



Ανοιχτά δεδομένα και ευφυής πόλη: από την ποσότητα στην ποιότητα



Βαγγέλης Νομικός
(Co-Founder & CTO)

Intelibility
Innovative Data Engineering

Σύρος,
07/07/2017



“Τα δεδομένα είναι το καύσιμο της ψηφιακής οικονομίας και η πρώτη ύλη για την κοινωνία της γνώσης.”



“Τα δεδομένα οδηγούν στην πληροφορία με
την οποία χτίζεται η γνώση.
Η γνώση είναι δύναμη. ”



“Το λογισμικό τρώει τον κόσμο.”



Η πρόκληση...

Δεξιότητες
στην
Επιστήμη
Δεδομένων

+

Δεξιότητες
στην
Ανάπτυξη
Λογισμικού

=>

Καινοτόμες
Εφαρμογές
και
Υπηρεσίες



Ποιοι είμαστε

- ✕ νέα εταιρεία αλλά με έμπειρη ομάδα
 - διαχείριση ανοιχτού πολιτιστικού περιεχομένου
 - ανάπτυξη αποθηκών δεδομένων και αποθετηρίων
 - τεχνικές ανάλυσης δεδομένων με εφαρμογές σε διάφορα πεδία (π.χ. πωλήσεις, “Metadata Analytics & Quality over OAI-PMH”)
 - ανάπτυξη λογισμικού
- ✕ προϊόντα και υπηρεσίες για την παγκόσμια αγορά

Repository Analytics (repolitics.com)

[Home](#)[About](#)[Features](#)[Contact](#)[Sign in](#)

Features



Measure metadata quality

Analyze metadata elements and values and derive quality metrics. Assess completeness, accuracy and distinct values to see whether the metadata conform to particular standards.



Data profiling

Analyze repository data and produce informative summaries about files. Collect statistics such as file type and file size distributions. Have a quick overview of your storage needs.



Real-time monitoring of user activity

Capture and analyze web-server logs to identify hidden access patterns. Configure the agent on your infrastructure and let repolitics produce useful metrics.



Multiple sources

Integrate data from multiple sources: OAI-PMH providers, upload archive files, open data portals (i.e., CKAN, DKAN).



Multiple metadata schemas

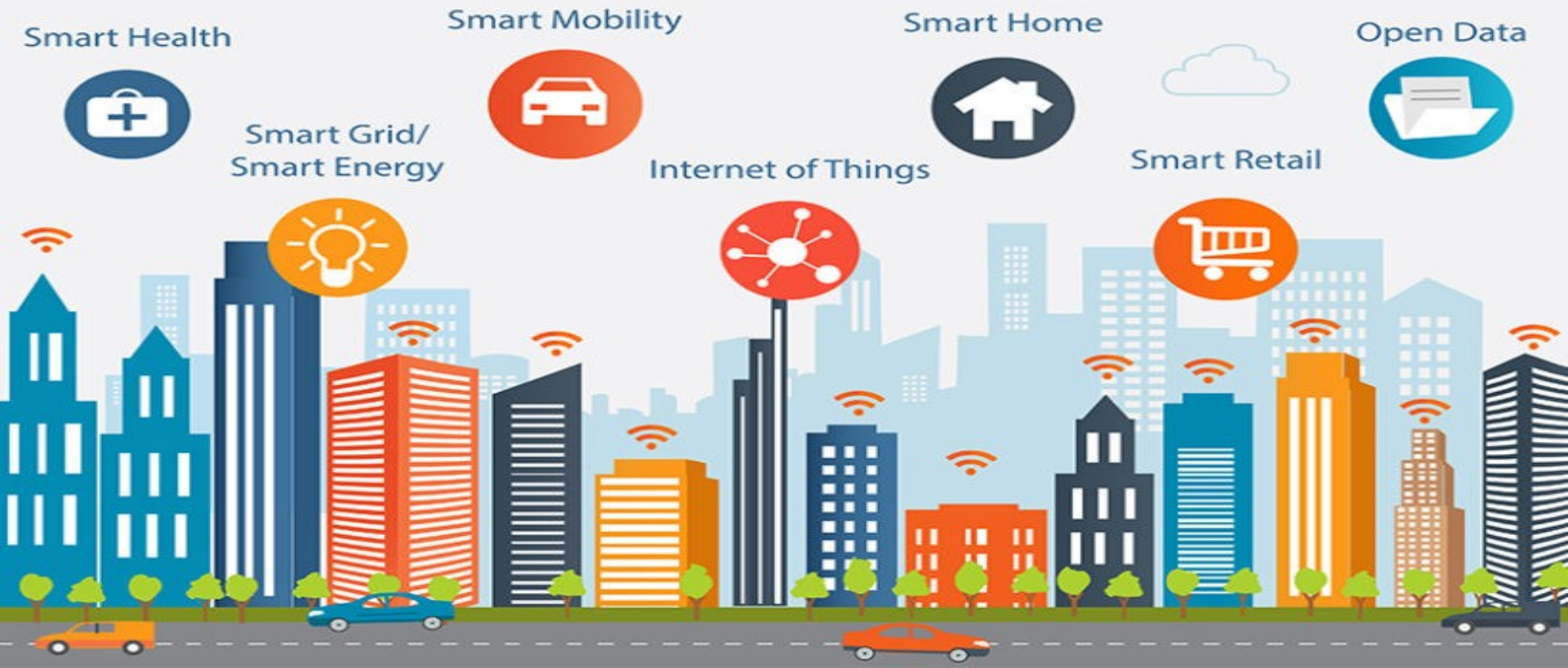
Easy integration of multiple metadata schemas. OAI_DC, ORE, METS are currently supported. On demand integration of other metadata schemas.



Analyze your subscription data

Upload your annual publishing subscription data and easily identify usage patterns and opportunities to increase the efficiency of your subscriptions.

Ανοιχτά Δεδομένα και Έξυπνες Πόλεις





Γιατί ανοιχτά δεδομένα;

- ✗ ανοιχτά δεδομένα + οι νέες τεχνολογίες $\neq$ έξυπνη πόλη
Η τεχνολογική μόνο προσέγγιση δεν είναι αρκετή.
- ✗ πρωτοβουλία των έξυπνων πόλεων και ένα “city asset”
- ✗ αναγκαία διάσταση για μια σύγχρονη, πολιτοκεντρική και εξωστρεφή τοπική αυτοδιοίκηση
 - αυξημένη διαφάνεια και δέσμευση
 - ενίσχυση της οικονομίας (επαναχρησιμοποίηση περιεχομένου, ανάπτυξη καινοτόμων εργαλείων και υπηρεσιών)



Κοινό ανοιχτών δεδομένων

- ✕ πολίτες (κάτοικοι, εργαζόμενοι, επισκέπτες)
- ✕ ερευνητές, δημοσιογράφοι που θέλουν να συγκεντρώσουν και να αναλύσουν τα δεδομένα για να πουν ιστορίες
- ✕ προγραμματιστές για ανάπτυξη εφαρμογών
- ✕ άλλοι φορείς και οργανισμοί



“Δεν μπορούμε να έχουμε
αναλογικές κυβερνήσεις που να
εξυπηρετούν μια ψηφιακή
οικονομία και μια ψηφιακή
κοινωνία.”

(Barbara Ubaldi - ΟΟΣΑ)



“Παρατηρούμε ότι συχνά οι άνθρωποι θεωρούν ότι τα ανοιχτά δεδομένα είναι ένα συγκεκριμένο “είδος” δεδομένων – κάτι διαφορετικό και μοναδικό από τα δεδομένα που χρησιμοποιούν καθημερινά στον οργανισμό ή την ομάδα τους – αντί για μια επιλογή σχετικά με τον τρόπο που οι άνθρωποι δημοσιεύουν δεδομένα.”

(Open Data Institute)



Διεθνής πρακτική



Categories

[Administration & Finance](#)[Buildings](#)[Community & Economic Development](#)[Education](#)[Environment & Sustainable Development](#)[Show All...](#)

View Types

[Calendars](#)[Charts](#)[Data Lens pages](#)[Datasets](#)[External Datasets](#)[Files and Documents](#)[Filtered Views](#)[Forms](#)

1157 Results

Sort by

Most Accessed

Current Employee Names, Salaries, and Position Titles

Administration & Finance

Dataset

This dataset is a listing of all current City of Chicago employees, complete with full names, departments, positions, employment status (part-time or full-time), frequency of hourly employee -where...

[More](#)Tags [personnel](#)[API Docs](#)Updated
May 2, 2017Views
893,305

Crimes - 2001 to present - Dashboard

Public Safety

Data Lens

This dataset reflects reported incidents of crime (with the exception of murders where data exists for each victim) that occurred in the City of Chicago from 2001 to present, minus the most recent seven days...

[More](#)Tags *No tags assigned*Updated
July 6, 2017Views
701,699

Beach Weather Stations - Automated Sensors - 2016 - Humidity

Parks & Recreation

Chart

The Chicago Park District maintains weather sensors at beaches along Chicago's Lake Michigan lakefront. These sensors generally capture the indicated measurements hourly while the sensors are in operation...

[More](#)Tags [water](#), [recreation](#), [parks & recreation](#), [open spaces](#), [chicago park district](#), and 2 moreUpdated
July 6, 2017Views
516,326

Building Permits

Buildings

Dataset

Permits issued by the Department of Buildings in the City of Chicago from 2006 to the present. The dataset for each year contains more than 65,000 records/rows of data and cannot be viewed in full In...

Updated
July 6, 2017

Views

1157
σύνολα
δεδομένων



Open Data

Brussel
Bruxelles
Brussels

[Home](#)[Verkennen](#)[Cartograaf](#)[API](#)[Licentie](#)[FAQ](#)[Nieuws](#)[Links](#)[Suggestions](#)[Contact](#)

Open Data Portaal

Raadpleeg vrij de gegevens van de Stad Brussel

Open Data Brussel

De Stad Brussel stelt - als onderdeel van haar Open Data initiatief - haar datasets ter beschikking van iedereen. U hebt zo toegang tot honderden gegevens rond cultuur, toerisme, mobiliteit, gemeentelijke installaties,...

Hoe zoeken?

Via Verkennen in het menu is het mogelijk om datasets te zoeken en te selecteren, ofwel via zoeken ofwel door te klikken op de 'facetten' (linkerkolom).

Tip: selecteer eerst de taal.

Wijzig en publiceer vrij

Deze gegevens zijn van u en het is dus mogelijk om ze in alle vrijheid op te halen, aan te passen en te publiceren. De open licentie van deze datasets geeft u het recht om te doen wat u wil met deze gegevens.

Ontdek één van onze 540 open datasets!

De laatste datasets

 [Parking spaces for disabled](#)

 Gewijzigd 3 maanden geleden

 [Parkeerplaatsen voor gehandicapten](#)

 Gewijzigd 3 maanden geleden

 [Parkings pour personnes handicapées](#)

 Gewijzigd 3 maanden geleden

 [Crèches & préguardiennats](#)

 Gewijzigd 3 maanden geleden

De populairste datasets

 [Villo!-stations & beschikbaarheid in real time](#)

 107279 downloads

 [Villo! stations & availability in real time](#)

 32786 downloads

 [Stations Villo! & disponibilités en temps réel](#)

 10477 downloads

 [Collecto stops](#)

 1516 downloads

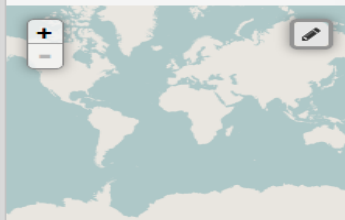
540
σύνολα
δεδομένων



Filter by location

[Clear](#)

Search Location...



Map © OpenStreetMap contributors
[Disclaimer](#)

Countries

Czech Republic (335937)

Germany (207965)

United Kingdom (42474)

Italy (33042)

France (27664)

Spain (26856)

Netherlands (20697)

Finland (9859)

Belgium (6039)

Greece (5536)

Datasets Feed

Save Query

Search datasets...



769,567 datasets found

Order by:

Last Modified ▾

Value added tax (VAT) registrations by sector.

VAT registrations by NACE code Sector.

Data.gov.ie | May 23, 2016

[CSV](#) [PDF](#)

Revenue net receipts by taxhead

Table showing net tax receipts by taxhead for years since 2003

Data.gov.ie | May 23, 2016

[CSV](#) [PDF](#)

Sheet 7072 — Haan — DGK5 tk25-nr. 4808/01

The German Land Map (DGK5) 1: 5000 is a topographic base map. It represents the earth's surface is geometrically, taking into account the caps. The continuation in the future...

GDI-DE | February 14, 2016

dop20 dop20rgb_32532_6010_1_sh_2012,

The ortho-photo dop20rgb_32532_6010_1_sh_2012 was calculated using the DGM10 The aerial survey took place in 2012. "The ortho-photo dop20rgb_32532_6010_1_sh_2012 issued by:...

769.567
σύνολα
δεδομένων



Ελληνική πραγματικότητα

- Κατηγορίες Περιεχομένου
- Σύνολο δεδομένων
- Θέματα
- Αστικός Σχεδιασμός (36)
- Δημόσια Διοίκηση (13)
- Περιβάλλον (13)
- Τουρισμός (6)
- Εκπαίδευση (5)
- Πολιτισμός (4)
- Δημόσια ασφάλεια (1)
- Οικονομία (1)
- Λέξεις κλειδιά
- Τύπος αρχείου
- Εκδότης
- Άδεια χρήσης

74 αποτελέσματα

Search

Αναζήτηση

Ταξινόμηση

Date changed

Φθίνουσα

Εφαρμογή

Καθαρισμός

Καταγραφή ημερήσιας δραστηριότητας των πλυντηρίων οχημάτων του Τμήματος Σταθμού Μεταφόρτωσης Απορριμμάτων του Δήμου Θεσσαλονίκης

Δήμος Θεσσαλονίκης

Περιβάλλον

Το σύνολο δεδομένων περιέχει την καταγραφή των οχημάτων (ημερομηνία, αριθμός και τύπος οχήματος) που πλύνονται καθημερινά στα πλυντήρια οχημάτων του Τμήματος Σταθμού Μεταφόρτωσης Απορριμμάτων της Διεύθυνσης Ανακύκλωσης & Διαχείρισης Αστικών Απ

27x ods

Ημερήσια καταγραφή ανά βάρδια εργασίας (από την αναφορά υπευθύνου βάρδιας) των κοντίνερ που διακινούνται από τον Σταθμό Μεταφόρτωσης Απορριμμάτων στο ΧΥΤΑ

Δήμος Θεσσαλονίκης

Περιβάλλον

Το σύνολο δεδομένων περιέχει την ημερήσια καταγραφή (αναφορά υπευθύνου βάρδιας) των κοντίνερ που διακινούνται από τον Σταθμό Μεταφόρτωσης Απορριμμάτων της Διεύθυνσης Ανακύκλωσης & Διαχείρισης Αστικών Απορριμμάτων του Δήμου Θεσσαλονίκης στο Χώ

12x ods 17x zip

74
σύνολα
δεδομένων



Αναζήτηση δεδομένων

π.χ. Περιβάλλον



Ετικέτες

Γενικά Αρχεία του Κ...

Αρχεία

Εργαλεία έρευνας

5.532
σύνολα
δεδομένων

Το data.gov.gr είναι ο κεντρικός κατάλογος των δημόσιων δεδομένων που παρέχει πρόσβαση σε βάσεις δεδομένων των φορέων της ελληνικής κυβέρνησης.



data.gov.gr στατιστικά

5,5k

σύνολα
δεδομένων

246

Φορείς

9

ομάδες

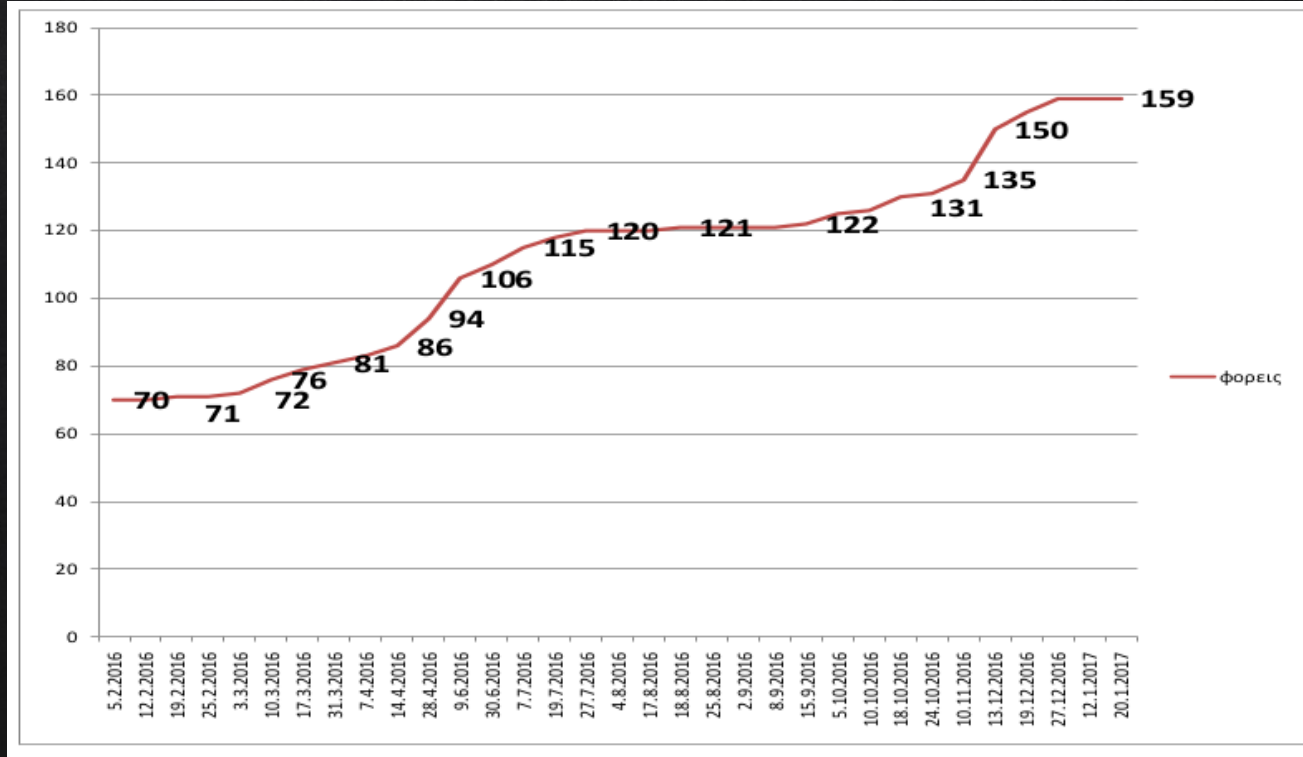
6

σχετικά στοιχεία

data.gov.gr links

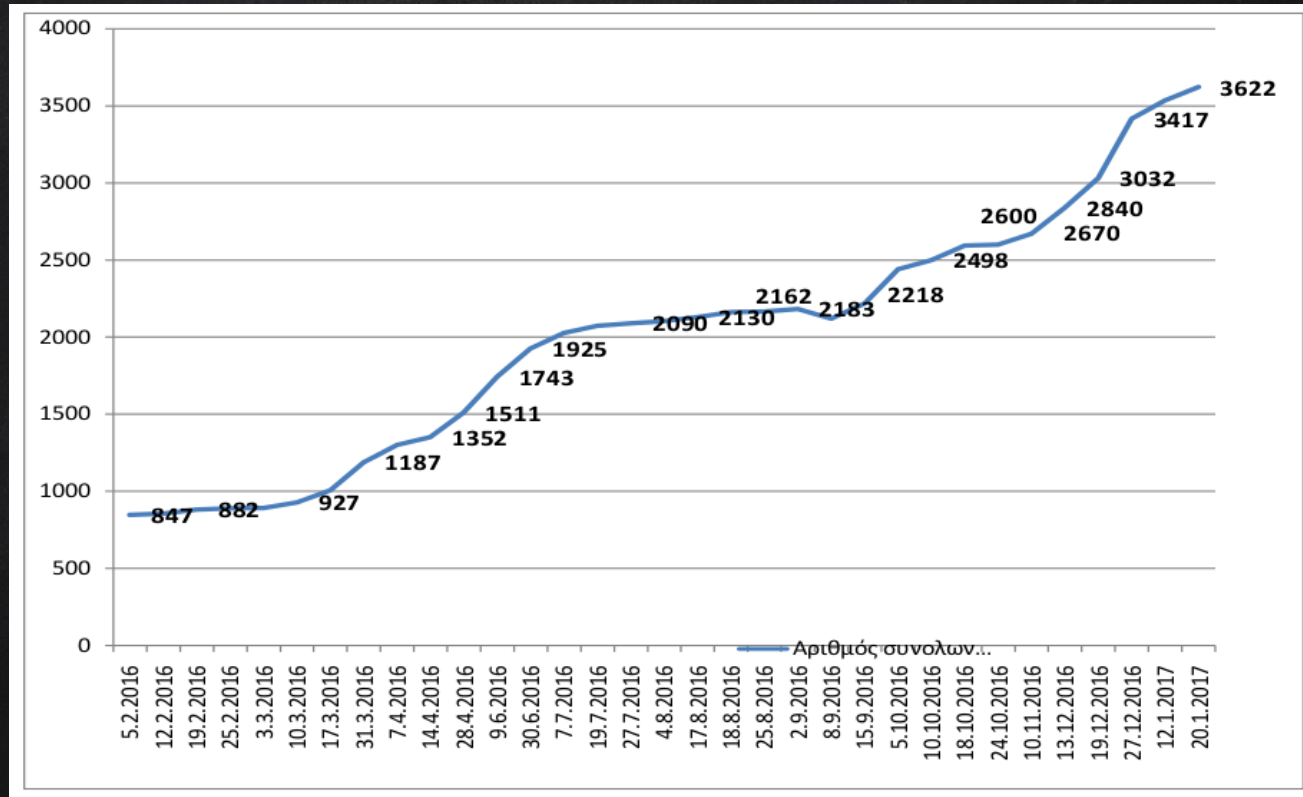
[Αποφάσεις διάθεσης δεδομένων](#)[Οδηγός Διαχειριστή Φορέα](#)[Ετήσια Έκθεση](#)

Εξέλιξη ένταξης φορέων



Πηγή: www.data.gov.gr (5.2.2016 – 20.1.2017)

Εξέλιξη συνόλων δεδομένων



Πηγή: www.data.gov.gr (5.2.2016 – 20.1.2017)

15%
δήμων
χώρας

31%
σύνολα
δεδομένων

Πηγή: www.data.gov.gr (5.2.2016 – 20.1.2017)

Σήμερα...

~24%
δήμων
χώρας
(78/325)

~33%
σύνολα
δεδομένων
(1823/5532)



Top 5

Δήμος	# σύνολα δεδομένων
Δήμος Ραφήνας - Πικερμίου	148
Δήμος Νίκαιας – Αγίου Ι. Ρέντη	95
Δήμος Ιλίου	86
Δήμος Θεσσαλονίκης	74
Δήμος Ηρακλείου	61

Ποσοτική πρόοδος



- ✘ Το κίνημα των ανοιχτών δεδομένων κινδυνεύει. Τα ανοιχτά δεδομένα είναι η εξαίρεση και όχι ο κανόνας (Open Data Barometer)
- ✘ Υπουργείο Διοικητικής Ανασυγκρότησης
 - Εγκύκλιος για επίσπευση ενεργειών για τη διάθεση και χρήση των ανοικτών δεδομένων (01/2017)
 - Ετήσια έκθεση για τη διάθεση και περαιτέρω χρήση των ανοικτών δεδομένων (03/2017)
- ✘ Απογραφή ανοικτών δεδομένων ελληνικών πόλεων 2016(OKF Greece)



Το πρόβλημα

- ✗ ποσοτική και όχι ποιοτική πρόοδος
- ✗ τα σύνολα δεδομένων που είναι ανοιχτά τείνουν να είναι ξεπερασμένα, κακής ποιότητας και δύσκολο να βρεθούν
- ✗ σπανίως είναι αδειοδοτημένα ως ανοιχτά
- ✗ φτωχά μεταδεδομένα
- ✗ προβλήματα στη μορφή και στη διάθεση
- ✗ χαμηλός βαθμός διάχυσης και χρήσης
- ✗ έλλειψη παραδειγμάτων και καλών πρακτικών
- ✗ δεν αρκεί η ανάρτηση σε μια ιστοσελίδα

Προβλήματα στη μορφή

ΜΟΡΦΟΤΥΠΟΣ	ΑΡΙΘΜΟΣ DATASETS	ΑΡΙΘΜΟΣ DATASETS	Μεταβολή (%)
	2016	2017	
XLS/XLSX	272	1074	294,85%
HTML	55	451	720,00%
CSV/XML/JSON	25	280	1020,00%
PDF	50	260	420,00%
ZIP/RAR	37	175	372,97%
DOC/DOCX	55	142	158,18%
KML/KMZ/SHP	11	57	418,18%
TXT	34	34	0,00%
JPEG/PNG/GIF/TIFF	13	21	61,54%

Πηγή: <https://goo.gl/5Pks3M>

Έλεγχος Μεταδεδομένων

```
{
  "help": "http://www.data.gov.gr/api/3/action/help_show?name=package_show",
  "success": true,
  "result": {
    "license_title": "Μη καθορισμένη άδεια",
    "maintainer": "Μπούσια Ιωάννα",
    "relationships_as_object": [],
    "private": false,
    "maintainer_email": "mpousia@melissia.gr",
    "num_tags": 2,
    "id": "02af136b-0fb7-4765-89b7-e6e8491e9e04",
```

```
{
  "help": "http://www.data.gov.gr/api/3/action/help_show?name=package_show",
  "success": true,
  "result": {
    "license_title": "Άλλο (Ανοιχτό)",
    "maintainer": "",
    "relationships_as_object": [],
    "private": false,
    "maintainer_email": "",
    "num_tags": 0,
    "id": "ffe8a86a-32bd-4588-96ab-b40005ff22ef",
    "metadata_created": "2016-10-06T10:05:22.998174",
    "metadata_modified": "2017-05-18T10:56:44.811334",
    "author": "",
    "author_email": "",
    "state": "active",
```



Ποιοτικά Ανοιχτά Δεδομένα

“Η αξία των δεδομένων έγκειται στη χρήση των δεδομένων και εξαρτάται από την ποιότητα τους για ένα συγκεκριμένο σκοπό.”



Απαιτήσεις σε πολλαπλά επίπεδα...

- ✗ πρακτικό (π.χ. συχνές ενημερώσεις για αλλαγές)
- ✗ νομικό (κατάλληλη άδεια χρήσης)
- ✗ τεχνικό (π.χ. μορφότυπος, δομή, κανάλια διάθεσης)
- ✗ κοινωνικό (π.χ. ανάπτυξη κοινοτήτων, υποστήριξη)



Δείκτες ποιότητας

- ✗ υπάρχουν αλλά δεν είναι πλήρεις
- ✗ αξιολογούν την ποιότητα πολύ διαφορετικά
- ✗ η ποικιλομορφία των υφιστάμενων δεικτών αποτρέπει μια στοχοθετημένη και στρατηγική προσέγγιση



Βελτίωση της ποιότητας

- ✕ βελτιώνοντας τον τρόπο παραγωγής
- ✕ χρήση προτύπων (community-driven)
 - 5* Open Data (Tim Berners-Lee)
 - Open Data Certificates (ODI)
- ✕ εργαλεία για καθαρισμό / επικύρωση δεδομένων
 - OpenRefine
 - DataCleaner
 - csvlint.io



Άλλες πρωτοβουλίες

- ✕ τα ανοιχτά δεδομένα πρέπει να είναι μια μακροπρόθεσμη δέσμευση
- ✕ ανάδειξη επιτυχημένων παραδειγμάτων χρήσης
- ✕ επαναχρησιμοποίηση από φορείς και υπηρεσίες της πόλης
- ✕ άνοιγμα δεδομένων με βάση τη ζήτηση
- ✕ ενίσχυση της κουλτούρας της ανοιχτότητας
- ✕ προώθηση “data literacy” και ανάπτυξη δεξιοτήτων
- ✕ εμπλοκή των πολιτών



“κάποια χρόνια πριν το κίνημα των ανοιχτών δεδομένων ήταν σε αχαρτογράφητα νερά, τώρα έχει μια παγκόσμια δυναμική. Αυτό που χρειάζεται είναι ένα “focus shift” από την ποσότητα στην ποιότητα.”



“Έξυπνες θα είναι οι πόλεις που θα βελτιώσουν τη λειτουργία τους, θα χρησιμοποιούν δεδομένα για τη λήψη αποφάσεων και θα χρησιμοποιούν την ψηφιακή τεχνολογία για να συμμετέχουν ενεργά οι πολίτες. ”



Ευχαριστώ!

Email:
nomikos@intelibility.com

Πηγές

- ✗ Open Data Barometer - <https://goo.gl/6qwV8p>
- ✗ Open Data Portal - <https://goo.gl/u2BncW>
- ✗ Εθνική Πύλη Ανοιχτών Δεδομένων - <http://data.gov.gr/>
- ✗ Εγχειρίδιο Ανοιχτών Δεδομένων - <https://goo.gl/J9pu0>
- ✗ Ποιότητα ανοιχτών δεδομένων: η επόμενη αλλαγή στα ανοιχτά δεδομένα - <https://goo.gl/CeIn3G>
- ✗ Δημόσιοι φορείς και ανοιχτά δεδομένα στην Ελλάδα - <https://goo.gl/5Pks3M>
- ✗ Απογραφή Ανοιχτών Δεδομένων Ελληνικών Πόλεων 2016 (OKF Greece) - <https://goo.gl/z7EXqn>