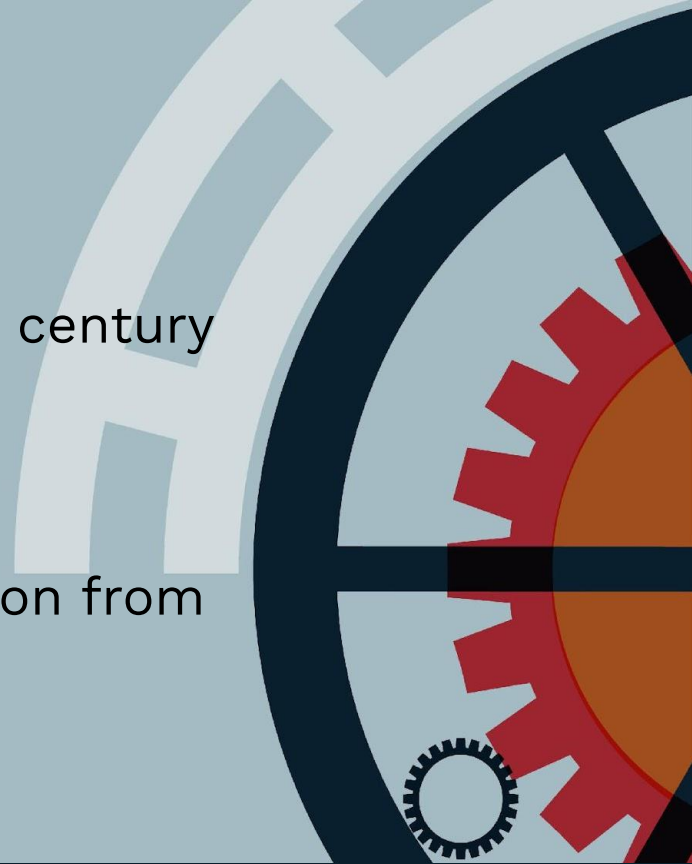
### Our Funders

Arts & Humanities  
Research Council

UK Research  
and Innovation

# Living with Machines

- Digitised collections from the long 19th century
- Computational methods
- Multidisciplinary approach
- Digitised Ordnance Survey map collection from the National Library of Scotland



## Our Partners

The  
Alan Turing  
Institute



UNIVERSITY OF  
CAMBRIDGE

UEA University of  
East Anglia

UNIVERSITY OF  
EXETER

Queen Mary  
University of London

## Our Funders

Arts & Humanities  
Research Council

UK Research  
and Innovation

# The Living with Machines Team

## Principal and Co-Investigators



**Ruth Ahnert**  
(QMUL)



**David Beavan**  
(Turing)



**Emma Griffin**  
(UEA)



**Timothy Hobson**  
(Turing)



**Jon Lawrence**  
(Exeter)



**Maja Maricevic**  
(British Library)



**Barbara McGillivray**  
(Turing / King's College London)



**Mia Ridge**  
(British Library)



**Alan Wilson**  
(Turing)

## Project team



**Claire Austin**  
Rights Assurance



**Kaspar Beelen**  
Digital Humanities Senior  
Research Associate



**Mariona Coll Ardanuy**  
Computational Linguistics  
Senior Research Associate



**André Piza**  
Research Project Manager



**Griffith Rees**  
Research Data Scientist



**Kalle Westerling**  
Research Software Engineer



**Lórlé Demertzi**  
Programme Coordinator



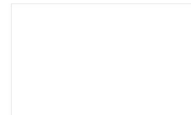
**Luke Hare**  
Research Data Scientist



**Sherman Lo**  
Research Data Scientist



**Daniel Wilson**  
History Senior Research  
Associate



**Rosie Wood**  
Research Data Scientist



**Katie McDonough**  
Senior Research Associate



**Federico Nanni**  
Senior Research Data  
Scientist



**Nilo Pedrazzini**  
Corpus-Based Digital  
Humanities Research  
Assistant

## Our Partners

The  
Alan Turing  
Institute



UNIVERSITY OF  
CAMBRIDGE

UEA University of  
East Anglia

UNIVERSITY OF  
EXETER

Queen Mary  
University of London

## Our Funders

Arts & Humanities  
Research Council

UK Research  
and Innovation



1888

# How do researchers work with maps?



**Digitized maps can be more than sheets to browse in a virtual reading room. But how?**

**Ordnance Survey** maps of England, Wales, and Scotland

6 inches to 1 mile  
1888-1913  
(2nd edition)

~15K sheets

# Scale: Making trustworthy historical claims based on *lots* of maps

**case studies** → ‘high resolution’ archival research or anecdotes from printed materials

**MAPS AS HUMANITIES DATA**

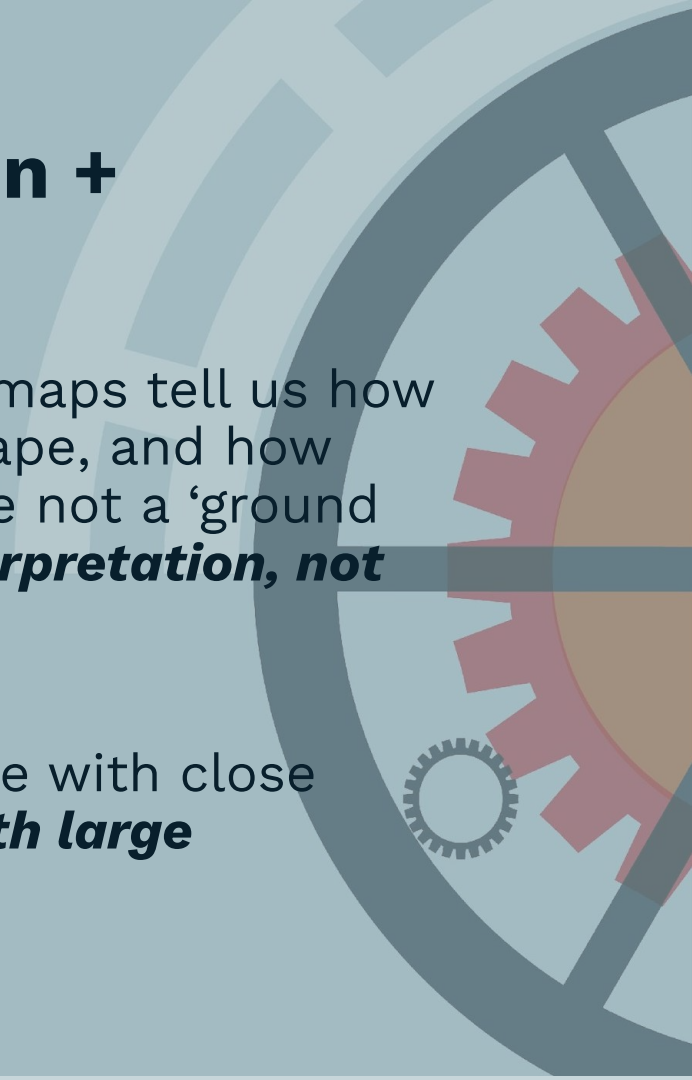
**aggregated statistics** → ‘low resolution’ regional/national



# Opportunity: Computer Vision + Machine Learning

**Iterative and reproducible:** Historical OS maps tell us how Victorians represented the British landscape, and how that landscape was changing, but they are not a 'ground truth'. We want to ***use CV to advance interpretation, not define truth.***

**Fast:** There are too many maps to examine with close reading. ML makes it possible to ***work with large collections quickly.***



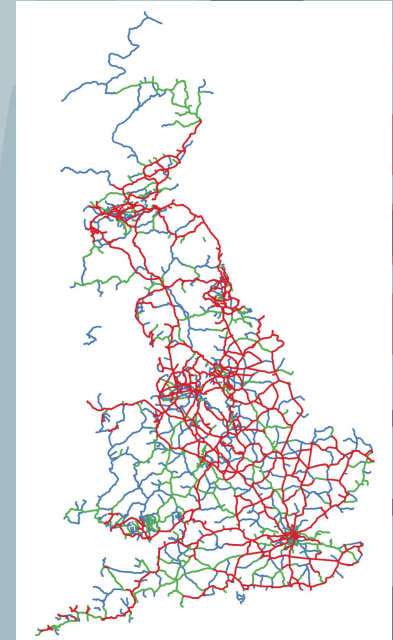
# AI for the humanities: Let's move on from data 'mining'





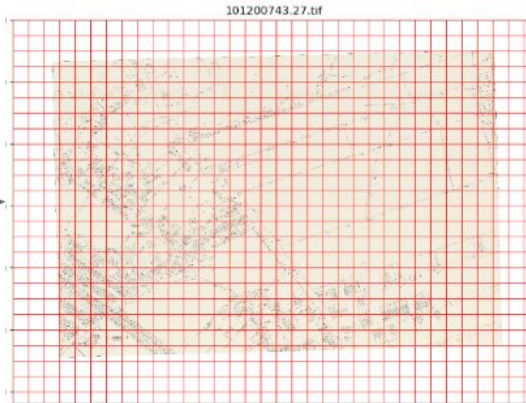
# Cultural Spatial Data Creation

Can we identify and search for new shapes as intentionally imprecise **visual signs** forming new **patterns** rather than reproducing manual vector data collection?



# Solution: 'Patches' as a new shape for historical research

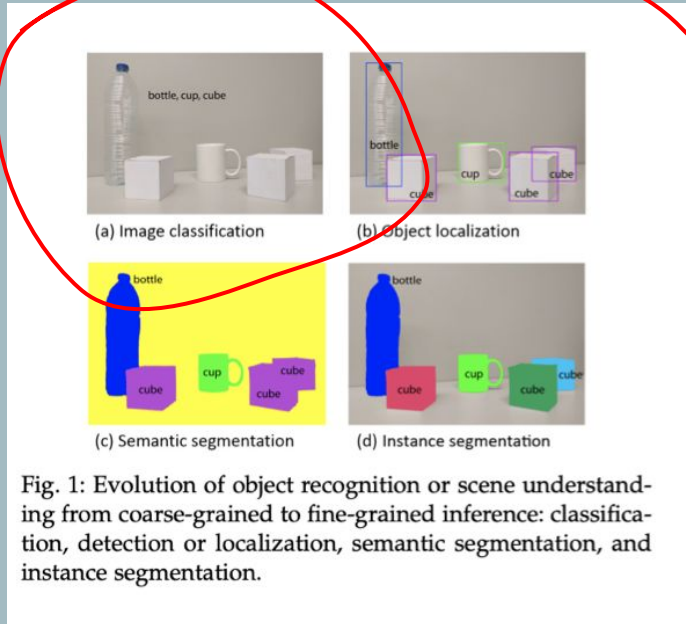
Parent image



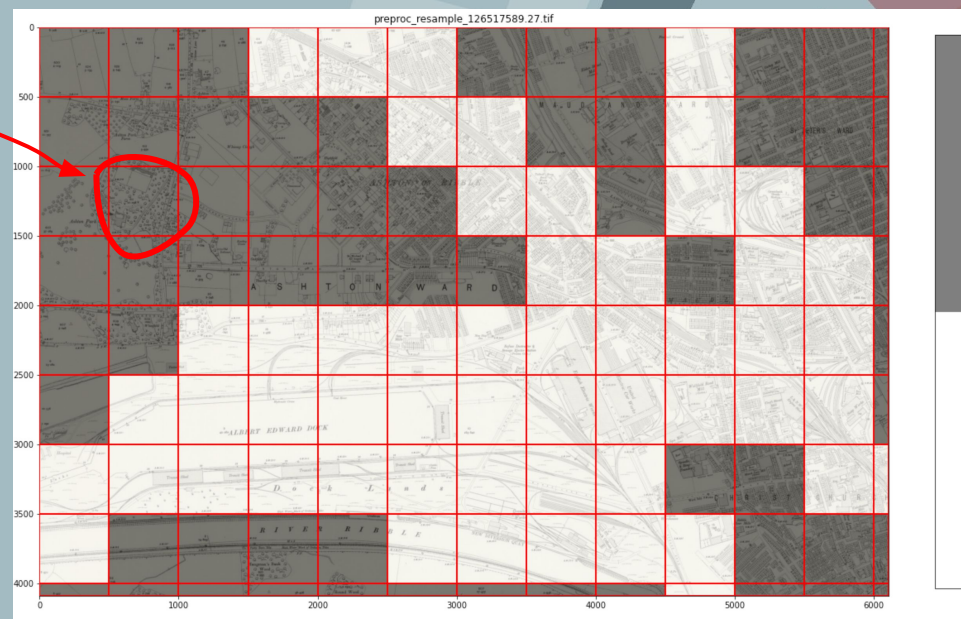
Patches



# Image Classification with Maps: Raster Patches



<https://arxiv.org/pdf/1704.06857.pdf>



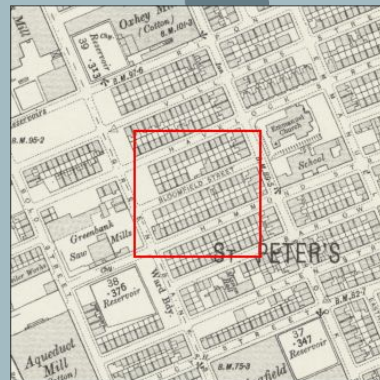
Map regions (patches) annotated as training data

# Patches as humanistic data

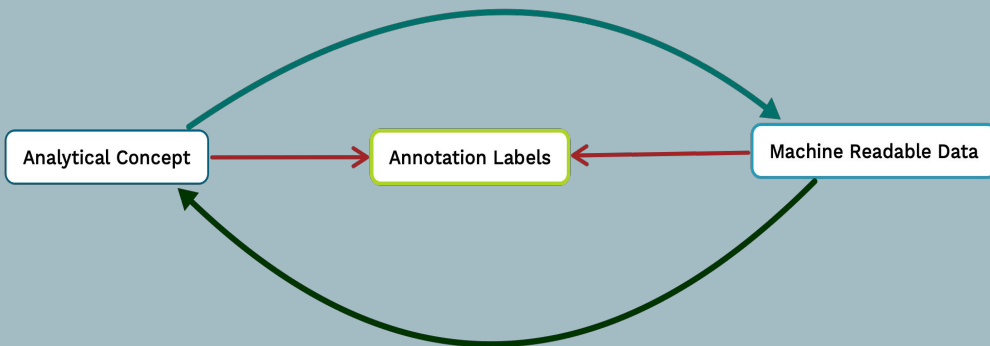
tile-5250-500-5500-750-#preproc\_resample\_126517589.27.tif#.PNG

Image classification tasks on 'patches' are common in medical imaging research.

Can we adapt this approach for humanistic inquiry and search?



# Annotating patches: What is a good label?



Rail Space<sup>1</sup> No Rail Space<sup>2</sup> ← back<sup>j</sup> → next<sup>k</sup>

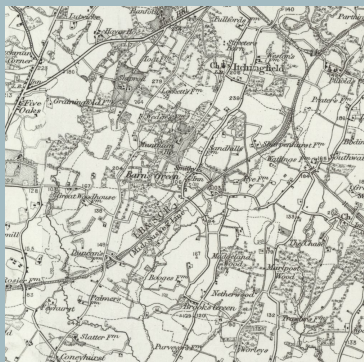
<Figure size 432x288 with 0 Axes>

tile-4500-2000-5000-2500-#preproc\_resample\_126517589.27.tif#.PNG

-----  
Additional info:  
URL to the NLS map: <https://maps.nls.uk/view/126517589>

# Patch sizes and label choices

E.g.



**1000 pixel**  
Too big



**200 pixel**  
Woods

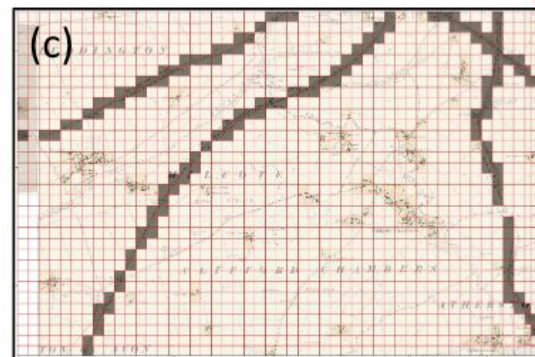
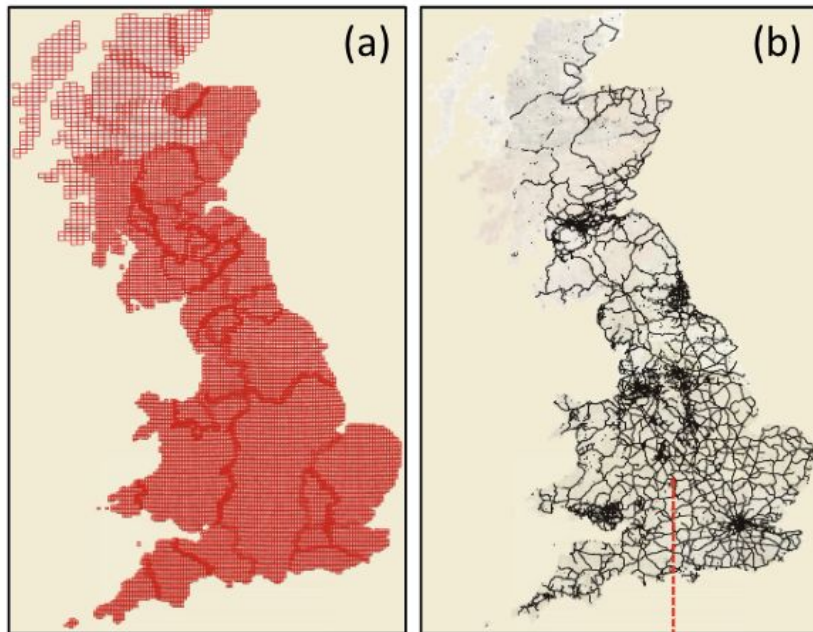


**100 pixel**  
Trees



# Railspace patches

- British **railspace** as predicted by MapReader from ~15K late nineteenth-century 6" OS maps (NLS)
- ~30.5 million patches total
- Railspace = the totality of rail infrastructure in the landscape
- Captured on 50m x 50m patches
- ~62k expert-annotated patch dataset now available on [Hugging Face](#)





# Living with Machines: railspace and building datasets

Download and explore data derived from NLS 2nd-edition six inch to one mile Ordnance Survey maps of England, Wales, and Scotland.

---

This dataset was created by members of the **Living with Machines** digital history project (The Alan Turing Institute/British Library) while developing the **MapReader** computer vision pipeline. Living with Machines is a research project that rethinks the impact of technology on the lives of ordinary people during the Industrial Revolution. MapReader creates datasets for humanities research using historical map scans and their metadata as input. Here we share gold standard annotations and outputs from the earliest MapReader experiments.



# Thank you

## Find out more:

<https://livingwithmachines.ac.uk/latest/@LivingwMachines>

<https://mapreader.readthedocs.io/>



### Our Partners

The  
Alan Turing  
Institute



UNIVERSITY OF  
CAMBRIDGE

UEA University of  
East Anglia

UNIVERSITY OF  
EXETER

Queen Mary  
University of London

### Our Funders

Arts & Humanities  
Research Council

UK Research  
and Innovation