

### 3<sup>η</sup> εργασία – XML

Στόχος των εργασιών είναι η εξοικείωση με θεωρητικά και πρακτικά προβλήματα των Βάσεων Δεδομένων, μέσα από χρηστικά παραδείγματα. Στην τρίτη εργασία θα ασχοληθούμε: α) με τη δομή και το validation των XML αρχείων, β) με τη υποβολή X-Path ερωτημάτων.

#### Ερώτημα 1ο – Έλεγχος εγκυρότητας και DTD ή XML-Schema

Για το αρχείο all.xml που σας δίνεται μαζί με την εκφώνηση

#### Ερώτημα 2ο – Βελτιστοποίηση ερωτήματος ανισότητας

Χρησιμοποιώντας ως βάση για τις ερωτήσεις σας το αρχείο all.xml να δώσετε σε xpath τα ερωτήματα για τις ακόλουθες περιπτώσεις:

- i) Υπολογίστε το πλήθος των ταινιών στο αρχείο
- ii) Υπολογίστε το πλήθος των έγχρωμων ταινιών στο αρχείο
- iii) Εμφανίστε τους τίτλους των έγχρωμων ταινιών στο αρχείο
- iv) Εμφανίστε τους σκηνοθέτες Ινδικών ταινιών
- v) Τα διαφορετικά είδη των ταινιών στις οποίες πρωταγωνιστεί ο Kenneth Connor
- vi) Τους τίτλους των ταινιών που έχουν διάρκεια μικρότερη από 20 λεπτά στις ΗΠΑ
- vii) Τους σκηνοθέτες Short και Animation ταινιών
- viii) Τους τίτλους ταινιών τύπου Comedy που γυρίστηκαν μεταξύ 2000 και 2006 (ανοικτό)
- ix) Τους τίτλους ταινιών που στην πλοκή τους περιέχουν τη λέξη sport
- x) Τους τίτλους των ταινιών που ο σκηνοθέτης τους είναι και ηθοποιός τους και σε ξεχωριστό ερώτημα ποιοι είναι οι σκηνοθέτες αυτοί

Μπορείτε να κάνετε χρήση και των ακόλουθων συναρτήσεων:

- count(path) - απαριθμεί τα elements που επιστρέφει το path
- sum(path) - αθροίζει τις ακέραιες τιμές που επιστρέφει το path
- path/text() - εμφανίζει το text περιεχόμενο ενός element που ταιριάζει στο path
- distinct-values(path) - εμφανίζει τα μοναδικά elements (ή τιμές) που επιστρέφει το path
- contains(path, 'sub-string') - εμφανίζει μόνο τα path που δείχνουν ένα string το οποίο περιέχει το συγκεκριμένο sub-string.

#### Β – Οδηγίες Παράδοσης

Η εργασία θα υλοποιηθεί από ομάδες των **3 ατόμων (το πολύ)**, αν και επιτρέπονται μικρότερες ομάδες. Θα πρέπει τελικά να ανεβάσετε ένα zip αρχείο με ονομασία τους ΑΜ των μελών της ομάδας: π.χ. **AM1-AM2-AM3.zip**

- Το zip θα περιλαμβάνει:
  - ένα αρχείο readme.txt
    - με τα ονοματεπώνυμα και τους ΑΜ των φοιτητών της ομάδας
  - το αρχείο word με την τελική εργασία. Τυχόν συνοδευτικά αρχεία dtd, txt κλπ.

#### Γ – Άλλες Οδηγίες

Όσες εργασίες δεν τηρούν τις οδηγίες παράδοσης, θα έχουν επίπτωση στο βαθμό.

Όσες εργασίες κριθούν ότι είναι **αντιγραφές θα μηδενίζονται**.

Ημερομηνία παράδοσης: **Στο e-class με οριστική τελική ημερομηνία 14-1-2022**

Όσες εργασίες παραδοθούν μετά το πέρας της ημερομηνίας και μέχρι τις 16-1-2022 θα έχουν μείωση 2 μονάδων στο βαθμό.