

**ΠΡΑΚΤΙΚΕΣ ΓΙΑ ΤΗΝ ΠΑΡΟΥΣΙΑΣΗ ΔΕΔΟΜΕΝΩΝ ΚΑΙ
ΜΕΤΑΔΕΔΟΜΕΝΩΝ ΜΕ ΒΑΣΗ ΤΗΝ ΕΚΘΕΣΗ ΤΟΥ
ΟΡΓΑΝΙΣΜΟΥ ΓΙΑ ΤΗΝ ΟΙΚΟΝΟΜΙΚΗ ΣΥΝΕΡΓΑΣΙΑ &
ΑΝΑΠΤΥΞΗ (ΟΟΣΑ - ΟΕCD)**

ΕΙΣΑΓΩΓΗ

Η παρούσα τεχνική έκθεση παρουσιάζει προτάσεις για το περιεχόμενο και τον τρόπο δημοσίευσης των μεταδεδομένων. Αποτελεί οδηγία για την αναφορά διάφορων τύπων στατιστικών δεδομένων, πρακτικών αναθεώρησης, λαθών, παραπομπών και αλλαγών βάσεων.

Τα στοιχεία που παρουσιάζονται προέρχονται από την τελευταία επίσημη έκθεση “Data and metadata reporting and presentation handbook”¹ του Οργανισμού για την Οικονομική Συνεργασία και Ανάπτυξη (ΟΟΣΑ) η οποία εκδόθηκε το 2007.

Η ιδιαίτερη χρησιμότητα της έκθεσης έγκειται σε δύο στοιχεία: α) Στην διασφάλιση ότι σε κάθε δημοσίευση αναφέρονται τα απαραίτητα μεταδεδομένα και β) στην ύπαρξη κοινού τρόπου παρουσίασης των μεταδεδομένων στα πλαίσια μιας εταιρίας ή υπηρεσίας. Ο κοινός τρόπος παρουσίασης των μεταδεδομένων, συντελεί στην εύκολη ερμηνεία και κατανόηση των δεδομένων και δίνει συνοχή στο σύνολο των δημοσιευμάτων.

Την παρούσα έκθεση σύνταξε και επιμελήθηκε ο επιστημονικός συνεργάτης του Παρατηρητηρίου, Σωκράτης Σεϊτανίδης, Msc χωροτάκτης - πολεοδόμος.

Τα περιεχόμενα και οι τυχόν απόψεις που φιλοξενούνται στις εκθέσεις και τα κείμενα εργασίας του Παρατηρητηρίου δεν αντανakλούν απαραίτητα την σύμφωνη γνώμη της εταιρείας ΕΓΝΑΤΙΑ ΟΔΟΣ Α.Ε. Επιτρέπεται η αναπαραγωγή για μη εμπορικούς σκοπούς με την προϋπόθεση ότι θα αναφέρεται η πηγή (Παρατηρητήριο - ΕΓΝΑΤΙΑ ΟΔΟΣ Α.Ε.).

Internet:

<http://www.egnatia.gr>

<http://observatory.egnatia.gr>

© Εγνατία Οδός Α.Ε. 2007

¹ Η πλήρης έκθεση του ΟΟΣΑ είναι διαθέσιμη στη διεύθυνση:

\\ERIDANI\intranet1\Intra2\Appl\Observatory\DOC\Library\metadata\OECD_metadata.pdf

ΠΕΡΙΕΧΟΜΕΝΑ

1. Μεταδεδομένα: Ορισμός	4
2. Χρησιμότητα των σταθεροτύπων (standards) στα μεταδεδομένα	4
3. Προτεινόμενες πρακτικές για τη συνεχή χρήση συγκεκριμένης ορολογίας και ορισμών	5
4. Οδηγίες για την αναφορά σε διάφορους τύπους στατιστικών δεδομένων	5
4.1 Για τα πρωτογενή δεδομένα (original data)	5
4.2 Για σταθμισμένες τιμές	6
4.3 Για ρυθμούς μεγέθυνσης (growth rates)	7
4.4 Για λόγους, αναλογίες, ποσοστά και ρυθμούς	7
4.5 Προτεινόμενες πρακτικές για την αναφορά διαφορετικών τύπων «δεδομένων με χρονική σειρά» (time series data)	8
5. Προτεινόμενες πρακτικές για την αναφορά και διάδοση των μεταδεδομένων	10
5.1 Οι ανάγκες των μεταδεδομένων	10
5.2 Η πρόσβαση στα μεταδεδομένα	10
5.3 Υιοθέτηση ενός κοινού συνόλου ορολογίας για την παρουσίαση των μεταδεδομένων	11
5.4 Ευδιάκριτη παρουσίαση παρόμοιων στατιστικών δεδομένων	11
6. Προτεινόμενες πρακτικές για την αναθεώρηση δεδομένων	12
6.1 Οι διαβουλεύσεις με τους χρήστες για τις πρακτικές αναθεώρησης	12
6.2 Παροχή καθαρών, εύκολα προσβάσιμων και σύντομων δηλώσεων προς τους χρήστες σχετικά με το πότε και το γιατί γίνονται αναθεωρήσεις	12
6.3 Η δημιουργία, όπου είναι δυνατόν, ενός σταθερού κύκλου ανανέωσης	12
6.4 Η εξισορόπηση της διαχρονικής σταθερότητας των σειρών των δεδομένων έναντι στην ανάγκη να αναθεωρηθούν οι σειρές λόγω της εισαγωγής νέων μεθοδολογιών, ιδεών, εννοιών, ταξινομήσεων κτλ.	12
6.5 Η εξασφάλιση ότι οι αναθεωρήσεις γίνονται και για τα προηγούμενα έτη, προκειμένου να δώσουν συνεπή διαχρονικά δεδομένα	12
6.6 Η παροχή κατάλληλης τεκμηρίωσης στις αναθεωρήσεις	13
6.7 Οι υπηρεσίες πρέπει να υπενθυμίζουν τους χρήστες για το μέγεθος των πιθανών αναθεωρήσεων με βάση αυτά που έγιναν στο παρελθόν	13
6.8 Διαφανής και έγκαιρη υποβολή έκθεσης των λαθών	13
6.9 Προτεινόμενες πρακτικές για την παρουσίαση σπασμένων σειρών	13
7. Παρουσίαση πληροφορίας για δειγματοληπτικά και μη δειγματοληπτικά λάθη	15
7.1 Δειγματοληπτικά λάθη	15
7.2 Μη δειγματοληπτικά λάθη	16
8. Προτεινόμενες πρακτικές για τις αλλαγές βάσεων	17
9. Προτεινόμενες πρακτικές για τον τρόπο παραπομπής	17
9.1 Παραπομπή συνόλου δεδομένων	17
9.2 Παραπομπή σε κείμενο	18
10. Προτεινόμενες πρακτικές για την παρουσίαση και την αναφορά δεδομένων από πηγές δημοσίων υπηρεσιών	19

1. Μεταδεδομένα: Ορισμός

Ο απλούστερος ορισμός των μεταδεδομένων είναι ότι αποτελούν «δεδομένα που αφορούν δεδομένα» και πιο συγκεκριμένα είναι πληροφορία (δεδομένα) για ένα συγκεκριμένο περιεχόμενο (δεδομένα).²

Ένα αντικείμενο μεταδεδομένων μπορεί να περιγράφει ένα αυτόνομο στοιχείο-δεδομένο ή μία συλλογή δεδομένων. Τα μεταδεδομένα χρησιμοποιούνται για να διευκολύνουν την κατανόηση και την διαχείριση των δεδομένων. Για το λόγο αυτό, τα μεταδεδομένα διαφέρουν ανάλογα με τον τύπο των δεδομένων και το πλαίσιο μέσα στο οποίο θα γίνει η χρήση τους. Για παράδειγμα, σε μία βιβλιοθήκη, όπου τα στοιχεία είναι το περιεχόμενο των τίτλων που υπάρχουν στα ράφια, τα μεταδεδομένα για έναν τίτλο θα μπορούσαν να περιλαμβάνουν μια περιγραφή του περιεχομένου, τον συγγραφέα, την ημέρα δημοσίευσης και τη φυσική θέση του βιβλίου. Στο πλαίσιο μιας φωτογραφικής μηχανής, όπου το δεδομένο είναι η φωτογραφική εικόνα, τα μεταδεδομένα θα μπορούσαν τυπικά να περιλαμβάνουν την ημερομηνία που λήφθηκε η φωτογραφία και λεπτομέρειες των ρυθμίσεων της φωτογραφικής μηχανής. Τα μεταδεδομένα για μια βάση δεδομένων, π.χ. για ένα αρχείο του υπολογιστή, μπορεί τυπικά να περιλαμβάνει το όνομα του αρχείου, τον τύπο του αρχείου και το όνομα του διαχειριστή των δεδομένων.¹

2. Χρησιμότητα των σταθεροτύπων (standards) στα μεταδεδομένα

Η ύπαρξη διεθνών standards (σταθερότυπα) στο περιεχόμενο και την παρουσίαση των μεταδεδομένων, συντελεί στα εξής:

A. Βελτιώνεται η ποιότητα των δεδομένων σε εθνικό και διεθνές επίπεδο και δημιουργούνται αποτελέσματα που ερμηνεύονται εύκολα και έχουν συνοχή. Τα τελευταία δέκα χρόνια, ένας αριθμός διεθνών οργανισμών και εθνικών στατιστικών υπηρεσιών (όπως Eurostat, IMF, OECD, Statistics Canada, Statistics Denmark, κ.α.) ανέπτυξαν πλαίσια ποιότητας ως μέρος των πρωτοβουλιών τους για την βελτίωση της ποιότητας των στατιστικών αποτελεσμάτων, τα οποία διαδίδονται σε εσωτερικούς και εξωτερικούς χρήστες.

B. Ελαχιστοποιείται το φορτίο των υποβαλλόμενων εκθέσεων. Η ύπαρξη των διεθνών σταθεροτύπων συντελεί στο να ελαχιστοποιηθεί ο φόρτος των εκθέσεων που παρέχεται από τις εθνικές υπηρεσίες στους διεθνείς οργανισμούς.

Τα σταθερότυπα των μεταδεδομένων αφορούν:

- Την αναθεώρηση των πρακτικών παρουσίασης.

² <http://en.wikipedia.org/wiki/Metadata>

- Την αναφορά πρακτικών που αφορούν την παρουσίαση των κενών και λαθών στα δεδομένα που παρουσιάζονται με χρονική σειρά.
- Τις πρακτικές για την αναφορά δειγματοληπτικών και μη δειγματοληπτικών λαθών.
- Τα έτη βάσης κατά την παρουσίαση ανηγμένων τιμών.
- Τις πρακτικές παραπομπής σε δεδομένα και μεταδεδομένα.
- Τις πρακτικές για αναφορά δεδομένων που προέρχονται από άλλες επίσημες πηγές.

3. Προτεινόμενες πρακτικές για τη συνεχή χρήση συγκεκριμένης ορολογίας και ορισμών

1. Να δημιουργηθεί μια στρατηγική διαχείρισης ορολογίας (terminology management strategy). Κύριο εργαλείο για το παραπάνω είναι η δημιουργία ενός εταιρικού λεξικού το οποίο θα είναι εύκολα προσβάσιμο από διαφορετικά μέρη της υπηρεσίας.
2. Να υπάρχει senior management ώστε να επιβεβαιωθεί ότι η χρήση της σωστής ορολογίας θα υιοθετηθεί και θα χρησιμοποιείται από ολόκληρη την υπηρεσία.
3. Το λεξικό να περιέχει τις ελάχιστες δομές που υπάρχουν στις διεθνείς οδηγίες-προδιαγραφές όπως αυτές που διαμορφώνονται από τους ISO, OMG και W3C.
4. Το λεξικό θα πρέπει για κάθε λύμα να περιέχει α) τίτλο (concept label), β) ορισμό, γ) λεπτομερείς πληροφορίες για τις πηγές από τις οποίες προέρχεται ο ορισμός, δ) σχετικούς όρους-έννοιες με το λύμα ε) επιπρόσθετη πληροφορία για το πως χρησιμοποιείται ο όρος σε ένα στατιστικό ή γεωγραφικό περιεχόμενο.
5. Πρέπει να παρέχονται δεσμοί (links) ή παραπομπές σε συγκεκριμένα κύρια γλωσσάρια που αναπτύχθηκαν σε εθνικό ή διεθνή επίπεδο στο Internet ή intranet.
6. Πρέπει να αποφεύγεται η χρήση ακριβώς του ίδιου τίτλου για διαφορετικούς όρους.

4. Οδηγίες για την αναφορά σε διάφορους τύπους στατιστικών δεδομένων

4.1 Για τα πρωτογενή δεδομένα (original data)

1. Τα πρωτογενή δεδομένα πρέπει να διατίθενται με τον τρόπο που να εξασφαλίζει τη μέγιστη δυνατή ευελιξία, ώστε οι χρήστες τους να μπορούν να

κάνουν εύκολα περαιτέρω ανάλυση ή να τα μετασχηματίσουν. Οι χρήστες θα πρέπει να έχουν πρόσβαση σε τουλάχιστον ορισμένα από τα σύνολα-κλειδιά των πρωτογενών δεδομένων, που μπορούν να υπάρχουν είτε απευθείας στην δημοσίευση, είτε με την παροχή παραπομπών ή υπερσυνδέσμων. Οι προϋποθέσεις κάτω από τις οποίες επιτρέπεται η πρόσβαση στα λεπτομερή, εμπιστευτικά δεδομένα, θα πρέπει να διέπονται από την πολιτική διάδοσης δεδομένων της εταιρίας.

2. Σε περιπτώσεις όπου τα πρωτογενή δεδομένα είναι γνωστό ότι έχουν σημαντικά, μη-δειγματοληπτικά λάθη, πρέπει να παρέχονται τα κατάλληλα μεταδεδομένα που να διευκολύνουν τη σωστή χρήση των δεδομένων.

3. Τα πρωτογενή δεδομένα συχνά μετασχηματίζονται σε ρυθμούς ανάπτυξης ή σταθμισμένες τιμές, για να διευκολυνθεί η ερμηνεία και κατανόησή τους. Το ιδανικό είναι, τα πρωτογενή δεδομένα να παρουσιάζονται επιπρόσθετα στις μετασχηματισμένες σειρές. Ωστόσο, αυτό μπορεί να μην είναι δυνατό, λόγω περιορισμένου χώρου. Σε αυτές τις περιπτώσεις, θα πρέπει να υπάρχουν σαφείς ενδείξεις της διαθεσιμότητας των στοιχείων και του πως είναι προσβάσιμα.

4.2 Για σταθμισμένες τιμές

Σύμφωνα με τη Στατιστική Υπηρεσία του Καναδά, προτείνεται η παροχή των παρακάτω μεταδεδομένων:

1. Ακριβείς ορισμοί των εννοιών που στοχεύουν να μετρήσουν οι δείκτες. Ιδιαίτερη σημασία θα πρέπει να δοθεί στους τυχόν περιορισμούς που υπάρχουν στη χρήση ή εφαρμογή του περιεχομένου.
2. Περιγραφές των μεθοδολογιών που χρησιμοποιούνται στην σύνταξη των δεικτών με ιδιαίτερη αναφορά σε:
 - Μεθόδους υπολογισμού του δείκτη σαν επακόλουθα της επιλογής της φόρμουλάς του. (e.g. Laspeyres, Paasche, Fisher) και η στρατηγική για την κατασκευή των σειρών του δείκτη (π.χ. είτε ως βασικές τιμές είτε ως σύνθετες τιμές σε σχέση με άλλους δείκτες).
 - Σύστημα στάθμισης που χρησιμοποιείται, αναθεωρήσεις των πρακτικών στάθμισης και η συχνότητα με την οποία αναθεωρούνται οι σταθμιστές.
 - Υπολογισμός σε διάφορα επίπεδα συνάθροισης.
 - Επιλογή του έτους βάσης.
 - Η συχνότητα της αλλαγής της επιλογής έτους βάσης.
 - Οι διαδικασίες με τις οποίες δικτυώνονται, συντίθενται ή/και συσχετίζονται οι δείκτες.

Οι μεθοδολογίες που εφαρμόζονται θα πρέπει να συγκρίνονται και να αιτιολογούνται με έμφαση στις έννοιες των δεικτών και στα αποτελέσματα των εξαγόμενων προϊόντων που περιγράφονται.

Τέλος, εάν αρκετή από την παρεχόμενη πληροφορία έχει ειδικό ενδιαφέρον σε συγκεκριμένους χρήστες, πρέπει να επιδιωχθεί να υπάρχουν διάφορα επίπεδα ενημέρωσης, που να στοχεύουν σε διαφορετικές ομάδες χρηστών. Υπάρχει διαφορετική επιθυμία από τους χρήστες για τη δομή των μεταδεδομένων, ανάλογα με το επίπεδο εξειδίκευσής τους και τις ανάγκες τους.

4.3 Για ρυθμούς μεγέθυνσης (growth rates)

Υπάρχουν δύο συστάσεις-κλειδιά όσον αφορά την ορολογία στους ρυθμούς μεγέθυνσης:

1. Να ελαχιστοποιηθεί ο κίνδυνος να κατανοηθούν λάθος από τους χρήστες οι ρυθμοί μεγέθυνσης μέσω της συνεχούς και επίμονης εφαρμογής της ορολογίας και την παροχή συνοδευτικών ορισμών.
2. Να αναπτυχθεί ένα σύνολο εκφράσεων, το οποίο χρησιμοποιείται στο αναλυτικό κείμενο για να περιγράψει τις αλλαγές σε ετήσιους και πολυετή ρυθμούς μεγέθυνσης και να εφαρμόζονται αυτές οι εκφράσεις συνεχώς και στην ίδια σειρά δημοσιευμάτων (π.χ. ενημερωτικά δελτία).

4.4 Για λόγους, αναλογίες, ποσοστά και ρυθμούς

1. Απαιτείται ακρίβεια στην παροχή πληροφορίας στον χρήστη, όσον αφορά την περίοδο στην οποία αναφέρονται, τη φύση του πληθυσμού που περιγράφεται και τον τύπο του φαινομένου που μετράται. Συγκεκριμένα: Οι όροι “ρυθμός”, “ποσοστό”, ή “λόγος” πρέπει να περιλαμβάνονται στον τίτλο, ώστε να είναι βέβαιο ότι ο χρήστης καταλαβαίνει πως τα πρωτογενή δεδομένα έχουν μετασχηματιστεί.
2. Για ρυθμούς μεταβολής όπου ο συνολικός πληθυσμός εκφράζεται σαν ενότητα, η μονάδα μέτρησης που χρησιμοποιείται για τον πληθυσμό πρέπει να περιλαμβάνεται στην επικεφαλίδα του πίνακα που αναφέρεται στον ρυθμό (π.χ. ανά 1.000 κατοίκους).
3. Τα ονόματα των στοιχείων που αποτελούν λόγο ή αναλογία πρέπει να βασίζονται όσο είναι δυνατόν πάνω στην διεθνή ορολογία, που αναφέρεται στις διεθνείς οδηγίες και προτάσεις ή γλωσσάρια διεθνών οργανισμών όπως το OECD’s Glossary of Statistical Terms, ή το Eurostat Concepts and Definitions Database (CODED), ή το UNSD’s Definitions for United Nations Common Database. Η απόκλιση από τις διεθνείς έννοιες, θα πρέπει να καταγράφεται στα μεταδεδομένα, που συνοδεύουν τον ρυθμό/ αναλογία.
4. Πρέπει να παρέχεται λεπτομερής πληροφορία για τις πηγές των στοιχείων που χρησιμοποιούνται στον υπολογισμό του λόγου ή του ρυθμού μεταβολής. Η

ελάχιστη πληροφορία πρέπει να περιλαμβάνει: τύπος της πηγής των δεδομένων (δημόσια διοίκηση, έρευνα ή απογραφή νοικοκυριών, έρευνα ή απογραφή εταιριών), περίοδος αναφοράς, πλήρης επίσημος τίτλος της σειράς στοιχείων, πλήρης όνομα του ιδρύματος ή της υπηρεσίας που παρέιχε τις πηγές.

5. Οι χρήστες θα πρέπει να έχουν πρόσβαση στα πρωτογενή δεδομένα που χρησιμοποιούνται για τον υπολογισμό του λόγου ή του ρυθμού μεταβολής. Αυτό μπορεί να γίνεται στην δημοσίευση όπου υπολογίζονται οι αντίστοιχοι δείκτες (π.χ. στους πίνακες του παραρτήματος) ή μπορεί να παρέχεται με παραπομπές και υπερσυνδέσμους στα πρωτογενή δεδομένα.

6. Στους χρήστες θα πρέπει να παρέχεται πληροφορία σχετικά με τις μεθοδολογίες που χρησιμοποιούνται. Η ποιότητα των ρυθμών/ λόγων που δημοσιεύονται, είναι καλή μόνο αν η ποιότητα των στοιχείων που χρησιμοποιούνται στον υπολογισμό και τα μεταδεδομένα επιτρέπει τους χρήστες να διαμορφώσουν μια κατανόηση για την ανάγκη-σκοπιμότητα της χρήσης των ρυθμών/ λόγων.

4.5 Προτεινόμενες πρακτικές για την αναφορά διαφορετικών τύπων «δεδομένων με χρονική σειρά» (time series data).

- **Μορφές των δεδομένων με χρονική σειρά (forms of time series) που παρουσιάζονται.**

Πρόταση 1: Οι δείκτες που αναφέρονται σε χρονική περίοδο μικρότερη του ενός έτους, θα πρέπει να είναι προσαρμοσμένοι εποχικά (seasonally adjusted). Το επίπεδο λεπτομέρειας στο οποίο θα προσαρμοστούν οι δείκτες θα πρέπει να επιλεγεί λαμβάνοντας ως κριτήρια την ζήτηση των χρηστών και το κόστος-αποτελεσματικότητα.

Πρόταση 2: Όποτε είναι δυνατό, θα πρέπει η εστίαση στα δημοσιεύματα που αφορά τέτοιους δείκτες, να έχει τη σωστή εποχιακή προσαρμογή. Θα πρέπει να παρέχεται στους χρήστες πρόσβαση στις σειρές των πρωτογενών δεδομένων, είτε στην δημοσίευση αν το επιτρέπει ο χώρος, είτε με παραπομπή. Όπου υπάρχει ζήτηση του χρήστη, η υπηρεσία μπορεί επίσης να παρέχει ενδιάμεσα στοιχεία της διαδικασίας με την οποία γίνεται η εποχιακή προσαρμογή. (π.χ. σειρές που προσαρμόστηκαν λόγω ημερολογιακών επιπτώσεων), αλλά θα πρέπει να είναι εμφανές ότι η εστίαση είναι στα εποχικά προσαρμοσμένα συμπεράσματα, όταν αυτό που προσελκύει το ενδιαφέρον είναι οι βραχυπρόθεσμες μεταβολές.

- **Μετασχηματισμοί ανάλυσης**

Πρόταση 3: Τα δημοσιεύματα τα οποία παρουσιάζουν εποχικά προσαρμοσμένες σειρές δεδομένων θα πρέπει κατ' ελάχιστο να παρέχουν ρυθμούς αύξησης για την τελευταία περίοδο και εάν το επιτρέπει ο χώρος σε διαφορετικά επίπεδα

(π.χ. μήνας με προηγούμενο μήνα, δεκαπενθήμερο με προηγούμενο δεκαπενθήμερο κτλ).

Πρόταση 4: Για μηνιαίους και δεκαπενθήμερους ρυθμούς μεταβολής, τα εποχιακά προσαρμοσμένα δεδομένα είναι ο καλύτερος τρόπος να παρουσιάσουμε πληροφορίες για μια σειρά δεδομένων, καθώς και για να παρουσιάσουμε βραχυπρόθεσμες αλλαγές.

Πρόταση 5: Για ρυθμό αλλαγών που αναφέρεται στην ίδια περίοδο προηγούμενων ετών, οι ετήσιες αλλαγές θα πρέπει να εφαρμόζονται σε δεδομένα που έχουν προσαρμοστεί σε ημερολογιακές διαφορές.

Πρόταση 6: Λόγω του ρίσκου να υπάρχουν παραπλανητικές ετικέτες-επικεφαλίδες, ιδιαίτερα όπου οι σειρές δεδομένων παρουσιάζουν μεγάλη αστάθεια³, δεν συνιστάται η παρουσίαση των αναγόμενων σε έτος αλλαγών, ιδιαίτερα ως επικεφαλίδα των σειρών. Όπου χρησιμοποιούνται αλλαγές ανηγμένες σε έτος, θα πρέπει στους χρήστες να παρέχεται η πληροφορία σχετικά με την πιθανότητα να παρερμηνευτούν τα ονόματα στις επικεφαλίδες των στηλών λόγω αστάθειας των σειρών.

Πρόταση 7: Οι ρυθμοί μεγέθυνσης που έχουν αναχθεί ανά έτος, δεν συνιστάται να χρησιμοποιούνται για την παρουσίαση τριμηνιαίων ή μηνιαίων αποτελεσμάτων.

Πληροφορίες που πρέπει να παρέχονται στους χρήστες σχετικά με την εποχική προσαρμογή.

Πρόταση 8: Οι στατιστικές υπηρεσίες θα πρέπει να παρέχουν μία μη τεχνική εξήγηση της εποχικής προσαρμογής των δεδομένων με τρόπο που θα συλληφθεί από το ευρύ κοινό και επίσης να αναφέρουν τα πλεονεκτήματα της εποχιακής προσαρμογής.

Πρόταση 9: Προς όφελος των χρηστών οι οποίοι ενδιαφέρονται για πληροφορία σχετικά με την εγκυρότητα της εποχιακής προσαρμογής που εφαρμόστηκε, θα πρέπει να παρέχεται μία ελάχιστη ποσότητα πληροφορίας που θα επιτρέπει μια αξιολόγηση της αξιοπιστίας κάθε εποχιακά προσαρμοσμένης σειράς.

Πρόταση 10: Θα πρέπει να διατηρούνται αρκετά μεταδεδομένα σχετικά με την εποχιακή προσαρμογή, ώστε να επιτρέπουν τους εξωτερικούς χρήστες να προσαρμόσουν εποχικά με παρόμοιο τρόπο άλλες σειρές δεδομένων με τον ίδιο τρόπο.

³ Σαν αστάθεια των σειρών εννοείται ότι σε μία σειρά δίνονται διαφορετικού τύπου πληροφορίες. Π.χ. μια σειρά του excel, η οποία περιέχει και τιμές που αφορούν το ΑΕΠ, και τιμές που αφορούν το σταθμισμένο ΑΕΠ το 1995.

5. Προτεινόμενες πρακτικές για την αναφορά και διάδοση των μεταδεδομένων

5.1 Οι ανάγκες των μεταδεδομένων

Όλες οι υπηρεσίες παροχής στοιχείων θα πρέπει:

1. Να συντάσσουν τα μεταδεδομένα που απαιτούνται ώστε οι χρήστες να καταλάβουν τα πλεονεκτήματα και τους περιορισμούς των στατιστικών στοιχείων που παρέχονται.
2. Να κρατάνε τα μεταδεδομένα τους ενήμερα, ενσωματώνοντας τις τελευταίες αλλαγές σε ορισμούς, ταξινομήσεις, μεθοδολογία κ.τ.λ.

5.2 Η πρόσβαση στα μεταδεδομένα

Οι συστάσεις-κλειδιά για αυτό το πεδίο περιλαμβάνουν:

1. Την επιβεβαίωση ότι οι χρήστες έχουν πρόσβαση σε μεταδεδομένα μέσω της παροχής τους με ένα ευρύ φάσμα διαφορετικών μέσων - δημοσιεύσεις συγγραμμάτων, CD-ROMs, κ.α. Παρόλα αυτά είναι σημαντικό όλα τα μεταδεδομένα να είναι διαθέσιμα στους χρήστες στο διαδίκτυο, δεδομένου ότι αποτελεί το πιο προσβάσιμο μέσο για την απόκτηση επικαιροποιημένων στοιχείων. Η πρακτική που συνιστάται για τα μεταδεδομένα είναι να δομούνται με τέτοιο τρόπο ώστε να ικανοποιούν τις ανάγκες ενός εύρους χρηστών με διαφορετικές απαιτήσεις και εξειδίκευση. Αυτό δεν σημαίνει απαραίτητα την παρουσίαση διαφορετικών μεταδεδομένων για κάθε κατηγορία χρηστών. Ωστόσο συνιστάται μία κλιμακωτή παρουσίαση των μεταδεδομένων, που να εξελίσσεται από περίληψη των μεταδεδομένων σε πιο λεπτομερείς πληροφορίες. Κάθε επίπεδο πληροφόρησης θα πρέπει να χρησιμοποιεί ακριβή και σαφές κείμενο.
2. Η διάδοση των μεταδεδομένων πρέπει να είναι χωρίς χρέωση στο διαδίκτυο. Υπάρχει ισχυρή υποστήριξη προς την άποψη ότι τα μεταδεδομένα που περιγράφουν στατιστικά αποτελέσματα είναι κατά σημαντικό βαθμό δημόσιο αγαθό και κατά συνέπεια θα πρέπει να διαδίδεται δωρεάν στο διαδίκτυο ακόμα και εάν τα κοινωνικά και οικονομικά στοιχεία που περιγράφουν υπόκεινται στο καθεστώς τιμών-κοστολόγησης της οργάνωσης που τα δημοσιεύει.
3. Ενεργός δεσμός των μεταδεδομένων με τους πίνακες και τα γραφήματα που περιγράφουν και το αντίστροφο.
4. Διαθεσιμότητα των μεταδεδομένων όχι μόνο στην εθνική γλώσσα, αλλά όπου οι πηγές το επιτρέπουν και στην αγγλική.
5. Δόμηση των μεταδεδομένων σε διαφορετικές στατιστικές ενότητες με βάση κάποια ιεραρχική κατάταξη. Θα μπορούσε να ληφθεί υπόψη η υιοθέτηση του

UNECE Classification of International Statistical Activities ως διεθνές πρότυπο για τα μεταδεδομένα⁴.

6. Η παροχή μιας τοπικής μηχανής αναζήτησης βασισμένη στη δωρεάν αναζήτηση κειμένου.

7. Η υλοποίηση της προτεινόμενης πρακτικής, για την επιβεβαίωση της σταθερότητας των URLs (Uniform Resource Locators) ή την παροχή δεσμών μεταξύ παλιών και νέων URL, που θα μεταφέρουν αυτόματα τους χρήστες στην καινούρια ηλεκτρονική διεύθυνση. Αυτό είναι ένα ζήτημα-κλειδί, δεδομένης της σημασίας των δεσμών μεταξύ των ιστοτόπων.

8. Η παροχή ονομάτων των προσώπων για επικοινωνία ή ηλεκτρονικών διευθύνσεων, όπου είναι δυνατό να αποκτηθεί περαιτέρω πληροφορία για τις έννοιες, τους ορισμούς, τις στατιστικές μεθοδολογίες. Σε ορισμένους οργανισμούς, αυτή η «επικοινωνία» μπορεί να είναι ένα γενικό - εταιρικό σημείο επαφής, ή σχετική υπηρεσία για όλες τις επιθυμίες των χρηστών-πελατών.

5.3 Υιοθέτηση ενός κοινού συνόλου ορολογίας για την παρουσίαση των μεταδεδομένων.

Σημαντικοί πόροι δαπανώνται συχνά από διεθνείς οργανισμούς, προκειμένου να σιγουρέψουν ότι οι μεθοδολογικές περιγραφές των στατιστικών υπηρεσιών των διαφόρων χωρών είναι όσο το δυνατό συνεπής. Ένας χρήσιμος μηχανισμός θα ήταν η αυστηρή χρήση της ορολογίας, συναρμοσμένη στις διάφορες διεθνείς στατιστικές οδηγίες και προδιαγραφές. Αυτό θα μπορούσε να διευκολυνθεί από τη χρήση γλωσσάριων που δημοσιεύονται από τους διάφορους διεθνείς οργανισμούς.

5.4 Ευδιάκριτη παρουσίαση παρόμοιων στατιστικών δεδομένων.

Σε αυτό το πεδίο υπάρχουν πέντε ευρείες προτάσεις:

1. Διαφορετικές σειρές δεδομένων οι οποίες μοιάζουν θα πρέπει να έχουν διαφορετικούς τίτλους, ώστε να διευκολύνεται ο καθαρός διαχωρισμός τους από τους χρήστες.

2. Οι διεθνείς οργανισμοί που διαδίδουν εθνικά στοιχεία θα πρέπει να είναι ενήμεροι και να δηλώνουν σαφώς στα μεταδεδομένα εάν τα στοιχεία που διαδίδουν και προέρχονται από εθνικές πηγές διαδίδονται επίσης και στην χώρα προέλευσης ή συντάσσονται/ μετασχηματίζονται με άλλο τρόπο προκειμένου να ανταποκρίνονται στις απαιτήσεις των διεθνών οργανισμών.

3. Οι διεθνείς οργανισμοί θα πρέπει να περιγράφουν ξεκάθαρα στα μεταδεδομένα τους ειδικές λεπτομέρειες κάθε μετασχηματισμού των εθνικών

⁴ Η έκδοση είναι διαθέσιμη στην σελίδα:

<http://www.uncece.org/stats/documents/ece/ces/bur/2005/5.e.pdf>

δεδομένων που πραγματοποιούν προκειμένου να κάνουν τα στοιχεία πιο συγκρίσιμα σε διεθνές επίπεδο. Πρέπει να δηλώνεται ξεκάθαρα ο τυχόν μετασχηματισμός των δεδομένων που μετασχηματίζονται από τους διεθνείς οργανισμούς, ιδιαίτερα (αλλά όχι μόνο) όπου δημοσιεύονται μαζί με διαφορετικά εθνικά στοιχεία. Τα δύο σύνολα στοιχείων πρέπει να είναι ξεκάθαρα διαφοροποιημένα στο μυαλό του χρήστη.

4. Το ακριβές όνομα της ταξινόμησης που χρησιμοποιείται στα στοιχεία που διαδίδονται από τις εθνικές υπηρεσίες και τους διεθνείς οργανισμούς (ιδιαίτερα όταν μετασχηματίζονται σε μια διεθνή ταξινόμηση για να βελτιωθεί η συγκρισιμότητα σε υπερεθνικό επίπεδο) πρέπει πάντα να δηλώνεται ξεκάθαρα.

5. Όταν ένα πεδίο δραστηριότητας ή δεδομένων καλύπτεται μερικά, πρέπει να δηλώνεται ξεκάθαρα, για παράδειγμα με έναν αστερίσκο ή με μια υποσημείωση.

6. Προτεινόμενες πρακτικές για την αναθεώρηση δεδομένων

6.1 Οι διαβουλεύσεις με τους χρήστες για τις πρακτικές αναθεώρησης

Στο πρώτο στάδιο της διαμόρφωσης της πρακτικής αναθεώρησης είναι σημαντικό να υπάρχει η συμβολή των κύριων χρηστών, για να είναι βέβαιο ποιες είναι οι ιδιαίτερες ανάγκες και οι προτεραιότητες. Θα μπορούσαν για παράδειγμα να συλλεχθούν οι απόψεις των χρηστών σχετικά με τις ιδιαίτερες ανάγκες για την επικαιρότητα των στοιχείων, τα προβλήματα που συνάντησαν λόγω των αναθεωρήσεων κ.α.)

6.2 Παροχή καθαρών, εύκολα προσβάσιμων και σύντομων δηλώσεων προς τους χρήστες σχετικά με το πότε και το γιατί γίνονται αναθεωρήσεις.

6.3 Η δημιουργία, όπου είναι δυνατόν, ενός σταθερού κύκλου ανανέωσης.

6.4 Η εξισορόπηση της διαχρονικής σταθερότητας των σειρών των δεδομένων έναντι στην ανάγκη να αναθεωρηθούν οι σειρές λόγω της εισαγωγής νέων μεθοδολογιών, ιδεών, εννοιών, ταξινομήσεων κτλ.

6.5 Η εξασφάλιση ότι οι αναθεωρήσεις γίνονται και για τα προηγούμενα έτη, προκειμένου να δώσουν συνεπή διαχρονικά δεδομένα.

Για να διατηρηθεί ένα ικανοποιητικό επίπεδο εξυπηρέτησης μετά από μεγάλης έκτασης αναθεωρήσεις, τα δεδομένα θα πρέπει να αναθεωρούνται προς τα προηγούμενα έτη λαμβάνοντας υπόψη και ισορροπώντας τις ανάγκες των χρηστών, το κόστος και την διαθεσιμότητα των πηγών. Οι σειρές των αναθεωρημένων στοιχείων, θα πρέπει να δημοσιεύονται ταυτόχρονα με τα αναθεωρημένα σύγχρονα στοιχεία ή σύντομα μετά από αυτά, και κατά προτίμηση με εύκολα προσβάσιμη μορφή στο διαδίκτυο. Οι αναθεωρημένες

σειρές θα πρέπει να έχουν αρκετή λεπτομέρεια ώστε οι χρήστες να μπορούν να εντοπίσουν τις πηγές προέλευσης στις οποίες έγινε η αλλαγή.

6.6 Η παροχή κατάλληλης τεκμηρίωσης στις αναθεωρήσεις.

Η τεκμηρίωση αυτή πρέπει να περιλαμβάνει:

1. Ξεκάθαρο προσδιορισμό των πρωταρχικών ή προσωρινών και αναθεωρημένων δεδομένων.
2. Ειδοποίηση για την εξέλιξη των σημαντικών αλλαγών στις έννοιες, τους ορισμούς, τις ταξινομήσεις και τις στατιστικές μεθόδους.
3. Τις πληροφορίες για τις πηγές αναθεώρησης, οι οποίες πρέπει να εξηγούνται όταν δημοσιεύονται οι αναθεωρημένες σειρές.
4. Πληροφορίες για τα κενά των σειρών δεδομένων, όταν συγκεκριμένες από αυτές δεν γίνεται να κατασκευαστούν πλήρως.
5. Πληροφορίες όσον αφορά την έκταση των πιθανών μελλοντικών αναθεωρήσεων, βασισμένες στην μέχρι τώρα ιστορία.

6.7 Οι υπηρεσίες πρέπει να υπενθυμίζουν τους χρήστες για το μέγεθος των πιθανών αναθεωρήσεων με βάση αυτά που έγιναν στο παρελθόν.

Στους χρήστες θα πρέπει να παρέχεται αρκετή πληροφορία ώστε να μπορούν να διαμορφώσουν άποψη για την αξιοπιστία και την ακρίβεια των πρωταρχικών ή προσωρινών δεδομένων.

6.8 Διαφανής και έγκαιρη υποβολή έκθεσης των λαθών.

6.9 Προτεινόμενες πρακτικές για την παρουσίαση σπασμένων⁵ σειρών.

1. Η υπηρεσία παροχής στοιχείων θα πρέπει να λάβει όλα τα δυνατά μέτρα για να αποφύγει και να ελαχιστοποιήσει τυχόν αλλαγές στους ορισμούς και στις ταξινομήσεις που χρησιμοποιούνται για τη συλλογή ή σύνταξη των στοιχείων. Θα πρέπει να αναπτύσσονται μεθοδολογίες με σκοπό τη μείωση της συχνότητας των αναθεωρήσεων.

Παρόλα αυτά, έρχεται μια στιγμή που οι χρονικές σειρές των δεδομένων μπορεί να αναδιαρθρώνονται, ακόμα κι όταν ξεπερασμένες ταξινομήσεις, έννοιες και ερωτηματολόγια διατηρούνται. Σε τέτοιες περιπτώσεις, ένα ολοκληρωτικό σπάσιμο-κενό στις σειρές μπορεί να προτιμάται από το να συνεχίζεται η συλλογή δεδομένων με βάση ξεπερασμένες ταξινομήσεις και έννοιες που δεν προσεγγίζουν την πραγματικότητα.

2. Όπου τα σπασίματα στις χρονικές σειρές είναι αναπόφευκτα, οι χρήστες θα πρέπει να ειδοποιούνται επαρκώς και εκ των προτέρων, δηλαδή αρκετά πριν

⁵ Δηλαδή σειρών με κενά ή λάθη.

το σπάσιμο των σειρών, με έμφαση στον χρόνο του σπασίματος και με λεπτομερή εξήγηση για τους λόγους των αλλαγών. “Εκ των προτέρων” σημαίνει όχι απλά πριν την εφαρμογή του σπασίματος, αλλά αρκετά πριν, ώστε οι χρήστες να έχουν τη δυνατότητα να εφαρμόσουν τροποποιήσεις στα συστήματά τους, τα προγράμματα και τις βάσεις δεδομένων τους και να ζητήσουν επιπλέον διευκρινίσεις αν είναι αναγκαίο. Μία κοινή πρακτική που χρησιμοποιείται από πολλές υπηρεσίες, είναι η έκδοση ενός λεπτομερούς discussion paper πολλούς μήνες πριν τις αλλαγές.

3. Οι αλλαγές των σειρών θα πρέπει να δηλώνονται ξεκάθαρα και στους πίνακες των στοιχείων και στα τυχόν συνοδευτικά διαγράμματα. Μια ποικιλία μεθόδων χρησιμοποιούνται από τις εθνικές υπηρεσίες και τους διεθνείς οργανισμούς για να δοθεί έμφαση στους πίνακες όπου έχει γίνει σπάσιμο σειρών. Σε αυτές συμπεριλαμβάνονται η εισαγωγή μιας γραμμής στον πίνακα στο σπασμένο σημείο, συνοδευόμενη με μία υποσημείωση, ή παρουσίαση σε μορφή πίνακα ως νέα σειρά. Όποια μέθοδος και να υιοθετείται, το κύριο ζήτημα είναι το κενό-σπάσιμο να είναι εμφανές στους χρήστες.

Θα πρέπει επίσης να παρέχονται οι ακόλουθες πληροφορίες, όπως πηγάζουν από τις οδηγίες της Eurostat:

- α. Η αναφορά στην περίοδο της έρευνας όπου εμφανίστηκε το σπάσιμο-κενό.
- β. Εάν η διαφορά που αναφέρεται έχει περιορισμένες επιπτώσεις στις χρονικές σειρές και εάν η αλλαγή οδήγησε σε εναρμόνιση σύμφωνα με κάποια πρότυπα/σταθερότυπα.
- γ. Ο ακριβής υπογραμμισμός των διαφορών σε έννοιες και μεθόδους μετρήσεων πριν και μετά το σπάσιμο των σειρών.
- δ. Η περιγραφή των αιτιών που έφεραν την αλλαγή (π.χ. αλλαγές στην ταξινόμηση, στη στατιστική μεθοδολογία, τον στατιστικό πληθυσμό, μέθοδοι μετασχηματισμού δεδομένων, έννοιες, διοικητικές διαδικασίες, διαδικασίες σχετικά με στατιστικά στοιχεία από διοικητικές πηγές).
- ε. Μια αξιολόγηση του μεγέθους της επίδρασης της αλλαγής, όπου είναι δυνατόν, με ένα ποσοτικό μέτρο. Πρέπει επίσης να παρέχονται δεσμοί και αναφορές σε πιο λεπτομερείς πληροφορίες.

4. Τα σημεία σε γραφήματα που έχουν γραμμές, δεν πρέπει να συνδέονται όταν έχουν ασυνέχεια στα δεδομένα. Ο λόγος του σπασίματος της σειράς θα πρέπει να εξηγείται σε υποσημείωση που συνοδεύει το γράφημα με κατάλληλους δεσμούς ή αναφορές σε πιο λεπτομερείς εξηγήσεις των αιτιών του σπασίματος.

5. Όταν εισάγονται μεθοδολογικές αλλαγές, γίνεται μια απόπειρα να αναθεωρηθούν οι ιστορικές σειρές όσο το δυνατόν προς τα πίσω όσο το επιτρέπουν οι πηγές. Ιδανικά, αυτή η αναθεώρηση προς τα πίσω πρέπει να

επεκταθεί κατά 2-3 χρόνια ώστε να δώσει μια αντανάκλαση της νέας μεθοδολογίας.

7. Παρουσίαση πληροφορίας για δειγματοληπτικά και μη δειγματοληπτικά λάθη

7.1 Δειγματοληπτικά λάθη

Για να διασφαλιστεί η διαφάνεια και να βεβαιωθεί η σωστή χρήση των δεδομένων, τα στατιστικά στοιχεία που προέρχονται από δειγματοληπτικές έρευνες θα πρέπει να συνοδεύονται από πληροφορία σχετικά με τα δειγματοληπτικά λάθη. Τέτοια πληροφορία θα πρέπει να παρέχετε σε όλα τα μέσα διάδοσης - βάσεις δεδομένων στο διαδίκτυο, ιστοσελίδες, άλλα ηλεκτρονικά προϊόντα, papers και εκδόσεις. Είναι επίσης σημαντικό για την πληροφορία να εκφράζεται σε μη τεχνικούς όρους, που είναι δυνατόν να γίνουν κατανοητοί από μη εξειδικευμένους χρήστες.

Ο τρόπος παρουσίασης και ποσότητα της λεπτομέρειας που παρέχεται, θα πρέπει να ανταποκρίνεται στις ιδιαίτερες ανάγκες ξεχωριστών κατηγοριών χρηστών. Θα πρέπει να παρέχεται η ακόλουθη πληροφορία σε συνδυασμό με εμφανείς δεσμούς σε τεχνικές υποσημειώσεις:

1. Ειδοποίηση των χρηστών ότι τα δεδομένα προέρχονται από τυχαίο ή μη τυχαίο δείγμα. Αν ισχύει το τελευταίο τότε οι επιπτώσεις του συμπεράσματος πρέπει να δηλωθούν σαφώς.
 2. Το δειγματοληπτικό λάθος θα πρέπει να δηλώνετε ως μια πηγή λάθους που πρέπει να εξηγείται και να αντιλαμβάνεται από τους χρήστες δεδομένων, μέσα από έναν σύντομο ορισμό του. Π.χ. έντονη προειδοποίηση για την αναξιοπιστία των δεδομένων με σημαντικά δειγματοληπτικά λάθη.
 3. Τα δειγματοληπτικά λάθη θα πρέπει να παρουσιάζονται στο συνολικό περιεχόμενο των λαθών της έρευνας. Σε αυτό, θα πρέπει οι χρήστες να αντιλαμβάνονται ότι το δειγματοληπτικό λάθος είναι συχνά όχι ο πιο σημαντικός παράγοντας που διαμορφώνει το συνολικό λάθος.
 4. Εάν σε μια έκθεση χρησιμοποιούνται στατιστικοί έλεγχοι, το επίπεδο σημαντικότητάς τους θα πρέπει να δηλώνεται ρητά.
 5. Τα δειγματοληπτικά λάθη για τους υπολογισμούς-κλειδιά θα πρέπει να είναι διαθέσιμα στον χρήστη είτε σε μορφή πίνακα στην δημοσίευση, είτε ως δεσμοί στο διαδίκτυο. Κάποια μορφή σημείωσης πρέπει επίσης να τοποθετηθεί αμέσως δίπλα στις εκτιμήσεις οι οποίες έχουν πολύ υψηλό δειγματοληπτικό ή μη λάθος.
- Τα δειγματοληπτικά λάθη θα πρέπει να παρουσιάζονται με έναν αριθμό διαφορετικών μορφών, για παράδειγμα:

- Ως απόλυτες τιμές του λάθους (standard error ή se);

- Ως σχετικές τιμές δηλαδή (standard error) / (εκτίμηση = estimate) = (rse)
- Υπό μορφή διαστημάτων πιθανότητας ή εμπιστοσύνης.

Η προτιμώμενη μορφή εξαρτάται από τη φύση της εκτίμησης και τους χρήστες. Στη δημοσίευση των Ηνωμένων Εθνών, *Sampling Errors in Household Surveys* (UNSD 1993) αναφέρεται μια λεπτομερή ανάλυση των διάφορων μορφών παρουσίασης και αρκετά παραδείγματα της προτιμώμενης πρακτικής. Αυτό που έχει σημασία είναι η επιλεγόμενη μέθοδος να περιγράφεται ξεκάθαρα και μη αμφιλεγόμενα και να παρουσιάζεται με συνοδευτικούς ορισμούς και σημειώσεις. Προκειμένου να είναι βέβαιη η συνεπής διάδοση αυτής της πληροφορίας, ορισμένες στατιστικές υπηρεσίες χρησιμοποιούν ένα standard σύνολο λέξεων που πρέπει να περιλαμβάνεται σε όλες τις σχετικές δημοσιεύσεις.

Όπου τα προβλήματα χώρου αποκλείουν την εισαγωγή λεπτομερών πληροφοριών, θα πρέπει να παρέχονται είτε παραπομπές είτε δεσμοί σε πιο λεπτομερείς τεχνικές εκθέσεις και εγχειρίδια. Αυτές οι πληροφορίες θα πρέπει να επιτρέπουν τους εξειδικευμένους χρήστες να αναλύσουν τα λεπτομερή στοιχεία ή να συνθέσουν νέες πινακοποιήσεις και να μπορούν:

- α. Να προσδιορίσουν την ιδιαίτερη μέθοδο που χρησιμοποιήθηκε για τον υπολογισμό του δειγματοληπτικού λάθους.
- β. Να παρέχουν υπολογισμούς του δειγματοληπτικού λάθους για διαφορετικούς τύπους εκτιμήσεων (π.χ. επίπεδα, ποσοστά, λόγοι, κινήσεις, μέσες και ενδιάμεσες τιμές) για έναν αριθμό μεταβλητών.

7.2 Μη δειγματοληπτικά λάθη

1. Όλα τα στατιστικά αποτελέσματα που διαδίδονται και έχουν πληροφορία για τα δειγματοληπτικά λάθη, θα πρέπει να συνοδεύονται και με πληροφορία για τα μη δειγματοληπτικά λάθη. Τέτοια πληροφορία πρέπει να είναι προσβάσιμη σε όλα τα μέσα διάδοσης των στοιχείων. Είναι σημαντικό να εκφράζεται σε μη τεχνικούς όρους, ώστε να είναι κατανοητή από τον μη εξειδικευμένο χρήστη.

2. Όπου είναι δυνατόν, θα πρέπει να παρέχονται ποσοτικές μετρήσεις των μη δειγματοληπτικών λαθών. Λόγω της δυσκολίας στην ποσοτικοποίηση ορισμένων από αυτά, οι υπηρεσίες θα πρέπει να παρέχουν ένα μείγμα ποσοτικών και ποιοτικών πληροφοριών, που επιτρέπει τον μη ειδικευμένο χρήστη να καταλάβει ξεκάθαρα τα πλεονεκτήματα και τους περιορισμούς των δεδομένων. Ιδιαίτερα, οι πληροφορίες στα μη δειγματοληπτικά λάθη θα πρέπει να μεταδίδουν στον χρήστη το γεγονός ότι τέτοια λάθη, είτε μεμονωμένα είτε συνολικά, μπορεί να έχουν μεγαλύτερη επίδραση στην αξιοπιστία των στοιχείων από ότι τα δειγματοληπτικά λάθη, και ότι η ποσοτική μέτρηση των δειγματοληπτικών λαθών δεν είναι απαραίτητα ένδειξη της σχετικής σημασίας τους.

3. Το σημαντικό είναι να αναπτυχθεί μία κουλτούρα κριτικής αξιολόγησης των στατιστικών αποτελεσμάτων και των βασικών πλεονεκτημάτων και αδυναμιών που τεκμηριώνονται και που διαδίδονται.

8. Προτεινόμενες πρακτικές για τις αλλαγές βάσεων

Προτείνεται η αλλαγή βάσης να διεξάγεται κάθε πέντε χρόνια και μέσα σε τρία χρόνια από το τέλος του έτους-βάση. Προτείνεται το έτος βάση να λήγει σε 5 ή 0, εκτός κι αν το έτος είναι ασυνήθιστο. Προκειμένου να παρέχεται αρκετή διαφάνεια στους χρήστες, σχετικά με μια αλλαγή βάσης, κάθε στοιχεία πρέπει να συνοδεύονται από τα ακόλουθα μεταδεδομένα:

1. Τη μεθοδολογική προσέγγιση που ακολουθήθηκε για την αλλαγή βάσης και ιδιαίτερα τις διαδικασίες που έγιναν κατά την αλλαγή βάσης π.χ. απλή αλλαγή αναφορών, εισαγωγή νέων βαρών κτλ.
2. Τη χρονιά δεσμού.
3. Το επίπεδο ταξινόμησης στο οποίο οι δείκτες ανάγονται ξανά σε βάσεις και διαδίδονται.
4. Η πολιτική στρογγυλοποίησης που ακολουθήθηκε στην αλλαγή βάσεων, ακόμα κι αν η στρογγυλοποίηση πρέπει μόνο να πραγματοποιηθεί στο έσχατο στάδιο για λόγους παρουσίασης.
5. Ένας πίνακας μετάβασης από το παλιό στο νέο σύστημα ταξινόμησης, εάν εισάγεται νέο.
6. Η περιγραφή κάθε νέου συστήματος στάθμισης και των επιπτώσεών του.

Όταν υιοθετείται η άμεση-ευθεία προσέγγιση, μία υποσημείωση είναι χρήσιμη για να ειδοποιήσει τους χρήστες ότι κάθε συνάθροιση δεικτών με αναθεωρημένη βάση χαράζεται την ενημέρωση των βαφών των προηγούμενων βάσεων.

9. Προτεινόμενες πρακτικές για τον τρόπο παραπομπής

9.1 Παραπομπή συνόλου δεδομένων

Οι υπηρεσίες παροχής στοιχείων θα πρέπει να:

1. Να διαμορφώσουν και να τοποθετήσουν την πολιτική παραπομπών σε μια εμφανή θέση στις ιστοσελίδες τους. Επιπλέον, αυτή η πολιτική θα πρέπει να συνοδεύεται με λεπτομερή δείγματα παραπομπών. Αυτό βοηθάει τους χρήστες να περιλαμβάνουν τη σωστή πληροφορία στις δικές τους παραπομπές. Παρακάτω φαίνεται το παράδειγμα παραπομπής της Στατιστικής Υπηρεσίας του Καναδά στο διαδίκτυο. Η οδηγία παραπομπής παρέχεται στο κάτω μέρος της σελίδας, με διαμόρφωση κατάλληλη για εκτύπωση.

Statistics Canada. 2002. 2001 Community Profiles, Released June 27, 2002. Last modified: 2005-11-30. Statistics Canada Catalogue no. 93F0053XIE.

2. Δεύτερον να προωθήσουν μία κουλτούρα παραπομπών. Αυτό μπορεί να ενισχυθεί επικοινωνώντας με όλους τους γνωστούς χρήστες των δεδομένων, όλους τους εκδότες που χρησιμοποιούν τα δεδομένα του οργανισμού κτλ, ζητώντας τους να ακολουθήσουν την πολιτική παραπομπής για τον οργανισμό στις μελλοντικές εκδόσεις.

Ένα απλό και αποτελεσματικό στυλ παραπομπής για τα σύνολα δεδομένων θα πρέπει να περιλαμβάνει τα παρακάτω στοιχεία:

- α. Αδιαμφισβήτητο όνομα του συνόλου δεδομένων.
- β. Συγγραφέας του συνόλου δεδομένων.
- γ. Υπηρεσία ή τμήμα της υπηρεσίας που είναι υπεύθυνο για την βάση δεδομένων.
- δ. Ημερομηνία του συνόλου δεδομένων (ή αριθμός έκδοσής του).
- ε. Λεπτομέρειες επικοινωνίας για ερωτήματα.
- στ. Διεύθυνση του αρχείου ή άλλος χώρος αποθήκευσης ή συστήματος πρόσβασης στα δεδομένα.
- ζ. Εκδότης (αν είναι διαφορετικό πρόσωπο από τον συγγραφέα).
- η. Αν είναι εφικτό και κατάλληλο, η παράγραφος, ο πίνακας ή ο αριθμός σελίδας.

Αυτό το στυλ παραπομπής θα πρέπει να ακολουθείται για όλα τα δεδομένα που δημοσιεύονται εσωτερικά ή εξωτερικά, όπως και για την τεκμηρίωση κάθε συνόλου δεδομένων που δημιουργούνται ή μορφοποιούνται. Η πραγματική σειρά των στοιχείων που παρατίθενται παραπάνω (π.χ. στίξη, χρήση πλάγιας γραμματοσειράς κτλ), είναι ζήτημα προσωπικής επιλογής ή επιλογής του οργανισμού.

9.2 Παραπομπή σε κείμενο

Η κύρια σύσταση για την παραπομπή σε κείμενο αφορά τη συστηματική χρήση όλων των μεταδεδομένων σε ένα από τα ευρέως αποδεκτά στυλ βιβλιογραφικής παραπομπής. Τα δύο πιο συχνά χρησιμοποιούμενα συστήματα για την παρουσίαση των παραπομπών σε κείμενο για μια βιβλιογραφία είναι το σύστημα Harvard και το Αριθμητικό (Numeric) σύστημα.

Οι παραπομπές μπορεί να αφορούν:

1. Αναφορές ή πηγές για ιδέες ή ποικίλους ορισμούς που χρησιμοποιούνται σε κάθε έκδοση (π.χ. ορισμοί που εμφανίζονται σε επεξηγηματικές σημειώσεις,

γλωσσάρια κτλ). Έτσι ο χρήστης μπορεί να προσδιορίσει εάν ένας συγκεκριμένος ορισμός προκύπτει ευθέως από τα υφιστάμενα διεθνή στατιστικά standards, εάν είναι μια μορφοποιημένη έκδοση που προσαρμόστηκε για συγκεκριμένη χρήση, ή ένας εξ ολοκλήρου νέος ορισμός.

2. Επαρκής αναφορά (παραπομπή) με πληροφορία που να δίνει τη δυνατότητα στον χρήστη να προσδιορίσει τη διαθεσιμότητα πιο λεπτομερών πληροφοριών και εννοιών, συλλογή μεθοδολογίας, κ.α.

10. Προτεινόμενες πρακτικές για την παρουσίαση και την αναφορά δεδομένων από πηγές δημοσίων υπηρεσιών.

Επειδή τα στατιστικά στοιχεία που προέρχονται από πηγές Δημόσιας Διοίκησης βασίζονται σε δεδομένα που δεν συντάχθηκαν ή παράχθηκαν για στατιστικούς σκοπούς και συχνά από άλλες μη στατιστικές υπηρεσίες, είναι ιδιαίτερα σημαντικό τέτοια δεδομένα να συνοδεύονται από τους ακόλουθους τύπους μεταδεδομένων:

1. Το όνομα της δημόσιας υπηρεσίας - πηγής των δεδομένων. Εάν περισσότερες από μία υπηρεσίες ή υπουργεία παρέχουν τις υπηρεσίες και αντιπαραβάλουν δεδομένα σε αυτές (π.χ. δεδομένα υγείας ή εκπαίδευσης παρέχονται από διάφορες υπηρεσίες σε ορισμένες χώρες) θα πρέπει να παρέχεται ιδιαίτερη πληροφορία για το αν τα δεδομένα προέρχονται από όλες τις υπηρεσίες, ή μόνο από την κύρια υπηρεσία και υπουργείο.
2. Μία λεπτομερής περιγραφή των λόγων για τους οποίους τα στατιστικά δεδομένα συντάχθηκαν και συλλέχθηκαν πρωτογενώς από την συγκεκριμένη Δημόσια Υπηρεσία.
3. Μία περίληψη των πλεονεκτημάτων και μειονεκτημάτων των δεδομένων σε όρους στατιστικής εφαρμογής. Ιδιαίτερη προσοχή πρέπει να δοθεί στην επίδραση που προκαλείται από τις διαφορές από τις διεθνείς στατιστικές οδηγίες, ιδιαίτερα στην χρήση των μη standard ταξινομήσεων και εννοιών.
4. Μία περιγραφή της επεξεργασίας ή του μετασχηματισμού που ακολουθεί την λήψη των δεδομένων από διοικητικές υπηρεσίες. Αυτή η διαδικασία μπορεί να επιχειρήσει να μειώσει ή να ελαχιστοποιήσει έμφυτες αδυναμίες στα πρωτογενή δεδομένα.
5. Περιγραφή της αξιοπιστίας των στοιχείων, με εμμονή στα διεθνή πρότυπα και τις διεθνείς συνήθειες πρακτικές και προειδοποιήσεις/ περιορισμούς στη χρήση των παρεχόμενων δεδομένων.