

Вовед во

# DATA

# PIPELINE





# The Data Pipeline

...како ја опишуваме нашата работа со податоци!

- Најди
- Земи
- Верификувај
- Исчисти
- Анализирај
- Презентирај



# Најди ги податоците

Може да се каде било

- Користејќи пребарувачи
  - Google, Bing, Yahoo! & напредни операции
- Мрежа на контакти/извори/вебсајтови
- Преку медиуми
- Други извори на податоци?



# Земи ги податоците

Claim what's yours

- Скрејпување од PDF фајлови & веб-страници
  - Автоматизирани скрипти
  - Бесплатни алатки
- Пристап до машински читливи фајлови
  - RSS, експортирање
- Закон за слободен пристап до информации



# Верификувај

Не им верувај на податоците

- Провери од каде потекнуваат
- Барај мета податоци
  - o Кој ги создал?
  - o Како? Кога?
- Провери ја методологијата на собирање
- Дали има документација?
- Дали треба да се чистат?

# Исчисти ги податоците

На повеќето им треба „туш“

- Дали податоците се создадени за луѓе или за машини?
- Дали има само еден вид податоци во една колона?
- Дали има печатни грешки?
- Дали сите колони имаат име?
- Дали има празни редови? Двојни редови? Колони?
- Дали е потребно да се направи подсет на податоци од оние кои ги користиш?



# Анализирај ги податоците

Кријат тајни, само треба да ги најдеме

- Третирајте ги како извор на информации
- Направете листа на прашања кои ќе ги поставите
- Користете алатки за анализа на податоци (Google Spreadsheets)
- Бидете ригорозни, не се дозволени грешки!
- Документирајте ги чекорите и проверете два пати
- Замолете ги колегите да ги прегледаат резултатите





# Презентирај ги податоците

Ова не се однесува [само] на визуелизации

- Побарајте помош од дизајнер и дивелопер
- Ова е последен чекор
- Поставете тешки прашања
- Одете на најкорисно, не на „најубаво“
- Најчесто најдобриот начин за презентација е наједноставниот
- Не ја игнорирајте комплексноста, но не дозволувајте да преовлада

# БЛАГОДАРАМ



The “Accountability, Technology and Institutional Openness Network in South East Europe - ACTION SEE” project is implemented by Metamorphosis Foundation, Westminster Foundation for Democracy, CRTA – Center for Research, Transparency and Accountability, Citizens Association Why not?, Center for Democratic Transition, Open Data Kosovo (ODK) and [Levizja Mjaft!](http://www.levizjajmjafti.org).

This project is funded by the European Union