Open Data Node

Technické aspekty publikovania otvorených údajov



Peter Hanečák <hanecak@opendata.sk>

Konferencia ZISS, Tatranská Lomnica, 6.5.2016

CC BY

Agenda

- základy:
  - publikujte to, čo máte
  - publikujte tak, ako to máte
  - publikovanie je proces
  - treba komunikovať
  - “keep it simple“
- podrobnosti
- eDemokracia / MOD

Základy

Publikujte to, čo máte.

Nie je potrebná nadpráca, takže:

- ak to stačí vám na prácu, bude to (aspoň na začiatku) plne postačovať aj na zverejňovanie otvorených údajov
- ak to nemáte v počítači, otvorené údaje z toho nebudú
  - najprv treba zdigitalizovať papierové procesy
- ak je v počítači len časť informácií,
  budú nekompletné aj zverejnené otvorené údaje

Základy

Publikujte tak, ako to máte.

Nie je potrebná nadpráca, takže:

- ak to stačí vám na prácu, bude to (aspoň na začiatku) plne postačovať aj na zverejňovanie otvorených údajov
- stačia “surové údaje”, netreba čistiť, opravovať, robiť štatistiky a pod.
  - iba ak by ste vylepšovali priamo interné informácie

Výnimky:

- formátová konverzia potrebná kvôli dodržaniu Výnosu 55/2014
- anonymizácia, ...

Základy

Publikovanie je proces.

- nestačí jednorázovo zavesiť nejaký súbor na web
- treba automatizovať
  - aby sa predišlo opakovaným manuálnym a nudným prácam
  - aby to nestálo zbytočne veľa
- treba to čo najoptimálnejšie začleniť do existujúcich procesov
  - ideálne je mierne upravovať existujúce procesy a informačné systémy

Základy

Treba komunikovať.

Hovorcov zvyčajne máte, takze v spolupráci s nimi:

- informovať / chváliť sa novými datasetmi
- včas hlásiť zmeny, dokumentovať prípadné chyby (a prijaté opatenia)
- reagovať na dopyty od občanov a vývojárov

Špecialita:

- pridávať informácie o datasetoch aj na data.gov.sk
  - lebo Výnos 55/2014

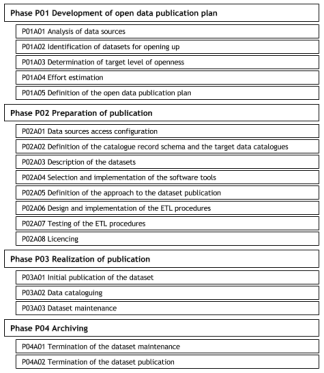
Základy

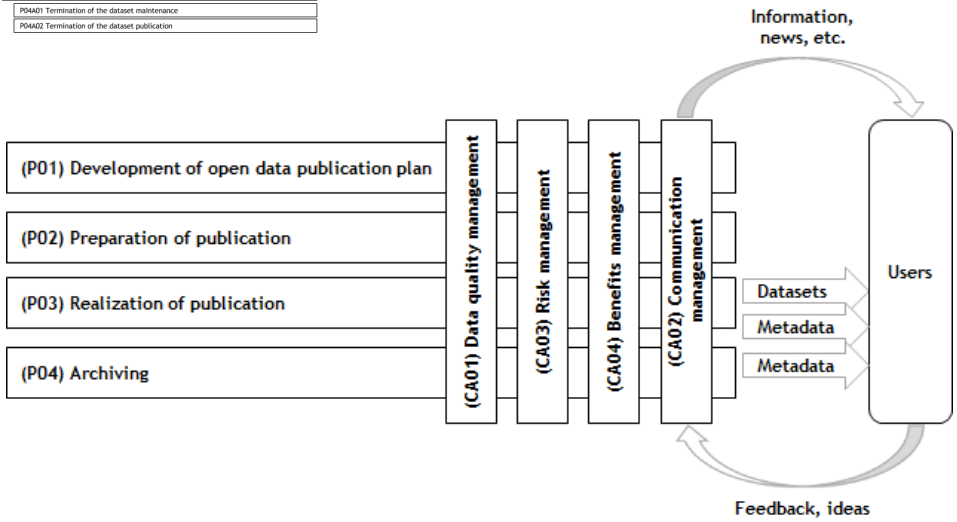
“Keep it simple.“

Komplikovaná je politika a byrokracia, netreba teda komplikovať ešte aj technické veci, Open Data sú preto o.i. aj o jednoduchých riešeniach:

- radšej REST než SOAP
- radšej pár CSV než komplikovaný XML
- atď.

Podrobnosti



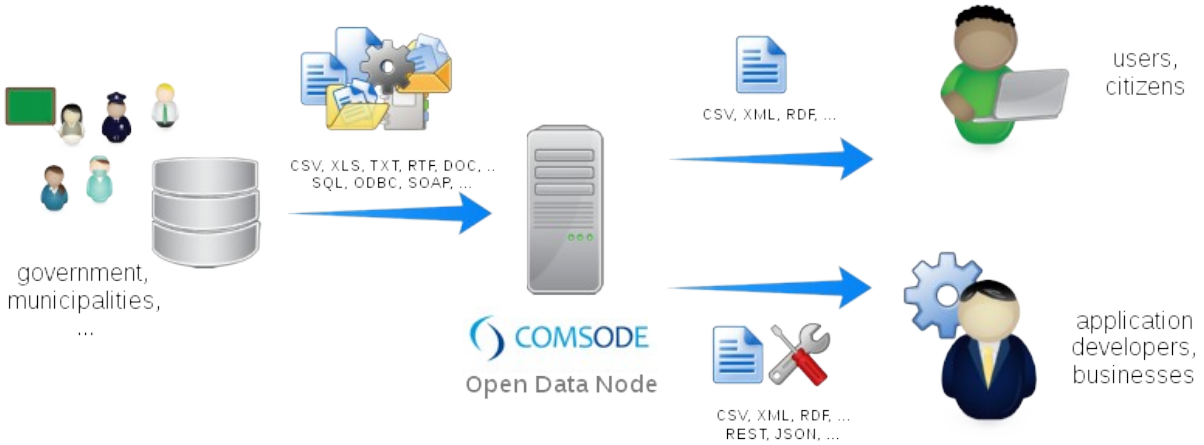


Použiť metodiku (napr. COMSODE alebo eDem):

- publikačný plán
- príprava publikovania
- realizácia publikovania
- archivovanie

reference: http://opendatanode.org/product/methodology-for-od-publishing/

Podrobnosti

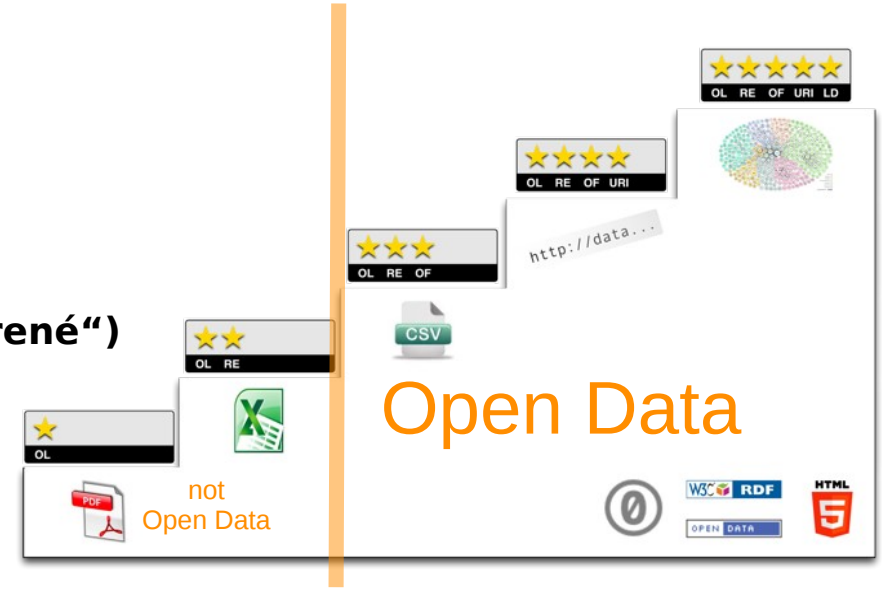


- **základ: ETL / automatizácia zverejňovania**
  - **bonus: menežment údajov**
  - **bonus: nástroje pre používateľov údajov**

Podrobnosti

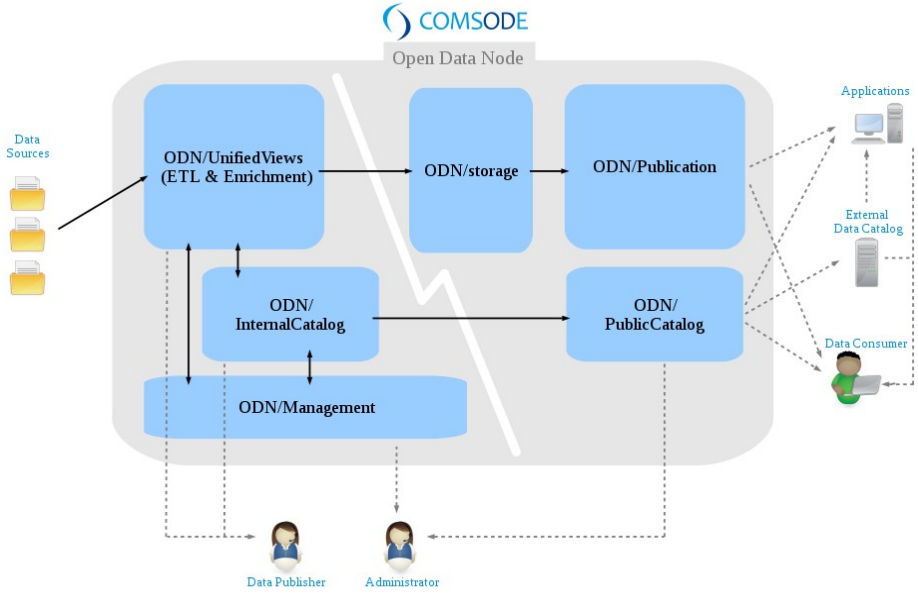
základný scánár: 2* -> 3*+

(t.j. preklopiť „uzavreté“ údaje na „otvorené“)



- **vstupy:** XLS, SQL DB, ...
- **transformácie:** XLS, SQL -> CSV, „zlé CSV“ -> CSV, CSV -> Linked Data
- **výstupy:**
  - tabuľkové/relačné údaje: CSV, REST API
  - Linked Data: RDF, SPARQL endpoint

Podrobnosti



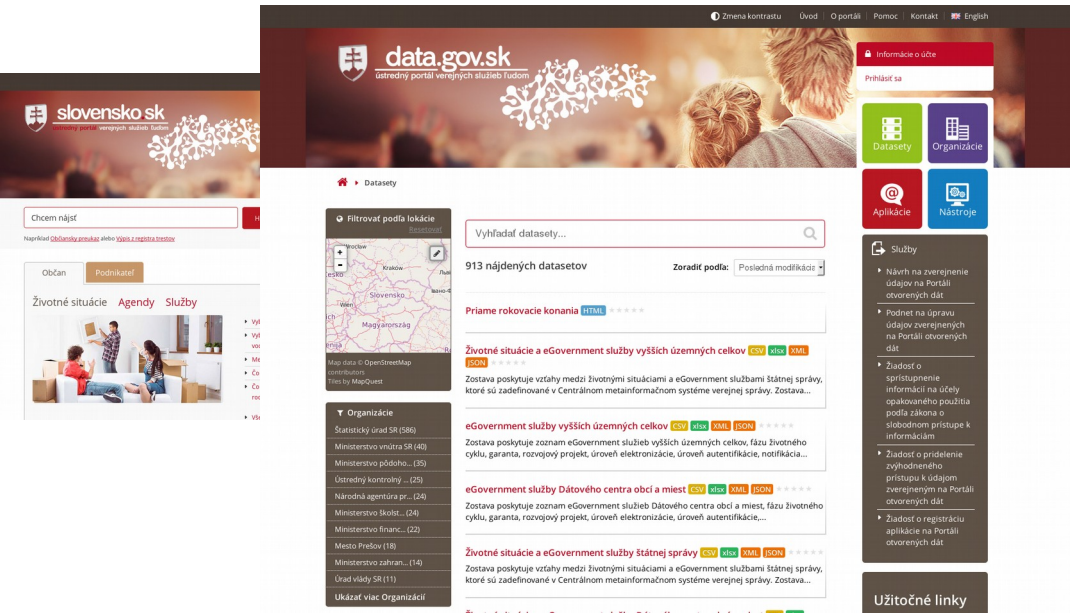
- **základ: ETL / automatizácia zverejňovania**
- **bonus: menežment údajov**
- **bonus: nástroje pre používateľov údajov**

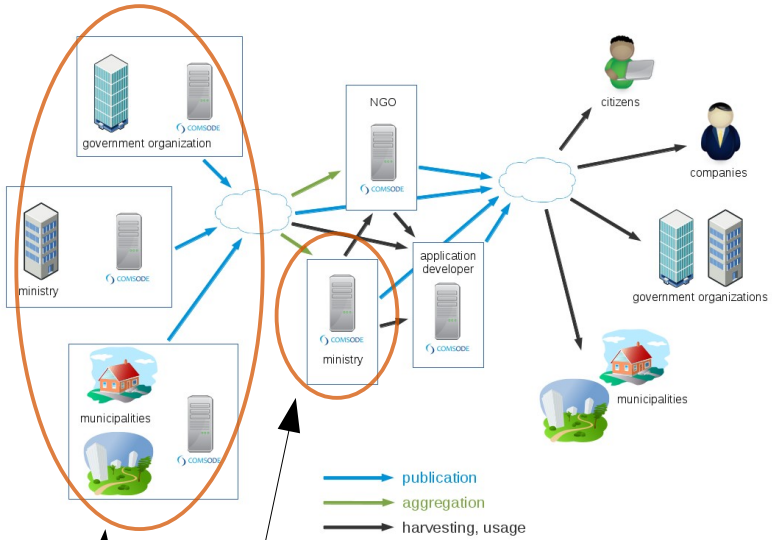
na ZISS už prezentované bolo:
https://www.comsode.eu/wp-content/uploads/201505-ODN-ZISS.pdf
resp. aktuálnejšia verzia:
http://opendatanode.org/wp-content/uploads/201604-SOIT_OSS_weekend-ODN-technical.pdf

eDemokracia / MOD

Projekt eDemokracia, Modul Open Data (MOD):

- **centrálny komponent**
- **decentralizovaný komponent**





centrálny komponent

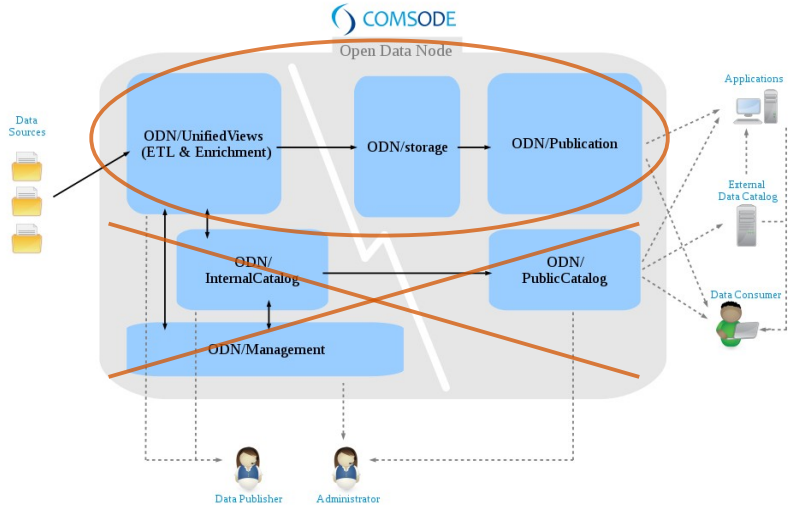
decentralizovaný komponent

eDemokracia / MOD

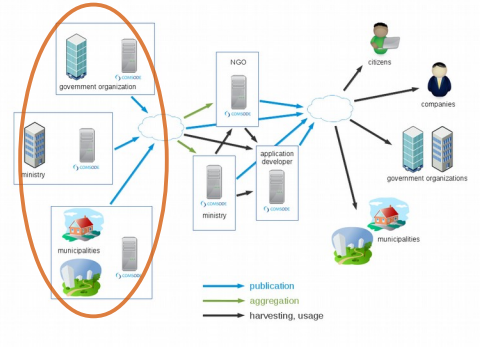
Centralizovaný komponent:

- data.gov.sk – dátový katalóg
- údajový back-end
- transformačné sluzby
- prevázkuje NASES
- ochota zdieľať aj so samosprávou





eDemokracia / MOD



Decentralizovaný komponent:

- tzv. Podporná aplikácia
- ODN upravený pre špecifické slovenské podmienky
- určený pre štátne inštitúcie ale aj samosprávu
- nápomocný pri komplikovanejších scénároch publikovania

