

ΑΝΟΙΚΤΑ ΔΕΔΟΜΕΝΑ & ΔΗΜΟΣΙΟΙ ΦΟΡΕΙΣ

ΜΕΡΟΣ Α | 2η ώρα: Ανοιχτά πρότυπα και σύστημα 5 αστέρων

Απόστολος Κρητικός
Open Data Certified Trainer, ODI

Βασίλης Χρυσός
Open Data Certified Trainer, ODI

Ατζέντα

1. *Εισαγωγή στα δεδομένα.*
2. *Εισαγωγή στα ανοικτά δεδομένα.*
3. *Τα ανοικτά δεδομένα ως μέσο επικοινωνίας μεταξύ δημόσιου φορέα και πολίτη.*
4. *Αδειοδότηση (licensing) ανοικτών δεδομένων*
5. *Τυποποίηση Δεδομένων*
6. *Δεδομένα & μηχανές*

Μαθησιακοί στόχοι

Στο τέλος της συνεδρίας οι συμμετέχοντες θα μπορούν να:

1. Περιγράψουν την ανάγκη για χρήση ανοιχτών προτύπων
2. Αναπαραγάγουν τον ορισμό ανοιχτού προτύπου
3. Απαριθμήσουν τις βασικές προκλήσεις για την εφαρμογή ανοιχτών προτύπων και να προτείνουν τρόπους αντιμετώπισης
4. Περιγράψουν το σύστημα των πέντε αστέρων για τη δημοσίευση δεδομένων
5. Απαριθμήσουν πλεονεκτήματα / μειονεκτήματα του κάθε επιπέδου

Τυποποίηση ανοιχτών δεδομένων (5/6)

Ερώτηση

Πιστεύετε ότι είναι απαραίτητη η τυποποίηση στα ανοιχτά δεδομένα;

Για ποιους λόγους;



Γιατί τυποποίηση;

Κανένα σύνολο δεδομένων δεν είναι απομονωμένη νησίδα.

Η πραγματική αξία των δεδομένων ξεδιπλώνεται όταν φέρνουμε τα σύνολα δεδομένων μαζί: για σύγκριση, αντιπαραβολή και ανάλυση.

Δεδομένα από:

- διαφορετικά τμήματα ενός οργανισμού, ή
- ίδια τμήματα διαφορετικών οργανισμών, ή
- διαφορετικά τμήματα διαφορετικών οργανισμών

Γιατί τυποποίηση;

Επιτρέπει ευκολότερη πρόσβαση και χρήση σε φορείς που δημοσιεύουν και / ή χρησιμοποιούν ανοιχτά δεδομένα.

(Ποιοι είπαμε ότι μπορεί να είναι τέτοιοι φορείς;)

Τυποποιημένα ανοιχτά δεδομένα

Επιτρέπουν τη **διαλειτουργικότητα των δεδομένων** παρέχοντας μια **κοινή βάση κατανόησης** για το τι είναι αυτό που περιγράφεται και με ποιο τρόπο.

“Βέλτιστη πρακτική: επαναχρησιμοποίηση τυποποιημένων λεξιλογίων δεδομένων για την κωδικοποίηση των δεδομένων και των μεταδεδομένων”¹

W3C

Σε τι αναφέρεται η τυποποίηση;

Μπορεί να αναφέρεται σε πολλά διαφορετικά πράγματα, όπως:

- αρχές και πολιτικές
- ορισμός του 'ανοιχτού'
- μορφότυποι αρχείων δεδομένων
- πολιτικές για συγκεκριμένες θεματικές
- γενικά τεχνικά πρότυπα
- τεχνικά πρότυπα για συγκεκριμένες θεματικές

Ορισμός Ανοιχτού Προτύπου από FSFE (1/2)

Ένα Ανοιχτό Πρότυπο αναφέρεται σε έναν τύπο αποθήκευσης ή ένα πρωτόκολλο

1. το οποίο υπόκειται σε πλήρη δημόσια αξιολόγηση και χρήση χωρίς περιορισμούς με τρόπο ισότιμα διαθέσιμο σε όλους τους ενδιαφερόμενους·
2. χωρίς τμήματα ή επεκτάσεις που να εξαρτώνται από τύπους αρχειοθέτησης ή πρωτόκολλα τα οποία δεν πληρούν τα ίδια τον ορισμό του Ανοιχτού Προτύπου·
3. ελεύθερο από νομικές/τεχνικές διατυπώσεις που περιορίζουν τη χρησιμότητά του από οποιονδήποτε ενδιαφερόμενο ή επιχειρηματικό μοντέλο·

Ορισμός Ανοιχτού Προτύπου από FSFE (2/2)

Ένα Ανοιχτό Πρότυπο αναφέρεται σε έναν τύπο αποθήκευσης ή ένα πρωτόκολλο

4. με διαχείριση και επιπλέον ανάπτυξη ανεξάρτητα από οποιονδήποτε μοναδικό προμηθευτή σε μια διαδικασία ανοιχτή στην ισότιμη συμμετοχή ανταγωνιστών και τρίτων·
5. διαθέσιμο σε πολλαπλές πλήρεις υλοποιήσεις από ανταγωνιστικούς προμηθευτές ή ως πλήρης υλοποίηση ισότιμα διαθέσιμο σε όλους τους ενδιαφερόμενους.

Η βασική πρόκληση

Ποια θα λέγατε ότι είναι η βασική πρόκληση για την εφαρμογή κάποιου ανοιχτού προτύπου;

Η βασική πρόκληση

Αφορά στον φορέα που δημοσιεύει τα ανοιχτά δεδομένα. Συγκεκριμένα είναι η απάντηση στην ερώτηση: “ποια πρότυπα να χρησιμοποιήσω;”

- (α) πού να ψάξω;
- (β) ποια από τα διαθέσιμα πρότυπα είναι κατάλληλα για το φορέα μου;

(α) Πηγές προτύπων

Μέσω εξειδικευμένων ιστοσελίδων, όπως:

- OpenReferral,
- Open Contracting Data Standard,
- Humanitarian Exchange Standard

(β) Προφίλ των προτύπων ανοιχτών δεδομένων

Πως επιλέγουμε πρότυπα κατάλληλα για το φορέα μας; Παράγοντες που επηρεάζουν αυτή την απόφαση:





- εφαρμόζεται το πρότυπο στην περίπτωση μας;
- ποιος άλλος το χρησιμοποιεί;
- σε τι επίπεδο λεπτομέρειας φτάνει;
- συντηρείται το πρότυπο τακτικά;

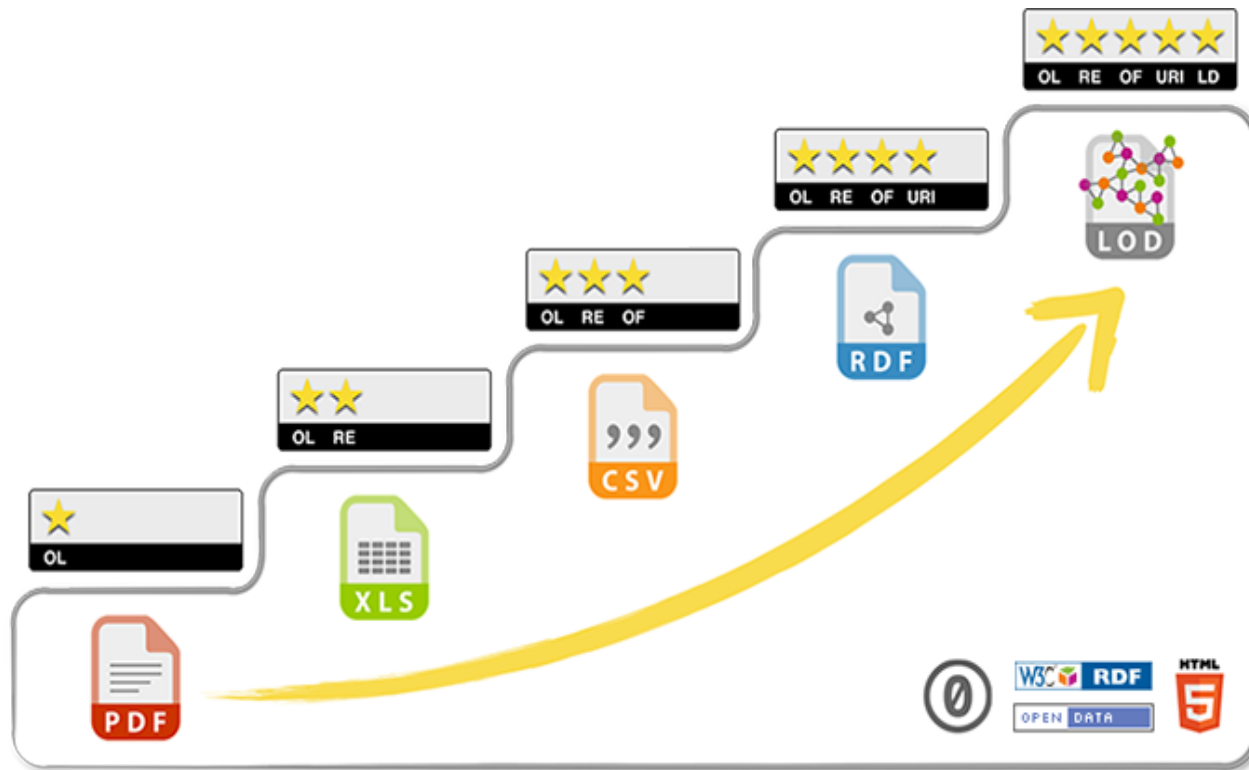
Αυτές οι πληροφορίες θα πρέπει να είναι διαθέσιμες μέσω κατάλληλων μετα-δεδομένων (πρότυπο μετα-δεδομένων για την περιγραφή των προτύπων!)

Πρωτοβουλίες για τα ανοιχτά πρότυπα

- **Geothink**, μια Καναδική πρωτοβουλία έρευνας σε ανοιχτά γεωχωρικά δεδομένα
- **The International Open Data Charter Technical Working Group**, δουλεύει πάνω σε μέθοδο για την ανακάλυψη διαθέσιμων προτύπων
- **The U.S. Data Federation**, έχει δημιουργήσει ένα κατάλογο πρωτοβουλιών που δουλεύουν πάνω σε θέματα προτυποποίησης
- **Joined-up Data Standards Thesaurus**, αποθετήριο για πρότυπα
- **Data.gov.uk**, μητρώο για μητρώα ανοιχτών προτύπων
- **FSFE.org**, εργασία για τον ορισμό των ανοιχτών προτύπων

Δεδομένα και μηχανές (6/6)

-  Raw
-  Pilot
-  Standard
-  Expert



- ★ τα δεδομένα διαθέσιμα προς χρήση στο Διαδίκτυο (σε οποιοδήποτε format)

- ★★ τα δεδομένα διαθέσιμα σαν δομημένα δεδομένα (π.χ. έγγραφο excel αντί για έναν σκαναρισμένο πίνακα)

- ★★★ τα δεδομένα διαθέσιμα σε μη κλειστό πρότυπο (π.χ. CSV αντί για έγγραφο excel)

- ★★★★ χρήση συνδέσμων (URLs) για να αναγνωρίζουμε πράγματα έτσι ώστε άλλοι χρήστες να μπορούν να αναφέρονται σε αυτά

- ★★★★★ σύνδεση των δεδομένων με τα δεδομένα άλλων χρηστών (συμφραζόμενα - context)

(πηγή: <https://inkdroid.org/2010/06/04/the-5-stars-of-open-linked-data/>)

Υπέρ και κατά του ★ επιπέδου

Ως “καταναλωτές” δεδομένων ...

- Μπορούμε να τα δούμε.
- Μπορούμε να τα εκτυπώσουμε.
- Μπορούμε να τα αποθηκεύσουμε τοπικά (σε σκληρό δίσκο ή USB stick).
- Μπορούμε να τα μεταφέρουμε σε ένα άλλο σύστημα.
- Μπορούμε να τα διαμοιράσουμε.

Σαν εκδότης (δημιουργός) των δεδομένων ...

- Είναι εύκολα ως προς τη δημοσίευση.
- Δεν χρειάζεται να εξηγούμε συνέχεια στους άλλους ότι μπορούν να χρησιμοποιήσουν τα δεδομένα μας.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★ επιπέδου

Ως “καταναλωτές” δεδομένων μπορούμε να κάνουμε όσα και στο ★ επίπεδο και επιπλέον:

- Μπορούμε να τα επεξεργαστούμε / αλλάξουμε, να τα ενοποιήσουμε με άλλα, να κάνουμε υπολογισμούς, οπτικοποίηση, κλπ. (χρησιμοποιώντας συνήθως κάποια εμπορική μορφή - MS Windows -).
- Μπορούμε να τα εξάγουμε σε κάποιο άλλο δομημένο format.

Σαν εκδότης (δημιουργός) των δεδομένων ...

- Είναι εύκολα ως προς τη δημοσίευση.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★☆☆ επιπέδου

Ως “καταναλωτές” μπορούμε να κάνουμε όσα και στο ★★ επίπεδο και επιπλέον:

- Μπορούμε να τα επεξεργαστούμε / αλλάξουμε, να τα ενοποιήσουμε με άλλα, να κάνουμε υπολογισμούς, οπτικοποίηση, κλπ. (χωρίς απαραίτητα την χρήση εμπορικού λογισμικού).

Σαν εκδότης (δημιουργός) των δεδομένων ...

- Μπορεί να χρειαστούν διαδικασίες μετατροπής δεδομένων για την εξαγωγή των δεδομένων από το τρέχων (εμπορικό) format.
- Είναι εύκολα ως προς τη δημοσίευση.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★★★★ επιπέδου

Ως “καταναλωτές” δεδομένων μπορούμε να κάνουμε όσα και στο ★★★ επίπεδο και επιπλέον:

- Μπορούμε να συνδεθούμε στα δεδομένα από οπουδήποτε θέλουμε.
- Μπορούμε να δημιουργήσουμε έναν σελιδοδείκτη στα δεδομένα.
- Μπορούμε να επαναχρησιμοποιήσουμε τμήματα των δεδομένων.
- Μπορούμε να επαναχρησιμοποιήσουμε εργαλεία και βιβλιοθήκες, ακόμη και αν αντιλαμβάνονται μόνον ένα τμήμα από το πρότυπο των δημοσιευμένων δεδομένων.
- Το να κατανοήσουμε τη δομή ενός γράφου RDF απαιτεί περισσότερη προσπάθεια από ότι δεδομένα σε πινακική (Excel/CSV) ή δενδρική (XML/JSON) δομή.
- Μπορούμε να συνδυάσουμε τα δεδομένα εύκολα με άλλα δεδομένα. Τα URI είναι ένα διεθνές σχήμα απεικόνισης.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★★★★ επιπέδου

Σαν εκδότης (δημιουργός) των δεδομένων ...

- Έχουμε τον έλεγχο των δεδομένων μας και μπορούμε να βελτιστοποιήσουμε την πρόσβαση σε αυτά.
- Άλλοι εκδότες δεδομένων μπορούν να συνδέονται με συνδέσμους στα δεδομένα μας.
- You typically invest some time slicing and dicing your data.
- Χρειάζεται να δημιουργείτε URI στα δεδομένα και να σκεφτούμε τον καλύτερο τρόπο να αναπαραστήσουμε τα δεδομένα μας.
- Χρειάζεται να ανακαλύπτουμε υπάρχοντα πρότυπα για να αναπαραριστούμε τα δεδομένα μας ή να ανακαλύπτουμε καινούρια.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★★★★ επιπέδου

Ως “καταναλωτές” δεδομένων μπορούμε να κάνουμε όσα και στο ★★★★★ επίπεδο και επιπλέον:

- Μπορούμε να ανακαλύπτουμε και άλλα συναφή δεδομένα καθώς καταναλώνουμε τα δεδομένα που μας ενδιαφέρουν.
- Μπορούμε απευθείας να ανακαλύπτουμε το σχήμα των δεδομένων.
- Θα πρέπει να έχουμε στο νου μας να διαχειριζόμαστε links που δε λειτουργούν, σφάλματα τύπου 404, κλπ.
- Το να χρησιμοποιούμε δεδομένα από τυχαία links και να τα παρουσιάζουμε σαν γεγονότα εμπεριέχει ρίσκο όπως το να χρησιμοποιούμε πληροφορίες που βρίσκουμε τυχαία σε sites στο Διαδίκτυο.

(πηγή: <http://5stardata.info>)

Υπέρ και κατά του ★★★★★ επιπέδου

Σαν εκδότης (δημιουργός) των δεδομένων ...

- Τα δεδομένα μας γίνονται ανιχνεύσιμα.
- Αυξάνεται η αξία των δεδομένων μας.
- Ο οργανισμός που εκπροσωπείτε λαμβάνει τα ίδια οφέλη όπως και οι καταναλωτές των δεδομένων.
- Θα πρέπει να επενδύσετε σε υποδομή για να μπορείτε να συνδέεται τα δεδομένα σας με άλλα σημεία του Διαδικτύου.
- Μπορεί να χρειαστεί να επιδιορθώσετε εσφαλμένα links.

(πηγή: <http://5stardata.info>)

Άσκηση

Θέλετε να δημοσιεύσετε κάποια δεδομένα του τμήματός σας ως ανοιχτά δεδομένα. Ποιο επίπεδο δημοσίευσης θα διαλέγατε; Ποια είναι τα υπέρ και ποια τα κατά της επιλογής σας;

Εργαστείτε σε ζεύγη.

Πηγές για περαιτέρω πληροφόρηση

1. <https://www.w3.org/TR/dwbp/#ReuseVocabularies>
2. <http://fsfe.org/activities/os/def.el.html>
3. <http://5stardata.info>
4. <https://inkdroid.org/2010/06/04/the-5-stars-of-open-linked-data/>
5. <http://opendefinition.org/>
6. https://en.wikipedia.org/wiki/Creative_Commons_license

Αναφορά σε πηγές

Παρακάτω αναφερόμαστε σε υλικό της παρουσίασης που προήλθε από συναδέλφους ή άλλες πηγές με ανοικτή αδειοδότηση για το οποίο τους ευχαριστούμε θερμά:

1. *Open Data Institute: 18-22, 35-43*
2. *Κατερίνα Λενάκη: 23-26*

(σημείωση: εξαιρούνται διαφάνειες που η αναφορά στον δημιουργό γίνεται επί τόπου)

ΕΥΧΑΡΙΣΤΩ ΓΙΑ ΤΗΝ ΠΡΟΣΟΧΗ ΣΑΣ

ερωτήσεις

Απόστολος Κρητικός
Open Data Certified Trainer, ODI

Βασίλης Χρυσός
Open Data Certified Trainer, ODI