**ΑΝΟΙΧΤΑ ΔΕΔΟΜΕΝΑ**

**ΑΡΧΕΣ, ΚΑΛΕΣ ΠΡΑΚΤΙΚΕΣ ΚΑΙ ΠΡΟΤΑΣΕΙΣ**



Περιεχόμενα

[1. Εισαγωγή 4](#_Toc500149917)

[2. Tο υφιστάμενο θεσμικό πλαίσιο 5](#_Toc500149918)

[3. Οι Βασικές Αρχές 7](#_Toc500149919)

[4. Μείωση της υστέρησης των ανοιχτών δεδομένων 7](#_Toc500149920)

[4.1. Η σημασία του ζητήματος 7](#_Toc500149921)

[4.2. Προτάσεις 7](#_Toc500149922)

[5. Μηχανισμός και ταχύτητα ανάρτησης 9](#_Toc500149923)

[5.1. Κύκλος Ζωής Ανοιχτών Δεδομένων: Συνιστώμενες πρακτικές 9](#_Toc500149924)

[6. Η Κρισιμότητα της Έγκαιρης Ανάρτησης 12](#_Toc500149925)

[7. Επικαιροποίηση & Διαλειτουργικότητα 13](#_Toc500149926)

[8. Σημασιολογία και Διαλειτουργικότητα 15](#_Toc500149927)

[8.1. Η κρισιμότητα του ζητήματος 15](#_Toc500149928)

[8.2. Μέθοδοι & Τεχνικές 15](#_Toc500149929)

[9. Οικονομική αποτίμηση των Δεδομένων (Data Monetization) 15](#_Toc500149930)

[9.1. Η σημασία του ζητήματος 15](#_Toc500149931)

[9.2. Προτάσεις και Παραδείγματα 16](#_Toc500149932)

[10. Η χρησιμότητα των συμβάσεων SLA 17](#_Toc500149933)

[11. Εκπαίδευση 18](#_Toc500149934)

[11.1. Άξονες Εκπαίδευσης 18](#_Toc500149935)

[11.2. Προτάσεις 19](#_Toc500149936)

[ΤΕΧΝΙΚΟ ΠΑΡΑΡΤΗΜΑ 21](#_Toc500149937)

[Αρχές έκδοσης (versionong) δεδομένων 21](#_Toc500149938)

# Εισαγωγή

O κόσμος παρακολουθεί έναν μετασχηματισμό στην παγκόσμια οικονομία και κοινωνία που δεν έχει προηγούμενο. Η αλλαγή αυτή οφείλεται στην φρενήρη ρυθμό εξέλιξης της τεχνολογίας. Αυτός ο μετασχηματισμός μπορεί εν δυνάμει να κάνεις τις κυβερνήσεις πιο αποδοτικές, πιο αποτελεσματικές και να αντιδρούν με μεγαλύτερη παχύτητα στα προβλήματα που αντιμετωπίζουν. Μπορεί επίσης να βελτιώσει την λογοδοσία στην οποία είναι υποχρεωμένες οι κυβερνήσεις να παρέχουν στους πολίτες.

Όμως ο μετασχηματισμός αυτός δεν επηρεάζει μόνον τις κυβερνήσεις. Αντίστοιχα επηρεάζονται οι επιχειρήσεις, η κοινωνία των πολιτών και οι μεμονωμένοι πολίτες.

Στην καρδιά αυτού του μετασχηματισμού είναι τα ανοιχτά δεδομένα. Τα ανοιχτά δεδομένα επιτρέπουν στις κυβερνήσεις, στις επιχειρήσεις, την κοινωνία των πολιών και στους πολίτες να χρησιμοποιούν την διαθέσιμη πληροφόρηση για να λαμβάνουν καλύτερες αποφάσεις. Η αποτελεσματική και έγκαιρη χρήση στα δεδομένα επιτρέπει την δημιουργία καλύτερης αντίληψης για τα πράγματα και την αποτελεσματικότερη λήψη αποφάσεων. Αυτό μπορεί να δημιουργήσει κοινωνικά και οικονομικά οφέλη σε κάθε χώρα που χρησιμοποιεί αποτελεσματικά τα ανοιχτά της δεδομένα.

Επιπλέον τα ανοιχτά δεδομένα μπορούν να προωθήσουν την οικονομική ανάπτυξη της χώρας ενισχύοντας τις υπάρχουσες αγορές και δημιουργώντας νέες. Είναι εμφανές ότι όσο καλύτερη είναι η ποιότητα των ανοιχτών δεδομένων και όσο πιο έγκαιρη η ανάρτησή τους, τόσο μεγαλύτερες είναι οι εν δυνάμει επιχειρηματικές ευκαιρίες. Όπως δείχνει και η διεθνής εμπειρία, το άνοιγμα και η διάθεση των δεδομένων του δημοσίου τομέα ελεύθερα και με οργανωμένο τρόπο θα συμβάλει ισχυρά στην καινοτόμο επιχειρηματικότητα, δημιουργώντας νέες υπηρεσίες προς τους πολίτες και τις επιχειρήσεις

Τα παραπάνω πλεονεκτήματα των ανοιχτών δεδομένων μπορούν να γίνουν πραγματικότητα αν τηρηθούν κάποιες αρχές και ακολουθηθούν ορισμένες καλές πρακτικές. Στο παρόν κείμενο αναλύονται αρχές και πρακτικές που πρέπει να διέπουν την δημιουργία, την χρήση και την συντήρηση των ανοιχτών δημόσιων δεδομένων, προκειμένου να μεγιστοποιηθούν τα οφέλη τα οποία μπορούν να προκύψουν.

# Tο υφιστάμενο θεσμικό πλαίσιο

Η Ευρωπαϊκή Επιτροπή έχει θεσπίσει ένα νομικό πλαίσιο που επιτρέπει την επαναχρησιμοποίηση πληροφοριών του δημόσιου τομέα μέσω της οδηγίας [Directive 2003/98/EC](http://eur-lex.europa.eu/legal-content/en/ALL/?uri=CELEX:32003L0098). Η παρούσα οδηγία βασίζεται σε δύο πυλώνες της εσωτερικής αγοράς: διαφάνεια και θεμιτό ανταγωνισμό. Επικεντρώνεται στις οικονομικές πτυχές της επαναχρησιμοποίησης των πληροφοριών. Η παραπάνω Οδηγία αναθεωρήθηκε με την οδηγία [Directive 2013/37/EU](http://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX:32013L0037), η οποία περιλαμβάνει το πεδίο εφαρμογής των μουσείων, των βιβλιοθηκών και των αρχείων επιπροσθέτως.

Το data.gov.gr αποτελεί το εργαλείο υλοποίησης της πολιτικής των ανοικτών δεδομένων κατ' εφαρμογή της σχετικής νομοθεσίας, μετά και την ενσωμάτωση της Οδηγίας [Directive 2013/37/EU](http://eur-lex.europa.eu/legal-content/FR/ALL/?uri=CELEX:32013L0037) (<https://ec.europa.eu/digital-single-market/en/open-data>).

Το θεσμικό πλαίσιο που διέπει τη συγκεκριμένη δράση είναι:

1. *Ν.* [*4305/2014*](http://www.et.gr/idocs-nph/search/pdfViewerForm.html?args=5C7QrtC22wEc63YDhn5AeXdtvSoClrL8oeKAuTKOuiV5MXD0LzQTLWPU9yLzB8V68knBzLCmTXKaO6fpVZ6Lx3UnKl3nP8NxdnJ5r9cmWyJWelDvWS_18kAEhATUkJb0x1LIdQ163nV9K--td6SIubMfH2r_a2DXjO6MJnF-5f9_LW7pRMszX0fGIINVmMIh) *(ΦΕΚ 237/Α΄) «Ανοικτή διάθεση και περαιτέρω χρήση εγγράφων, πληροφοριών και δεδομένων του δημόσιου τομέα, τροποποίηση του ν. 3448/2006 (Α΄ 57), προσαρμογή της εθνικής νομοθεσίας στις διατάξεις της Οδηγίας 2013/37/ΕΕ του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου, περαιτέρω ενίσχυση της διαφάνειας, ρυθμίσεις θεμάτων Εισαγωγικού Διαγωνισμού Ε.Σ.Δ.Δ.Α. και άλλες διατάξεις».*
2. *Ν.* [*3448/2006*](http://www.et.gr/idocs-nph/search/pdfViewerForm.html?args=5C7QrtC22wFGQ40gSLPFOXdtvSoClrL8KoqaERu8wDDtIl9LGdkF53UIxsx942CdyqxSQYNuqAGCF0IfB9HI6qSYtMQEkEHLwnFqmgJSA5WIsluV-nRwO1oKqSe4BlOTSpEWYhszF8P8UqWb_zFijM7pZ0uAVJnEIRDT8gfgjDrktnVRzo5d-fi1RvYPzP-4) *(ΦΕΚ 57/Α΄) «Για την περαιτέρω χρήση πληροφοριών του δημόσιου τομέα και τη ρύθμιση θεμάτων αρμοδιότητας Υπουργείου Εσωτερικών, Δημόσιας Διοίκησης και Αποκέντρωσης».*
3. [*Αιτιολογική έκθεση*](http://www.hellenicparliament.gr/UserFiles/2f026f42-950c-4efc-b950-340c4fb76a24/a-diafeg-eis-sunolo.pdf) *Ν. 4305/2014.*
4. *Αριθ. ΔΗΔ/Φ.40/407/8.1.2015 εγκύκλιος με θέμα «Εφαρμογή των διατάξεων του Κεφαλαίου Α’ του ν. 4305/2014 (ΦΕΚ 237/Α΄ ) σχετικά με την «ανοικτή διάθεση και περαιτέρω χρήση εγγράφων, πληροφοριών και δεδομένων του δημόσιου τομέα, την τροποποίηση των διατάξεων του πρώτου κεφαλαίου του ν. 3448/2006, προσαρμογή της εθνικής νομοθεσίας στις διατάξεις της οδηγίας 2013/37 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου και την περαιτέρω ενίσχυση της διαφάνειας στο δημόσιο τομέα». (ΑΔΑ:* [*ΩΩΡΜΧ-ΜΒΛ*](https://diavgeia.gov.gr/decision/view/%CE%A9%CE%A9%CE%A1%CE%9C%CE%A7-%CE%9C%CE%92%CE%9B)*).*
5. *Αριθ. ΔΗΔ. Φ./ 19710/16.6.2015 εγκύκλιο (ΑΔΑ:*[*7ΧΩΨ465ΦΘΕ-Β2Ι*](https://diavgeia.gov.gr/decision/view/7%CE%A7%CE%A9%CE%A8465%CE%A6%CE%98%CE%95-%CE%922%CE%99)*) με θέμα «Ανοικτή διάθεση και περαιτέρω χρήση εγγράφων, πληροφοριών και δεδομένων του δημόσιου τομέα σύμφωνα με το κεφ. Α΄ του Ν. 4305/2014».*
6. *Αριθμ. Πρωτ.: ΔΗΔ/Φ.40/2369/24.1.2017 εγκύκλιο (ΑΔΑ:* [*6Γ07465ΧΘΨ-473*](https://diavgeia.gov.gr/decision/view/6%CE%9307465%CE%A7%CE%98%CE%A8-473)*) με θέμα «Επίσπευση ενεργειών από τους υπόχρεους φορείς για τη διάθεση και περαιτέρω χρήση των ανοικτών δεδομένων σε εφαρμογή του ν. 4305/2014 (ΦΕΚ 237 Α΄)».*

# Οι Βασικές Αρχές

Ο φορέας Open Government Partnership δημοσίευσε, τον Οκτώβριο του 2015 έναν Καταστατικό Χάρτη[[1]](#footnote-1) με 6 θεμελιώδεις αρχές για τα ανοιχτά δεδομένα:

1. Η διάθεση των δεδομένων είναι εκ προοιμίου ανοιχτή (open by default)
2. Τα δεδομένα διατίθενται έγκαιρα και περιεκτικά
3. Τα δεδομένα είναι προσβάσιμα και χρήσιμα (accessible and usable)
4. Τα δεδομένα δίνουν την δυνατότητα διενέργειας συγκρίσεων και διαλειτουργικότητας (comparable and interoperable)
5. Τα δεδομένα δίνουν την δυνατότητα βελτίωσης της διακυβέρνησης και εμπλοκής των πολιτών (For Improved Governance and Citizen Engagement),
6. Τα δεδομένα επιτρέπουν την συμμετοχική ανάπτυξη και καινοτομία (For Inclusive Development and Innovation).

Οι παραπάνω αρχές διέπουν τις καλές πρακτικές και τις προτάσεις που παρουσιάζονται σε αυτό το κείμενο.

# Μείωση της υστέρησης των ανοιχτών δεδομένων

# Η σημασία του ζητήματος

Σε εγκύκλιο του Υπουργείου Διοικητικής Ανασυγκρότησης με ημερομηνία 24 Ιανουαρίου2017 και θέμα «Επίσπευση ενεργειών από τους υπόχρεους φορείς για τη διάθεση και περαιτέρω χρήση των ανοικτών δεδομένων σε εφαρμογή του ν. 4305/2014» η οποία απευθύνεται προς όλους τους φορείς του δημοσίου που εμπίπτουν στο πεδίο εφαρμογής του ν. 4305/2014 αναφέρεται «Παρά τη σημαντική πρόοδο η οποία έχει σημειωθεί στο τελευταίο χρονικό διάστημα, από τις πιο σημαντικές αδυναμίες παραμένει η ανταπόκριση κρίσιμων φορέων της δημόσιας διοίκησης, η οποία θα μπορούσε να χαρακτηριστεί ως εξαιρετικά περιορισμένη ενώ, όπως διαπιστώνεται στο πλαίσιο συγκριτικών μελετών σε διεθνές επίπεδο οι επιδράσεις της υιοθέτησης της πολιτικής των ανοικτών δεδομένων στις επιχειρήσεις και στην Κοινωνία των Πολιτών (impact) παραμένουν περιορισμένες.»

# Προτάσεις

Θεωρούμε ότι για να αντιμετωπιστεί η διαπιστούμενη υστέρηση στη ποσότητα αλλά και στη ποιότητα των ανοιχτών δεδομένων θα πρέπει να ενδυναμωθεί η υιοθετηθούν δράσεις όπως:

* Να ορισθεί, ανά φορέα του δημοσίου που εμπίπτει στο πεδίο εφαρμογής του ν. 4305/2014, υπεύθυνος για θέματα ανοικτών δεδομένων και να δημοσιοποιηθεί σχετική ηλεκτρονική διεύθυνση αλληλογραφίας (e-mail) ώστε να μπορούν να επικοινωνούν στοχευμένα ενδιαφερόμενοι φορείς τόσο αναφορικά με αιτήματα διάθεσης δεδομένων όσο και για παρατηρήσεις αναφορικά με τη ποιότητα των σχετικών δεδομένων.
* Να δημιουργηθεί στο Υπουργείο Διοικητικής Ανασυγκρότησης δομή “διαμεσολάβησης” αναφορικά με τα ανοικτά δεδομένα η οποία θα υποστηρίζει ενδιαφερόμενους φορείς τόσο αναφορικά με αιτήματα διάθεσης δεδομένων όσο και για παρατηρήσεις αναφορικά με τη ποιότητα των σχετικών δεδομένων οι οποίοι δεν έχουν βρει ανταπόκριση από τους παραγωγούς – διαχειριστές των σχετικών δεδομένων.
* Να οριστεί ότι κάθε φορέας που δημοσιεύει δεδομένα πρέπει να τα ταυτοποιεί με χρήση ημερομηνίας δημοσίευσης, έκδοσης (version number) και να συμπεριλαμβάνει κατάλογο αλλαγών από την προηγούμενη έκδοση (changelog). Φυσικά αυτό δεν αφορά περιπτώσεις on-line δεδομένων. Επίσης σε περίπτωση που κάποιοι φορείς παρέχουν δεδομένα που περιλαμβάνουν ή συσχετίζουν και άλλα σύνολα δεδομένων (για παράδειγμα αντιστοίχιση ΚΕΝ με ICD-10) θα πρέπει να περιγράφουν την ημερομηνία δημοσίευσης και τον αριθμό έκδοσης (version) αυτών των δεδομένων. Έτσι θα διασφαλιστεί η ιχνηλασιμότητα και θα μειώνονται τα λάθη.
* Να μεγιστοποιηθεί η ένταξη του data.gov.gr στη παραγωγική διαδικασία της διάθεσης των δεδομένων από φορείς, ορίζοντας ότι η σχετική διάθεση θα γίνεται είτε αποκλειστικά μέσω του data.gov.gr είτε από διαδικτυακές διευθύνσεις των φορέων οι οποίες είναι υποχρεωτικά ορισμένες στο data.gov.gr. Σήμερα παρατηρείται το φαινόμενο κάποιοι φορείς να δημοσιεύουν δεδομένα τα οποία αργότερα αναρτούν και στο data.gov.gr. Έτσι δημιουργούνται αντίγραφα με προφανή προβλήματα συγχρονισμού, αναντιστοιχίας και διαχείρισης.
* Να οριστεί στο data.gov.gr πίνακας με σύνολο δεδομένων στο οποίο να περιγράφεται και ο “ιδιοκτήτης – διαχειριστής” του κάθε συνόλου. Φορείς που στα πλαίσια των δραστηριοτήτων τους αναδημοσιεύουν σύνολα δεδδομένων που διαχειρίζονται άλλοι φορείς θα πρέπει να παρέχουν Link προς το κατάλληλο σημείο του data.gov.gr και να μην δημιουργούν αντίγραφα γιατί μια τροποποίηση του αρχικού συνόλου δεδομένων δεν διαχέεται εύκολα και γρήγορα και έτσι δημιουργούνται διαφορετικές εκδόσεις του ίδιου συνόλου δεδομένων με διαφοροποιήσεις στο περιεχόμενο. Ενδεικτικά στα πλαίσια των δραστηριοτήτων τους φορείς όπως Υπουργείο Υγείας, ΕΟΠΥΥ, ΗΔΙΚΑ κλπ χρησιμοποιούν κωδικοποιήσεις όπως ICD-10, ΚΕΝ κλπ. Θα πρέπει να είναι ορισμένο ποιος φορέας είναι υπεύθυνος για τη διαχείριση του κάθε συνόλου δεδομένων και να παρέχεται link στην ίδια διεύθυνση του data.gov.gr

# Μηχανισμός και ταχύτητα ανάρτησης

# Κύκλος Ζωής Ανοιχτών Δεδομένων: Συνιστώμενες πρακτικές

Ο κύκλος ζωής των ανοιχτών δεδομένων[[2]](#footnote-2) περιλαμβάνει τα βήματα της συλλογής, της προετοιμασίας, της δημοσίευσης και της συντήρησης των δεδομένων. Με τη σειρά του κάθε ένα από αυτά τα βήματα αναλύεται σε επιμέρους διαδικασίες ως ακολούθως:

1. *Συλλογή*. Αφορά την αναγνώριση και συλλογή των δεδομένων και μπορεί να προσεγγιστεί τόσο από τη σκοπιά της εξασφάλισης γρήγορων λύσεων όσο και εκείνη της υιοθέτησης μιας ενδελεχούς διαδικασίας για τη διαχείριση των δεδομένων, ενώ εξαρτάται από τις επιλογές του οργανισμού σε επίπεδο υποδομών (ενδεικτικά κεντρική διαχείριση ή επεξεργασία των δεδομένων σε διάφορες μονάδες, κλπ.). Περιλαμβάνει τα κάτωθι βήματα:

* *Χαρτογράφηση/καταγραφή των ήδη διαθέσιμων δεδομένων*. H εν λόγω διαδικασία παρέχει μια εποπτική εικόνα για τα δεδομένα που είναι ήδη διαθέσιμα σε έναν οργανισμό.
* *Προτεραιοποίηση των διαθέσιμων δεδομένων*. Όλα τα δεδομένα δεν είναι κατάλληλα προς δημοσιοποίηση. Προκειμένου να γίνει προτεραιοποίηση αυτών πρέπει να εξεταστούν μια σειρά από κριτήρια. Ενδεικτικά:
  + Μπορούν τα δεδομένα να δημοσιευθούν (από νομική, πολιτική σκοπιά ή τη σκοπιά του οργανισμού);
  + Έχουν ήδη την απαιτούμενη ποιότητα ή χρειάζονται περαιτέρω επεξεργασία πριν τη δημοσίευσή τους;
  + Έχει ληφθεί μέριμνα για τον καθαρισμό, την ανωνυμία και τη μορφοποίησή τους;
  + Εντάσσονται σε θεματική υψηλού ενδιαφέροντος;

Με τη διαδικασία αυτή είναι δυνατό να καταστεί σαφές τι μπορεί να δημοσιευθεί, τι όχι και που πρέπει να δοθεί προτεραιότητα. Σε δεύτερο χρόνο επιπλέον κριτήρια μπορούν να χρησιμοποιηθούν για την προτεραιοποίηση των δεδομένων ανάλογα με τις ανάγκες του οργανισμού.

* *Εφαρμογή*. Το βήμα αυτό συνίσταται στην εφαρμογή και αναθεώρηση της διαδικασίας συλλογής δεδομένων με στόχο την προτυποποίηση και τεκμηρίωσή της και την αναγνώριση και υιοθέτηση βέλτιστων πρακτικών.
* *Επιλογή των δεδομένων προς δημοσίευση*. Τα δεδομένα δύνανται να καλύπτουν διάφορα θέματα και κατηγορίες μεγαλύτερου ή μικρότερου ενδιαφέροντος. Για την επιλογή των συνόλων δεδομένων που έχει μεγαλύτερη αξία να δημοσιευθούν ενθαρρύνεται η αναφορά στον **Καταστατικό Χάρτη των G8 για τα Ανοιχτά Δεδομένα** και συγκεκριμένα τις κατηγορίες και τα επιμέρους σύνολα δεδομένων που ο τελευταίος αναγνωρίζει ως μείζονος ενδιαφέροντος.
* *Θεματική κατηγοριοποίηση των δεδομένων*. Η δημοσίευση των δεδομένων μπορεί εν τέλει να λάβει χώρα είτε με οδηγό τις προτεινόμενες κατηγορίες του προαναφερθέντος καταστατικού χάρτη είτε με την προσθήκη επιπλέον κατηγοριών ανάλογα με τις ανάγκες του οργανισμού.

1. *Προετοιμασία*. Αφορά στην προετοιμασία των συγκεντρωθέντων δεδομένων και κατ’ επέκταση το μετασχηματισμό τους από ακατέργαστα δεδομένα, σε σύνολα δεδομένων προς δημοσίευση. Περιλαμβάνει με τη σειρά της τα ακόλουθα βήματα:

* *Διασφάλιση της ποιότητας των δεδομένων*. Η ποιότητα των δεδομένων αφορά στις όψεις της ποιότητας του περιεχομένου τους, της επικαιρότητας και της συνοχής τους. Η χρησιμότητα των δεδομένων πράγματι καθορίζεται από την ποιότητά τους, η οποία με τη σειρά της συναρτάται από παράγοντες όπως η πληρότητα, η καθαρότητα και η ακρίβειά τους. Η όψη της επικαιρότητας απαιτεί το σχετικό έλεγχο των δεδομένων και την εφαρμογή κατάλληλης διαδικασίας για την ενημέρωσή τους, ενώ συνδέεται στενά και με το βήμα της συντήρησής τους. Η όψη της συνοχής τέλος απαιτεί τη χρήση προτύπων και είναι μείζονος σημασίας για την εξασφάλιση ης δυνατότητας επαναχρησιμοποίησης των δεδομένων.
* *Προετοιμασία των δεδομένων από τεχνική σκοπιά*. Το βήμα αυτό έγκειται στον εμπλουτισμό των δεδομένων με τη βοήθεια εργαλείων, όπως τα «συνδεδεμένα δεδομένα» (Linked Data), οι οντολογίες, τα συναφή πρότυπα (π.χ. RDF), κλπ.
* *Προσθήκη μετα-δεδομένων*. Η προσθήκη μεταδεδομένων είναι ουσιώδης για την εξασφάλιση της δυνατότητας ανίχνευσης (discoverability) των δεδομένων, καθώς επίσης και για την κατανόηση της φύσης και δομής τους αλλά και την αναγνώριση της προέλευσης και των όρων χρήσης τους. Τα μεταδεδομένα περιγράφουν το ίδιο το σύνολο των δεδομένων (π.χ. ημερομηνία, τίτλος, περιεχόμενο, συγγραφέας, τύπος, μέγεθος, κλπ.) και μπορούν να προστεθούν σε καταλόγους για να διευκολύνουν την ανακάλυψή τους. Ο εμπλουτισμός των μεταδεδομένων ενός συνόλου δεδομένων με URIs επιτρέπει τη δημοσίευση των δεδομένων ως linked data και αυξάνει σημαντικά την ανιχνευσιμότητα αλλά και τη διαλειτουργικότητά τους. Αποτελεί ως εκ τούτου **βέλτιστη πρακτική**. **Για έναν πληρέστερο κατάλογο βέλτιστων πρακτικών και κατευθυντήρων γραμμών ως προς τη χρήση μεταδεδομένων (δημοσίευση των μεταδεδομένων μαζί με τα ίδια τα σύνολα δεδομένων, χρήση machine-readable μορφών, χρήση πρότυπων όρων, περιγραφή του συνόλου των χαρακτηριστικών ενός συνόλου δεδομένων, παροχή πληροφοριών για τοπικές παραμέτρους, π.χ. ημερομηνία, γλώσσα, την άδεια και τους όρους χρήσης, την ποιότητα των δεδομένων κλπ.) προτείνεται η αναφορά στο διαδικτυακό τόπου του φορέα W3C**. Η δημοσίευση των μεταδεδομένων περιλαμβάνει με τη σειρά της τα βήματα της συλλογής και καταγραφής των μεταδεδομένων.
* *Προετοιμασία των δεδομένων από νομική σκοπιά*. Η ελεύθερη διάθεση των δεδομένων από νομική άποψη αποτελεί βασική αρχή των ανοιχτών δεδομένων, καθώς διαφορετικά τα τελευταία είναι αδύνατον να επαναχρησιμοποιηθούν. Η διάθεση των ανοιχτών δεδομένων πρέπει να είναι εναρμονισμένη με την ισχύουσα νομοθεσία τόσο σε εθνικό όσο και υπερεθνικό επίπεδο, ενώ τα δεδομένα προς δημοσίευση πρέπει να συνοδεύονται από αντίστοιχη άδεια (open license), αποκτηθείσα μέσω της εγκαθίδρυσης κατάλληλης διαδικασίας αδειοδότησης. Σημειώνεται ότι αποτελεί ευθύνη του εκδίδοντος οργανισμού να είναι ενημερωμένος σε σχέση με τους ισχύοντες κανόνες και νομικούς περιορισμούς.

1. *Δημοσίευση*. Η δημοσίευση των δεδομένων στο διαδίκτυο τα καθιστά διαθέσιμα προς επαναχρησιμοποίηση από τον καθένα και για κάθε χρήση. Οι πρακτικές δημοσίευσης περιλαμβάνουν την ανάρτηση αρχείων σε μια ιστοσελίδα, την ανάρτηση απευθείας σε κάποιο διαδικτυακό τόπο ή τη δημοσίευση μέσω ενός API. Η ανάρτηση αρχείων αφορά συνήθως μικρό όγκο δεδομένων, η ανάρτηση σε μια διαδικτυακή πύλη αποτελεί την πλέον χρησιμοποιούμενη επιλογή, ενώ η τρίτη πρακτική εξαρτάται σε μεγάλο βαθμό από το χρησιμοποιούμενο λογισμικό και απαιτεί επιπλέον τη δημοσίευση των προδιαγραφών του API στην ιστοσελίδα του οργανισμού.
2. *Συντήρηση*. Αφορά στην εγκαθίδρυση κατάλληλης διαδικασίας συντήρησης, η οποία είναι απαραίτητη για τη διαχείριση τυχόν μεταβολών στα δεδομένα και την αποφυγή του να καταστούν αυτά παρωχημένα.

* *Τακτική συντήρηση των δεδομένων και των μεταδεδομένων αυτών*. Ενημερώσεις, μεταβολές και άλλες συνθήκες μπορεί να καταστήσουν τόσο τα δεδομένα όσο και τα μεταδεδομένα τους παρωχημένα, γεγονός που με τη σειρά του θα οδηγήσει στον περιορισμό της ανιχνευσιμότητάς τους. Η ενημέρωση των δεδομένων και των μεταδεδομένων τους θα πρέπει να λαμβάνει χώρα σε τακτική βάση με τη συχνότητα αυτής να εξαρτάται από τον τρόπο με τον οποίο παράγονται τα δεδομένα, τη σχετικότητα των μεταβολών και τις δυνατότητες επαναχρησιμοποίησης του συνόλου δεδομένων. **Σύσταση κλειδί σε ό,τι αφορά την εν λόγω διαδικασία αποτελεί η αναφορά της ημερομηνίας ενημέρωσης στο πλαίσιο των μεταδεδομένων**.
* *Έλεγχος των διευθύνσεων URIs και URLs*. Ο παγκόσμιος ιστός (World Wide Web) είναι δυναμικός. Απαιτείται ως εκ τούτου ο τακτικός έλεγχος των URIs και URLs τόσο από όσο και προς τα σύνολα δεδομένων, προκειμένου να διασφαλίζεται ότι ένας ιστότοπος που αναφέρεται σε ένα σύνολο δεδομένων δεν ανακατευθύνει τους χρήστες σε ανύπαρκτες ιστοσελίδες.
* Έλεγχος των παρατηρήσεων των χρηστών (user feedback) και συνεχής βελτίωση. Οι παρατηρήσεις των χρηστών δύνανται να αυξήσουν την ποιότητα των δημοσιευμένων δεδομένων ενός οργανισμού: οι χρήστες μπορεί να παρέχουν σχόλια για διάφορες όψεις των δεδομένων και η ενσωμάτωση αυτών στις διαδικασίες του οργανισμού μπορεί να βελτιώσει τη χρηστικότητα και την ανιχνευσιμότητα των δεδομένων. Για το λόγο αυτό, κρίνεται σκόπιμο να παρέχεται η δυνατότητα συλλογής των παρατηρήσεων των χρηστών.

***Συνοψίζοντας ο κύκλος ζωής των ανοιχτών δεδομένων περιλαμβάνει τα βήματα***

* ***της συλλογής, και κατά συνέπεια της καταγραφής, προτεραιοποίησης, επιλογής και κατηγοριοποίησης των δεδομένων,***
* ***της προετοιμασίας αυτών σε όρους διασφάλισης ποιότητας, εμπλουτισμού από τεχνική σκοπιά, προσθήκης μεταδεδομένων και εναρμόνισης με το νομικό πλαίσιο,***
* ***της δημοσίευσης/ανάρτησης των δεδομένων,***
* ***και φυσικά της συντήρησης αυτών σε τακτικό επίπεδο μέσω της εφαρμογής κατάλληλων ενημερώσεων, ελέγχων όπως και της παρακολούθησης των παρατηρήσεων των χρηστών.***

***Η εφαρμογή των παραπάνω βημάτων δύναται να οδηγήσει στην εγκαθίδρυση ενός δομημένου μηχανισμού για τη δημοσίευση/ανάρτηση δεδομένων από έναν οργανισμό. Η υλοποίηση των βημάτων αυτών εξαρτάται από τις εσωτερικές διαδικασίες του οργανισμού και είναι σημαντικό να συμπεριληφθεί ρητά και στη στρατηγική του, ενώ συνιστάται να υιοθετεί, όπου αυτό είναι δυνατό, βέλτιστες πρακτικές, όπως η ενσωμάτωση μεταδεδομένων και η δημοσίευση των δεδομένων ως linked data.***

# Η Κρισιμότητα της Έγκαιρης Ανάρτησης

Η έγκαιρη (timely) και πλήρης (comprehensive) διάθεση των δεδομένων αποτελεί μια από τις έξι αρχές του Καταστατικού Χάρτη για τα Ανοιχτά Δεδομένα. Στο πλαίσιο της εν λόγω αρχής, ο Καταστατικός Χάρτης αναγνωρίζει ότι:

* είναι δυνατόν να απαιτούνται τόσο χρόνος όσο και ανθρώπινοι και τεχνικοί πόροι για τον προσδιορισμό των δεδομένων προς δημοσίευση.
* η διαβούλευση με τους χρήστες των δεδομένων, συμπεριλαμβανομένων των πολιτών, άλλων κυβερνητικών φορέων και οργανώσεων της κοινωνίας των πολιτών όσο και του ιδιωτικού τομέα, είναι κρίσιμης σημασίας για την προτεραιοποίηση των δεδομένων προς δημοσίευση ή διαβούλευση.
* προκειμένου να έχουν τα δεδομένα αξία για τις κυβερνήσεις, τους πολίτες, τους φορείς του ιδιωτικού τομέα, τα δεδομένα πρέπει να είναι πλήρη, ακριβή και υψηλής ποιότητας.

Επιπροσθέτως, προωθεί:

* τη δημιουργία, συντήρηση και διαμοιρασμό δημόσιων, περιεκτικών καταλόγων δεδομένων με στόχο την υποστήριξη της διεξαγωγής δημόσιας διαβούλευσης αναφορικά με την προτεραιοποίηση των δεδομένων, τη δημοσίευση τους και τις ημερομηνίες δημοσίευσης.
* την έγκαιρη «απελευθέρωση» δεδομένων υψηλής ποιότητας χωρίς αδικαιολόγητη καθυστέρηση. Τα δεδομένα αυτά πρέπει να είναι πλήρη και ακριβή και να διατίθενται σύμφωνα με την ιεράρχηση που έχει προκύψει από διαβούλευση με τους χρήστες των δεδομένων.
* την αντίληψη ότι στο μέτρο του δυνατού, τα δεδομένα πρέπει να διατίθενται στην αρχική, μη επεξεργασμένη μορφή τους και να συνδέονται με σχετικές οδηγίες, τεκμηρίωση, οπτικοποιήσεις ή άλλες αναλύσεις.
* την αντίληψη ότι στο μέτρο του δυνατού, στα δεδομένα πρέπει να διατηρείται ο διαχωρισμός σε επιμέρους κατηγορίες, συμπεριλαμβανομένων εκείνων του φύλου, της ηλικίας, του εισοδήματος κλπ.
* τη δυνατότητα των χρηστών να παρέχουν συνεχή ανατροφοδότηση (feedback) για την αναθεώρηση και τη βελτίωση των δεδομένων.
* την εφαρμογή συνεπών πρακτικών διαχείρισης του κύκλου ζωής των δεδομένων και τη διασφάλιση ότι τα ιστορικά αντίγραφα των συνόλων δεδομένων διατηρούνται, αρχειοθετούνται και διατηρούνται προσβάσιμα εφόσον διατηρούν την αξία τους.
* τη διαβούλευση με τους χρήστες δεδομένων για την εφαρμογή σημαντικών αλλαγών στη δομή ή την παροχή δεδομένων, προκειμένου να ελαχιστοποιηθεί ο αντίκτυπος για χρήστες που έχουν δημιουργήσει εργαλεία, βασισμένα σε ανοιχτά δεδομένα.
* τη διαφάνεια γύρω από τις διαδικασίες συλλογής και δημοσίευσης δεδομένων και τα χρησιμοποιούμενα πρότυπα μέσω της διάθεσης της τεκμηρίωσής τους online.

***Η έγκαιρη (timely) διάθεση των δεδομένων αποτελεί μία από τις έξι αρχές του Καταστατικού Χάρτη για τα Ανοιχτά Δεδομένα και είναι ιδιαίτερα σημαντική για τη διατήρηση της αξίας τους. Η εν λόγω αρχή είναι ουσιώδης για την εγκαθίδρυση κατάλληλου μηχανισμού ανάρτησης των δεδομένων. Εξίσου ουσιώδεις για τον ίδιο σκοπό, όπως φαίνεται από τα βήματα του κύκλου ζωής των ανοιχτών δεδομένων, είναι και οι αρχές της προσβασιμότητας και χρηστικότητας και της διαλειτουργικότητας .***

# Επικαιροποίηση & Διαλειτουργικότητα

Για να διατηρούν την αξία τους, τα ανοικτά δεδομένα, όπως αναφέρθηκε και παραπάνω, θα πρέπει να διατίθενται επικαιροποιημένα. Η επικαιροποίηση του κάθε συνόλου ανοικτών δεδομένων, ανάλογα με τη φύση του, θα πρέπει να μπορεί να πραγματοποιείται σε τακτική βάση (λιγότερο ή περισσότερο) ή ακόμα και σε πραγματικό χρόνο. Για παράδειγμα οι τοποθεσίες των στάσεων στις αστικές και υπεραστικές συγκοινωνίες είναι ένα σύνολο δεδομένων που έχει υψηλή αντοχή στο χρόνο. Αντίθετα, ένα σύνολο δεδομένων κυκλοφοριακού φόρτου ενός αστικού οδικού δικτύου χάνει την αξία του αν αυτό δεν επικαιροποιείται ανά τακτά χρονικά διαστήματα (της τάξης των μερικών λεπτών της ώρας) ή ακόμα και σε πραγματικό χρόνο.

Σπουδαίο ρόλο στην επικαιροποίηση των ανοικτών δεδομένων κατέχει η διαλειτουργικότητα μεταξύ των πληροφοριακών συστημάτων των φορέων που κατέχουν τα σύνολα δεδομένων (πάροχοι ανοικτών δεδομένων) και των πληροφοριακών συστημάτων των φορέων που αξιοποιούν αυτά τα σύνολα (καταναλωτές ανοικτών δεδομένων). Η Ευρωπαϊκή Κοινότητα, στις 23 Μαρτίου 2017, δημιούργησε το νέο Ευρωπαϊκό Πλαίσιο Διαλειτουργικότητας (https://ec.europa.eu/isa2/eif\_en), το οποίο, εκτός των άλλων, θα διευκολύνει τους «καταναλωτές» να αξιοποιούν σύνολα δεδομένων από «παρόχους» σε οποιοδήποτε ευρωπαϊκό κράτος με τον ίδιο τρόπο. Έτσι, στο ανωτέρω παράδειγμα με τον κυκλοφοριακό φόρτο, ένας «καταναλωτής» θα μπορεί να αναπτύξει μια υπηρεσία, η οποία θα αξιοποιεί, με τον ίδιο τρόπο, επικαιροποιημένα δεδομένα από όλες τις ευρωπαϊκές χώρες. Σκοπός του Ευρωπαϊκού Πλαισίου Διαλειτουργικότητας είναι η ανάπτυξη της ψηφιακής αγοράς, η οποία αξιοποιώντας ταυτόχρονα την ιδέα των ανοικτών δεδομένων (και τα ίδια τα ανοικτά δεδομένα) μπορεί να παίξει καθοριστικό ρόλο στη βελτίωση της καθημερινής ζωής όλων μας.

# Σημασιολογία και Διαλειτουργικότητα

# Η κρισιμότητα του ζητήματος

Μια από τις μεγαλύτερες προκλήσεις, αν όχι η μεγαλύτερη, στον κόσμο των ανοιχτών δεδομένων που διαρκώς διευρύνεται, είναι ο καθορισμός μηχανισμών που θα επιτρέπουν την κατανόηση, χρήση και ανταλλαγή τους με έναν κοινό τρόπο, κατά μήκος διαφόρων και ετερογενών ομάδων χρηστών (ανθρώπων ή/και εφαρμογών).

Η απάντηση στο παραπάνω ερώτημα είναι η επίτευξη της διαλειτουργικότητας των δεδομένων και είναι διττή. Το πρώτο και πιο ανεπτυγμένο σκέλος έχει να κάνει με τον καθορισμό της δομής και της σύνταξης των δεδομένων (π.χ. πρότυπα XML, βάσεις δεδομένων, τύποι δεδομένων). Το δεύτερο σκέλος έχει να κάνει την σημασία και κοινή κατανόηση των δεδομένων (semantics).

Η σημασιολογία δίνει νόημα στα ανοιχτά σύνολα δεδομένων συσχετίζοντας μια περιοχή ενός συνόλου δεδομένων με έναν ορισμό για αυτή την περιοχή. Αποτελεί με αυτό τον τρόπο μια σύνδεση του πραγματικού κόσμου με τον τρόπο που αποτυπώνονται τα ανοιχτά σύνολα δεδομένων. Ουσιαστικά έχει σαν σκοπό να δώσει ερμηνεία στα δεδομένα και κατά συνέπεια μια κοινή συνισταμένη στη μέθοδο με την οποία τα δεδομένα μεταφράζονται σε πληροφόρηση για όλους τους εμπλεκόμενους, χρήστες ή/και δημιουργούς.

# Μέθοδοι & Τεχνικές

Οι μέθοδοι και οι τεχνικές που κάνουν πράξη αυτή την απαίτηση, σχετίζονται συνήθως με:

1. Γλωσσάρια και οντολογίες, τα οποία δίνουν τους ορισμούς των βασικών οντοτήτων που συμμετέχουν σε σύνολα ανοιχτών δεδομένων
2. Καταλόγους που κατηγοριοποιούν τα ανοιχτά δεδομένα και κοινοποιούν τη διαθεσιμότητα τους
3. Επιχειρηματικά λογικά μοντέλα, που πολλές φορές συνδυάζουν τόσο την σημασιολογική όσο και την συντακτική διαλειτουργικότητα των δεδομένων σε μια συγκεκριμένη περιοχή (π.χ. το πρότυπο MISMO για τα στεγαστικά δάνεια).

Τα παραπάνω αποτελούν απαραίτητη προϋπόθεση προκειμένου να γίνει εφικτή η αξιοποίηση, πολύ δε περισσότερο για εμπορικούς/επιχειρηματικούς σκοπούς, των ανοιχτών δεδομένων καθώς συνεισφέρουν στην καλύτερη κατανόηση των στοιχείων που έχουν συλλεχτεί, και ενθαρρύνουν την εύρεση και τη συνδυαστική χρήση συνόλων ανοιχτών δεδομένων με άλλες πηγές επιχειρηματικής πληροφόρησης (π.χ. Data Warehouse, Big Data).

# Οικονομική αποτίμηση των Δεδομένων (Data Monetization)

# Η σημασία του ζητήματος

Οι κυβερνήσεις συγκεντρώνουν πλήθος πληροφορίας για τους πολίτες και τους οργανισμούς με τους οποίους συναλλάσσονται. Επιπλέον σήμερα γίνεται όλο και περισσότερο φανερό ότι τα δεδομένα από μόνα τους είναι ένα από τα πιο σημαντικά περιουσιακά στοιχεία για κάθε οργανισμό. Η πρόκληση που αντιμετωπίζουν είναι πώς ακριβώς θα μπορέσουν να χρησιμοποιήσουν τα δεδομένα τους προς όφελος τους. Για τους δημόσιους οργανισμούς, η πώληση των δεδομένων ως περιουσιακών στοιχείων θα μπορούσε να είναι ένας αποτελεσματικός τρόπος για αύξηση των δημοσίων εσόδων.

# Προτάσεις και Παραδείγματα

Η εκποίηση των δεδομένων από δημόσιους οργανισμούς απαιτεί καταρχήν να γίνει μια καταγραφή όλων των δεδομένων που έχουν οι δημόσιες υπηρεσίες. Αυτό θα βοηθήσει να κατανοηθεί ενδελεχώς η χρηστικότητα των δεδομένων και οι διαφορετικές δυνατότητες εκμετάλλευσής τους. Στη συνέχεια θα πρέπει να σχεδιαστεί το επιχειρηματικό μοντέλο βάση του οποίου θα γίνει η χρήση τους από τρίτες εταιρίες καθώς επίσης και να βρεθούν τρόποι ώστε να συστηματοποιηθεί η συλλογή των δεδομένων. Αυτό δε θα βοηθήσει μόνο στο να αποκτηθεί μια πιο πλήρης εικόνα αλλά επιπλέον θα συνεισφέρει στη μείωση κόστους για την συντήρηση τους.

**Παραδείγματα:**

1. Copenhagen’s data Xchange

Η πόλη κατευθύνεται προς το να χρησιμοποιεί δεδομένα συλλεγμένα από τις κυβερνήσεις, ώστε να κινητοποιήσει τους κατοίκους, εταιρείες και πανεπιστήμια στο να δημιουργήσουν έσοδα παρέχοντας τα δικά τους στοιχεία πληροφοριών.

Το Copenhagen Solutions Lab (CSL), το οποίο επιβλέπει το έργο ανοικτών δεδομένων της πόλης, επένδυσε τον προηγούμενο χρόνο στο "City Data Exchange", μία διαδικτυακή αγορά δεδομένων που διαμορφώθηκε έπειτα από την ανοικτή πύλη δεδομένων της πόλης. Το 2014, υποβλήθηκε αίτηση υποβολής προτάσεων ύψους 1.2 εκ. δολαρίων ώστε να δημιουργηθεί ένα βιώσιμο επιχειρηματικό μοντέλο για την διαδικτυακή αγορά.

Οι πωλήσεις δεδομένων μίας εταιρείας σε μία άλλη, δεν είναι καινούργιο φαινόμενο. Εταιρείες τρίτου μέρους ευημερούν δεκαετίες τώρα. Ωστόσο, το City Data Exchange είναι μοναδικό, καθιστώντας αυτές τις συναλλαγές διαφανείς και συγκεντρωμένες γύρω από τον εγγενώς "τοπικό" χώρο μιας συγκεκριμένης πόλης. Στο RFP, η CSL διατύπωσε ένα σαφές σύνολο προσδοκιών για το έργο - συγκεκριμένα, η προσφορά έπρεπε να καλύψει τις μοναδικές ανάγκες των νεοσύστατων επιχειρήσεων και των ακαδημαϊκών ερευνητών της Κοπεγχάγης.

Καθώς προχωράει αυτό το έργο, θα δοκιμαστεί πόσο επιτυχώς μπορεί ένα δημόσιο ίδρυμα να αντλήσει πόρους από τρίτους προμηθευτές ώστε να ενισχύσει το δυναμικό ανάλυσης δεδομένων των τοπικών οργανισμών, προωθώντας παράλληλα άλλες προτεραιότητες της πόλης, όπως η εξασφάλιση της προστασίας της ισότητας και της ιδιωτικής ζωής.

<http://datasmart.ash.harvard.edu/news/article/how-cities-can-help-local-institutions-monetize-their-data-970>

1. NHS Care.data

Στο Ηνωμένο Βασίλειο, τα σχέδια πώλησης η ευρύτερης δωρεάν χρήσης δεδομένων της Εθνικής Υπηρεσίας Υγείας (NHS) της Αγγλίας αντιμετώπισαν μια τεράστια αντίδραση. Σκοπός της υπηρεσίας ήταν η αποστολή των αρχείων των ασθενών από όλους τους γενικούς ιατρούς της Αγγλίας σε μια κεντρική βάση δεδομένων, η οποία θα χρησιμοποιηθεί για την ιατρική έρευνα τόσο από το NHS όσο και από εταιρείες όπως οι φαρμακευτικές εταιρείες.

Το NHS Αγγλίας δήλωσε ότι τα δεδομένα θα χρησιμοποιηθούν για τον εντοπισμό περαιτέρω δυνατοτήτων καθώς και βελτίωσης της αποτελεσματικότητας στο NHS και ότι μόνο εγκεκριμένες εταιρείες θα έχουν πρόσβαση στα δεδομένα. Ωστόσο, πολλοί ανησυχούν - όχι για τους στόχους και τις δυνατότητες της πρωτοβουλίας, αλλά για την προστασία της ιδιωτικής ζωής των ασθενών, την έλλειψη σαφήνειας και προώθησης των σχεδίων και ακόμη με τις υπερβολικά χαμηλές τιμές πώλησης των δεδομένων στις ιδιωτικές εταιρίες.

<http://www.bbc.com/news/health-26259101>

<http://www.computerweekly.com/news/450299728/Caldicott-review-recommends-eight-point-consent-model-for-patient-data-sharing>

# Η χρησιμότητα των συμβάσεων SLA

Τα Ανοικτά Δεδομένα δηλώνουν την ύπαρξή τους ως δεδομένα που παρέχονται όπως είναι (as is). Ο κάθε εκδότης δεν μπορεί να εγγυηθεί για το λόγο αυτό την ορθότητα, την επικαιρότητα, την παροχή δεδομένων σε συγκεκριμένη μορφή ή τη μελλοντική διαθεσιμότητά τους. Οτιδήποτε υπόκειται σε μελλοντική αλλαγή ή διόρθωση είναι επίσης υπό ενδεχόμενη αίρεση. Οι οργανισμοί που βασίζονται σε δημόσια δεδομένα μπορεί μερικές φορές να απαιτούν ισχυρότερες δεσμεύσεις από τον εκδότη των δεδομένων αυτών. Τέτοιες απαιτήσεις μπορούν να επιτευχθούν μόνο μέσω συμφωνιών επιπέδου υπηρεσιών (SLA).

Τα SLA είναι συμβάσεις που μπορεί να καλύψουν πτυχές όπως:

* Μεγαλύτερη συχνότητα ενημέρωσης
* Δυνατότητες φιλτραρίσματος των δεδομένων
* Προ-επεξεργασία των δεδομένων
* Σταθερότητα, π.χ. συνεπείς και σταθερές μορφές, σχήματα και χρήση προτύπων.
* Υποστήριξη, π.χ. πρόσβαση στο τεχνικό προσωπικό για την αντιμετώπιση προβλημάτων.

Παρόλο που η καλύτερη πρακτική είναι να δημοσιεύονται ανοιχτά δεδομένα χωρίς χρέωση, μπορεί δυνητικά να δικαιολογηθεί η χρέωση ενός τέλους SLA εάν απαιτηθεί από τον εκδότη δεδομένων να παρέχει πρόσθετες υπηρεσίες δεδομένων ή / και υποδομή πάνω από τα αρχικά δεδομένα.

# Εκπαίδευση

Η διεθνής εμπειρία έχει αναδείξει την ανάγκη για εκπαίδευση στα στελέχη φορέων του δημοσίου. Η πλειονότητα των στελεχών του δημοσίου συχνά δεν έχει την απαραίτητη γνώση για την υλοποίηση του νέου θεσμικού πλαισίου που έχει διαμορφωθεί από την ψήφιση του Νόμου 4305/2014 για τα Ανοικτά Δεδομένα και είναι αρνητική ως προς το άνοιγμα δεδομένων, αγνοώντας τα οφέλη που απορρέουν από αυτά.

Η ανάγκη εκπαίδευσης γύρω από τα Ανοικτά Δεδομένα γίνεται εύκολα αντιληπτή παρατηρώντας την ποσότητα και κυρίως την ποιότητα των δεδομένων που έχουν αναρτηθεί στο data.gov.gr. Τα σύνολα δεδομένων που έχουν αναρτηθεί έχουν στην πλειοψηφία τους χαμηλή ποιότητα. Πρόκειται για αρχεία χωρίς μηχαναγνώσιμη μορφή (format) τα οποία δεν συνοδεύονται από κατάλληλα μεταδεδομένα (τίτλος, περιγραφή, άδεια χρήσης κλπ).

# Άξονες Εκπαίδευσης

Η εκπαίδευση μπορεί να κινηθεί γύρω από τους παρακάτω άξονες.

**Εισαγωγικές έννοιες και ορισμοί**

* Τι είναι ανοικτά δεδομένα.
* Άδειες χρήσης ανοικτών δεδομένων
* Ποιότητα ανοικτών δεδομένων
* Μηχαναγνώσιμα πρότυπα δεδομένων

**Αναγνώριση συνόλων δεδομένων στο πλαίσιο του οργανισμού τους**

* Απομόνωση συνόλων δομών δεδομένων μέσα από τις διαδικασίες του οργανισμού τους.

**Καθαρισμός δεδομένων**

* Ανωνυμοποίηση συνόλων δεδομένων.
* Καθαρισμός δεδομένων από περιττές πληροφορίες

**Δημοσίευση δεδομένων**

* Χρήση μηχαναγνώσιμων προτύπων
* Βέλτιση χρήση της πλατφόρμας data.gov.gr
* Σύγχρονες τεχνικές δημοσίευσης δεδομένων (REST APIs κλπ)

**Συντήρηση συνόλων δεδομένων**

* Ανάλυση και ορισμός διαδικασιών επικαιροποίησης συνόλων δεδομένων.

Τέλος προτείνεται η ενημέρωση των στελεχών γύρω από την οικονομία των δεδομένων (data economy). Παρουσίαση παραδειγμάτων (case studies), διεθνών πρακτικών και λύσεων με μετρήσιμο οικονομικό αντίκτυπο.

Σκοπός είναι τα στελέχη του δημοσίου να αντιληφθούν τα οικονομικά μεγέθη γύρω από τα ανοικτά δεδομένα και το πως ο ίδιος ο οργανισμός τους μπορεί να επωφεληθεί από το άνοιγμά τους (με όρους προϋπολογισμού, βελτιωμένων διαδικασιών αλλά και κερδισμένων εργατοωρών).

# Προτάσεις

Τα παραπάνω συγκεντρώνονται στις προτάσεις που ακολουθούν

**Σχεδιασμός εκπαιδευτικών σεμιναρίων στοχευμένων για στελέχη του δημοσίου**

Σεμινάρια διάρκειας μίας έως τριών ημερών οι οποίοι μπορεί να καλύψουν τους παρακάτω άξονες:

* Εισαγωγή στα Ανοικτά Δεδομένα (Open Data in a Day) (1η ημέρα)
* Τεχνικές λύσεις για τον καθαρισμό και τη δημοσίευση των Ανοικτών Δεδομένων (2η ημέρα)
* Εργασία σε ενδεικτικά σύνολα δεδομένων του εκάστοτε οργανισμού οπου οι συμμετέχοντες θα εργαστούν με σκοπό την δημοσίευση ποιοτικών συνόλων δεδομένων. (3η ημέρα).

Τα σεμινάρια μπορούν να απευθύνονται τόσο σε μεμονωμένους οργανισμούς (στελέχη από έναν οργανισμό) όσο και σε οριζόντιες δράσεις (ομοειδή τμήματα από διαφορετικούς οργανισμούς). Ανάλογα με τους συμμετέχοντες το σεμινάριο μπορεί να προσαρμόζεται ώστε να καλύπτει ιδανικά την ιδιότητα και το επίπεδο των συμμετεχόντων.

Τα σεμινάρια θα πρέπει να εκτελούνται σε μικρές ομάδες συμμετεχόντων (κάτω των 20 ατόμων) με σκοπό την επίτευξη του καλύτερο δυνατού αποτελέσματος.

**Σχεδιασμός πλατφόρμας εκπαίδευσης στελεχών δημοσίου - eToolkit**

Ηλεκτρονική πλατφόρμα ασύγχρονης εκπαίδευσης για στελέχη του δημοσίου. Η πλατφόρμα θα χωρίζεται σε δυο μέρη.

* Μαθήματα (courses) με εκπαιδευτικό περιεχόμενο γύρω από τις βασικές αρχές των Ανοικτών Δεδομένων.
* Έναν οδηγό με βήμα προς βήμα διαδικασία για το καθαρισμό και τη δημοσίευση των δεδομένων. Ο οδηγός θα περιέχει παραδείγματα καλών πρακτικών και σύντομους οδηγούς αξιοποίησης ανοικτών εργαλείων (ΕΛ/ΛΑΚ) για την εκτέλεση των προτεινόμενων διαδικασιών.

# ΤΕΧΝΙΚΟ ΠΑΡΑΡΤΗΜΑ

# Αρχές έκδοσης (versionong) δεδομένων

Παρακάτω διατυπώνονται κάποιες βασικές αρχές της έκδοσης (versioning) δεδομένων.

1. Η πρώτη αρχή της έκδοσης δεδομένων είναι ότι οι αλλαγές στα δεδομένα έχουν πηγές ή εξηγήσεις. Ένα σύστημα έκδοσης δεδομένων πρέπει να είναι σε θέση να συνδέει τις τιμές δεδομένων (data values), τη δομή (structure) και τα μεταδεδομένα (metadata) καθώς και τις μεταβολές αυτών των χαρακτηριστικών, με εξηγήσεις αυτών των τιμών ή με τις μεταβολές των τιμών στο επίπεδο της αξίας (και όχι στο επίπεδο μεταβλητών, παρατηρήσεων ή αρχεία).
2. Η δεύτερη αρχή είναι ότι η έκδοση δεδομένων πρέπει να βασίζεται στην αξία (value-based), και όχι με βάση τη μεταβλητή ή την παρατήρηση (observation-based). Ένα σύστημα δεν μπορεί να προνοήσει παρατηρήσεις πάνω σε μεταβλητές ή μεταβλητές πάνω από τις παρατηρήσεις. Μια αλλαγή σε μια παρατήρηση είναι απαραίτητα μια αλλαγή σε μια μεταβλητή, και αντίστροφα.
3. Η τρίτη αρχή της έκδοσης δεδομένων είναι ότι τα δεδομένα υπάρχουν ανεξάρτητα από τη μορφή τους. Αν κάποιος αλλάξει μια τιμή δεδομένων σε ένα CSV σε σχέση με ένα JSON tree, αυτές είναι αλλαγές ισοδύναμες ως προς το περιεχόμενο. Ως εκ τούτου, οποιοδήποτε σύστημα ελέγχου έκδοσης θα πρέπει να επιτρέπει στους χρήστες δεδομένων να αλληλεπιδρούν με τα δεδομένα σε οποιαδήποτε μορφή αρχείου επιλέγουν, χωρίς απαραίτητα να χρησιμοποιούν τη υποκείμενη μορφή αποθήκευσης δεδομένων.
4. Η τέταρτη αρχή της έκδοσης δεδομένων είναι ότι η συνεργασία είναι απαραίτητη για την παραγωγή και την επεξεργασία δεδομένων. Ένα σύστημα έκδοσης δεδομένων πρέπει να είναι εγγενώς συνεργατικό και λογικά να καταγράφει ποιος παράγει και τροποποιεί δεδομένα.
5. Η πέμπτη αρχή της έκδοσης δεδομένων είναι ότι οι αλλαγές στη δομή δεδομένων θα πρέπει να καταγράφονται ανεξάρτητα από τις τιμές δεδομένων. Η σειρά ταξινόμησης των παρατηρήσεων, η διάταξη των στηλών / μεταβλητών και η διευθέτηση των σειρών ως περιπτώσεων έναντι των περιπτώσεων κ.λπ. (δηλαδή "ευρείες" έναντι "μεγάλων" διευθετήσεων) αποτελούν δομικά χαρακτηριστικά ενός συνόλου δεδομένων και όχι χαρακτηριστικά των δεδομένων αφ ΄εαυτά. Αυτές είναι σημαντικές, αλλά μια διαδικασία εκδόσεων δεδομένων θα πρέπει να είναι σε θέση να διακρίνει μια αλλαγή στο περιεχόμενο ενός συνόλου δεδομένων από μια αλλαγή στην οργάνωση ή τη δομή ενός συνόλου δεδομένων και, στην τελευταία περίπτωση, να αναγνωρίζει σωστά ως ταυτόσημο ένα σύνολο δεδομένων που είναι διατεταγμένο με δύο διαφορετικούς τρόπους.
6. Η έκτη αρχή της έκδοσης δεδομένων είναι ότι τα μεταδεδομένα έχουν σημασία αλλά, όπως και η δομή, θα πρέπει να αντιμετωπίζονται ξεχωριστά από τις αλλαγές στα δεδομένα. Δύο πανομοιότυπα σύνολα δεδομένων με διαφορετικά μεταδεδομένα πρέπει να αναγνωρίζονται ως πανομοιότυπα ως προς το περιεχόμενο.

Τα δεδομένα λοιπόν πρέπει να αποθηκεύονται με ένα key-value τρόπο, όπου ένα αυθαίρετο κλειδί διατηρεί μια συγκεκριμένη τιμή δεδομένων. Μια χαρτογράφηση (mapping) συνδέει αυτές τις συγκεκριμένες τιμές δεδομένων με τις παρατηρήσεις και τις μεταβλητές, έτσι ώστε οποιαδήποτε αξιολόγηση των αλλαγών στα δεδομένα να είναι ανεξάρτητη από τη μορφή και ανεξάρτητη από τη δομή. Έτσι, μια μεταβολή σε μια τιμή καταγράφεται πρώτα ως αλλαγή σε μια τιμή, αλλά μπορεί δευτερευόντως να αναγνωριστεί ως ταυτόχρονη αλλαγή τόσο σε παρατήρηση όσο και σε μεταβλητή.

Οποιαδήποτε διεπαφή σε ένα τέτοιο key-value store θα πρέπει να έρχεται σε μια οικεία και ευέλικτη μορφή: οι χρήστες θα πρέπει να αλληλεπιδρούν με ένα σύστημα εκδόσεων δεδομένων με οποιονδήποτε τρόπο χρησιμοποιούν δεδομένα (π.χ. επεξεργαστή κειμένου, λογισμικό ανάλυσης δεδομένων, εφαρμογή υπολογιστικού φύλλου, κλπ. ). Οι αλλαγές πρέπει να καταγράφονται σε ένα κύριο αρχείο το οποίο μπορεί να εισάγει και να εισάγει άμεσα από και να εξάγει σε οποιαδήποτε μορφή αρχείου δεδομένων (συμπεριλαμβανομένων των απεριόριστων αρχείων, υπολογιστικών φύλλων, XML, JSON κλπ.).

Τα συστήματα έκδοσης δεδομένων πρέπει να διαθέτουν ένα πιο εξελιγμένο σύστημα δέσμευσης από εκείνο που παρέχεται από τα τρέχοντα συστήματα ελέγχου εκδόσεων με γνώμονα το λογισμικό. Συγκεκριμένα, οι δεσμεύσεις θα πρέπει να καταγράφουν όχι μόνο την αλλαγή σε μια τιμή δεδομένων, τη δομή ή τα μεταδεδομένα αλλά επίσης και τις δομημένες πληροφορίες που εξηγούν αυτήν την αλλαγή, συμπεριλαμβανομένου του λόγου της αλλαγής, της πηγής της αξίας των δεδομένων και του συντάκτη της αλλαγή. Στην ουσία, τόσο οι αλλαγές όσο και οι καταστάσεις των πλήρων δεδομένων θα πρέπει να είναι πλήρως αξιόπιστα και να περιέχουν μεταδεδομένα σχετικά με τις παραπομπές

1. Open data Charter, <http://opendatacharter.net/> [↑](#footnote-ref-1)
2. Ο εν λόγω κύκλος ζωής περιγράφεται στο πλαίσιο **πρακτικού οδηγού** που διατίθεται από την Ευρωπαϊκή Διαδικτυακή Πύλη Δεδομένων (European Data Portal) με στόχο τη διαχείριση ανοιχτών δεδομένων και την επιτυχή υλοποίηση σχετικών πρωτοβουλιών από ενδιαφερόμενους οργανισμούς.

   <https://www.europeandataportal.eu/en/providing-data/goldbook>

   <https://www.europeandataportal.eu/en/providing-data/goldbook/putting-place-open-data-lifecycle> [↑](#footnote-ref-2)