ontotext

Finding Open Data sources is a walk in the park: a simple search leads to hundreds of pages of datasets. Governments, NGOs and organizations keep on aggregating and publishing Open Data and more and more businesses and developers use data analytics to gain insights, predict trends and make data-driven decisions.

Why Open Data?

In recent years, Open Data has opened the doors to easier and more efficient ways of finding and analyzing huge datasets. Yet, the true value of Open Data sources is finding ways and using tools to explore, analyze and reuse all that content to minimize efforts and maximize returns.

Once an organization has decided which open datasets it will use, Semantic Technology and more specifically semantic graph databases can help it enrich and classify entities with Linked Open Data, identify relationships between concepts, and disambiguate one concept from another. Linking Open Data is increasingly turning into a means for organizations to stay ahead of competitors.

Government Open Data Sources

Still, the first step to analyzing Open Data is to use as reliable sources of datasets as possible.

Our search for Open Data sources begins with government data.

Data.gov

The US Government's Open Data portal Data.gov has 195,000-plus federal and local datasets on topics ranging from agriculture and education to finance and climate.

However, sometimes datasets need more visuals and more user-friendly experience to become easier to search and time- and cost-efficient to use.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

Data USA

In early April 2016 the MIT Media Lab, in cooperation with Deloitte and Datawheel, launched the Data USA website, a “visualization engine of public US Government data that tells stories about America”, as the developers had put it. Users can search for any of the four categories – locations, industries, occupation or education.

For example, if you type Boston or Pittsburgh in the search field, the website shows an aerial photo of the city with its population, median household income and median age as main statistics. Links below lead to five other categories for each city: economy, demographics, education, housing & living, and health & safety.

The website also features profiles on cross-topics such as ‘Most Common Universities for Computer Science’ or ‘Gender Pay Gap in Connecticut’.

data.gov.uk

In Europe, we find the UK government’s data.gov.uk website, which has aggregated Open Data on topics such as environment, society, towns, and business & economy from government bodies and agencies.

European Data Portal

Europe-wide, we have the European Data Portal developed by the European Commission. This source has nearly 430,000 datasets tagged under various categories, including environment, economy & finance, education, culture & sport.

Global Organizations

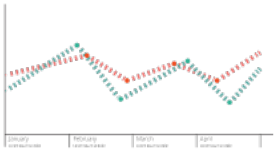
On a supra-governmental and supra-continent level, the World Health Organization and UNICEF provide Open Datasets with statistics on hunger, diseases, deaths, children and women’s health. The World Bank has a free and open access to data about development in countries around the globe, with economy & growth, health, education, environment and climate change featured. So does the OECD data portal.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

If you are not sure where you to begin, OpenDataSoft has compiled a list of more than 2,500 Open Data source portals by country.

Linked Open Data and semantic technology boost data analytics by building ranking reports, trend lines and extending analytics.



Googling Open Data Sources

Still, browsing all these portals each at a time is sometimes tedious and always time- and resource-consuming. Google Public Data – though not as comprehensive as the separate statistics websites – has aggregated some of the most popular and reliable official sources and key economic and health indicators across the world.

For example, a random try search for wages in the US in the ‘Metrics’ menu shows the ‘Compensation of Employees’ report by the U.S. Bureau of Economic Analysis, with charts and comparisons by region or by industry.

Google Trends gives info on the search habits, traffic and interest over time on searches, with historical data dating back to 2004. It also contains infographics on the search interest in global trending topics such as the US Presidential Elections, the Panama Papers, the Brussels attacks or the Zika virus.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext

Crowdsourcing for Open Data

Users and developers are not only browsing for data, they are actively contributing to creating open databases, DBpedia and GeoNames being the most notable examples.

Corporate Data as Open Data Sources

Businesses and organizations may be increasingly using Open Data sources to support decisions but they are reluctant to publish their proprietary data except for statutory filings. If you want to have some basic company info, browsing the websites of the US Securities and Exchange Commission (SEC) or UK's Financial Conduct Authority (FCA), to name just these two, this is a dull and often unproductive task.

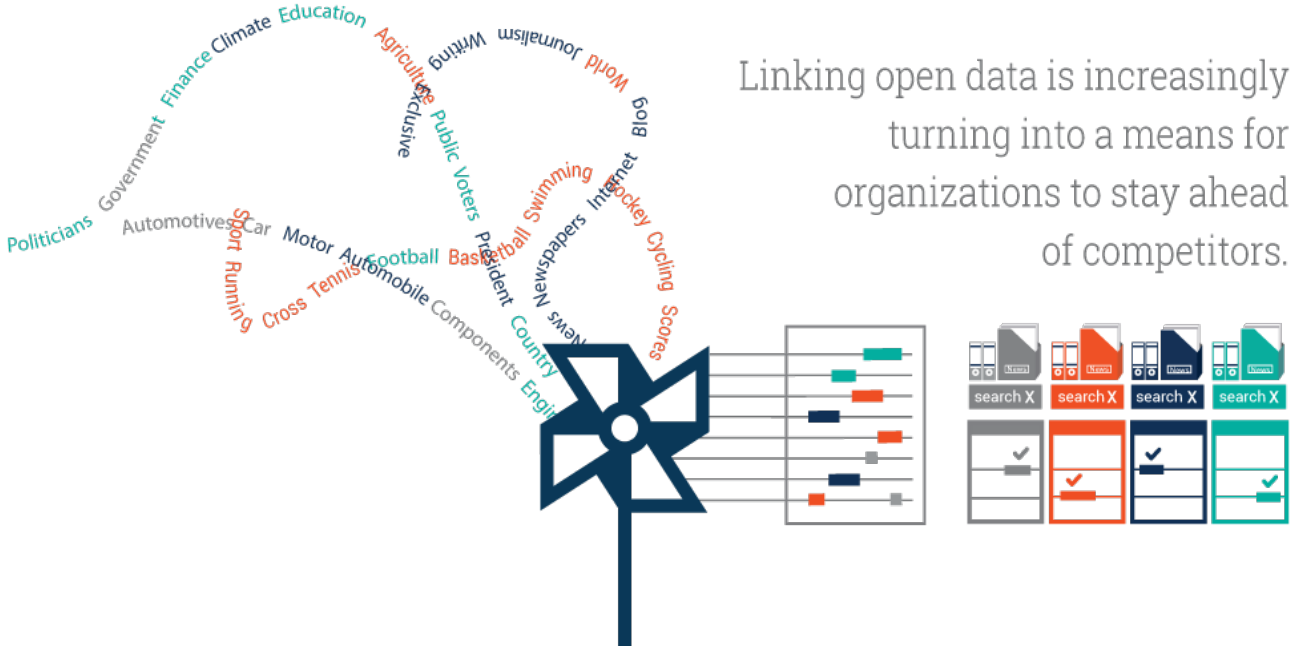
OpenCorporates, a company based at the Open Data Institute (ODI), contains basic data on almost 100,000,000 companies around the world. OpenCorporates has also designed visuals using several sources: filings to the SEC, banking data held by the National Information Center of the Federal Reserve System in the US, and information about individual shareholders published by the official New Zealand corporate registry.

The visualizations show all (the thousands of) subsidiaries in all countries of BP, Bank of America, Citigroup, Goldman Sachs, Morgan Stanley, JP Morgan and Wells Fargo.

Another organization, Berlin-based OpenOil, has collected more than 1 million corporate filings related to the oil, gas and mining industries. It has indexed the full text of contracts, company disclosures, news articles and government reports, which allows users to simultaneously check documents from different sources.

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com

ontotext



Gaining Insights from Open Data

The number of open datasets is only set to grow and so is the need for organizations to have tools to rapidly analyze data in order to have the upper hand in a fierce competitive environment. Linked Open Data and Semantic Technology help organizations boost data analytics by building ranking reports, viewing topics linked implicitly, drawing trend lines, and extending analytics with additional data sources. Share on X

Apart from generating economic and social value, Open Data creates new business models and opportunities. More and more organizations are and will be embracing smart analytics to create additional value for their stakeholders, users and customers.

Download Primer

Copyright: Ontotext New York | Sofia | Basel
Reach out to us: info@ontotext.com