



Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών
Υπολογιστών
Εξάμηνο 7ο

Τελική εργασία στο μάθημα Τεχνολογία Πολυμέσων

Σύνοψη

Στόχος της συγκεκριμένης εργασίας είναι η ανάπτυξη μιας *desktop* εφαρμογής σε *Java* με δυνατότητες δημιουργίας και επεξεργασίας *PDF* αρχείων με βάση ένα σύνολο απλών εντολών τυποποίησης (*markup syntax*) και τη βιβλιοθήκη *PDFBox* [1].

Περιγραφή Εφαρμογής

Αντικείμενο της εργασίας αποτελεί ο σχεδιασμός και υλοποίηση μιας εφαρμογής για τη δημιουργία και επεξεργασία *PDF* αρχείων. Η εφαρμογή θα λειτουργεί ως ένας επεξεργαστής κειμένου που θα επιτρέπει στους χρήστες να περιγράφουν το περιεχόμενο και τη δομή των αρχείων *PDF*, μέσω της χρήσης ορισμένων βασικών εντολών μορφοποίησης. Οι περιγραφές θα πρέπει να αποθηκεύονται σε αρχεία κειμένου που θα έχουν την ειδική κατάληξη *mlab*.

Επιπλέον, οι χρήστες θα μπορούν να διαχειρίζονται διάφορα αρχεία περιγραφής (*mlab*), να δημιουργούν τα αντίστοιχα *PDF* καθώς και να επεξεργάζονται *PDF* αρχεία μέσω της υποστήριξης των λειτουργιών της συνένωσης, του διαχωρισμού αρχείων καθώς και της εξαγωγής συγκεκριμένων σελίδων.

Για λόγους απλότητας θεωρούμε ότι η εφαρμογή υποστηρίζει μόνο μερικές πολύ βασικές δυνατότητες μορφοποίησης κειμένου για τη δημιουργία *PDF*. Πιο συγκεκριμένα, έχουμε τα παρακάτω βασικά στοιχεία μορφοποίησης:

- **Επικεφαλίδα:** Ξεκινά πάντα από μια νέα σειρά, έχει μαύρο χρώμα, είναι **bold** και πρέπει να έχει μία σαφή απόσταση από το προηγούμενο στοιχείο. Ο χρήστης μπορεί να επιλέξει μόνο τον τύπο της γραμματοσειράς. Υπάρχουν 4 τύποι που αντιστοιχούν σε 4 προκαθορισμένα μεγέθη γραμματοσειράς (20,18,16,14).
- **Παράγραφος:** Αντιστοιχεί σε μία ή περισσότερες προτάσεις, πρέπει πάντα να ξεκινά από μια νέα γραμμή και να έχει μια σαφή απόσταση από το προηγούμενο στοιχείο. Ο χρήστης μπορεί να επιλέξει τη γραμματοσειρά, το μέγεθος, το χρώμα καθώς και τον τύπο (**bold**, *italic*, **bold & italic**). Οι επιλογές αυτές ισχύουν για όλα τα περιεχόμενα μιας παραγράφου.

- **Τροποποίηση Μορφοποίησης:** Εφαρμόζεται σε οποιοδήποτε επιμέρους τμήμα μιας παραγράφου και επιτρέπει την αλλαγή των παραμέτρων μορφοποίησης για το κείμενο που περιλαμβάνεται στο συγκεκριμένο τμήμα. Ο χρήστης μπορεί να επιλέξει τη γραμματοσειρά, το μέγεθος, το χρώμα καθώς και τον τύπο (bold, italic, bold & italic).
- **Εικόνα:** Ξεκινά πάντα από μια νέα σειρά και πρέπει έχει μια σαφή απόσταση από το προηγούμενο στοιχείο. Ο χρήστης μπορεί να ορίσει επιπλέον το σχετικό μέγεθος της εικόνας μέσω της παραμέτρου κλιμάκωσης ως ποσοστό του αρχικού μεγέθους.
- **Διατεταγμένη λίστα:** Περιλαμβάνει ένα ή περισσότερα στοιχεία κειμένου, με το κάθε στοιχείο να ξεκινά από μια νέα γραμμή, να έχει περισσότερο περιθώριο σε σχέση με τις γραμμές καθώς και να είναι αριθμημένο κατά αύξουσα σειρά. Ο χρήστης μπορεί να επιλέξει τη γραμματοσειρά, το μέγεθος, το χρώμα καθώς και τον τύπο (bold, italic, bold & italic). Οι επιλογές αυτές είναι κοινές για όλα τα στοιχεία της λίστας.
- **Μη διατεταγμένη λίστα:** Περιλαμβάνει ένα ή περισσότερα στοιχεία κειμένου, με το κάθε στοιχείο να ξεκινά από μια νέα γραμμή, να έχει περισσότερο περιθώριο σε σχέση με τις γραμμές καθώς και να αρχίζει με το χαρακτήρα «-». Ο χρήστης μπορεί να επιλέξει τη γραμματοσειρά, το μέγεθος, το χρώμα καθώς και τον τύπο (bold, italic, bold & italic). Οι επιλογές αυτές είναι κοινές για όλα τα στοιχεία της λίστας.
- **Νέα γραμμή:** Προσθέτει μια νέα κενή γραμμή με το ύψος που αντιστοιχεί στη γραμματοσειρά Times Roman με μέγεθος 12.

Σε κάθε στοιχείο μορφοποίησης αντιστοιχεί μια εντολή μορφοποίησης που το προσδιορίζει και περιλαμβάνει τις επιμέρους επιλογές που αυτό υποστηρίζει. Κάθε εντολή μορφοποίησης αποτελείται από δύο μέρη: (i) ετικέτα εντολής και παράμετροι (ii) περιεχόμενο (στην περίπτωση της κενής γραμμής δεν υπάρχει περιεχόμενο). Στο αρχείο περιγραφής (mlab) για τη δημιουργία PDF θεωρούμε ότι κάθε μέρος ξεκινά πάντα από μια νέα σειρά. Τέλος, ένα αρχείο περιγραφής μπορεί να περιλαμβάνει οποιοδήποτε συνδυασμό από τα βασικά στοιχεία μορφοποίησης, ενώ οι κενές γραμμές στο mlab θα πρέπει να αγνοούνται κατά τη δημιουργία του PDF. Στη συνέχεια, περιγράφεται η σύνταξη κάθε εντολής καθώς και τα αναγνωριστικά των επιμέρους παραμέτρων.

Στοιχείο	Εντολή Μορφοποίησης
Επικεφαλίδα Τύπος 1	&;Heading1 <i>fontType:font_id</i> Text here
Επικεφαλίδα Τύπος 2	&;Heading2 <i>fontType:font_id</i> Text here
Επικεφαλίδα Τύπος 3	&;Heading3 <i>fontType:font_id</i> Text here
Επικεφαλίδα Τύπος 4	&;Heading4 <i>fontType:font_id</i> Text here
Νέα Γραμμή	&;NewLine
Εικόνα	&;Image scale: value File url
Παράγραφος	&;Paragraph <i>fontSize:num</i> fontType: type_id fontStyle: style_id fontColor: color_id

	Text Here
Τροποποίηση	&;Format <i>fontSize:num</i> fontType:type_id fontStyle:style_id fontColor:color_id Text Here
Μη Διατεταγμένη λίστα	&;UnorderedList <i>fontSize:num</i> fontType:type_id fontStyle:style_id fontColor:color_id Element 1 Element N
Διατεταγμένη λίστα	&;OrderedList <i>fontSize:num</i> fontType:type_id fontStyle:style_id fontColor:color_id Element 1 Element N

Πίνακας 1 Εντολές μορφοποίησης

Font ID	Font	Type ID	Font Style	Color ID	Font Color
1	Times	1	Regular	1	Black
2	Helvetica	2	Bold	2	Blue
3	Courier	3	Italic	3	Red
		4	Bold italic	4	Yellow

Πίνακας 2 Αναγνωριστικά παραμέτρων

Στη σελίδα εργαστηρίου μπορείτε να βρείτε το αρχείο “examples.rar” που περιλαμβάνει ένα αρχείο mlab με όλες τις υποστηριζόμενες παραμέτρους μορφοποίησης καθώς και το αντίστοιχο PDF.

A Μέρος (75 %)

Θα πρέπει να σχεδιαστεί και να υλοποιηθεί ένα αυτόνομο κομμάτι λογισμικού σε μορφή βιβλιοθήκης με δυνατότητες διαχείρισης ενός ή πολλαπλών (όχι ταυτόχρονα) αρχείων κειμένου που περιέχουν τις εντολές μορφοποίησης, δημιουργίας και επεξεργασίας PDF καθώς και σύνδεσης με την υπηρεσία OMDb [2] για τη λήψη περιεχομένου. Η βιβλιοθήκη αυτή θα στηρίζεται στη βιβλιοθήκη PDFBox (έκδοση 1.8.10) [1] για τη δημιουργία και διαχείριση των αρχείων PDF.

A.1. Διαχείριση αρχείων mlab για τη δημιουργία PDF (20%)

Το συγκεκριμένο τμήμα της βιβλιοθήκης περιλαμβάνει την υλοποίηση των μεθόδων που είναι απαραίτητες για τη διαχείριση των αρχείων περιγραφής. Η βιβλιοθήκη θα πρέπει να παρέχει τις κατάλληλες μεθόδους για τις ακόλουθες λειτουργίες:

- Δημιουργία ενός αρχείου mlab (New)
- Άνοιγμα αρχείου mlab και ανάκτηση περιεχομένου (Open)
- Αποθήκευση δεδομένων σε αρχείο mlab (Save)
- Αποθήκευση περιεχομένων ενός αρχείου mlab με άλλο όνομα (Save as)



- Κλείσιμο αρχείου mlab (Close)
- Διαγραφή αρχείου mlab (Delete)
- Παροχή πληροφοριών κατάστασης για ένα αρχείο mlab (Status): μέγεθος, ημερομηνία δημιουργίας, ημερομηνία τελευταίας επεξεργασίας, όνομα χρήστη

Σε κάθε περίπτωση, η υλοποίηση που θα ακολουθήσετε θα πρέπει να εξασφαλίζει τη σωστή χρήση και αποδέσμευση των σχετικών πόρων.

A.2. Δημιουργία και διαχείριση αρχείων PDF (35%)

Στο συγκεκριμένο τμήμα της βιβλιοθήκης καλείστε να υλοποιήσετε ένα σύνολο κατάλληλων κλάσεων και τις αντίστοιχες μεθόδους που θα εξασφαλίζουν ένα επίπεδο αφαίρεσης στη χρήση της βιβλιοθήκης PDFBox για τη δημιουργία PDF με βάση τις διάφορες εντολές μορφοποίησης καθώς και την επεξεργασία αρχείων PDF.

Επιπλέον, ιδιαίτερη σημασία θα πρέπει να δοθεί όχι μόνο στην αντικειμενοστραφή οργάνωση της βιβλιοθήκης αλλά και στη σωστή διαχείριση του περιεχομένου, καθώς οι μέθοδοι που θα υλοποιήσετε θα πρέπει να χειρίζονται σωστά τα ζητήματα που σχετίζονται με τη σελιδοποίηση.

Πιο συγκεκριμένα, η βιβλιοθήκη θα πρέπει να παρέχει τις κατάλληλες μεθόδους για τις ακόλουθες λειτουργίες:

- Δημιουργία PDF με συγκεκριμένο όνομα.
- Προσθήκη νέας σελίδας σε συγκεκριμένο αρχείο PDF.
- Δημιουργία ενός αρχείου PDF με βάση τα περιεχόμενα ενός mlab αρχείου περιγραφής.
- Προσθήκη κειμένου ως επικεφαλίδα τύπου 1 με δυνατότητα προσδιορισμού της γραμματοσειράς (fontType).
- Προσθήκη κειμένου ως επικεφαλίδα τύπου 2 με δυνατότητα προσδιορισμού της γραμματοσειράς (fontType).
- Προσθήκη κειμένου ως επικεφαλίδα τύπου 3 με δυνατότητα προσδιορισμού της γραμματοσειράς (fontType).
- Προσθήκη κειμένου ως επικεφαλίδα τύπου 4 με δυνατότητα προσδιορισμού της γραμματοσειράς (fontType).
- Προσθήκη νέας γραμμής.
- Προσθήκη εικόνας με δυνατότητα ορισμού της παραμέτρου κλιμάκωσης (scale) για το μέγεθος.
- Προσθήκη ενός ή περισσοτέρων προτάσεων ως διατεταγμένη λίστα, με δυνατότητα καθορισμού όλων των παραμέτρων που υποστηρίζει.
- Προσθήκη ενός ή περισσοτέρων προτάσεων ως μη διατεταγμένη λίστα, με δυνατότητα καθορισμού όλων των παραμέτρων που υποστηρίζει.
- Προσθήκη κειμένου ως νέα παράγραφο (Paragraph εντολή μορφοποίησης), με δυνατότητα καθορισμού όλων των παραμέτρων που υποστηρίζει ο συγκεκριμένος τύπος μορφοποίησης.

- Προσάρτηση κειμένου στην τελευταία διαθέσιμη παράγραφο με δυνατότητα προσδιορισμού εναλλακτικής μορφοποίησης (Format εντολή μορφοποίησης), με δυνατότητα καθορισμού όλων των παραμέτρων που υποστηρίζει ο συγκεκριμένος τύπος μορφοποίησης.
- Συνένωση δύο ή περισσότερων αρχείων PDF σε ένα νέο αρχείο.
- Εξαγωγή όλων των σελίδων ενός αρχείου PDF και αποθήκευση κάθε σελίδας σε ξεχωριστό αρχείο.
- Εξαγωγή από ένα αρχείο PDF συγκεκριμένων συνεχόμενων σελίδων και αποθήκευση σε νέο αρχείο. Πρέπει να γίνεται έλεγχος εγκυρότητας για το ζητούμενο εύρος σελίδων.
- Διαχωρισμός ενός αρχείου PDF σε δύο επιμέρους αρχεία με βάση κάποια σελίδα, η οποία θα πρέπει να αποτελεί την τελευταία σελίδα στο πρώτο αρχείο. Πρέπει να γίνεται έλεγχος εγκυρότητας για τον αριθμό σελίδας.

A.3. Περιεχόμενο από web υπηρεσία (5%)

Τέλος, η βιβλιοθήκη θα πρέπει να παρέχει τη δυνατότητα λήψης περιεχομένου από τη web υπηρεσία OMDb [2]. Πιο συγκεκριμένα, η βιβλιοθήκη θα πρέπει να υποστηρίζει τις παρακάτω λειτουργίες:

- Λήψη με βάση το όνομα μιας σειράς ή ταινίας: του επίσημου τίτλου (*Title*), της χρονολογίας (*Year*), του είδους (*Genre*), της βαθμολογίας στο IMDB (*imdbRating*) και του πλήρους σεναρίου (*Plot*).
- Δημιουργία ενός PDF χρησιμοποιώντας τις μεθόδους που αναπτύξατε στο A.2 το οποίο θα έχει τη δομή: Τίτλος (Header 1), Τύπος – Χρονολογία – Βαθμολογία IMDB (Header 3), και Σενάριο (Paragraph).

A.4 Λοιπές απαιτήσεις (15%)

- Η βιβλιοθήκη θα πρέπει να ακολουθεί τις αρχές σχεδίασης αντικειμενοστραφούς προγραμματισμού (OOP design principles) [3].
- Κάθε public μέθοδος θα πρέπει να είναι τεκμηριωμένη [4].
- Η βιβλιοθήκη θα πρέπει να ενσωματωθεί σε ένα jar αρχείο [5].

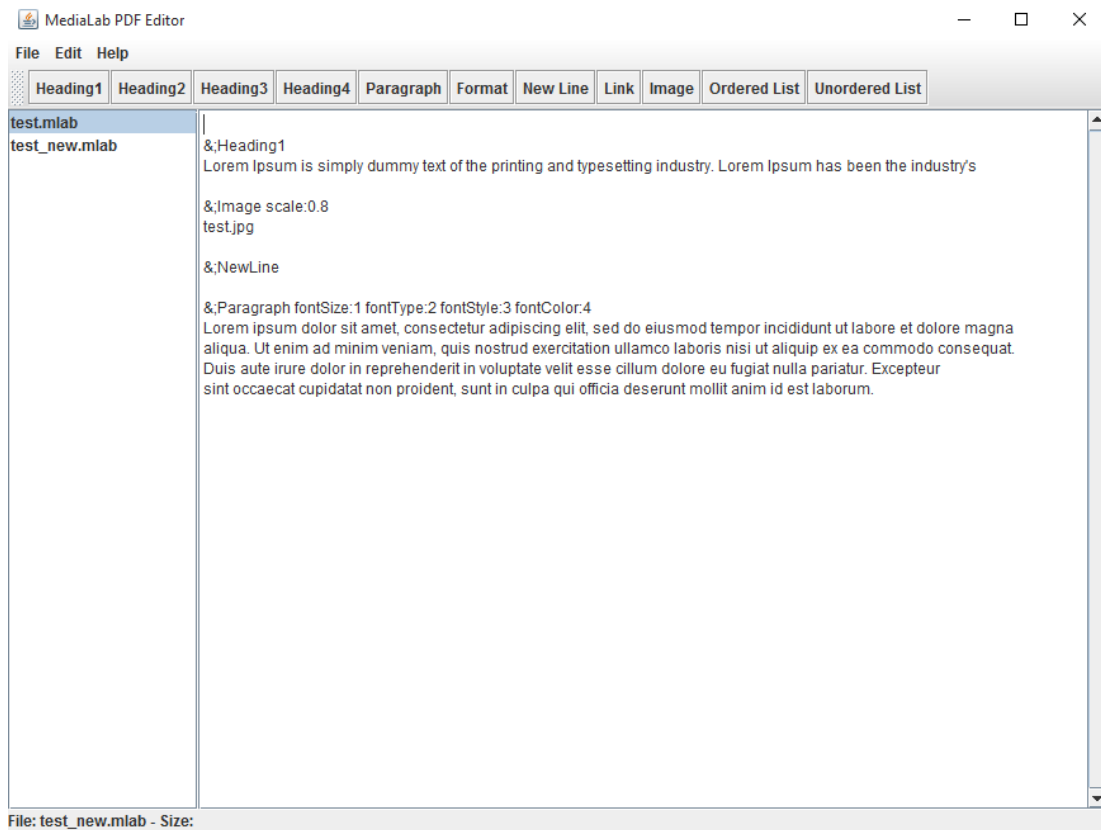
B Μέρος (25 %)

Το δεύτερο μέρος αποτελείται από το σχεδιασμό και την ανάπτυξη μιας γραφικής διεπαφής χρήστη (Graphical User Interface - GUI) χρησιμοποιώντας το swing, για την αλληλεπίδραση ενός χρήστη με τη βιβλιοθήκη που κατασκευάστηκε στο Α Μέρος.

Το GUI θα είναι πλήρως ανεξάρτητο από τη βιβλιοθήκη: το μόνο που θα κάνει είναι να χρησιμοποιεί τις μεθόδους που αυτή παρέχει. Καμία λειτουργικότητα σχετικά με τη

διαχείριση των αρχείων mlab και PDF που θα έπρεπε να υλοποιείται στη βιβλιοθήκη, δεν πρέπει να υλοποιηθεί στο GUI.

Το GUI θα μοιάζει με το παρακάτω πρότυπο:



Το GUI θα αποτελείται από τα εξής components:

1. Ένα menu bar που θα αποτελείται από:
 - a. File
 - i. New: Δημιουργία ενός νέου αρχείου mlab. Θα πρέπει να υπάρχει ένα popup παράθυρο για τον ορισμό του ονόματος, ενώ θα πρέπει να ενημερώνεται κατάλληλα η λίστα των διαθέσιμων αρχείων.
 - ii. Open: ένα ή πολλαπλά αρχεία mlab και εμφάνιση του περιεχομένου στο component επεξεργασίας. Θα πρέπει να υπάρχει ένα φίλτρο για την αποδοχή μόνο mlab αρχείων. Τα επιλεγμένα αρχεία πρέπει να τοποθετούνται στη λίστα των διαθέσιμων αρχείων. Στην περίπτωση πολλαπλών αρχείων να φορτώνονται τα περιεχόμενα από το αρχείο που τοποθετήθηκε τελευταίο στη λίστα.
 - iii. Save: Αποθήκευση των περιεχομένων του component επεξεργασίας στο τρέχον ανοιχτό αρχείο.
 - iv. Save As: Αποθήκευση των περιεχομένων του component επεξεργασίας σε ένα νέο αρχείο mlab. Θα πρέπει να υπάρχει ένα popup παράθυρο για τον ορισμό του ονόματος και θα πρέπει να ενημερώνεται η λίστα των διαθέσιμων αρχείων.

- v. Close: Θα πρέπει να κλείνει το τρέχον αρχείο που επεξεργαζόμαστε και να αποδεσμεύονται οποιοδήποτε πόροι.
 - vi. Exit: Η εφαρμογή θα τερματίζει. Κάθε πόρος που χρησιμοποιούνταν θα πρέπει να αποδεσμεύεται.
- b. Edit
- i. Create PDF: Δημιουργία αρχείου PDF με βάση την περιγραφή μορφοποίησης που περιέχει το αρχείο mlab που είναι ανοιχτό στην εφαρμογή.
 - ii. Online Content: Μέσω ενός popup παραθύρου θα γίνεται ο καθορισμός του τίτλου για την υπηρεσία OMDb, στη συνέχεια θα πρέπει το περιεχόμενο να φορτώνεται στο component επεξεργασίας με βάση τις εντολές μορφοποίησης και τη δομή που περιγράφεται στο Α.3.
 - iii. Online PDF: Μέσω ενός popup παραθύρου θα γίνεται ο καθορισμός του τίτλου για την υπηρεσία OMDb, στη συνέχεια θα πρέπει να δημιουργείται αυτόματα ένα PDF με βάση το περιεχόμενο που λαμβάνουμε το οποίο θα ακολουθεί τη δομή που περιγράφεται στο Α.3.
 - iv. Merge PDFs: Μέσω ενός popup παραθύρου θα γίνεται η επιλογή δύο ή περισσότερων αρχείων, ο καθορισμός του τελικού ονόματος και η εκτέλεση της συνένωσης.
 - v. Extract All Pages: Μέσω ενός popup παραθύρου θα γίνεται η επιλογή ενός αρχείου PDF και η εκτέλεση της εξαγωγής όλων των σελίδων και αποθήκευσης κάθε μίας σε ξεχωριστό αρχείο.
 - vi. Extract Pages: Μέσω ενός popup παραθύρου θα γίνεται η επιλογή ενός αρχείου PDF, το εύρος των συνεχόμενων σελίδων που θέλουμε να εξαχθούν καθώς και η εκτέλεση της λειτουργίας.
 - vii. Split PDF: Μέσω ενός popup παραθύρου θα γίνεται η επιλογή ενός αρχείου PDF, η σελίδα βάσει της οποίας θα γίνει ο διαχωρισμός σε δύο επιμέρους αρχεία καθώς και η εκτέλεση της λειτουργίας.
- c. Help (προαιρετικό)
- i. Θα εμφανίζει ένα παράθυρο about, με πληροφορίες για την εφαρμογή και το δημιουργό της.
2. Ένα tool bar που θα περιλαμβάνει ένα κουμπί για κάθε εντολή μορφοποίησης (Πίνακας 1) με σκοπό να γράφουν τη γενική σύνταξη της εντολής στο component επεξεργασίας.
 3. Ένα list για την εμφάνιση των διαθέσιμων προς επεξεργασία αρχείων mlab: (a) τα περιεχόμενα του θα πρέπει να ανανεώνονται κατάλληλα για τις λειτουργίες Open, New, Save As (b) όταν επιλέγουμε κάποιο αρχείο θα πρέπει να γίνεται το προς επεξεργασία αρχείο και τα περιεχόμενα του να εμφανίζονται στο αντίστοιχο component.
 4. Ένα κατάλληλο component για την εμφάνιση και επεξεργασία των περιεχομένων ενός αρχείου mlab.
 5. Ένα label για την εμφάνιση πληροφοριών σχετικά με το αρχείο που επεξεργάζεται κάθε φορά ο χρήστης και μια ένδειξη για την τελευταία ενέργεια.(π.χ. «File test.mlab created», «File test.mlab saved. Size: 1.2Kb».

Για την ανάπτυξη του GUI δεν συνιστάται (αλλά και δεν απαγορεύεται) η χρήση κάποιου εργαλείου για την αυτοματοποίηση του σχεδιασμού του.

Παραδοτέα

- Το project (του IDE της επιλογής σας) με τη βιβλιοθήκη του Α μέρους.
- Το jar αρχείο της βιβλιοθήκης.
- Το project (του IDE της επιλογής σας) με το GUI και το jar αρχείο που δημιουργήθηκε από το πρώτο project. Σε αυτό το project δεν πρέπει να υπάρχει κώδικας της βιβλιοθήκης (μόνο binaries).
- Μια σύντομη αναφορά με το τι κάνατε και πως καθώς και τι δεν κάνατε και γιατί.

Αναφορές

- [1] <https://pdfbox.apache.org/index.html>
- [2] <http://omdbapi.com/>
- [3] <http://www.oodeesign.com/design-principles.html>
- [4] <http://www.oracle.com/technetwork/java/javase/documentation/index-137868.html>
- [5] <https://docs.oracle.com/javase/tutorial/deployment/jar/build.html>
- [6] <https://docs.oracle.com/javase/tutorial/uiswing/components/index.html>