DundeeData Meetup

Exploring Open Data Scotland and Example Projects

Tuesday 26 September 2023, 6-8pm cyberQuarter, Abertay University

cyberQuarter | WAbertay University.





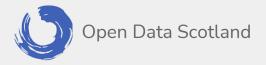






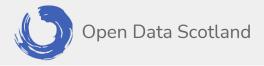
Open Data Scotland

A resource for community, built by community



"Open data is data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike."

as defined by the Open Definition (https://opendefinition.org)



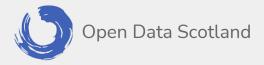
What is Open Data Scotland?

A 100% volunteer-run, collaborative community promoting open data through open source.

We aim to help people

- locate Open Data better
- understand what we have in Scotland
- and promote projects which use these Open Data

All of this is to help the public understand what Open Data is available in Scotland.



The problem we're trying to solve

Users of open data don't know where to find data

- Sources are disparate and decentralised
- Very reliant on knowing who to ask and where to look
- High barrier to access for the occasional user

There have been attempts to collate sources before

- Well intentioned, but one-off events
- Huge effort to maintain
- Sources become aged and irrelevant

Some people don't recognise the term "open data"

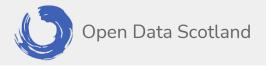
We wanted to build a community of like-minded people across Scotland and further!

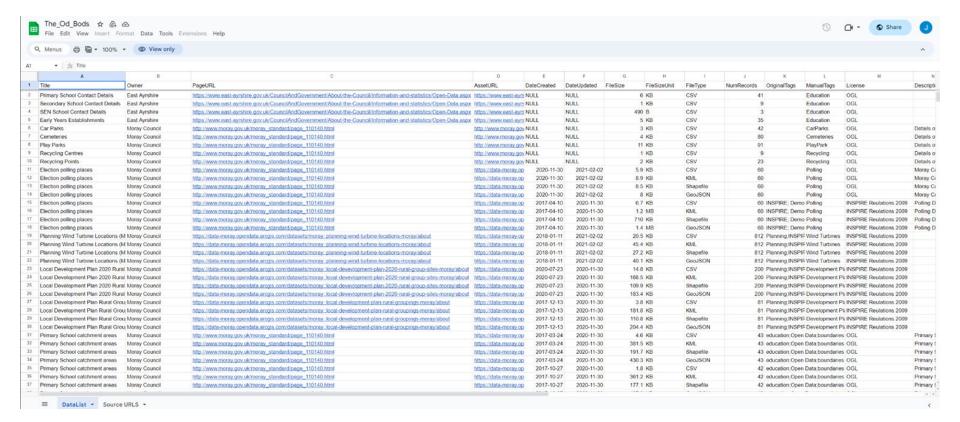
How we started

427 datasets

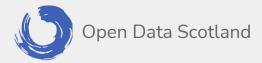
manually* gathered in a Google sheet











OD Bods



We've catalogued 427 datasets so far. With a (conservative) estimate that each council would have 150 datasets there should be around 4800 for the 32 Scottish councils, so 4373 are still MIA...

See the analysis of what we've catalogued so far.

Datasets

Local Development Plan 2020 Transportation Improvements (TSP) (Moray)

Local Development Plan 2020 Transportation Improvements (TSP) (Moray)

Created: 2020-07-23, Updated: 2021-06-12

Planning INSPIRE

FEATURE LAYER

From: Moray Council

Local Development Plan 2020 Special Landscape Areas (SLA) (Moray)

Local Development Plan 2020 Special Landscape Areas (SLA) (Morav)

Created: 2020-07-23, Updated: 2021-06-12

Planning INSPIRE

FEATURE LAYER

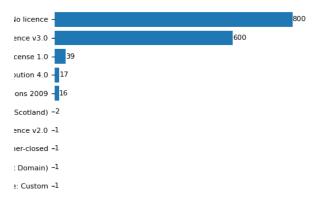
From: Moray Council

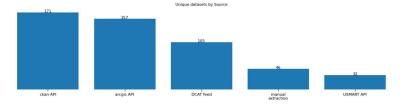
Private Water Supplies Source (Moray)

Private Water Supplies Source (Moray)Reg 2 (previously Cat A) are supplies which are commercial (including private lets) or more than 50 persons OTE - The location data we hold for the private water supplies is not 100% accurate (easting's and northing's can be o are currently working on improving these through GPS, unfortunately I cannot provide a figure on the accuracy, some maybe just in a general area of the props serves). The spreadsheet may contain multiple records for the same source reference, this may be where either multiple source abstractions points are located are other private water supply infrastructure located eg storage tanks. The location Type indicates the type. A or A1 etc. = Source, B or B1 etc. = supply infrastr or U1 etc. = unused supply infrastructure. Sites marked with "" confirmed are for sites where we have visited site and obtained more accurate locations. The li not include historic private water supplies. The data is provided based on the best information available to the council at the time of writing.

Created: 2020-12-22, Updated: 2021-06-11

Number of data files by License







How do we gather the data?



100% Python-based pipeline
11 scraper scripts
2 scripts for merging and cleaning



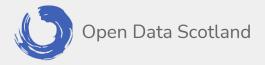


Some organisations have APIs CKAN, ArcGIS, USMART



Some organisations have static web pages Manual gathering Automate using scripting - Beautiful Soup





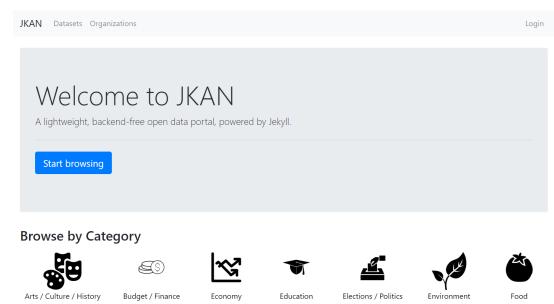
How do we display the end result?

Health / Human Services

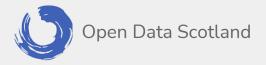
Parks / Recreation

Planning / Zoning

- JKAN https://jkan.io
- Static site generator using Jekyll (Ruby-based)
- Uses Markdown/YAML files for data storage
- Markdown converts to HTML files on site publish
- Hosted on GitHub pages (for free!)
- Open source project

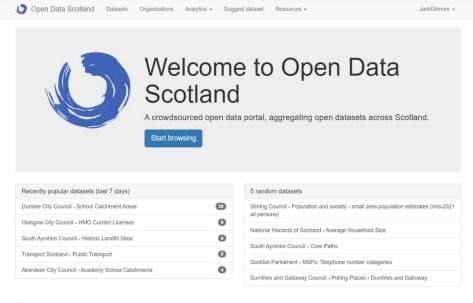


Public Safety



opendata.scot

- Soft launched November 2021
- Average 700+ visitors a month
- 60+ organisations with published open data
- 75% published with OGL licence
- Site analytics using Plausible.io
- Dataset analytics and stats using Chart.js and Leaflet.js



Browse by Category





















Budget / Finance Business and Economy











Health and Social Care

Housing and Estates

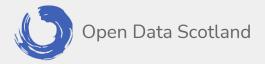
Law and Licensing

Parks / Recreation

inning and Development

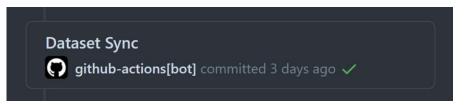
Public Safety

Sport and Leisur

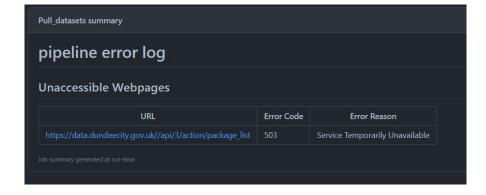


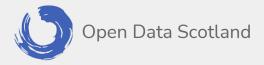
Automation

- GitHub Actions
- YAML based CI/CD platform
- Open source == free usage!
- Weekly updates
- Runs automatically
- Error handling and reporting





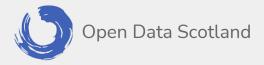




Where we are now

2,000+ datasets
automatically scraped each week

Live demo!



Publishers

<u>Current</u>

31 Local Authorities

Cycling Scotland

National Library Scotland

Public Health Scotland

Research Data Scotland

Scottish Forestry

Scottish Qualifications Authority (SQA)

Spatial Hub

Scottish Government Statistics

Scottish Parliament

Next up

National Records Scotland

Scottish Environmental

Protection Agency (SEPA)

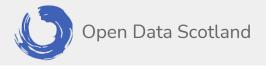
Marine Scotland

Public Contracts Scotland

Stagecoach

Scottish Council for Voluntary

Organisations (SCVO)



Publishers

<u>Current</u>

31 Local Authorities

Cycling Scotland

National Library Scotland

Public Health Scotland

Research Data Scotland

Scottish Forestry

Scottish Qualifications Authority (SQA)

Spatial Hub

Scottish Government Statistics

Scottish Parliament

Next up

National Records Scotland Scottish Environmental

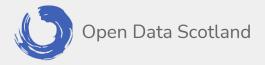
Protection Agency (SEPA)

Marine Scotland

Public Contracts Scotland

Stagecoach

Scottish Council for Voluntary Organisations (SCVO)



Publishers

<u>Current</u>

31 Local Authorities
Cycling Scotland
National Library Scotland
Public Health Scotland
Research Data Scotland
Scottish Forestry

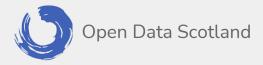
Spatial Hub
Scottish Government Statistics

Scottish Qualifications Authority (SQA)

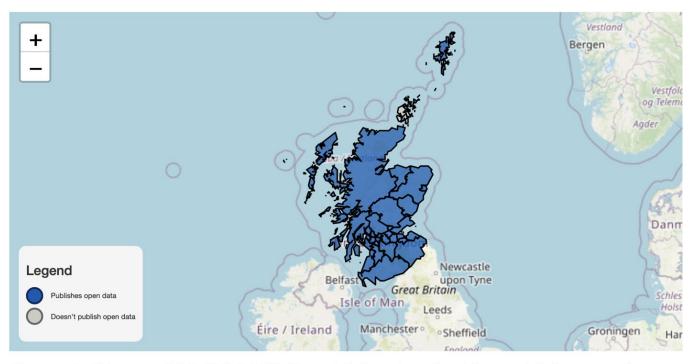
Scottish Parliament

Next up

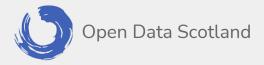
National Records Scotland
Scottish Environmental
Protection Agency (SEPA)
Marine Scotland
Public Contracts Scotland
Stagecoach
Scottish Council for Voluntary
Organisations (SCVO)



Local authority coverage

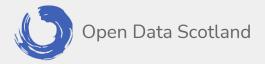


Currently 31 out of 32 (96.88%) Scottish local authorities publish open data



Top 10 Local Authorities by volume

Local authority	# datasets
Stirling Council	231
Glasgow City Council	132
North Ayrshire Council	131
Perth & Kinross Council 🗙	84
Moray Council	70
Dundee City Council 🗙	66
South Ayrshire Council	62
City of Edinburgh Council	58
Renfrewshire Council	51
Angus Council 🏡	44



Top 10 categories by volume

Council and Government	818
Planning and Development	776
Elections / Politics	454
Housing and Estates	422
Food and Environment	414
Business and Economy	366
Transportation	321
Health and Social Care	291
Sport and Leisure	273
Education	224

Angus Council Fish Smoking Areas

Angus Council

Fish Smoking Area in Arbroath for the production of the Arbroath Smokie



Bi-annual complaints by category

Dundee City Council

The percentage of complaints recorded by the Council in each six month period in each of the categories included in the model Complaint Handling Procedure in Local Authorities.



Wind Turbines

Perth & Kinross Council

Planning applications for wind turbines with relevant planning application information. PLEASE NOTE that the same site can have more than one planning application related to it.

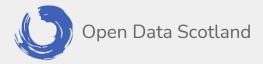




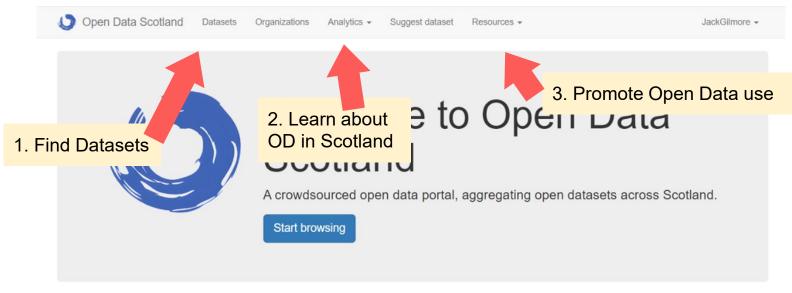


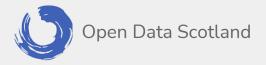




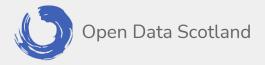


Service Areas

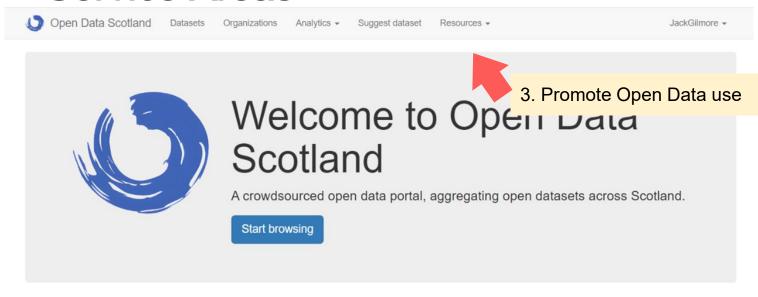


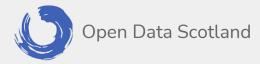


What's the point?



Service Areas





Projects using Open Data - Scotland

Travelling Tabby Scotland Coronavirus Tracker

https://www.travellingtabby.com/scotland-coronavirus-tracker/local

source data: Public Health Scotland

John is a student at the University of the Highlands and Islands (Argyll College) and made a very useful dashboard over lockdown which has been very successfully received and relied on by thousands of individuals. This is a very well-laid project and the section on additional information at the bottom of the dashboard is well worth a visit on its own.

Street Art Aberdeen

https://streetartaberdeen.org/

A collaborative project created by Andy Gaskell which collects and display locations of street art in Aberdeen, helping the public find the artworks so they can be appreciated as they were intended - in real life

The Courier: King David and the death of Karen: Scotland's baby name trends for past 50 years

https://www.thecourier.co.uk/fp/courier-investigations/3135287/baby-names-scotland/

source data: National Records Scotland/ The Courier Github Repo

An excellent (and entertaining) example of data storytelling by the data team at The Courier

Interactive SIMD Map

https://simd.scot/#/simd2016/BTTTFTT/9/-4.0000/55.9000/

source data: Scottish Government

Explore the Scottish Index of Multiple Deprivation (SIMD) for 2012, 2016 and 2020 by geography. Map created and maintained by Scottish Government.

SIMD Visualised by Local Authority

https://twitter.com/northerniamie/status/1225484162263547905?s=21

source data: Scottish Government

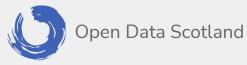
A effective and eye-catching single-page visual by Jamie Whyte displaying the distribution of the deprivation index, within and across Local Authorities

Survey of Scottish Witchcraft (1563 - 1736)

http://www.shca.ed.ac.uk/Research/witches/

source data: Wikidata

A queryable database of all indicviduals in Scotland accused of Witchcraft, compiled by researchers at Edinburgh University, and using data stored in Wikidata.



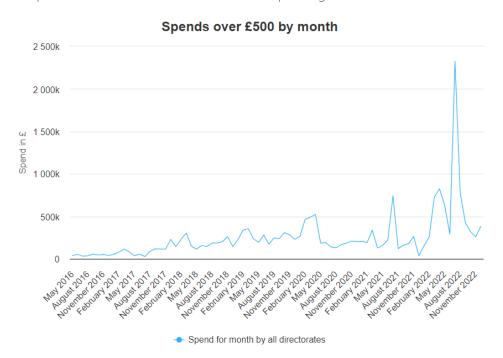


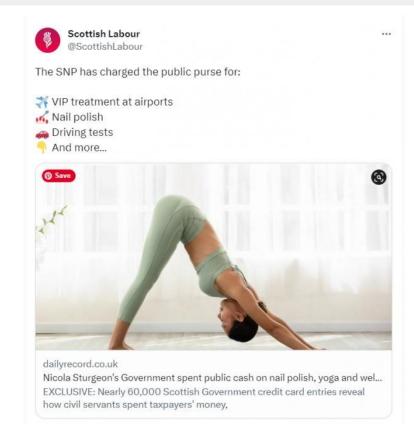


Source: https://www.glasgowchamberofcommerce.com/29659

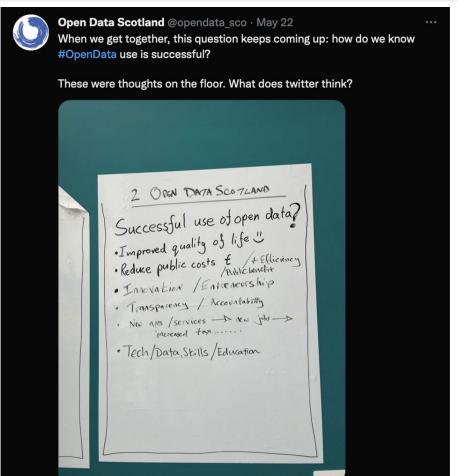
Welcome to Open Data Scotland GovSpend

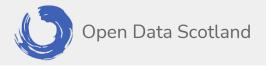
An open data source of Scottish Government spending data



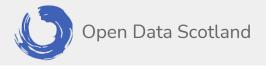


Source: https://govspend.opendata.scot

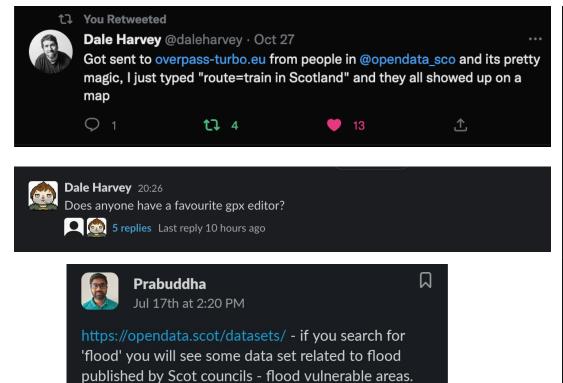


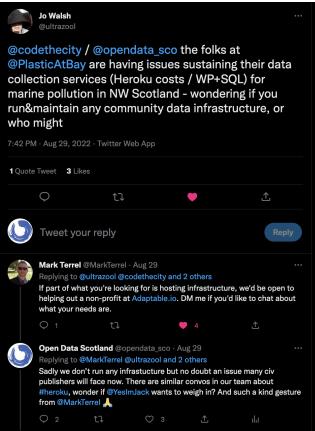


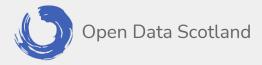




are those relevant?

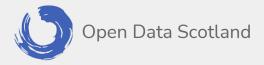






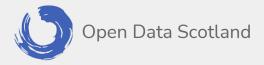
The challenges

- Inconsistent standards
 - Across different publishers
 - Within individual publishers
- Some metadata doesn't exist
 - Do we ignore, interpolate or other?
- Working over a long time period
 - How do you keep momentum and engagement over 12 months?
 - Keeping collaborators informed in non-real time
 - How do you remember what you did 2 months ago?
- Breaking tasks into bite-sized chunks
- Getting people to know this service exists



Future features and plans

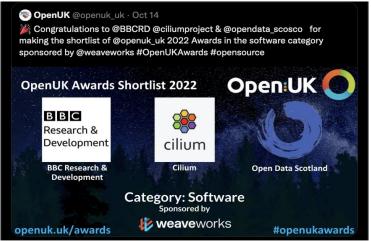
- Twitter bots?
 - Alerts for when sources go offline
 - Weekly key trends: pageviews + most popular dataset of the week
- Smarter categorisation of datasets?
- Dataset 5-star rating and review system?
- Recommender engine: "see other datasets like this"?
- MORE DATA!



The wins

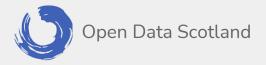
- Seeing active users on the service
- Broad range of skills and expertise available
- Opportunity to work with people you wouldn't normally work with
- Hitting 2000+ dataset listings





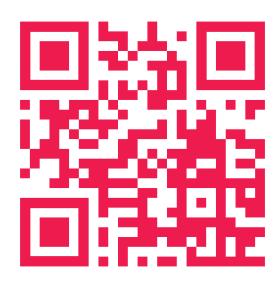
Key Messages

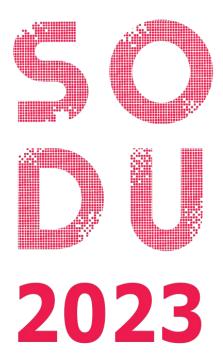
- 1. Know that opendata.scot is a resource
 - a. To find datasets and projects
 - b. But also to work on as a contributor
- 1. Talk about Open Data use
- 1. Know that you can do a lot with very little



Scottish Open Data Unconference

- 🔳 18/19th November 2023
- ONE Tech Hub Aberdeen
- 🗐 https://sodu.live





Thank you Questions?

Jack Gilmore







@jackgilmore@mstdn.social

@YesImJack

@JackGilmore

opendata.scot



@OpenDataScotland

opendata.scot/slack