

Vyhledávání dat

Co hledáme?

- „hotová“ data
- Informaci

Co hledám

- Téma
- Geografická lokace
- Časová lokace
- Úroveň – volič, obec, stát
- Dataset x výsledky
- Frekvence dat

Vyhledávání

- Google
 - Strategie záleží na tom co hledáme
 - Co nejpřesnější specifikace
 - správná klíčová slova
 - ne obecná slova a termíny bez jednoznačného významu
 - konkrétní popisná slova.
 - Různý dotaz v různých zdrojích
 - Jazyk
 - <https://datasetsearch.research.google.com/>
 - <https://www.zive.cz/clanky/tipy-pro-google-umi-mnohem-vic-nez-jen-vyhledavat-zkuste-nektere-figle/sc-3-a-196992/default.aspx#part=1>
 - <https://web.archive.org/>

Příklad - vyhledání seznamu starostů obcí ČR

Seznam starostů v ČR

Seznam starostů v ČR filetype:xls

Seznam starostů v ČR filetype:xlsx

Starosta obec filetype:xlsx

starosta obec filetype:xlsx site:mvcr.cz

Data vztažená k publikovaným textům

- Sekundární odkazy
 - Články v databázích
 - wiki

Datové archivy

- Inter-University Consortium for Political and Social Research
 - <https://www.icpsr.umich.edu/web/pages/>
- re3data.org: Registry of Research Data Repositories
 - <https://www.re3data.org/>
- <https://dataverse.harvard.edu/>
- <https://data.mendeley.com/>
- <https://data.europa.eu/>
- github: <https://github.com/search?q=datasets>
- Kaggle: <https://www.kaggle.com/datasets?search=elections>

Příklad exit poll belgických lokálních voleb 2012

Záludnosti nalezených dat

- Věrohodnost
 - Pravdivost
 - Srovnatelnost
 - Úplnost
-
- Kontrola metadat: čas a způsob sběru, instituce provádějící sběr, někdy i zdroj financování
 - Srovnání s jinými daty na dané téma (v jiném čase nebo zemi)

- SVĚT EXISTUJE I MIMO INTERNET a někdy i mimo počítače (archivy, knihy, místa, lidé, ...)