



Styles of data access to foster competition and innovation

Jack Hardinges
Programme Lead, Open Data Institute

2020-12-08

ODI

Founded in 2012, the Open Data Institute (ODI) is an international, independent and not-for-profit organisation based in London, UK.



Sir Nigel Shadbolt
Chairman



Sir Tim Berners-Lee
President




Dr Jeni Tennison
Vice President



Louise Burke
Managing Director

Our mission

**We work with
companies and
governments to
build an open,
trustworthy data
ecosystem**



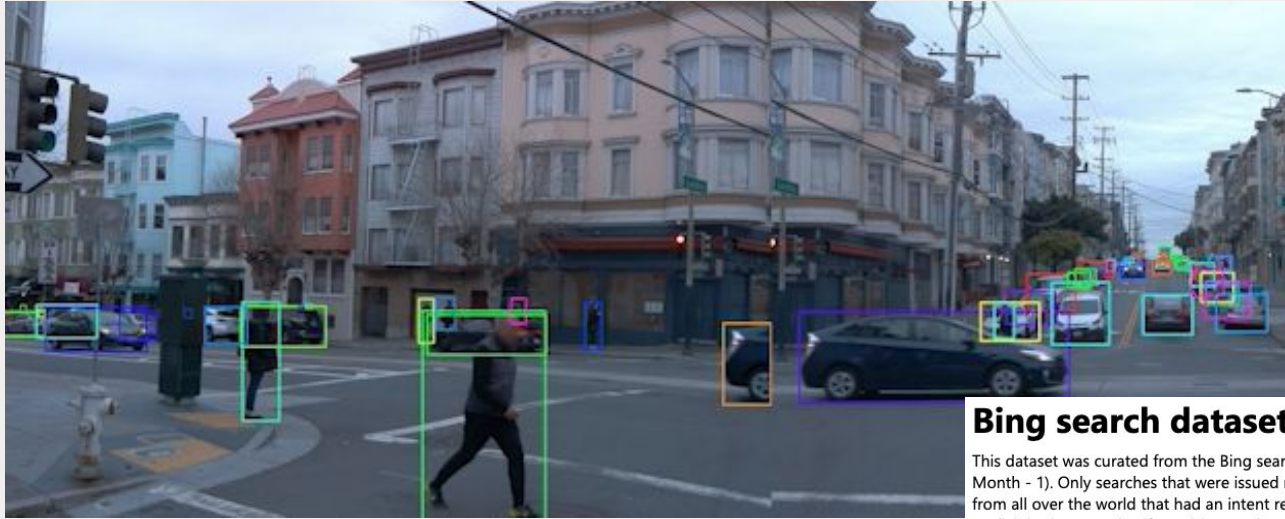
Why is access to data important?

- to improve transparency and accountability.
- **to support research.**
- **to stimulate innovation.**
- **to promote competition.**
- to support policy-making, planning and decision-making.
- to undertake law enforcement and preserve national security.
- ...



Some styles of data access

Public or open data



Bing search dataset for Coronavirus Intent

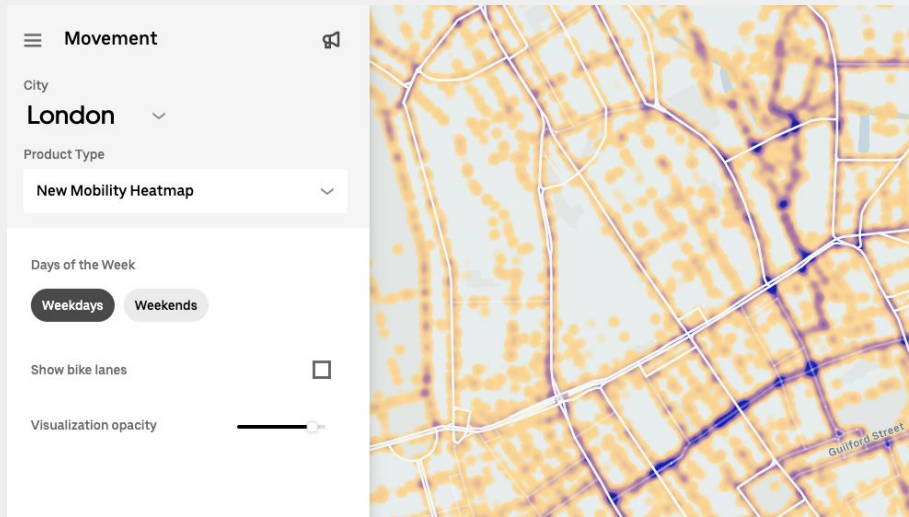
This dataset was curated from the Bing search logs (desktop users only) over the period of Jan 1st, 2020 – (Current Month - 1). Only searches that were issued many times by multiple users were included. Dataset includes queries from all over the world that had an intent related to the Coronavirus or Covid-19. In some cases this intent is explicit in the query itself, e.g. "Coronavirus updates Seattle" in other cases it is implicit, e.g. "Shelter in place". Implicit intent of search queries (e.g. Toilet paper) were extracted by using Random walks on the click graph approach as outlined in the following paper by Nick Craswell et al at Microsoft Research: <https://www.microsoft.com/en-us/research/wp-content/uploads/2007/07/craswellszummer-random-walks-sigr07.pdf> All personal data was removed. Contact BingCoronaVirusTeam@microsoft.com for all enquiries.

Category:
health care, genomics

Project URL:
<https://github.com/microsoft/BingCoronavirusQuerySet>

License:
[Open Use of Data Agreement v1.0](#)

Public tools, dashboards and dataviz



Ad Library

The Ad Library provides advertising transparency by offering a comprehensive, searchable collection of all ads currently running from across Facebook apps and services, including Instagram.



Search ads

Choose a category to start your search.

Issues, elections or politics

Search all

Search ads from candidates for public office, as well as ads about any election, ballot initiative or social issue.

Search for ads by keyword or advertiser name

More ways to search and filter ads about social issues, elections or politics

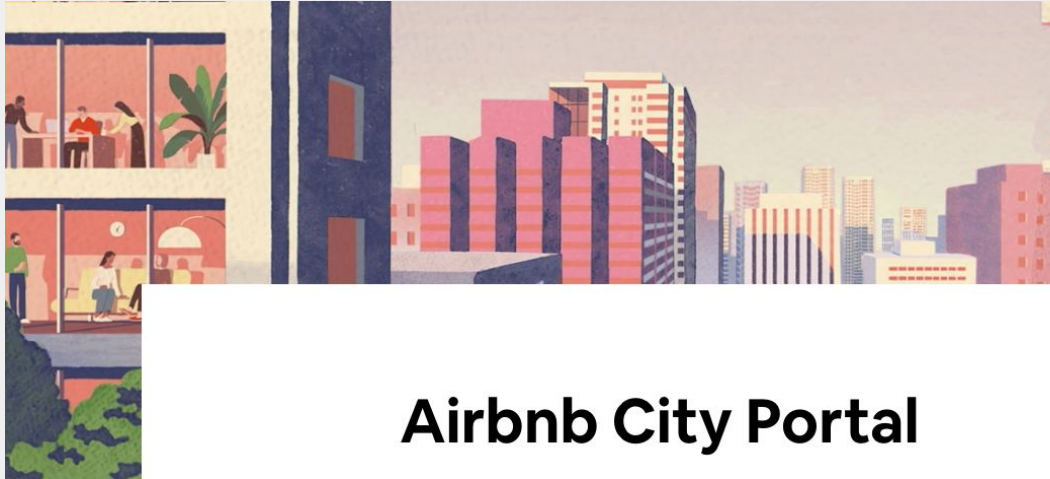
Report

Explore, filter and download data for ads about social issues, elections or politics. Also find the spending tracker and spending totals by advertiser and location.

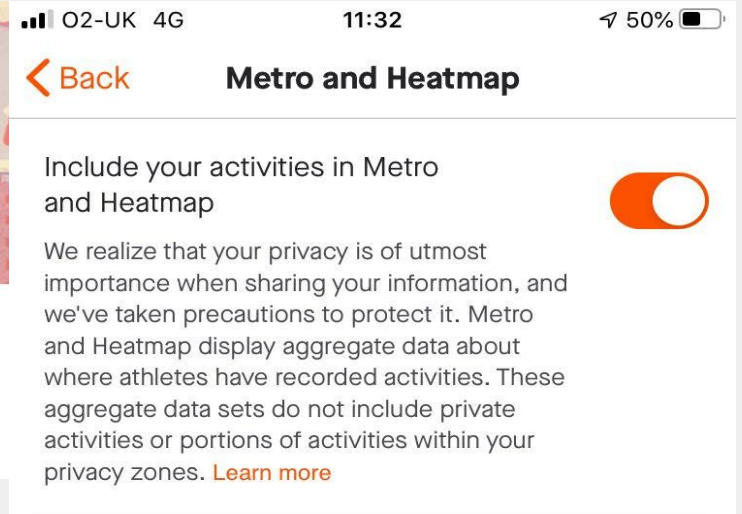
API

Perform customised keyword searches of active and inactive ads about social issues, elections or politics.

Data 'portals' for certain users



Airbnb City Portal



Why contribute?

Because Strava Metro and the Global Heatmap make running and riding in cities better. Metro displays aggregate Strava data to inform urban planners and advocacy groups about human-powered transportation trends. The Global Heatmap, powered exclusively by contributions from athletes like you, is one of the world's best free resources for route and trail discovery.

'Research gateways'

Request for Proposals: Facebook URLs Dataset

Social Science One

February 2020

Please direct all questions about this program to help@socialscience.one, with "URLs RFP" in the subject line.

Overview

Social Science One is a [new kind of partnership](#) between academic researchers and private industry to advance the goals of social science in understanding and solving society's greatest challenges. Our mutually incentive-compatible approach enables academics to analyze and use the increasingly rich troves of information amassed by companies to address societal issues, while protecting their respective interests and ensuring the highest standards of privacy and data security.

URLs dataset

Through Social Science One, researchers can apply for access to a unique Facebook dataset to study questions related to the effect of social media on democracy. The dataset contains approximately an exabyte (a quintillion bytes, or a billion gigabytes) of raw data from the platform, a total of more than 10 trillion numbers that summarize information about 38 million URLs shared more than 100 times publicly on Facebook (between 1/1/2017 and 7/31/2019). It also includes characteristics of the URLs (such as whether they were fact-checked or flagged by

Challenges, prizes and sandboxes

NETFLIX

Netflix Prize

Home Rules Leaderboard Register Update Submit Do

Leaderboard

Rank	Team Name	Best Score	% Impr
No Grand Prize candidates yet			
Grand Prize - RMSE <= 0.8563			
1	PragmaticTheory	0.8584	
2	BellKor in BiqChaos	0.8590	
3	Grand Prize Team	0.8593	
4	Dace	0.8604	
5	BiqChaos	0.8613	
Progress Prize 2008 - RMSE = 0.8616 - Winning Team: BellKor			
6	BellKor	0.8620	
7	Gravity	0.8634	
8	Opera Solutions	0.8640	
9	xivector	0.8640	
10	BruceDengDaoCiyiYou	0.8641	
11	Ces	0.8642	9.17
12	majja2	0.8642	9.17
13	xiangliang	0.8642	9.17
14	Feeds2	0.8647	9.11
15	Just a quv in a garage	0.8650	9.08
16	Team ESP	0.8653	9.05
17	pengpengzhou	0.8654	9.04
18	NewNetflixTeam	0.8657	9.01
19	J.Dennis.Su	0.8658	9.00
20	Vandelay Industries !	0.8658	9.00

data-pitch
INNOVATION PROGRAMME

47
companies

Number of selected companies in seven rounds of applications so far

13
Countries

Number of different countries of origin of the accelerated companies

112
Jobs

Number of additional jobs created by Data Pitch companies

€14.8m
Impact

Total amount of sales, investments and efficiencies unlocked by accelerated companies

Reference APIs/developer platforms

CREATE AWESOME STUFF WITH THE WORLD'S GREATEST COMIC API

The Marvel Comics API allows developers everywhere to access information about Marvel's vast library of comics— from what's coming up, to 70 years ago.

GET STARTED

Product Advertising API 5.0 Documentation

Type to search

Introduction

License Agreement

Register for Product Advertising API

▼ Quick Start

Using cURL

Using Scratchpad

Prime Eligible Products

You can advertise Prime Eligible products using Product Advertising API.

SearchItems

Add the `Prime` filter under `DeliveryFlags` parameter in `SearchItems` to every item contains atleast one offer which is eligible for prime shipping.

Data portability for users

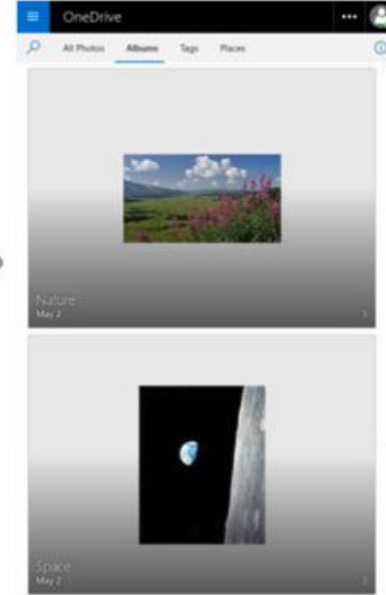
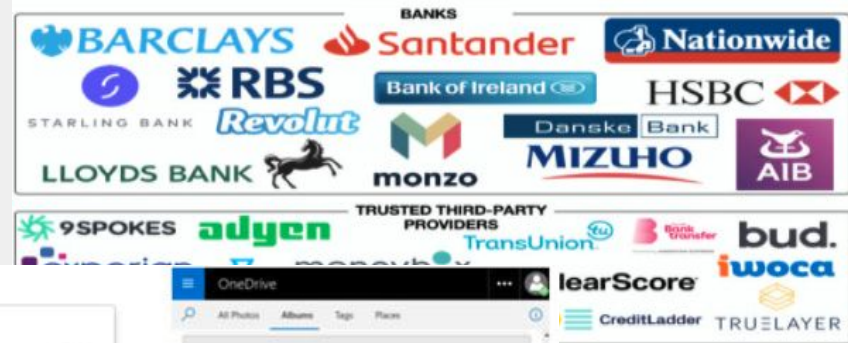
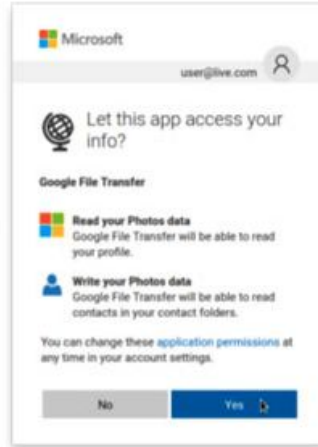
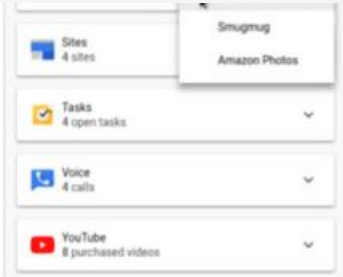
📶 O2-UK 4G 11:32 50%

Settings Data Permissions

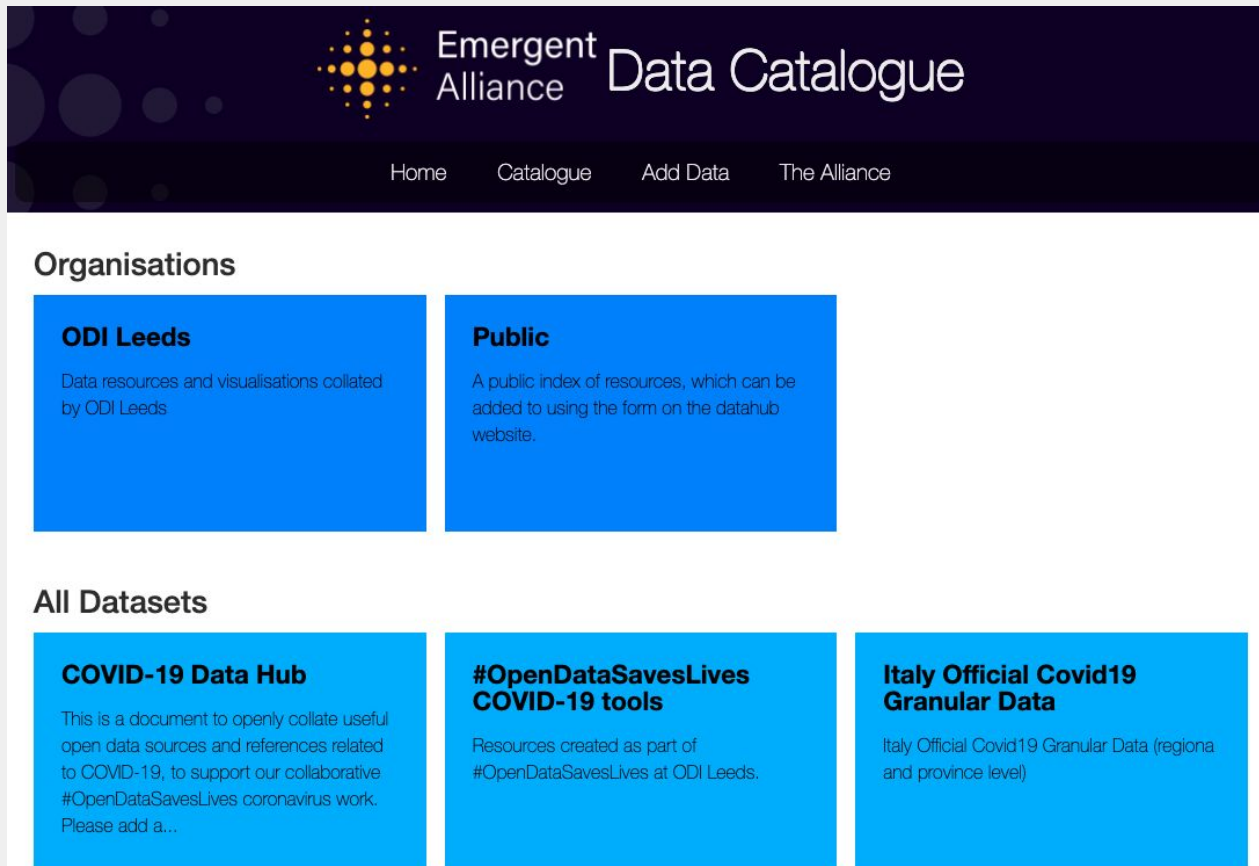
Health-Related Data Select one >

Strava collects and uses health data from paired devices, like a heart rate monitor, to give you interesting and useful performance analysis. We collect this data only from sensors or devices you've connected to Strava. We do not share it without your consent.

[Learn more](#)



Collaborative projects and missions



Emergent Alliance Data Catalogue

Home Catalogue Add Data The Alliance

Organisations

ODI Leeds
Data resources and visualisations collated by ODI Leeds

Public
A public index of resources, which can be added to using the form on the datahub website.

All Datasets

COVID-19 Data Hub
This is a document to openly collate useful open data sources and references related to COVID-19, to support our collaborative #OpenDataSavesLives coronavirus work. Please add a...

#OpenDataSavesLives COVID-19 tools
Resources created as part of #OpenDataSavesLives at ODI Leeds.

Italy Official Covid19 Granular Data
Italy Official Covid19 Granular Data (regional and province level)



**What does this mean
for policy measures?**

Incentives to increase access to data

- **Generate direct commercial value**
e.g. Amazon Product API
- **Create new products/services**
e.g. Strava permissions
- **Mutual value creation**
e.g. data challenges and prizes
- **Respond to consumer demand**
e.g. APIs between mobile apps
- **Philanthropy/‘for good’**
e.g. Google mobility maps
- **Self-regulation**
e.g. FB Ad Library
- **Compliance**
e.g. UK open banking
- ...

Policy measures to increase access to data

- 1. Contextual, based on type of data, actors involved and appropriate access 'style'**

Measures to increase data portability will be different to those needed to increase publication of open data

- 2. Options for different intensities**

e.g. European data sharing support centre (low) vs UK Open Banking mandate (high)

- 3. Relevant at sector, national and international levels**

- a. e.g. Health Data Research UK, UK National Data Strategy, EU Data Strategy

Thank you

If you would like to talk to us about collaborating, partnering, supporting our work, or anything else, we'd love you to get in touch.

jack.hardinges@theodi.org

[@ODIHQ](https://twitter.com/ODIHQ) / [@jhardinges](https://twitter.com/jhardinges)

