

Συλλογικοί κατάλογοι: επιπτώσεις και λόγοι δημιουργίας πολλαπλών εγγραφών

Ανέστης Σίτας
Αριστοτέλειο Πανεπιστήμιο Θεσσαλονίκης

Ο εντοπισμός και η αφαίρεση των διπλών (πολλαπλών) εγγραφών απασχολεί πολλές μεμονωμένες βιβλιοθήκες και δίκτυα βιβλιοθηκών. Στον κατάλογο μιας βιβλιοθήκης ή ενός συλλογικού καταλόγου, θεωρητικώς, η κάθε εγγραφή αντιπροσωπεύει μοναδικά τα τεκμήρια της συλλογής της. Η δυσκολία δεν εντοπίζεται τόσο στην εύρεση των απόλυτα διπλών εγγραφών, όσο στον εντοπισμό αυτών που αναφέρονται στο ίδιο τεκμήριο παρουσιάζοντας ταυτόχρονα κάποιες διαφορές. Σε ορισμένες περιπτώσεις οι όμοιες εγγραφές μπορούν να οριστούν με σαφήνεια. Σε άλλες πάλι, εμφανίζονται μικρές διαφορές οι οποίες μπορεί και να είναι ουσιαστικές, ανάλογα με τον ορισμό της έννοιας των διπλών εγγραφών που δίνει η συγκεκριμένη βιβλιοθήκη.

Από το συμβατικό/παραδοσιακό στο αυτοματοποιημένο περιβάλλον

Κατά το παρελθόν, ο εντοπισμός των διπλών εγγραφών ενός καταλόγου γινόταν από το προσωπικό των βιβλιοθηκών με μη αυτοματοποιημένο τρόπο. Οι βιβλιοθηκονόμοι που ταξιδεύουν τα νέα δελτία στους καταλόγους, εντόπιζαν τις διπλές εγγραφές και αποφάσιζαν για την ομοιότητά τους. Στις περιπτώσεις των βιβλιοθηκών που λειτουργούσαν αυτόνομα, χωρίς να συνεργάζονται άμεσα με άλλες βιβλιοθήκες, ο εντοπισμός των διπλών εγγραφών ήταν μια διαδικασία που μπορούσε να γίνει με μεγαλύτερη ευκολία αλλά και με μεγαλύτερη ασφάλεια. Όταν η απόφαση για την ομοιότητα δύο ή περισσότερων εγγραφών δεν ήταν δυνατό να ληφθεί με βάση μόνο τις βιβλιογραφικές περιγραφές τους, συνήθως υπήρχε η δυνατότητα προσφυγής στα τεκμήρια αυτά καθ' αυτά. Η ακεραιότητα και η λειτουργικότητα όμως του νέου αυτοματοποιημένου και συνεργατικού περιβάλλοντος, βασίζεται στην ποιότητα και τον έλεγχο των εγγραφών και των αρχείων καθιερωμένων τύπων.

Καθορισμός διπλών εγγραφών

Στο περιβάλλον των βιβλιογραφικών βάσεων, ως διπλές εγγραφές ορίζονται δύο ή περισσότερα τεκμήρια που αναπαριστούν την ίδια βιβλιογραφική οντότητα ή φέρουν το ίδιο περιεχόμενο. Στην περίπτωση του βιβλιακού υλικού για παράδειγμα, εάν οι αρχικές στοιχειοθεσίες είναι ίδιες όσον αφορά τα κύρια σημεία τους, τότε θεωρούνται πως αντιπροσωπεύουν το ίδιο βιβλιογραφικό τεκμήριο. Οι δευτερεύουσες παραλλαγές, όπως η βιβλιοδεσία και η διόρθωση των τυπογραφικών λαθών, δεν σηματοδοτούν από μόνες τους ένα νέο τεκμήριο. Οποιοσδήποτε άλλες παραλλαγές, υπαγορεύουν τη δημιουργία μιας ξεχωριστής βιβλιογραφικής εγγραφής.

Εάν η βιβλιοθήκη μπορούσε να εξασφαλίσει ένα αυστηρό ποιοτικό έλεγχο, τότε ίσως το πρόβλημα να ήταν ποσοτικά μικρότερο. Πάντα όμως θα υπάρχουν περιπτώσεις που ο ποιοτικός έλεγχος δεν θα μπορεί να αναστέλλει το φαινόμενο των διπλών εγγραφών. Δεν υπάρχουν πολλά λογισμικά, για παράδειγμα, που να απαιτούν μοναδικότητα στην μορφή καταχώρησης της ονομασίας του εκδοτικού οίκου ή να προλαμβάνουν την καταχώρηση λανθασμένων αριθμητικών δεδομένων στο πεδίο της φυσικής περιγραφής. Εάν ο ποιοτικός έλεγχος εφαρμόζεται σωστά, αυτές οι περιπτώσεις θα ελαττωθούν. Παρόλα αυτά όμως, υπάρχει μεγάλη ποικιλία διπλών

εγγραφών για τις οποίες πρέπει να αποφανθούμε εάν όντως αντιπροσωπεύουν το ίδιο τεκμήριο, ιδιαίτερα καθώς προστίθεται στη βάση νέο υλικό σε ποικίλες μορφές.

Επιπτώσεις ύπαρξης διπλών εγγραφών

Η ποιότητα των βάσεων απασχολεί ιδιαίτερα τις βιβλιοθήκες αφού μόνο έτσι μπορούν να εκπληρώσουν τον στόχο τους προσφέροντας ακριβείς πληροφορίες σε συνδυασμό με την επίτευξη όσο το δυνατόν μεγαλύτερης ταχύτητας απόκρισης. Οι διπλές εγγραφές σε μια βιβλιογραφική βάση μπορούν να προκαλέσουν προβλήματα στους διαχειριστές της βάσης, στους καταλογογράφους και στους τελικούς χρήστες. Το πρόβλημα γίνεται εντονότερο με την εμφάνιση της συνεργατικής καταλογογράφησης μέσω των αυτοματοποιημένων συστημάτων, καθώς κάθε βιβλιοθήκη-μέλος ενός συστήματος μπορεί να χρησιμοποιήσει τις εγγραφές των άλλων μελών. Οι διαχειριστές πρέπει να βελτιώσουν την ποιότητα της βάσης και να τη διατηρήσουν λειτουργική και «καθαρή».

Υπερφόρτωση πληροφοριών και μειωμένη αποτελεσματικότητα συστήματος

Η ανάκτηση πάρα πολλών εγγραφών από μια βιβλιογραφική βάση μπορεί να προκαλέσει «υπερφόρτωση πληροφοριών» (information overload). Παρουσιάζονται στον χρήστη περισσότερες πληροφορίες από όσες μπορεί να απορροφήσει. Οι χρήστες μπερδεύονται, κουράζονται και εγκαταλείπουν την έρευνα όταν πρέπει να χειριστούν μεγάλους όγκους δεδομένων, τα οποία συχνά μπορεί να είναι λανθασμένα, άχρηστα, ταυτόσημα ή ομοειδή. Το μέγεθος της «υπερφόρτωσης» αυξάνεται με την αύξηση των βιβλιογραφικών εγγραφών, αλλά και με τον αριθμό των βάσεων στις οποίες γίνεται ταυτόχρονη αναζήτηση. Το ζήτημα αυτό αποτελεί ένα από τα σημαντικότερα μειονεκτήματα των συστημάτων ανάκτησης.

Χαμηλή παραγωγικότητα καταλογογράφησης

Οι διπλές εγγραφές σε μια βάση, δυσκολεύουν επίσης την αναζήτηση και τη λήψη αποφάσεων καταλογογράφησης. Ο καταλογογράφος αφιερώνει πολύτιμο εργασιακό χρόνο έως ότου αποφασίσει για την καταλληλότητα της εγγραφής που θα χρησιμοποιήσει, εάν όντως υπάρχει κάποια, για ένα νεοαποκτηθέν τεκμήριο. Υπάρχει η πιθανότητα να δημιουργήσει μια νέα εγγραφή ή να χρησιμοποιήσει μια ήδη υπάρχουσα. Η δεύτερη περίπτωση δεν τον απαλλάσσει από επιπλέον εργασία. Συνήθως η χρήση έτοιμων εγγραφών απαιτεί τη διόρθωση και τη βελτίωσή τους, με βάση τις ανάγκες και τα πρότυπα της βιβλιοθήκης.

Αύξηση κόστους λειτουργίας και συντήρησης συστήματος

Η ύπαρξη των διπλών εγγραφών επιδρά στο κόστος συντήρησης του αυτοματοποιημένου συστήματος της βιβλιοθήκης μέσω της έμμεσης αύξησης του χρόνου (π.χ. αναζητήσεις και περιττές καταχωρήσεις), της άμεσης αύξησης των δαπανών (π.χ. αναζητήσεις), αλλά και των αμοιβών του προσωπικού. Η καταπολέμηση των διπλών εγγραφών είναι μια ακριβή διαδικασία, αφού αφορά την εφαρμογή πολύπλοκων και δαπανηρών μη αυτοματοποιημένων διαδικασιών ή την ύπαρξη ενός αυτοματοποιημένου συστήματος εντοπισμού τους ή ακόμη συχνότερα, απαιτεί το συνδυασμό τους. Οι διπλές εγγραφές υπολογίζεται πως αυξάνουν το κόστος της καταλογογράφησης κατά 5-15%.

Λόγοι ύπαρξης διπλών εγγραφών

Η ύπαρξη των διπλών εγγραφών, είναι ένα δυσεπίλυτο ζήτημα που απασχολεί έντονα τις βιβλιογραφικές βάσεις, συχνά μάλιστα αποδεικνύεται και άλυτο, εξαιτίας

ελλιπών, ατελών, λανθασμένων ή αντικρουόμενων πληροφοριών. Η αντιμετώπισή τους γίνεται ακόμη πιο πολύπλοκη στο περιβάλλον μιας συνεργατικής βάσης δεδομένων, λόγω των πρακτικών που ακολουθούνται στους κανόνες καταλογογράφησης και των παραλλαγών που εμφανίζονται ως αποτέλεσμα της διαφορετικής ερμηνείας τους.

Διπλές εγγραφές δημιουργούνται όταν ο καταλογογράφος δεν βρίσκει την εγγραφή που ήδη υπάρχει στη βάση ή δεν μπορεί να αναγνωρίσει πως η συγκεκριμένη εγγραφή αντιπροσωπεύει το τεκμήριο που επιθυμεί να καταλογογραφήσει, με αποτέλεσμα να την καταχωρεί ακόμη μία φορά. Στην περίπτωση της batch loading, διπλές εγγραφές δημιουργούνται όταν οι παραλλαγές μεταξύ των εγγραφών που καταχωρούνται και οι αντίστοιχες εγγραφές της βάσης αποτρέπουν το σύστημα από την αναγνώριση των διπλών εγγραφών (στην περίπτωση βέβαια που το σύστημα υποστηρίζει τέτοιες διεργασίες).

Λεπτομερέστερες μελέτες που διεξήχθησαν, έδειξαν πως εκτός από τους παραπάνω λόγους, υπάρχουν και άλλοι παράγοντες που συνεισφέρουν στην ύπαρξη διπλών εγγραφών. Αυτοί οι παράγοντες είναι οι εξής:

- Τοπικές πρακτικές
- Ασυνέπειες στην καταλογογράφηση
- Απροσεξίες καταχώρησης
- Λάθη σύνταξης του MARC format

Τοπικές πρακτικές

Όπως όλοι οι κανόνες έτσι και οι AACR2, παρόλο που επιδιώκουν την ομοιομορφία, δεν μπορούν να αποτρέψουν την ποικιλότητα των ερμηνειών τους. Επιπλέον το γεγονός πως προέρχονται από την βιβλιοθηκονομική κοινότητα, δεν συνεπάγεται και την αυτόματη αποδοχή τους από την εκδοτική κοινότητα. Οι εκδότες παράγουν τα τεκμήρια δίνοντάς τους τα χαρακτηριστικά και τις πληροφορίες που οι ίδιοι επιθυμούν. Οι επιθυμίες των καταλογογράφων και των εκδοτών, δεν είναι υποχρεωτικό να ταυτίζονται. Η ποικιλομορφία των δημοσιευμάτων οδηγεί την κάθε βιβλιογραφική υπηρεσία ή ακόμη και τον κάθε μεμονωμένο καταλογογράφο, να δίνουν διαφορετικές ερμηνείες σε φαινόμενα που συναντούν. Στο σύγχρονο περιβάλλον, που βασίζεται στην αυτοματοποίηση και τη συνεργασία μεταξύ των βιβλιοθηκών, οι ακολουθούμενες πρακτικές και πολιτικές των βιβλιοθηκών-μελών, ορίζονται λαμβάνοντας υπόψη όχι πλέον τις ιδιαίτερες τοπικές ανάγκες, αλλά κυρίως αυτές που μπορούν να συγκεράσουν και να ικανοποιήσουν το σύνολο των μελών της κοινοπραξίας ή της συνεργατικής υπηρεσίας. Ο καταλογογράφος είναι υποχρεωμένος να λαμβάνει υπόψη του τις πολιτικές που ακολουθεί το συνεργατικό σχήμα για τη δημιουργία νέων εγγραφών.

Ασυνέπειες καταλογογράφησης

Οι πληροφορίες που προέρχονται από ένα τεκμήριο μπορούν να ερμηνευτούν διαφορετικά. Ιδιαίτερα δε όσον αφορά περιπτώσεις που πρέπει να αποφασιστεί εάν ένα σύνολο ομοειδών τεκμηρίων πρέπει να καταχωρηθεί κάτω υπό μία κοινή εγγραφή ή το κάθε μέρος του πρέπει να καταχωρηθεί ξεχωριστά. Πρέπει δηλαδή να ληφθεί η απόφαση για το εάν τα τεκμήρια αυτά αποτελούν η όχι διαφορετικές εκδόσεις, ανατυπώσεις, κυκλοφορίες, κτλ. Για την διάκριση μεταξύ διαφορετικών εκδόσεων πρωταρχικό ρόλο παίζει ο ορισμός της έννοιας της έκδοσης.

Ορισμός έκδοσης (edition)

Ο ορισμός της όρου “έκδοση” διαδραματίζει πρωταρχικό ρόλο και αποτελεί το σημαντικότερο παράγοντα δημιουργίας διπλών εγγραφών εκ μέρους των καταλογογράφων. Η σπουδαιότητα της διευκρίνησής του αντανάκλαται στην συνεχή προσπάθεια ορισμού και συνεχούς προσαρμογής του. Έτσι κατά χρονολογική σειρά παρουσιάζονται οι εξής ορισμοί:

- Ο αριθμός των αντιγράφων ενός βιβλίου δημοσιευμένα την ίδια χρονική στιγμή στην ίδια μορφή (Cutter, 1904)
- Ο συνολικός αριθμός των αντιγράφων που εκτυπώθηκαν με την ίδια στοιχειοθεσία (ALA, 1941) (IFLA, 1974) (AACR2, 1988).
- Όλα τα αντίγραφα μιας πηγής που παράχθηκαν από την ίδια μήτρα και δημοσιεύτηκαν ή κυκλοφόρησαν από μια συγκεκριμένη υπηρεσία ή ομάδα υπηρεσιών (ISBD-ER, 2002).

Διάκριση μεταξύ εκδόσεων και εκτυπώσεων

Έχει διαπιστωθεί πως **Σφάλμα! Λανθασμένη σύνδεση.** της διάκρισης μεταξύ εκδόσεων, αποτελεί την κύρια αιτία παρουσίας διπλών εγγραφών, γεγονός που το αποδίδεται στον τρόπο εφαρμογής και ερμηνείας των κανόνων καταλογογράφησης. Οι απόπειρες της επεξήγησης και διευκρίνησης του όρου *έκδοση (edition)* είναι ενδεικτικές της προσπάθειας να συναντηθεί η καταλογογράφηση με τις νέες τεχνολογικές εξελίξεις που αφορούν τη μετάβαση από την τύπωση βιβλίων με μεθόδους στοιχειοθεσίας στην παραγωγή ποικίλων ειδών τεκμηρίων, με άλλους τρόπους και νέες μεθόδους αναπαραγωγής. Η κατανόηση της ιστορίας της τεχνολογίας της εκτύπωσης μπορεί να συνεισφέρει στην ικανότητα του καταλογογράφου να διακρίνει μια “έκδοση” η οποία καθορίζεται λαμβάνοντας υπόψη το περιεχόμενο του κειμένου και την μορφή της στοιχειοθεσίας. Η τεχνολογία της εκτύπωσης διαιρείται σε τρεις περιόδους:

- *Hand press, 1500-1800.* Οι έννοιες “έκδοση” και “εκτύπωση” ήταν ταυτόσημες.
- *Machine press, 1800-1950.* Η εφαρμογή αυτής της τεχνολογίας οδήγησε στη διάκριση των όρων και των διαδικασιών “έκδοση” και “εκτύπωση”
- *Σύγχρονες τεχνολογίες: φωτοστοιχειοθεσία και υπολογιστές.* Συνεχίζεται η ίδια κατάσταση με την προηγούμενη περίοδο.

Στις σύγχρονες μορφές εκτύπωσης μπορεί και πάλι να εμφανιστεί δυσκολία διάκρισης μεταξύ παραλλαγών εκδόσεων. Η αλλαγή που αφορά στη μορφή του κειμένου ή/και στο περιεχόμενό του, είναι ένα ακόμη κριτήριο διάκρισης μεταξύ “εκτύπωσης” και “έκδοσης”. Είναι αρκετά συνηθισμένο ένα συγκεκριμένο έργο να έχει αντιγραφεί σε έναν αριθμό διαφορετικών μορφών:

- Μπορεί να δημοσιευθεί με σκληρό δέσιμο ή χαρτόδετο,
- Μπορεί να είναι δημοσιευμένο σε μια χώρα από έναν εκδοτικό οίκο και σε κάποιον άλλο, με δυσδιάκριτες τις διαφορές του περιεχομένου του. Ένα βιβλίο μπορεί να τυπωθεί σε μια τρίτη χώρα και να κυκλοφορήσει σε διάφορες αγγλόφωνες ή γαλλόφωνες χώρες με διαφορετικά εκδοτικά χαρακτηριστικά (εξώφυλλα, σελίδες τίτλου, στοιχειοθεσία, κτλ).
- Μπορεί να κυκλοφορήσουν αντίγραφα που είναι μέρος ενός τυπωμένου τεκμηρίου σε microfilm, microfiche, CD, DVD ή σε αρχείο υπολογιστή και στην συνέχεια να είναι «σωσμένο» σε πολλά διαφορετικά formats (doc, pdf, jpeg, etc.). Οι δυνατότητες της ηλεκτρονικής έκδοσης, αλλά και οι νέες

τεχνολογίες που εφαρμόζονται στις έντυπες εκδόσεις, οδηγούν στη συνεχή κυκλοφορία ενημερωμένων εκδόσεων.

Οι παραλλαγές στην εξωτερική μορφή δημοσίευσης ενός τεκμηρίου, όπως π.χ. στη βιβλιοδεσία και η ταυτόχρονη διάθεσή του σε πολλαπλά formats απαιτούν τη λήψη αποφάσεων όσον αφορά το αν πρόκειται για μια έκδοση ή ανατύπωση, αν τα δύο τεκμήρια είναι ταυτόσημα στο περιεχόμενο και εν τέλει αν μπορεί να χρησιμοποιηθεί η υπάρχουσα εγγραφή (αυτούσια ή με τις ανάλογες τροποποιήσεις) ή απαιτείται η δημιουργία μιας νέας εγγραφής.

Εκτός από τον ορισμό της έκδοσης ενός τεκμηρίου, άλλα ζητήματα και πεδία της καταλογογράφησης, ο καθορισμός των οποίων οδηγεί στην δημιουργία διπλών εγγραφών, είναι οι εξής:

Κύρια πηγή καταλογογράφησης

Η κύρια πηγή καταλογογράφησης είναι η σελίδα τίτλου του τεκμηρίου. Ο χρόνος που μπορεί να διαθέσει ένας καταλογογράφος για την εξέταση ενός document συνήθως είναι περιορισμένος. Η συνηθέστερη πρακτική που εφαρμόζεται είναι το “τεχνικό διάβασμα”, που συνίσταται στη μελέτη κυρίως της σελίδας τίτλου και των προκαταρκτικών μιας έκδοσης, των περιεχομένων και τέλος σύντομων εδαφίων του κειμένου. Ακόμη πιο συνηθισμένη πρακτική στο σύγχρονο περιβάλλον εργασίας, είναι η προσφυγή σε βάσεις δεδομένων για την ανεύρεση της ίδιας ή παρόμοιων εγγραφών. Συνεξετάζοντας τα δεδομένα της σελίδας τίτλου, της σελιδαρίθμησης και των διαστάσεων, ο καταλογογράφος αποφασίζει εάν δύο τεκμήρια είναι εκδηλώσεις της ίδιας ή δύο διαφορετικών εκδόσεων. Δυστυχώς όμως, τα δεδομένα αυτά δεν αποτελούν πάντα αξιόπιστες πηγές διάκρισης διαφορετικών εκδόσεων και η σύγκριση των κειμένων για να διαπιστωθεί η ομοιότητα δύο εγγραφών είναι στην πράξη αδύνατη.

Λάθη στην κύρια αναγραφή - καθιέρωση ονομάτων

Ο καθορισμός της κύριας αναγραφής αποτελεί μια από τις σημαντικότερες αποφάσεις της καταλογογράφησης. Στις περισσότερες εγγραφές η κύρια αναγραφή αφορά το όνομα του συγγραφέα. Οι πρακτικές καταχώρησης της καθιερωμένης μορφής ονομάτων συγγραφέων ή συλλογικών οργάνων, παρουσιάζουν πολλές παραλλαγές και για το λόγο αυτό αποτελούν μια ιδιαίτερη πηγή προβλημάτων. Το ίδιο όνομα μπορεί να καταχωρηθεί με πολλούς τρόπους που εξαρτώνται και από την πηγή από την οποία προέρχονται οι καθιερώσεις, αλλά και από τις πρακτικές και το επίπεδο ανάλυσης που επιθυμεί η κάθε βιβλιοθήκη.

Τα ασυνήθιστα ονόματα, επίσης, προκαλούν συχνά προβλήματα, π.χ. ονόματα που περιέχουν τα «*Jr.*», «*Mac*», «*Mc*», ονόματα με αριθμητικά, ονόματα αγίων και τιτλούχων της εκκλησίας ή του κράτους, αλλά και οι κύριες αναγραφές-ενιαίοι τίτλοι και τέλος οι περιπτώσεις που αφορούν συλλογικά όργανα και ονόματα συνεδρίων, όπου το περιθώριο «αυτοσχεδιασμού» του καταλογογράφου είναι μεγαλύτερο, προκαλούν ακόμη περισσότερες παραλλαγές.

Μνεία έκδοσης/πνευματικής υπευθυνότητας (edition)

Έχει διαπιστωθεί πως η μνεία έκδοσης σε συνδυασμό με άλλα πεδία, ευθύνεται για ένα πολύ μεγάλο ποσοστό διπλών εγγραφών. Η παρουσία λέξεων όπως, έκδοση, κυκλοφορία ή ανατύπωση είναι πολλές φορές ενδεικτική για το ποια πολιτική καταλογογράφησης θα εφαρμοστεί για το συγκεκριμένο έργο. Στις περιπτώσεις αυτές, επιστρατεύεται η κρίση και η εμπειρία του καταλογογράφου. Οι AACR2, δεν είναι πάντα σε θέση να δίνουν την λύση σε τέτοιου είδους αμφισβητούμενα ζητήματα

και προτρέπουν στη διάκριση των εκδόσεων κάθε φορά που ανακύπτει κάποια αμφιβολία. Αυτό με τη σειρά του, οδηγεί στη δημιουργία πολλαπλών εγγραφών που περιγράφουν ταυτόσημα τεκμήρια.

Εκδοτικός Οίκος

Πολλές φορές δύο τεκμήρια μπορεί να ταυτίζονται παρά τις διαφορές που μπορούν να εμφανίζουν στο όνομα του εκδότη. Είναι συχνό φαινόμενο η ταυτόχρονη παρουσία παραλλαγών της ονομασίας ενός εκδοτικού οίκου. Στις περιπτώσεις αυτές, οι εγγραφές συνήθως θεωρούνται ταυτόσημες και το μόνο που απαιτείται είναι η γνώση από τον καταλογογράφο της ύπαρξης των παραλλαγών αυτών. Όταν όμως δύο documents παρουσιάζονται κατά τα άλλα όμοια, αλλά με εντελώς διαφορετικό τον εκδοτικό οίκο, αυτό είναι μια ένδειξη πως πιθανότατα θα υπάρχουν και άλλες διαφορές, στη σελιδαρίθμηση, τη στοιχειοθεσία, κτλ. και κατ' επέκταση ίσως και στο περιεχόμενο. Στην τελευταία περίπτωση το σύνηθες είναι να μην προχωρά ο καταλογογράφος σε περαιτέρω σύγκριση των τεκμηρίων και να δημιουργεί δύο διαφορετικές εγγραφές.

Χρονολογία έκδοσης

Η αλλαγή, εκ μέρους των εκδοτών της χρονολογίας έκδοσης ενός έργου, συνήθως αποτελεί ένδειξη πως το συγκεκριμένο έργο συνεχίζει να είναι διαθέσιμο στο εμπόριο και ταυτόχρονα ένδειξη της επικαιρότητάς του. Οι AACR2 δίνουν τη δυνατότητα στον καταλογογράφο να αγνοήσει την χρονολογία που παρουσιάζεται στην κύρια πηγή της καταλογογράφησης, και να χρησιμοποιήσει κάποια άλλη αναφερόμενη χρονολογία εφόσον κρίνει ότι αυτή ανταποκρίνεται περισσότερο στο συγκεκριμένο τεκμήριο. Οι διαφορές που εμφανίζονται στην παράθεση διαφορετικών χρονολογιών στο τεκμήριο, π.χ. χρονολογία έκδοσης, copyright και εκτύπωσης, συνήθως δεν αντανakλούν κάποια πραγματική διαφορά στο περιεχόμενο. Απλώς είναι μια πρακτική που εφαρμόζεται από τους εκδοτικούς οίκους και αφορά τρόπους προώθησης των εκδόσεών τους και τονισμό της επικαιρότητάς τους. Έχει διαπιστωθεί πως η χρονολογία έκδοσης ευθύνεται για το 50,7% των διπλών εγγραφών.

Σελιδαρίθμηση

Πολύ συχνά για τη διάκριση των εκδόσεων, η σελιδαρίθμηση αποδεικνύεται πως αποτελεί πιο αξιόπιστο στοιχείο από ότι η διαφορά που παρουσιάζεται στη σελίδα τίτλου. Οι διαφορές που παρουσιάζονται στην καταχώρηση του αριθμού των σελίδων, μπορεί να προέρχονται είτε από λάθος στην καταχώρηση είτε και από την διαφορετική πρακτική στον τρόπο καταλογογράφησης. Ο αριθμός των σελίδων μπορεί σε κάποιες περιπτώσεις να αποτελέσει την μοναδική πηγή διάκρισης μεταξύ δύο τεκμηρίων. Αποτελεί μια ασφαλή ένδειξη αλλαγής της στοιχειοθεσίας, η οποία πιθανά να συνοδεύεται από αλλαγές στο περιεχόμενο του κειμένου.

Πεδία σημειώσεων

Επειδή οι σημειώσεις καταχωρούνται σε ελεύθερο κείμενο, η χρησιμότητά τους στην ανίχνευση διπλών εγγραφών είναι γενικά περιορισμένη. Αν και στο μορφότυπο MARC υπάρχει μια σειρά από εξειδικευμένα πεδία σημειώσεων, πολλές βιβλιοθήκες προτιμούν να καταχωρούν όλες τις σημειώσεις στο πεδίο γενικών σημειώσεων, με αποτέλεσμα να καθίσταται δύσκολος έως αδύνατος ο αυτοματοποιημένος εντοπισμός των διπλών εγγραφών. Σε μερικές περιπτώσεις όμως οι πληροφορίες των πεδίων αυτών αποτελούν το μόνο τρόπο διάκρισης μεταξύ δύο παρόμοιων τεκμηρίων.

Συντομογραφίες

Οι συντμήσεις είναι μια ακόμη πηγή προβλημάτων που προκαλούνται από τις διαφορές στην καταλογογράφηση. Τα ονόματα των περιοδικών, των ονομάτων συνεδρίων, των δηλώσεων έκδοσης (έκδοση - εκδ., κτλ.), συντμούνται ή περικόπτονται με πολλούς μη πρότυπους τρόπους. Σε γενικές γραμμές η λύση των προβλημάτων που προκαλούνται από αυτήν την πηγή, είναι η σταθερότητα στην εφαρμογή προτυποποιημένων συντμήσεων και αποφυγή από τους καταλογογράφους συντετμημένων πληροφοριών όταν αυτό δεν επιβάλλεται από τις ίδιες τις πληροφορίες του τεκμηρίου.

Απροσεξίες: παράλειψη πληροφοριών και λάθη καταχώρησης

Οι ελλειψείς, οι μη ολοκληρωμένες και οι ανακριβείς πληροφορίες αποτελούν έναν ακόμη βασικό παράγοντα δημιουργίας διπλών εγγραφών. Η ανάλυση των διπλών εγγραφών στη βάση του OCLC, έδειξε πως τα στοιχεία που συχνότερα παραλείπονται είναι οι πληροφορίες στην περιοχή τίτλου και των εκδοτικών οίκων όταν είναι περισσότεροι από ένας. Υπάρχουν περιπτώσεις που οδηγούν στην καταγραφή μόνο όσων στοιχείων θεωρούνται αναγκαία σε τοπικό επίπεδο. Σε άλλες πάλι περιπτώσεις προέρχονται αποκλειστικά και μόνο από τον τρόπο εργασίας και το επίπεδο καταγραφής των διαθέσιμων λεπτομερειών από τον καταλογογράφο.

Αυτό το είδος των λαθών δημιουργείται κατά την καταχώρηση των βιβλιογραφικών δεδομένων στην εγγραφή. Τα ορθογραφικά και τα τυπογραφικά λάθη αποτελούν μια κατά βάση ακούσια και ευεξήγητη περίπτωση. Τα λάθη μεταγραφής, το λανθασμένο γράψιμο και τα τυπογραφικά λάθη, είναι συνηθισμένα. Συμπεριλαμβάνουν λάθη όπως, διαφορές στα κεφαλαία γράμματα, στα διαστήματα, στη στίξη και στα διακριτικά σημάδια. Αν και εκ πρώτης όψεως δεν φαίνονται ιδιαίτερα σημαντικά, εν τούτοις, δεν πρέπει να ξεχνάμε ότι η χρήση και η «ερμηνεία» τους γίνεται από ένα μηχάνημα, γεγονός που σημαίνει πως δημιουργούν πρόβλημα στην ανάκτηση, αφού στις περισσότερες εφαρμογές είναι δύσκολη η παράβλεψή τους από τον υπολογιστή. Έχει παρατηρηθεί πως ο αριθμός τους είναι ανάλογος του μήκους του πεδίου. Τα περισσότερα λάθη εμφανίζονται στα πεδία του συγγραφέα, του τίτλου, και του εκδότη. Τα περισσότερα από τα λάθη αυτά είναι διαφορές ενός χαρακτήρα, π.χ.

- αντικατάσταση γραμμάτων,
- προσθήκη επιπλέον γραμμάτων,
- ελλειψείς τίτλοι (παράληψη κάποιων λέξεων)

Λάθη σύνταξης του MARC format

Ο τελευταίος παράγοντας δημιουργίας διπλών εγγραφών είναι τα λάθη που οφείλονται στον χειρισμό του MARC format. Είναι τα λάθη που οφείλονται στην σύνταξή του και την ασυνέπεια στην καταχώρηση μεταξύ των πεδίων σταθερού και μεταβλητού μήκους (fixed and variable fields).

Λανθασμένα πεδία και υποπεδία

Το είδος αυτό των λαθών αφορά την τοποθέτηση πληροφοριών σε λανθασμένο πεδίο ή υποπεδίο μιας βιβλιογραφικής εγγραφής και όχι στα προοριζόμενα για αυτές πεδία και υποπεδία, καθώς και στην έλλειψη ή λανθασμένη χρήση των δεικτών του MARC format. Αυτά τα λάθη εμπίπτουν στην κατηγορία των λαθών απροσεξίας και παράβλεψης. Μπορούν να παρουσιαστούν στις εξής περιπτώσεις, όσον αφορά το UNIMARC MARC:

- ένας συγγραφέας που αποτελεί το θέμα του τεκμηρίου και καταχωρείται αντί των πεδίων 70XX, στο 6XX
- μνεία ευθύνης (πνευματικής υπευθυνότητας), αντί του υποπεδίου #f του πεδίου 200, καταχωρείται στο υποπεδίο #e ως υπότιτλος, και
- εκδοτικός οίκος, αντί του υποπεδίου #c του πεδίου 210, καταχωρείται στο υποπεδίο #a ως τόπος έκδοσης, κτλ

Στην κατηγορία αυτή ανήκουν και τα λάθη που εμφανίζονται στους δείκτες. Οι δείκτες, στην περίπτωση των εγγραφών UNIMARC, ιδιαίτερα στο πεδίο 200 είναι αυτοί που ορίζουν τον αριθμό των αρχικών χαρακτήρων που δεν λαμβάνονται υπόψη στην ταξιθέτηση(ή αλλιώς δεν λαμβάνονται υπόψη στον ευρετηριασμό που κάνει το πρόγραμμα του υπολογιστή). Αποτελεί συχνό φαινόμενο η λανθασμένη χρήση τους. Το αποτέλεσμά τους είναι η αδυναμία εντοπισμού των σχετικών τεκμηρίων σε μια βάση και η καταχώρηση διπλών εγγραφών.

Ασυνέπειες μεταξύ των πεδίων σταθερού και μεταβλητού μήκους

Οι πληροφορίες που περιέχει μία βιβλιογραφική εγγραφή καταχωρούνται σε πεδία το μήκος των οποίων είναι μεταβλητό. Στα πεδία αυτά καταχωρούνται όλες οι βιβλιογραφικές πληροφορίες ενός document με τον τρόπο που ορίζεται από τους AACR2, ανεξάρτητα από το μέγεθος που αυτές καταλαμβάνουν. Αντιθέτως, τα πεδία σταθερού μήκους στις εγγραφές MARC δεν περιέχουν πληροφορίες που παράγονται απευθείας από το document, αλλά περιέχουν πληροφορίες που κωδικοποιούν το σύνολο της καταλογογραφημένης βιβλιογραφικής εγγραφής. Για το λόγο αυτό πολύ συχνά είτε περιέχουν αντικρουόμενες πληροφορίες, είτε παραμένουν κενοί. Αντικρουόμενες πληροφορίες μπορούν να υφίστανται όταν π.χ. το τεκμήριο που καταγράφεται είναι microfilm και στο πεδίο σταθερού μήκους καταχωρείται ως έντυπο, ή ακόμη πιο συχνά η χρονολογία έκδοσης του τεκμηρίου παρουσιάζεται διαφορετική από αυτή που είναι καταχωρημένη στο πεδίο μεταβλητού μήκους. Τα λάθη αυτού του είδους εμπίπτουν στην κατηγορία των ασυνεπειών στην καταχώρηση των εγγραφών.

Συμπεράσματα

Το πλήθος των διπλών εγγραφών που βρίσκονται αποθηκευμένες σε μια βιβλιογραφική βάση, είναι μια πηγή ενόχλησης το μέγεθος και το είδος της οποίας εξαρτάται από το είδος της βιβλιοθήκης και έχει επιπτώσεις στην απόδοση του ευρετηρίασης, στην αύξηση του κόστους αποθήκευσης και διατήρησής τους, δυσχεραίνει την αποτελεσματική αναζήτησή τους και προκαλεί σοβαρά προβλήματα στον ποιοτικό έλεγχο και στους χρήστες (εσωτερικούς και εξωτερικούς). Για να μειωθούν οι επιπτώσεις αυτές, η κάθε εγγραφή της βάσης θα έπρεπε να απεικονίζει με ακρίβεια και μοναδικότητα τις συγκεκριμένες πληροφορίες που αντιπροσωπεύουν την κάθε ξεχωριστή βιβλιογραφική οντότητα, τα αντίτυπά της και τα μέρη της.

Η λύση στο πρόβλημα της ύπαρξης διπλών εγγραφών είναι ένα ζήτημα πολύπλοκο και δυσεπίλυτο. Αφενός τα λάθη - είτε προέρχονται από την ερμηνεία των κανόνων καταλογογράφησης είτε από αβλεψία στην καταχώρηση - εφ' όσον εξαρτώνται από τον ανθρώπινο παράγοντα και την κρίση του, θα συνεχίσουν να υπάρχουν. Αφετέρου, η εφαρμογή αυτοματοποιημένων διαδικασιών για την πρόληψη, διαχείριση και περιορισμό του φαινομένου απαιτεί επιλογές και λύσεις προσαρμοσμένες στην εκάστοτε περίπτωση. Σε κάθε περίπτωση πάντως θα πρέπει να είμαστε ανεκτικοί σε κάποιες περιπτώσεις χαμένων και λανθασμένων ταυτοποιήσεων.

Το ζητούμενο είναι η διαχείριση μεγάλου όγκου δεδομένων και η επίλυση του προβλήματος, σε δικτυωμένα περιβάλλοντα σε πραγματικό χρόνο. Η τεχνολογική πρόκληση που παρατηρούμε να εμφανίζεται, είναι η δημιουργία και η εφαρμογή ειδικών αλγορίθμων εντοπισμού και αφαίρεσης διπλών εγγραφών που να ικανοποιούν τις νέες λειτουργίες των καταλόγων με στόχο τη συγκέντρωση και ομαδοποίηση σχετικών εκδηλώσεων των τεκμηρίων, και η ταυτόχρονη εφαρμογή σύγχρονων αρχιτεκτονικών με βάση τα κυρίαρχα πρωτόκολλα επικοινωνίας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Ayres, F. H. (1990). Duplication and other manifestations: a new approach to the presentation of bibliographic information. *Journal of Librarianship*, 22 (4), 236-251.
2. OCLC glossaries: batch processing glossary. Retrieved from the World Wide Web: <http://www.oclc.org/support/documentation/glossary/default.html>.
3. O' Neill, Edward & Oskins, Mike W. (July 1990). Duplicate records in the Online Union Catalog [internal report], OCLC Office of Research.
4. O' Neil, Edward T. & Vizine-Goetz, Diane (1988). Quality-control in online database. *Annual Review of Information Science and Technology*, 23, 125-156.
5. Salim, Maud (2003). Information overload and its implications for a corporate library: as perceived by eight researchers at AstraZeneca. Magisteruppsats i Biblioteks, BHS. Retrieved from the World Wide Web: <http://www.hb.se/bhs/slutversioner/2003/03-112.pdf>.
6. Wanninger, Patricia Dwyer (1982). Is the OCLC database too large? a study of the effects of duplicate records in the OCLC system. *Library Resources and Technical Services*, 26, 353-361.
7. O' Neil, Edward T., Rogers, Sally A. & Oskins, Michael W. (1993). Characteristics of duplicate records in OCLC's Online Union Catalog. *Library Resources & Technical Services*, 37 (1), 59-71.
8. Anglo-american cataloguing rules (1988 revision). 2nd ed. Chicago: American Library Association.
9. Yee, Martha M. (1994). Manifestations and near-equivalents: theory, with special attention to moving-image materials. *Library Resources & Technical Services*, 38 (3), 227-255.
10. IFLA Multilingual Dictionary of Cataloging Terms and Concepts. Retrieved from the World Wide Web: <http://subito.biblio.etc.tu-bs.de/muldicat/page.php?urG=ENT&urS=edition>.
11. Jones, Barbara & Kastner, Arno (1983). Duplicate records in the bibliographic utilities: a historical review of the printing versus edition problem. *Library Resources and Technical Services*, 27, 211-220.