

“Open by Default”

Developing reproducible, computational research

Dani Arribas-Bel

@darribas

Martin Fleischmann

@martinfleis



**The
Alan Turing
Institute**



Geographic
Data Science
Lab



Open research

- Open! Open? “Open”; *Open*; Open
 1. Transparent
 2. Accessible
 3. Transferrable
- Open data is *not* necessary, *nor* sufficient

“Open by default”

- Like happiness, it’s about the journey not (only) the destination
- Treat research artifacts as OSS from Day 1
- Use (and nurture) the *existing* OSS infrastructure

For example

The Alan Turing Institute



☰ Menu

Home + Research + Research projects

Urban Grammar

Learning an urban grammar from satellite data through AI

Learn more ↓

Project status
Ongoing

Related
programmes

Urban analytics



Urban Grammar

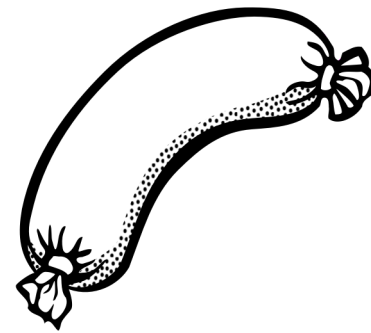
*A characterisation of space based on form and function
designed to understand urban environments*

+ Satellite imagery & AI

Needs



- Data, data, data
- (New) methods, (new) code
- Infrastructure



- (Academic) dissemination
- Open Data Products
- Journaling/reporting

The kitchen

Data

- Standard (formats) better than niche
- Open better than proprietary
- Static better than DB

E.g., parquet, COGs

New methods

Urban Grammar
AI
research project

Contribution to tobler: Speeding up areal interpolation

Have you ever needed to link two sources of data, each attached to a different geometry? In our work in the [WP2](#), we do. We have to transfer data from various sources, linked to output areas, urban blocks or other spatial units to our own bespoke set of geographies. Therefore, we often need to do areal interpolation to correctly map data from one layer to another. Luckily, the open-source Python ecosystem can help.

Tobler, a part of **PySAL family**, is a library for areal interpolation and dasymetric mapping which already offered what we needed. However, our data tends to be large, up to 15 million rows on which we need to interpolate several hundreds of thousands of rows of input data. That can take a while, so each performance improvement can help a lot.

GitHub - martinfleis/clustergram

https://github.com/martinfleis/clustergram

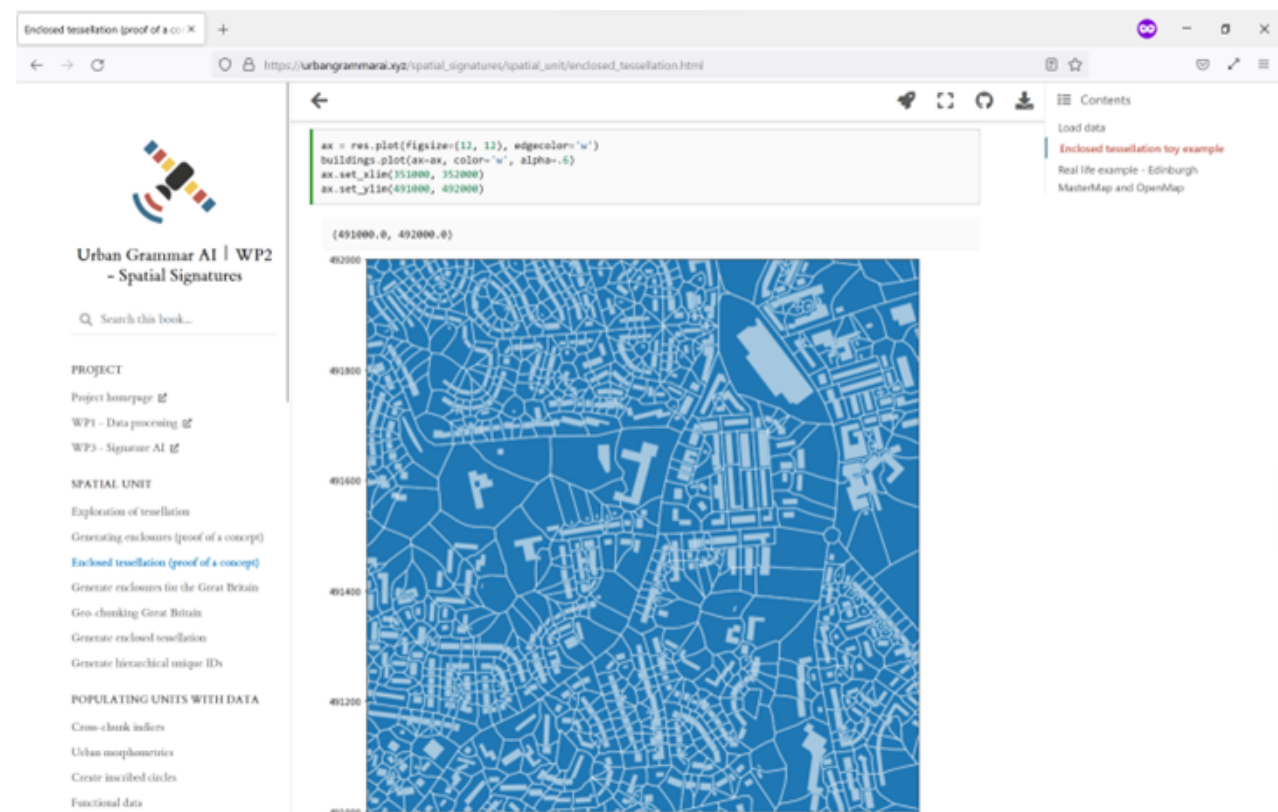
README.md

Clustergram

Visualization and diagnostics for cluster analysis

DOI 10.5281/zenodo.4750483

New code




Infrastructure

GDS Home | gds_env

← → ↻

https://darribas.org/gds_env/

☑ ↗ ⋮



Home


Stacks

Guides

Contributing

FAQ

A containerised platform for Geographic Data Science: `gds_env`



- Dani Arribas-Bel [[@darribas](#)]






The `gds_env` (short for “GDS environment”) provides a modern platform for Geographic Data Science. The project is a [Jupyter](#)-based stack that includes state-of-the-art **geospatial** libraries for **Python** and **R**. The `gds_env` is based on **container** technology to make it a transferrable platform for reproducibility. The source code is released under an [open source license](#) and the build process is transparent.

The `gds_env` extends the official [Jupyter Docker Stack](#) to include geospatial functionality in both Python and R. To offer more flexibility, this extension is provided in three different flavours, or stacks (to): `gds_py`, `gds` and `gds_dev`. Each of them builds on each other and adds further functionality. Please check the [Stacks section](#) for more information.

The goal of the `gds_env` is to make using Python and R for geospatial easy to set up in a large variety of contexts. The `gds_env` can support research and teaching activities, but is also suitable for data scientists using Python and R “in the field”. The stacks can be used in a range of environments, including: Windows/Mac/Linux laptops and desktops, servers, compute clusters, supercomputers or in the cloud (e.g. you can deploy them on [Binder](#)). For more information on how to build or install any of the stacks, check the [Guides section](#).

Building blocks

The `gds_env` stands on the shoulders of giants. Here are the core open technologies it is built with:



This site uses [Just the Docs](#), a documentation theme for Jekyll.

... and the “sausage”

Dissemination (I): Papers


urbangrammarai/spatial_signatu X

← → ↺



https://github.com/urbangrammarai/spatial_signatures_concept

☆





📄 ⚙️ 🔗 ↗️ ☰

 Search or jump to... /

[Pull requests](#) [Issues](#) [Marketplace](#) [Explore](#)

 + 


urbangrammarai / spatial_signatures_concept Private





 Edit Pins  Unwatch 1  Fork 0  Star 0

[Code](#) [Issues](#) 1 [Pull requests](#) [Actions](#) [Projects](#) [Security](#) [Insights](#) [Settings](#)

master 5 branches 0 tags

Go to file Add file Code




 **actions-user Build PDF** 9e4378e on 14 Feb 245 commits

| | | |
|---|----------------------|---------------|
|  .github/workflows | split for submission | 10 months ago |
|  code | add more comments | 5 months ago |
|  paper | Build PDF | 3 months ago |
|  .gitignore | git structure | 17 months ago |

Add a README with an overview of your project. [Add a README](#)

About

Conceptual paper on Spatial Signatures

 0 stars  1 watching  0 forks


Releases


No releases published [Create a new release](#)

Packages

No packages published [Publish your first package](#)

Contributors 2

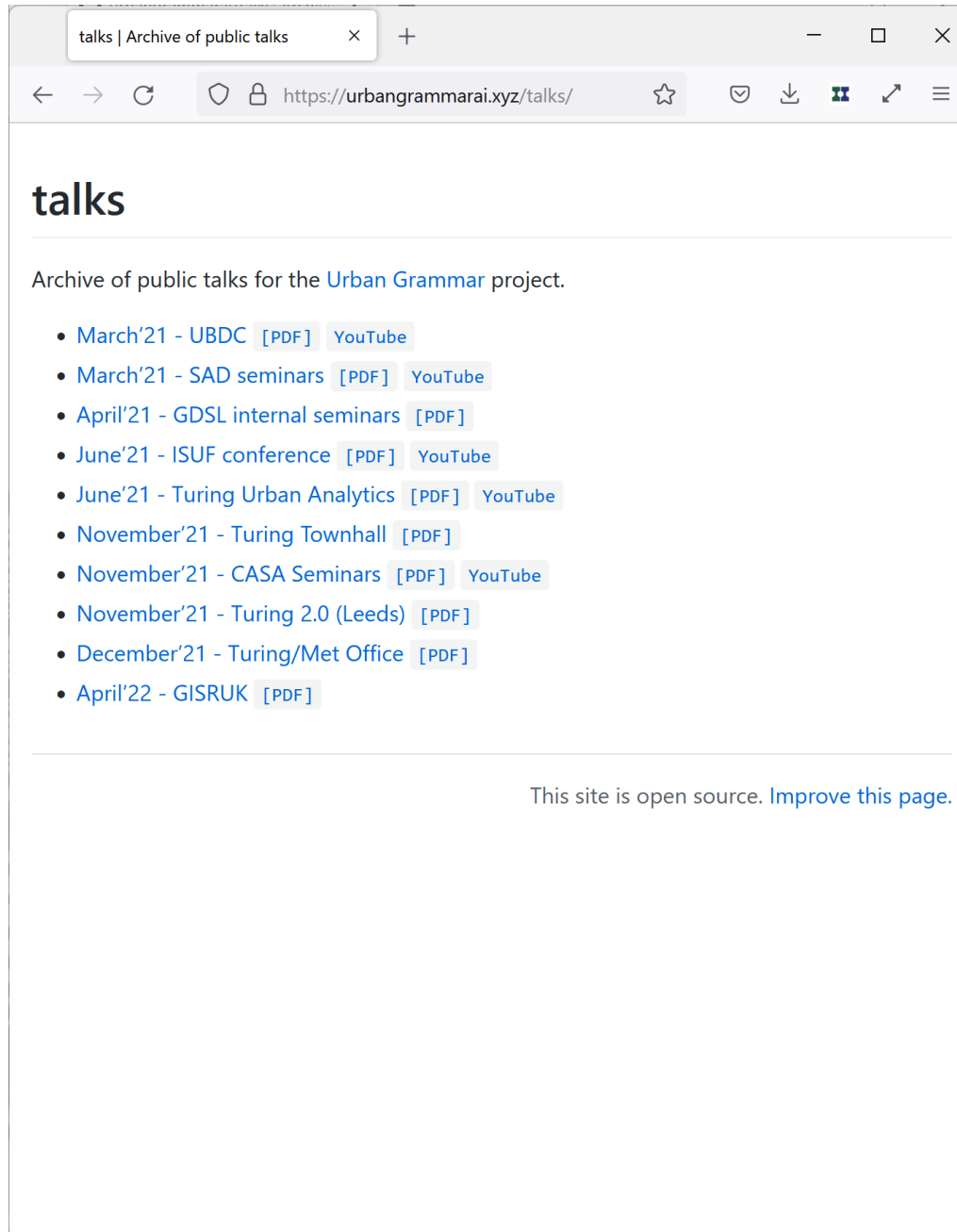
 **martinfleis** Martin Fleischmann

 **darribas** Dani Arribas-Bel

Languages

https://github.com/urbangrammarai/spatial_signatures_concept/projects?type=beta

Dissemination (II): Talks



A screenshot of a web browser showing the 'talks | Archive of public talks' page. The browser's address bar displays 'https://urbangrammarai.xyz/talks/'. The page title is 'talks'. Below the title, it says 'Archive of public talks for the Urban Grammar project.' A list of talks is provided, each with a date, title, and links to PDF and YouTube versions. At the bottom, a message states 'This site is open source. Improve this page.'

talks | Archive of public talks

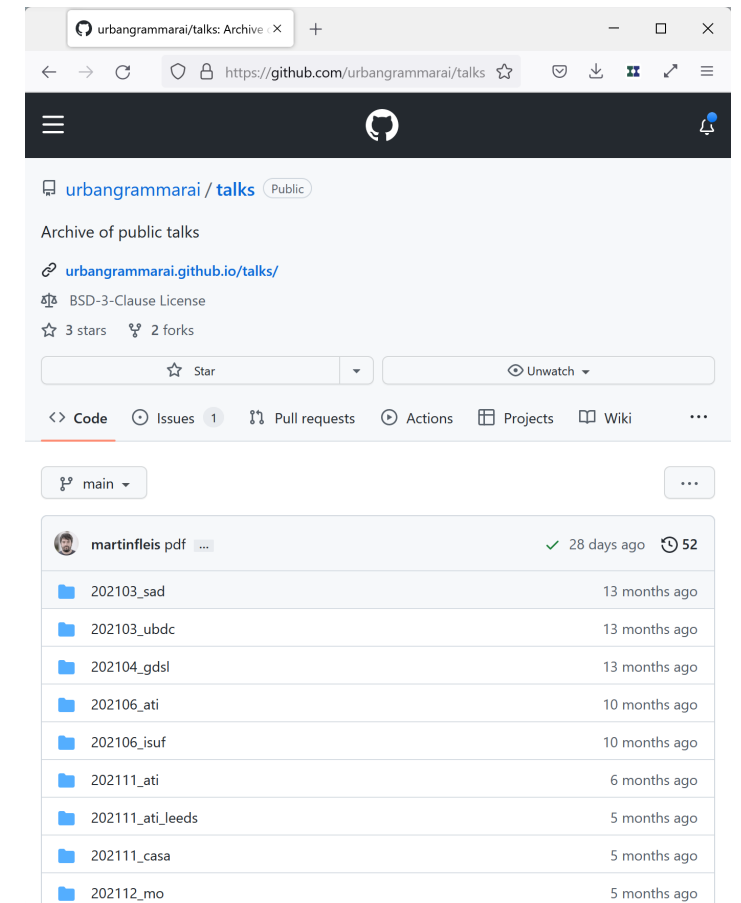
https://urbangrammarai.xyz/talks/

talks

Archive of public talks for the [Urban Grammar](#) project.

- [March'21 - UBDC](#) [PDF] [YouTube](#)
- [March'21 - SAD seminars](#) [PDF] [YouTube](#)
- [April'21 - GDSL internal seminars](#) [PDF]
- [June'21 - ISUF conference](#) [PDF] [YouTube](#)
- [June'21 - Turing Urban Analytics](#) [PDF] [YouTube](#)
- [November'21 - Turing Townhall](#) [PDF]
- [November'21 - CASA Seminars](#) [PDF] [YouTube](#)
- [November'21 - Turing 2.0 \(Leeds\)](#) [PDF]
- [December'21 - Turing/Met Office](#) [PDF]
- [April'22 - GISRUK](#) [PDF]

This site is open source. [Improve this page.](#)



A screenshot of the GitHub repository for 'urbangrammarai/talks'. The repository is public and has 3 stars and 2 forks. It is licensed under BSD-3-Clause. The repository contains a list of folders representing different talks, each with a date and a timestamp indicating when it was last updated.

urbangrammarai/talks: Archive

https://github.com/urbangrammarai/talks

urbangrammarai / talks (Public)

Archive of public talks

[urbangrammarai.github.io/talks/](#)


BSD-3-Clause License

3 stars 2 forks

Star Unwatch

Code Issues 1 Pull requests Actions Projects Wiki

main

| | |
|--|----------------|
|  martinleis pdf | 28 days ago 52 |
| 202103_sad | 13 months ago |
| 202103_ubdc | 13 months ago |
| 202104_gdsl | 13 months ago |
| 202106_ati | 10 months ago |
| 202106_isuf | 10 months ago |
| 202111_ati | 6 months ago |
| 202111_ati_leeds | 5 months ago |
| 202111_casa | 5 months ago |
| 202112_mo | 5 months ago |

Dissemination (III): Other


All posts — Urban Grammar AI doc

← → ↺

https://urbangrammarai.xyz/blog/archive

☆

🔒 ⬇️ 🛠️ ↗️ ☰



Urban Grammar AI

research project

🔍 Search the docs ...

Project overview

Updates

Public talks

WP1 - Data processing

WP2 - Spatial Signatures

WP3 - Signature AI

Posted in 2022

11 April 2022 - [GISRUK talk](#)

29 March 2022 - [Contribution to toblor once again: Areal interpolation even faster](#)

18 March 2022 - [Third Advisory Board](#)

10 February 2022 - [The Urban Grammar on MapScaping podcast!](#)

Posted in 2021

17 December 2021 - [British Spatial Signatures at Towards urban analytics 2.0](#)

10 December 2021 - [Third Advisory Board](#)

10 December 2021 - [British Spatial Signatures talk at CASA](#)

09 December 2021 - [Urban Grammar in the Turing Catch Up](#)

03 August 2021 - [xyzservices: a unified source of XYZ tile providers in Python](#)

09 July 2021 - [Signatures for \(all of\) GB at Turing Urban Analytics](#)

05 July 2021 - [Form-based Signatures at ISUF 2021](#)

04 June 2021 - [Google Summer of Code](#)

29 April 2021 - [Second Advisory Board](#)

27 April 2021 - [Clustergam: visualisation of cluster analysis](#)

26 April 2021 - [Spatial Signatures at the Spatial Analytics + Data Seminar Series](#)

08 April 2021 - [Spatial Signatures at UBDC](#)

17 February 2021 - [Visual style and a graphics package](#)

urbangrammarai/urbangrammarai

urbangrammarai.github.io

Public

urbangrammarai.xyz

BSD-3-Clause License

4 stars 2 forks

☆ Star

👁 Watch

<> Code

🔔 Issues

🔗 Pull requests 1

🔄 Actions

📁 Projects

📖 Wiki

⋮

📌 master

⋮

🚀 actions-user Build website ...

✓ 23 days ago

🕒 175

View code

README.md

📄

urbangrammarai.github.io

Repository for a project website.

Note:

- keep source in `src` folder and setup GitHub action to build from there to `docs` for publishing.

Open Data Products (I): Data

Spatial Signatures of Great Britain X

https://data.cdrc.ac.uk/dataset/spatial-signatures-great-britain

Spatial Signatures of Great Britain

Population & Mobility

Transport & Movement

Geographical Characterisation of British Urban Form and Function using the Spatial Signatures Framework

Spatial signatures characterise space based on form and function in a way designed to understand urban environments. This data product, part of the Urban Grammar project, contains a typology of spatial signatures in Great Britain. Each type has a distinct character capturing what the place looks like (form) and how it is used (function).


The data product contains bespoke Signature geometry with signature type, summary of input variables per each geometry and per each type, interpolation of signature types to OA and LSOA geometry and short pen portraits for the typology, shorthand descriptions of the characteristics of each signature type.

The interactive map showing the typology is available at <https://urbangrammarai.xyz/great-britain/>. More details about the project can be found at the project website <https://urbangrammarai.xyz>.

Related publications:

Fleischmann, Martin; Arribas-Bel, Daniel (2021): Geographical Characterisation of British Urban Form and Function using the Spatial Signatures Framework. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.16691575.v1>


Data Extent



Leaflet | Map data © OpenStreetMap

License

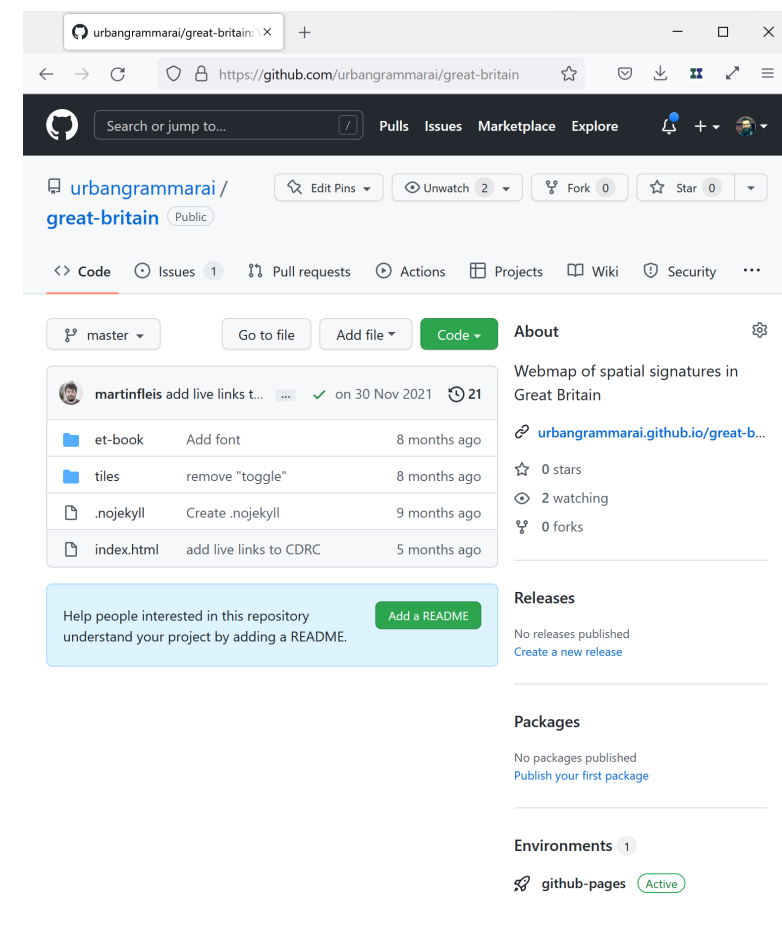
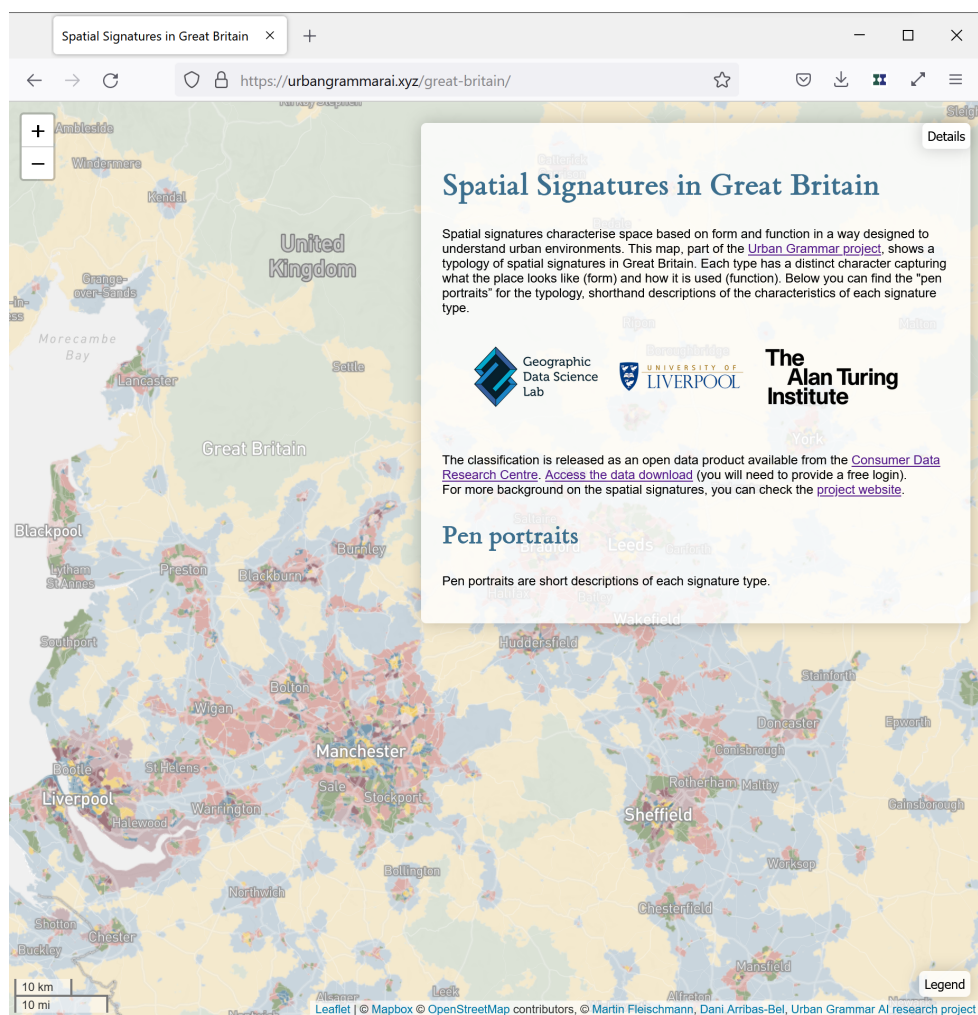
UK Open Government Licence (OGL)

Geographic Data Science Lab

UNIVERSITY OF LIVERPOOL

The Alan Turing Institute

Open Data Products (II): Exploratory tools



The takeaway

- “Release early, release often”
- Make open a feature, not a bug
- *Share your knowledge, it is a way to achieve immortality* (Dalai Lama)

A bit more to read

Rey, S., Arribas-Bel, D., & Wolf, L. (*in press*). Computational Tools for Geographic Data Science. Chapter 2 in *Geographic Data Science with Python*. Available at:
<https://geographicdata.science/book>

Boeing, G., & Arribas-Bel, D. (2021). GIS and Computational Notebooks. *Geographic Information Science & Technology Body of Knowledge*, 2021(Q1).

Arribas-Bel, D., Green, M., Rowe, F., & Singleton, A. (2021). Open data products—A framework for creating valuable analysis ready data. *Journal of Geographical Systems*, 23(4), 497–514.

<https://urbangrammarai.xyz>

“Open by Default”

Developing reproducible, computational research

Dani Arribas-Bel

@darribas

Martin Fleischmann

@martinfleis



**The
Alan Turing
Institute**



Geographic
Data Science
Lab