

## ΤΟ ΥΠΟ ΑΝΑΠΤΥΞΗ ΠΡΟΤΥΠΟ PDF/A ΑΝΑΚΑΛΥΠΤΟΝΤΑΣ ΤΑ ΟΦΕΛΗ ΤΗΣ ΤΕΧΝΟΛΟΓΙΑΣ PDF ΣΤΗΝ ΑΡΧΕΙΟΘΕΤΗΣΗ ΕΓΓΡΑΦΩΝ

**Ευάγγελος Κόκκορης, B.A., M.Sc.**

Εργαστηριακός Συνεργάτης, Τμ. Φωτ/φίας, Σχ. Γρ. Τεχν. & Καλ/ών Σπουδών, Τ.Ε.Ι. Αθήνας

### ΠΕΡΙΛΗΨΗ

Μια πρόσφατη πρωτοβουλία εμπλεκόμενων φορέων προτίθεται να αναπτύξει ένα διεθνές πρότυπο που θα καθορίζει τη χρήση του Portable Document Format (PDF) ως μέσου για την αρχειοθέτηση και τη συντήρηση των εγγράφων. Το πρότυπο αυτό, καλούμενο PDF/A, θα καλύψει την αυξανόμενη ανάγκη να αρχειοθετηθούν ηλεκτρονικά, έγγραφα τα οποία θα μπορούν να περιέχουν κείμενα, εικόνες και διανυσματικά γραφικά.

Πολλοί άνθρωποι συγχέουν το PDF-τη μορφή αρχείου-με την πλατφόρμα λογισμικού Adobe Acrobat. Θα διευκρινίσουμε ότι το PDF είναι μια ελεύθερα διαθέσιμη προδιαγραφή και θα δείξουμε πώς η Adobe ενθαρρύνει τους ενδιαφερόμενους φορείς να αναπτύξουν λύσεις εκτός της οικογένειας προϊόντων Acrobat. Σκοπεύουμε να παρουσιάσουμε την ιστορία του PDF και να προσδιορίσουμε το PDF ως de facto πρότυπο για τη διαχείριση και την αρχειοθέτηση όλων των τύπων εγγράφων.

Θεωρούμε το PDF ως ένα αποτελεσματικό πλαίσιο για την δημιουργία και διαχείριση αξιόπιστων, προσβάσιμων και προστατευόμενων ψηφιακών συλλογών εγγράφων και αρχείων. Περιγράφουμε τα κύρια χαρακτηριστικά ενός εγγράφου PDF το οποίο αποτελείται από μια ακολουθία σελίδων όπου κάθε σελίδα συμπεριλαμβάνει το κείμενο, τα τυπογραφικά χαρακτηριστικά, την οπτική διαμόρφωση και τα γραφικά στοιχεία. Μπορεί επίσης να συμπεριλάβει άλλα στοιχεία όπως μεταδεδομένα, συνδέσεις υπερκειμένου, και δυναμικά πεδία. Στοιχεία μεταδεδομένων καθορίζονται στο προτεινόμενο πλαίσιο εργασίας Dublin Core μπορούν επίσης να συμπεριληφθούν. Με αυτό το διαθέσιμο σύνολο πληροφοριών ένα αρχείο PDF μπορεί να απεικονίζεται αξιόπιστα στην οθόνη και στις συσκευές εκτύπωσης. Τα αρχεία PDF μπορούν επίσης να υποστηρίξουν κρυπτογράφηση δεδομένων και να διαθέτουν προσαρμοσμένα χαρακτηριστικά ελέγχου της πρόσβασης και των επιτρεπόμενων ενεργειών από τρίτους.

Αυτό το έγγραφο προσδιορίζει τις απαιτήσεις για τη δημιουργία ηλεκτρονικών συλλογών εγγράφων και αρχείων, καθώς και το ρόλο των μεταδεδομένων σε αυτόν τον τομέα και εξηγεί πώς το PDF μπορεί να χρησιμοποιηθεί ως σχήμα αρχειοθέτησης για τα υπάρχοντα έντυπα και για πολλές μορφές ψηφιακού περιεχομένου (συμπεριλαμβανομένων αρχείων ήχου, εικόνας και βίντεο). Οδηγίες δημιουργίας και μετατροπής εγγράφων θα παρασχεθούν.

**Λέξεις κλειδιά:** *Portable Document Format (PDF), ψηφιακές αρχειακές συλλογές, PDF/A, διεθνές πρότυπο, μακροπρόθεσμη φύλαξη, προστασία, ευρετηρίαση, αναζήτηση, μεταδεδομένα, Extensible Metadata Platform (XMP)*

## **PDF/A AS AN EMERGING ISO STANDARD UNVEILING THE BENEFITS OF PDF TECHNOLOGY FOR DOCUMENT ARCHIVE AND PRESERVATION**

**Evangelos Kokkoris, B.A., M.Sc.**

Part-time Lecturer at the Dept. of Photography, Faculty of Graphic Arts & Design,  
T.E.I. of Athens

### **ABSTRACT**

A new joint activity has been initiated to develop an international standard that defines the use of the Portable Document Format (PDF) for archiving and preserving documents. The project, currently referred to as PDF/A, will address the growing need to electronically archive documents that may contain a mixture of text, raster images and vector graphics.

Many people confuse PDF, the data format, with Adobe Acrobat, the software suite that Adobe sells to create, view and enhance PDF documents. We shall clarify that PDF is a publicly available specification and will show how Adobe encourages third-parties to further develop solutions outside of the Acrobat product family. We intend to present the history of PDF and identify PDF as a de facto standard for managing and archiving of all types of documents and forms.

We present the PDF as an effective framework for establishing and maintaining a reliable, accessible, and secure digital archive. We describe the main characteristics of a PDF document; it consists of a sequence of pages, with each page including the text, font specifications, margins, layout, graphical elements, and background and text colours. It can also include other items such as metadata, hyperlinks, and form fields. Dublin Core metadata elements can be contained within PDF files. With all this information present the PDF file can be imaged accurately for the screen and the printing device. PDF files can also be secured with password protection and encryption.

This paper also identifies the requirements for electronic archives, as well as the role of metadata in this field, and how PDF can be used as an archiving format for legacy paper documents and all types of digital content (including voice, images and video). Document acquisition and conversion guidelines will be provided.

**Keywords:** Portable Document Format (PDF), electronic archives, PDF/A, International Standard, long-term preservation, security, metadata, indexing, searching, Extensible Metadata Platform(XMP)

## 1. Εισαγωγή

Το PDF είναι ένα ψηφιακό σχήμα για την αναπαράσταση των εγγράφων, είτε αυτά δημιουργούνται εγγενώς σε μορφή PDF, μετατρέπονται σε PDF από άλλες ηλεκτρονικές μορφές, ή ψηφιοποιούνται απευθείας από έντυπα ή μικροφίλμ. Η ονομασία Portable Document Format, ή PDF, επιλέχθηκε για να διευκρινίσει ότι ένα αρχείο που ακολουθεί αυτήν την προδιαγραφή μπορεί να απεικονισθεί και να εκτυπωθεί σε οποιοδήποτε πλατφόρμα-όπως Unix, Mac OS, Microsoft Windows, καθώς και σε διάφορες συσκευές χειρός-με την ίδια πιστότητα. (Adobe, 2003c)

Η προέλευση του Portable Document Format και της οικογένειας προϊόντων Acrobat χρονολογείται στις αρχές του 1990. (Adobe, 1993b) Κατά τη διάρκεια των προηγούμενων δέκα ετών, υποβοηθούμενο από την εκρηκτική εξάπλωση του Διαδικτύου, το PDF έχει γίνει το de facto πρότυπο για την ηλεκτρονική ανταλλαγή των εγγράφων<sup>1</sup>. Εκατομμύρια αντίγραφα της δωρεάν εφαρμογής Adobe Reader έχουν διανεμηθεί σε όλο τον κόσμο, κάτι που διευκολύνει την αποτελεσματική διακίνηση του ψηφιακού περιεχομένου<sup>2</sup>. Καθώς μεγάλες εταιρίες, κυβερνητικές υπηρεσίες, και εκπαιδευτικά ιδρύματα εκσυγχρονίζουν τις διαδικασίες τους και αντικαθιστούν την παραδοσιακή-βασισμένη στο έντυπο-ροή εργασίας με την ηλεκτρονική ανταλλαγή των πληροφοριών, ο αντίκτυπος και οι χρήσεις της τεχνολογίας PDF θα συνεχίσουν να αυξάνονται με γρήγορο ρυθμό.

## 2. Συλλογή και φύλαξη της πληροφορίας

### 2.1. Ζητήματα διαχείρισης των ψηφιακών εγγράφων

Ιδιώτες, επιχειρήσεις, κυβερνητικές υπηρεσίες, βιβλιοθήκες, υπηρεσίες αρχείων, και άλλοι φορείς σε όλο τον κόσμο χρησιμοποιούν το PDF για να προσφέρουν αξιοσημείωτο όγκο σημαντικών πληροφοριών. (Sinha, 2003) Ένα μεγάλο μέρος αυτών των πληροφοριών πρέπει να διατηρηθεί για μεγάλα χρονικά διαστήματα και ένα επιπλέον μέρος πρέπει να διατηρηθεί μόνιμα. Αυτά τα έγγραφα πρέπει να παραμείνουν αξιοποιήσιμα και να είναι προσβάσιμα μέσω των σημερινών αλλά εξίσου και των μελλοντικών εφαρμογών της τεχνολογίας. Οι μελλοντικές χρήσεις και η πρόσβαση σε αυτά τα έγγραφα εξαρτώνται από τη διατήρηση της δυνατότητας να αναπαραχθεί η οπτική εμφάνιση τους καθώς επίσης και από άλλες ιδιότητες, όπως είναι η λογική οργάνωση των σελίδων, τμημάτων, και παραγράφων, η δυνατότητα ανάκτησης του κειμένου με αυτόματα μέσα, και η υποστήριξη μεταδεδομένων με σκοπό τη διαχείριση, τη συντήρηση, και την περιγραφή της υπάρχουσας πληροφορίας. Για τον αναγνώστη που θέλει να εμβαθύνει, η Adobe διαθέτει ένα εξαιρετικό έγγραφο, το οποίο πραγματεύεται ζητήματα σχετικά με την ψηφιακή αρχειοθέτηση. (Adobe, 2003b)

Υπάρχουν πολλές ηλεκτρονικές μορφές και τεχνολογίες για να επιλέξει κανείς ως λύση αρχειοθέτησης. Αυτές περιλαμβάνουν το ASCII (για το κείμενο), το TIFF, το PDF, και τη γλώσσα XML-χωρίς να αναφερθούμε σε εφαρμογές επεξεργασίας κειμένου, λογιστικών φύλλων, και σε άλλες μορφές. Η εγγενής φύση μερικών από αυτές τις ηλεκτρονικές μορφές οδηγεί στην αμφιβολία για το εάν μπορούν να εγγυηθούν πως θα συνεχίσουν να υφίστανται σε βάθος χρόνου. Μόνο μια από αυτές τις μορφές δείχνει να ικανοποιεί σημαντικά την ανάγκη διατήρησης τόσο της εμφάνισης, όσο και του περιεχομένου κατά τη διάρκεια μιας μεγάλης χρονικής περιόδου. Το PDF συλλαμβάνει όχι μόνο τις πληροφορίες που περιλαμβάνονται στο έγγραφο αλλά και την ακριβή

μορφή που το έγγραφο πήρε. Το αρχείο μπορεί να απεικονίζεται χωρίς την εφαρμογή που το δημιούργησε. Στην πραγματικότητα, δέκα χρόνια από τώρα, στο μέλλον, οι χρήστες θα είναι ακόμα σε θέση να δουν το αρχείο ακριβώς όπως δημιουργήθηκε<sup>3</sup>. Με την προσθήκη των μεταδεδομένων μορφής XML στο αρχείο PDF, μπορούμε να διατηρήσουμε εξίσου την πιστότητα και την πρόσβαση στο περιεχόμενο. Και επειδή το PDF είναι μια δημόσια διαθέσιμη προδιαγραφή, οι πληροφορίες για αυτή τη μορφή αρχείου θα είναι πάντα προσπές, κάτι που κάνει το PDF μια πολύ ελκυστική επιλογή για την οργάνωση ηλεκτρονικών αρχείων.

## **2.2. Τρέχουσες διεθνείς δραστηριότητες**

Μια νέα κοινή δραστηριότητα έχει αρχίσει μεταξύ του NPES<sup>4</sup>, "The Association for Suppliers of Printing, Publishing and Converting Technologies", και του Association for Information and Image Management, International για να αναπτύξει ένα διεθνές πρότυπο που καθορίζει τη χρήση του Portable Document Format (PDF) για την αρχειοθέτηση και τη συντήρηση των εγγράφων. (AIIM, 2002) Αυτή η μικτή επιτροπή θα προσδιορίσει τα ζητήματα που αντιμετωπίζονται, καθώς επίσης θα συνοψίσει τις προτεινόμενες λύσεις, και θα αναπτύξει ένα σχέδιο εργασίας που θα παρουσιαστεί έπειτα σε μια κοινή ομάδα εργασίας του διεθνούς οργανισμού τυποποίησης (ISO) για περαιτέρω επεξεργασία και έγκριση ως διεθνές πρότυπο. Ήδη, η πέμπτη έκδοση ενός προσχεδίου με την ονομασία ISO/WD 19005, "Document management-Electronic document file format for long-term preservation-Use of PDF (PDF/A)", είναι διαθέσιμη για αναθεώρηση και σχόλια. (ISO, 2003)

Κάθε ενδιαφερόμενος, που είναι ικανός να συμβάλει στο έργο της επιτροπής, καλείται και ενθαρρύνεται για να συμμετέχει ενεργά σε αυτήτην εργασία. Το πρόγραμμα, αυτή την περίοδο καλούμενο PDF/A, θα καλύψει την αυξανόμενη ανάγκη να αρχειοθετηθούν ηλεκτρονικά τα έγγραφα με έναν τρόπο που θα εξασφαλίσει συντήρηση του περιεχομένου τους κατά τη διάρκεια μιας εκτεταμένης χρονικής περιόδου, και περαιτέρω θα εξασφαλίσει ότι εκείνα τα έγγραφα θα είναι σε θέση να ανακτηθούν και να δοθούν με ένα συνεπές και προβλέψιμο αποτέλεσμα στο μέλλον. Αυτή η ανάγκη υπάρχει σε έναν συνεχώς αυξανόμενο αριθμό διεθνών κυβερνητικών οργανισμών και επιχειρήσεων του ιδιωτικού τομέα, συμπεριλαμβανομένων νομικών υπηρεσιών, βιβλιοθηκών, εφημερίδων, και άλλων.

## **3. Η τεχνολογία Adobe Portable Document Format (PDF)**

### **3.1. Ιστορική ανασκόπηση**

Στη δεκαετία του '80, η Adobe επινόησε ένα ισχυρό μοντέλο γραφικής απεικόνισης που με την πάροδο του χρόνου αποτέλεσε τη βάση για τις τεχνολογίες Adobe PostScript. Αυτές οι τεχνολογίες-έναν συνδυασμός μεταξύ της γλώσσας PostScript και των βασισμένων σε αυτή εφαρμογών γραφικής σχεδίασης και μορφοποίησης κειμένου, των οδηγών συσκευών, και των συστημάτων απεικόνισης-έχουν αλλάξει για πάντα τον κόσμο των γραφικών τεχνών και των εκδόσεων προκαλώντας εξελίξεις που χαρακτηρίστηκαν ως επαναστάσεις της επιτραπέζιας τυπογραφίας και της ηλεκτρονικής σελιδοποίησης.

Εκείνη την περίοδο, δεν υπήρξε κανένα αποτελεσματικό μοντέλο περιγραφής σελίδας, οι περισσότερες γραμματσειρές στοιχειοθετούνταν με τη χρήση εξειδικευμένου εξοπλισμού, η πάρα-

γωγή υψηλής ποιότητας εκδόσεων περιοριζόταν στους ειδικούς, και το κόστος παραγωγής για τα περισσότερα έντυπα ήταν απαγορευτικό. Η εφεύρεση της γλώσσας PostScript βοήθησε αρκετά στην επίλυση αυτού του προβλήματος. Αποτέλεσε επιπλέον, έναν από τους καταλυτικούς παράγοντες που οδήγησαν τους επαγγελματίες και τις επιχειρήσεις να κάνουν τη μετάβαση από τις γραφομηχανές στους προσωπικούς υπολογιστές.

Μια γλώσσα περιγραφής σελίδων όφειλε να επιτρέπει στις εφαρμογές να παράγουν αρχεία που να είναι σχετικά συμπαγή για την αποθήκευση και τη μετάδοση, καθώς και να είναι ανεξάρτητα της οποιασδήποτε συσκευής εκτύπωσης. Η γλώσσα PostScript είναι μια απλή ερμηνευτική γλώσσα προγραμματισμού με ισχυρές ικανότητες διαχείρισης γραφικών. (Adobe, 1993b) Η κύρια εφαρμογή της είναι να περιγράψει την εμφάνιση του κειμένου, των γραφικών, και των εικόνων που συνθέτουν το προς εκτύπωση έγγραφο, σύμφωνα με το μοντέλο γραφικής απεικόνισης της Adobe. Ένα πρόγραμμα σε αυτήν την γλώσσα μπορεί να διαβιβάσει μια περιγραφή ενός εγγράφου από ένα σύστημα σχεδίασης σε ένα σύστημα εκτύπωσης ή να ελέγξει την εμφάνιση του κειμένου και των γραφικών σε μια οθόνη. Η περιγραφή είναι υψηλού επιπέδου και ανεξάρτητη συσκευής. Αυτό επέτρεψε στους εκτυπωτές γραφείου να αναπαράγουν σύνθετες μορφές κειμένων και γραφικών. Για πρώτη φορά, οποιοσδήποτε διαθέτετε έναν υπολογιστή θα μπορούσε να παράγει υψηλών προδιαγραφών εκδοτικά αποτελέσματα.

Με την υιοθέτηση τους, οι τεχνολογίες PostScript επέτρεψαν πρωτοφανή έλεγχο των εργασιών της εκτυπωτικής παραγωγής, ενώ επέφεραν σημαντικές αλλαγές στις πρακτικές σχεδίασης και εκτύπωσης των εγγράφων. Σήμερα, περισσότερο από ποτέ, κατασκευαστές εφαρμογών λογισμικού και συστημάτων ψηφιακής απεικόνισης παρέχουν υποστήριξη μέσω των προϊόντων τους, καθιερώνοντας έμπρακτα την PostScript ως πρότυπο στο χώρο των επαγγελματικών εκτυπώσεων.

### **3.2. Οι σύγχρονες ανάγκες**

Εν τω μεταξύ, οι προσδοκίες παραγωγής ψηφιακών εγγράφων αυξάνονταν και σε χώρους πέραν των επαγγελματικών εκδόσεων. Το χαμηλό κόστος των προσωπικών υπολογιστών, έχει οδηγήσει στη ραγδαία εξάπλωση ικανών συστημάτων στο σπίτι και στο καθημερινό περιβάλλον γραφείου. Η προσομοίωση των γραφομηχανών (παραγωγή απλού κειμένου σε μια ενιαία γραμματοσειρά) δεν είναι πλέον επαρκής. Οι χρήστες θέλουν να δημιουργούν, να απεικονίζουν, και να εκτυπώνουν έγγραφα που συνδυάζουν τυπογραφικά στοιχεία και περίπλοκα γραφικά. Επιπλέον, η ταυτόχρονη ανάπτυξη αποτελεσματικών πρωτοκόλλων επικοινωνίας, οδήγησε στην άνθιση του Διαδικτύου, κάτι που ενίσχυσε την επιθυμία της αξιόπιστης ανταλλαγής αυτών των ψηφιακών εγγράφων.

"Το πρόβλημα συνοψίζεται στη δυνατότητα μας να επικοινωνήσουμε το οπτικό υλικό μεταξύ των διαφορετικών εφαρμογών και συστημάτων υπολογιστών. Το συγκεκριμένο πρόβλημα είναι ότι τα περισσότερα προγράμματα τυπώνουν σε ένα ευρύ φάσμα εκτυπωτών, αλλά δεν υπάρχει κανένας καθολικός τρόπος να μεταβιβαστούν και να απεικονιστούν αυτές οι πληροφορίες ηλεκτρονικά!...] Φανταστείτε να μπορεί κάποιος να στείλει πλήρη έγγραφα κειμένου και γραφικών (εφημερίδες, άρθρα περιοδικών, τεχνικά εγχειρίδια κλπ.) μέσα από τα δίκτυα διανομής ηλεκτρονικού ταχυδρομείου. Αυτή η ικανότητα θα άλλαζε αληθινά τον τρόπο που διαχειριζόμαστε την πληροφορία". (Warnock, 1991)

Τον προβληματισμό, ακολουθεί η πρώτη περιγραφή των ιδιοτήτων και χαρακτηριστικών αυτού που

έμελλε να γίνει γνωστό ως PDF. Αξίζει να διαβάσει κανείς, για να διαπιστώσει πόσο διεισδυτική, ήταν η πρόβλεψη του συγγραφέα -περισσότερο από δέκα χρόνια πριν- ως προς τις σύγχρονες ανάγκες διαχείρισης των εγγράφων. Δέκα χρόνια αργότερα, εφημερίδες όπως η Ισπανική EL PAIS, και η δική μας Καθημερινή, δημοσιεύουν καθημερινά ηλεκτρονικά φύλλα σε μορφή PDF. (IHT, Αγν.)

### **3.3. Η προτεινόμενη λύση**

Στο παράδειγμα της PostScript, αυτό το πρόβλημα της "απεικόνισης και εκτύπωσης οπουδήποτε", έχει ήδη αντιμετωπιστεί και επιλυθεί. Η γενικευμένη αποδοχή της PostScript ως επαγγελματικού προτύπου, την ταξινομεί ως έναν υποψήφιο φορέα για αυτή την ανταλλαγή ηλεκτρονικών εγγράφων. Ο λόγος που οι τεχνολογίες PostScript δεν είναι μια συνολική λύση στο σημερινό τοπίο είναι ότι αυτή η λύση απαιτεί ισχυρούς υπολογιστές γραφείου και εκτυπωτές PostScript, κάτι που για το μέσο χρήστη συνεχίζει να συνεπάγεται υψηλό κόστος. Η Adobe έχει καθορίσει ένα άλλο σχήμα, αποκαλούμενο PDF, για τη διακίνηση των ηλεκτρονικών εγγράφων.

Το Adobe Portable Document Format (PDF) είναι μια μορφή αρχείου για την αναπαράσταση των εγγράφων με έναν τρόπο ανεξάρτητο από τις εφαρμογές λογισμικού, τις συσκευές, και τα λειτουργικά συστήματα που χρησιμοποιούνται για τη δημιουργία τους, καθώς και από τις επιμέρους συσκευές εξόδου στις οποίες πρόκειται να απεικονιστούν ή να τυπωθούν. Ένα αρχείο PDF σχεδιάζεται για να είναι φορητό σε όλες τις πλατφόρμες και τα λειτουργικά συστήματα. Το PDF και η γλώσσα PostScript μοιράζονται το ίδιο μοντέλο γραφικής απεικόνισης της Adobe. Στην καρδιά του PDF είναι η δυνατότητα του να περιγράψει την εμφάνιση περίπλοκων γραφικών και τυπογραφικών χαρακτηριστικών. Αυτό επιτυγχάνεται μέσω της χρήσης του μοντέλου γραφικής απεικόνισης της Adobe, της ίδιας υψηλού επιπέδου, ανεξάρτητης από συσκευή αναπαράστασης που χρησιμοποιείται στη γλώσσα περιγραφής σελίδων PostScript. Ένα έγγραφο μπορεί να μετατραπεί απευθείας μεταξύ της μορφής PDF και της γλώσσας PostScript, οι δύο αναπαραστάσεις παράγουν το ίδιο αποτέλεσμα όταν τυπώνεται. Ωστόσο, ένα έγγραφο PDF διαθέτει προσαρμοσμένα χαρακτηριστικά, που του επιτρέπουν να συνεργάζεται και με μη επαγγελματικές συσκευές, όπως είναι αυτές που συναντώνται στο σπίτι, ή στο καθημερινό περιβάλλον γραφείου.

### **3.4. Βασικά χαρακτηριστικά των εγγράφων PDF**

Οι σελίδες ενός εγγράφου μπορούν να περιέχουν οποιοδήποτε συνδυασμό κειμένου, γραφικών, και εικόνων. Εκτός από την περιγραφή της στατικής εμφάνισης των σελίδων, ένα έγγραφο PDF μπορεί να συμπεριλάβει διαδραστικά στοιχεία που είναι δυνατά μόνο σε μια ηλεκτρονική αναπαράσταση, όπως είναι οι ηλεκτρονικές φόρμες. Το PDF υποστηρίζει επίσης εκτεταμένες λειτουργίες σχολιασμού όπως σημειώσεις, επισημάνσεις, και υποδείξεις διορθώσεων κειμένου. Επιπλέον, παρέχει τη δυνατότητα εισαγωγής επισυναπτόμενων αρχείων και υπερσυνδέσεων, καθώς επίσης προσφέρει εκτεταμένη υποστήριξη αναπαραγωγής ήχου και βίντεο. Ένα έγγραφο μπορεί να καθορίσει δυναμικά το περιβάλλον εργασίας μέσα από το οποίο θα απεικονίζεται, ενώ το πληκτρολόγιο και το ποντίκι μπορούν να προκαλέσουν ενέργειες που καθορίζονται στο PDF. Το έγγραφο μπορεί να περιέχει δυναμικά πεδία που συμπληρώνονται από το χρήστη, και μπορεί να εξαγάγει τις τιμές αυτών των πεδίων, ή να τις εισαγάγει από άλλες εφαρμογές. Τέλος, ένα έγγραφο PDF μπορεί να περιέχει υψηλότερου επιπέδου πληροφορίες -όπως είναι η ενσωμάτω-

ση μεταδεδομένων- που είναι χρήσιμες για την ανταλλαγή του περιεχομένου μεταξύ των εφαρμογών. Εκτός από την περιγραφή και την απόδοση της εμφάνισης, το περιεχόμενο ενός εγγράφου μπορεί να ενσωματώσει λογικές πληροφορίες που επιτρέπουν λειτουργίες ευρετηρίασης, αναζήτησης, ανάκτησης, ή ακόμη και εξαγωγής της πληροφορίας για την αξιοποίηση της αλλού. Για να προστατευτεί αυτό το εύρος πληροφοριών, το PDF διαθέτει ισχυρή κρυπτογράφηση και υποστηρίζει λειτουργίες διαχείρισης δικαιωμάτων πρόσβασης.

### **3.5. Η προδιαγραφή PDF**

Πολλοί άνθρωποι συγχέουν το PDF, τη μορφή αρχείου, με το Adobe Acrobat<sup>6</sup>, την εφαρμογή λογισμικού που η Adobe διαθέτει προς πώληση και με την οποία κάποιος μπορεί να παράγει, να αναγνώσει, και να επεξεργαστεί έγγραφα της μορφής PDF. Στην πραγματικότητα η οικογένεια προϊόντων Adobe Acrobat αριθμεί περισσότερα από δώδεκα μέλη, που καλύπτουν ένα εξαιρετικά ευρύ φάσμα επαγγελματικών αναγκών<sup>7</sup>.

Η προδιαγραφή της μορφής αρχείου PDF δημοσιεύθηκε αρχικά το 1993, συγχρόνως τα πρώτα προϊόντα Acrobat εμφανίστηκαν. (Adobe, 1993a) Από τότε, ενημερωμένες εκδόσεις της προδιαγραφής είναι διαθέσιμες από την Adobe μέσω του Διαδικτύου<sup>8</sup>. Το PDF είναι μια δημόσια διαθέσιμη προδιαγραφή, ανεξάρτητα από το γεγονός ότι η Adobe την δημιούργησε και συνεχίζει να την εξελίσει μέσω των ενημερωμένων εκδόσεων. Η τέταρτη έκδοση της προδιαγραφής περιγράφει την έκδοση PDF 1.5, η οποία τεκμηριώνεται στο Portable Document Format Reference Manual<sup>9</sup>. Όπως και τα προηγούμενα, αυτό το βιβλίο είναι απολύτως ανεξάρτητο, συνδυάζοντας πλήρεις περιγραφές του θεμελιώδους μοντέλου γραφικής απεικόνισης μαζί με τα ειδικά χαρακτηριστικά που προσδιορίζουν την έκδοση 1.5 της προδιαγραφής PDF.

### **3.6. Υποστήριξη και εργαλεία ανάπτυξης**

Η πρόθεση της Adobe είναι να διατηρηθεί η ακεραιότητα της προδιαγραφής του Portable Document Format. Εντούτοις, η Adobe επιθυμεί να προωθήσει τη χρήση του Portable Document Format για την ανταλλαγή πληροφοριών μεταξύ διαφορετικών συστημάτων και εφαρμογών<sup>10</sup>. Για την εξυπηρέτηση αυτού του σκοπού, επιτρέπεται υπό όρους η χρήση των-προστατευόμενων από πνευματικά δικαιώματα-δομών δεδομένων (Data Structures) και δειγμάτων κώδικα (Sample Code) που τεκμηριώνονται στο εγχειρίδιο της προδιαγραφής. Κατ' αυτόν τον τρόπο οποιοσδήποτε μπορεί να αναπτύξει εφαρμογές λογισμικού το οποίο θα μπορεί να παρέχει τις σχετικές λειτουργίες δημιουργίας, απεικόνισης, εκτύπωσης, καθώς και οποιασδήποτε άλλης μορφής παρέμβασης στο περιεχόμενο των αρχείων PDF. Επιπλέον, για να εξασφαλίσει πως η προδιαγραφή μπορεί να αξιοποιηθεί από τρίτους κατασκευαστές λογισμικού, η Adobe διαθέτει ένα Kit Ανάπτυξης Εφαρμογών (SDK).

Ολοκληρωμένες λύσεις μπορούν να αναπτυχθούν εκτός της οικογένειας προϊόντων Acrobat, ή ακόμη τα ίδια τα προϊόντα Acrobat μπορούν να ενισχυθούν με την ανάπτυξη κατάλληλων επεκτάσεων (plug-ins). Κατασκευαστές λογισμικού, έχουν χρησιμοποιήσει την προδιαγραφή ακόμη και για τη δημιουργία των δικών τους εφαρμογών δημιουργίας PDF<sup>11</sup>. Επιπλέον, υπάρχουν περισσότεροι από 1.800 προμηθευτές που προσφέρουν λύσεις βασισμένες σε PDF, συμπεριλαμβανομένων των συμβουλευτικών υπηρεσιών, καθώς και των υπηρεσιών κατάρτισης και υποστήριξης<sup>12</sup>.

## **4. Οργάνωση και διαχείριση της πληροφορίας**

### **4.1. Διαχείριση μεταδεδομένων**

Οι βιβλιοθήκες σε όλο τον κόσμο ψηφιοποιούν πολύτιμο περιεχόμενο για περισσότερο από δέκα χρόνια τώρα, και υπάρχουν πολλά παραδείγματα που θα μπορούσαν να αναφερθούν. (Deegan, 2001) Προκειμένου να δημιουργηθούν αποτελεσματικές ψηφιακές αρχειακές συλλογές, έμφαση πρέπει να δοθεί στη δημιουργία των μεταδεδομένων. Τα έγγραφα PDF δημοσιεύονται συχνά στο Διαδίκτυο ή σε άλλα περιβάλλοντα, όπου ειδικά εργαλεία χρησιμοποιούνται για να εξετάζουν, να καταχωρούν, και να ταξινομούν έγγραφα. Αυτά τα εργαλεία πρέπει να είναι σε θέση να καταλάβουν την ανεξάρτητη περιγραφή του εγγράφου ακόμα κι αν δεν καταλαβαίνουν τη δομή του αρχείου PDF. Για την εξυπηρέτηση αυτής της ανάγκης η προτεινόμενη λύση εστιάζει στη χρήση των μεταδεδομένων.

Τα μεταδεδομένα είναι "στοιχεία για τα στοιχεία". Παραδείγματος χάριν, ένας βιβλιογραφικός κατάλογος αποτελεί μεταδεδομένα, εφόσον ουσιαστικά παρέχει περιγραφές των αρχειοθετημένων εκδόσεων. Τα μεταδεδομένα γίνονται όλο και περισσότερο σημαντικά στην παραγωγή, τη διαχείριση, και τη δημοσίευση των κάθε μορφής εγγράφων. Τα έγγραφα που περιέχουν μεταδεδομένα μπορούν να βελτιώσουν κατά πολύ τη χρησιμότητά τους, παρέχοντας ευκολίες πρόσβασης στην πληροφορία που περικλείουν. Ένα παράδειγμα αποτελεσματικής χρήσης των μεταδεδομένων θα ήταν ένα αρχείο εικόνας-ιδανικά για την περίπτωση μας σε μορφή PDF-που περιέχει μεταδεδομένα όπως ο τίτλος του θέματος της εικόνας, η περιγραφή της εικόνας, μια μικρογραφία αυτής, καθώς και στοιχεία δικαιωμάτων πνευματικής ιδιοκτησίας. Χωρίς την παρουσία των μεταδεδομένων, θα ήταν εξαιρετικά δύσκολο να ταυτίσει κανείς τις εικόνες με τις ονομασίες των αρχείων, να αναζητήσει εικόνες ανά θεματική ενότητα, και να διαπιστώσει εάν είναι ελεύθερα διαθέσιμες προς χρήση, ή αποτελούν πνευματική ιδιοκτησία τρίτων.

Ενώ αυτές οι διαδικασίες αντιμετωπίζονται συνήθως ικανοποιητικά από τις διάφορες εφαρμογές<sup>13</sup>, είναι η ανταλλαγή των μεταδεδομένων-για χρήση σε διαφορετικές ροές εργασίας και συστήματα-αυτή η οποία δεν είναι και τόσο εύκολη. Παραδείγματος χάριν, οι πρώτες προδιαγραφές του PDF δεν προέβλεπαν την ύπαρξη ενός τυποποιημένου προτύπου για την αναπαράσταση των μεταδεδομένων. Αυτό είχε ως αποτέλεσμα, άλλες εφαρμογές που αγνοούσαν το PDF, να μην είναι σε θέση να αναγνωρίσουν και να αξιοποιήσουν πληροφορίες μεταδεδομένων-όπως Τίτλος, Συγγραφέας, Θέμα, κα.-που ήταν ήδη καταγεγραμμένες.

### **4.2. Η αρχιτεκτονική XMP**

Η Adobe έχει πλέον τυποποιήσει την υποστήριξη των μεταδεδομένων που συνοδεύουν τα έγγραφα PDF πάνω στη γλώσσα XML Η Extensible Markup Language (XML) είναι μια απλή και ιδιαίτερα ευέλικτη γλώσσα περιγραφής, που επιτρέπει την ανταλλαγή πληροφοριών. Τα τελευταία χρόνια έχει αναδειχθεί σε πολύτιμο εργαλείο για μια σειρά εφαρμογών<sup>14</sup>. Το σχήμα της XML που αναπαριστά τα μεταδεδομένα καθορίζεται σε ένα πλαίσιο αποκαλούμενο XMP και περιγράφεται στο έγγραφο της Adobe με τίτλο XMP. Extensible Metadata Platform. Αυτό το πλαίσιο παρέχει τις απαιτούμενες δομές για τη χρήση της XML για την αναπαράσταση των μεταδεδομένων, που περιγράφουν τα έγγραφα και το περιεχόμενό τους, και προορίζεται να υιοθετηθεί από μια ευρύτερη κατηγορία: εφαρμογών πέρα από αυτές που επεξεργάζονται PDF.



Η προδιαγραφή XMP σχεδιάστηκε με στόχο να αποτελέσει το πρότυπο επιλογής για την αναπαράσταση και την ανταλλαγή των μεταδεδομένων μεταξύ των εφαρμογών. Αντί της εισαγωγής ενός ακόμη νέου προτύπου, φάνηκε λογικό να υιοθετηθεί το μοντέλο RDF<sup>16</sup>, και η ανάπτυξη της XMP να ωφεληθεί από την υπάρχουσα τεκμηρίωση, τα εργαλεία, και την αποκτηθείσα εμπειρία<sup>17</sup>. Με αυτό τον τρόπο εξασφαλίζεται πως ένα βασικό σύνολο ιδιοτήτων μεταδεδομένων θα μπορεί να είναι αξιοποιήσιμο σε ένα ευρύ φάσμα εφαρμογών, συμπεριλαμβανομένων της οικογένειας προϊόντων Adobe Acrobat, σχεδόν του συνόλου των λοιπών προϊόντων της Adobe, καθώς επίσης και εφαρμογών από μια ευρεία γκάμα τρίτων κατασκευαστών λογισμικού.

#### **4.3. Μεταδεδομένα και έγγραφα PDF**

Τα αρχεία PDF περιέχουν μεταδεδομένα που περιγράφουν συγκεκριμένες ιδιότητες του εγγράφου-όπως Τίτλος, Συγγραφέας, Θέμα, κ.α.-αλλά δεν περιορίζονται εκεί. Επειδή τα μεταδεδομένα είναι με το σχήμα XML, στοιχεία μεταδεδομένων καθοριζόμενα στο πλαίσιο Dublin Core<sup>18</sup>, ή αλλιώς μπορούν επίσης να συμπεριληφθούν. Επιπλέον, οι χρήστες μπορούν να εισάγουν, να τροποποιήσουν, και να εξάγουν τα μεταδεδομένα, σύμφωνα με τις ανάγκες τους.

### **5. Αναζήτηση και ανάκτηση της πληροφορίας**

#### **5.1. Πρόσβαση στις ψηφιακές συλλογές**

Το χαρακτηριστικό της ενσωμάτωσης μεταδεδομένων της μορφής XML στα έγγραφα PDF εξασφαλίζει ότι τα μεταδεδομένα θα είναι αναγνώσιμα και προσβάσιμα στο μέλλον. Επιπλέον, τα αρχεία PDF προσφέρονται για αναζητήσεις τόσο στα μεταδεδομένα, όσο και στο σώμα του κειμένου που είναι πιθανό να περιέχουν. Πολλοί γνωστοί κατασκευαστές λογισμικού, συμπεριλαμβανομένων των Verity, Hummingbird, και Convera, έχουν ενσωματώσει υποστήριξη για τα αρχεία PDF στις μηχανές αναζήτησης τους εδώ και πολλά χρόνια<sup>19</sup>. Μεταξύ αυτών, η δωρεάν υπηρεσία αναζήτησης του Google<sup>20</sup>, επιστρέφει αποτελέσματα περιεχομένου των αρχείων PDF<sup>21</sup>.

Φυσικά, η οικογένεια προϊόντων Acrobat, παρέχει λειτουργίες αναζήτησης απευθείας μέσα από το περιβάλλον των εφαρμογών. Γενικά, μπορείτε να ψάξετε για λέξεις στο κείμενο, τα σχόλια, τους σελιδοδείκτες, τις πληροφορίες και τα μεταδεδομένα του εγγράφου, και αλλού. Μπορείτε επίσης να διεξάγετε μια σύγχρονη αναζήτηση σε πολλαπλά έγγραφα, χωρίς να χρειάζεται να ανοίξετε τα αρχεία προηγουμένως. Μπορείτε να χρησιμοποιήσετε τις λειτουργίες αναζήτησης για να βρείτε λέξεις σε αρχεία PDF που βρίσκονται αποθηκευμένα στο σκληρό σας δίσκο ή στο τοπικό δίκτυο, σε έτοιμους καταλόγους, και σε αρχεία PDF στο διαδίκτυο.

Για τη διευκόλυνση των χρηστών, το Adobe Acrobat διαθέτει μια λειτουργία δημιουργίας καταλόγων, οι οποίοι μπορούν να χρησιμοποιηθούν για γρήγορες αναζητήσεις σε μεγάλο αριθμό αρχείων PDF. Ένας κατάλογος είναι ένα ειδικά προετοιμασμένο αρχείο που αναλύει και ευρετηριάζει πολλαπλά αρχεία PDF. Αποτελεί στην ουσία μια αλφαβητική ταξινόμηση όλων των λέξεων που χρησιμοποιούνται σε ένα καθορισμένο σύνολο εγγράφων, και ορίζει τους κατάλληλους δείκτες με τις συντεταγμένες της κάθε λέξης. Εάν ένας τέτοιος κατάλογος είναι διαθέσιμος για ένα σύνολο εγγράφων PDF, μπορείτε να ψάξετε στον κατάλογο για μια λέξη παρά να διεξάγετε την αναζήτηση σε κάθε μεμονωμένο έγγραφο ξεχωριστά. Η έρευνα μέσω του καταλόγου είναι πολύ

γρηγορότερη από την απευθείας αναζήτηση του κειμένου στα έγγραφα. Επιπλέον, επιτρέπει αναζητήσεις με λογικά (Boolean) ερωτήματα<sup>22</sup>. Μπορείτε να διανείμετε σε CD-ROM ή να δημοσιεύσετε σε τοπικά δίκτυα, τους καταλόγους μαζί με το σύνολο των αρχείων σας PDF για να διευκολύνετε τους χρήστες να βρουν τις πληροφορίες που χρειάζονται<sup>23</sup>.

Για πιο ολοκληρωμένα σενάρια, που βασίζονται στη χρήση του διαδικτύου, η Adobe διαθέτει το Adobe PDF Filter. Πρόκειται για μια δωρεάν εφαρμογή που σας επιτρέπει την καταλογογράφηση των εγγράφων PDF με τη αξιοποίηση των εργαλείων της Microsoft, ώστε οι χρήστες να μπορούν με ευκολία να αναζητούν κείμενο στα έγγραφα PDF μέσω του διαδικτύου. (Adobe, Αγν.) Ένα IFilter αποτελεί ένα αρχείο DLL (Dynamic Link Library) το οποίο καλείται για να εξάγει λέξεις από μια δεδομένη μορφή αρχείου<sup>24</sup>. Το Adobe PDF IFilter διαθέτει τον απαραίτητο κώδικα για την κατανόηση της μορφής αρχείου PDF από τον Microsoft Index Server. Επιστρέφει κείμενο στη σχετική έκκληση, το οποίο προστίθεται στον κατάλογο που δημιουργείται, ή ενημερώνεται.

Η δυνατότητα αναζήτησης, η ικανότητα του να συμπεριλάβει σχεδόν όλες τις μορφές του ψηφιακού περιεχομένου-όπως κείμενο, φωτογραφίες, γραφικά, ήχος, βίντεο-και η δωρεάν διάθεση του Adobe Reader<sup>25</sup>, κ.α., καθιστούν το PDF ως την ιδανική μορφή για την οργάνωση των ψηφιακών αρχειακών συλλογών.

## **6. Συμπεράσματα και προτάσεις**

### **6.1. Τα οφέλη του PDF συνοπτικά**

- Διατήρηση της πιστότητας των εγγράφων κατά την διακίνηση τους, ανεξάρτητα από τη συσκευή, την πλατφόρμα, και το λογισμικό
- Συγχώνευση περιεχομένου από διαφορετικές πηγές-ιστοσελίδες, εφαρμογές επεξεργασίας κειμένου και λογιστικών φύλλων, ψηφιοποιημένα έντυπα, φωτογραφίες, και μακέτες γραφικών-σε ένα περιεκτικό έγγραφο, διατηρώντας ταυτόχρονα την ακεραιότητα του περιεχομένου όλων των αρχικών εγγράφων
- Συλλογική επεξεργασία των εγγράφων και συνεργασία σε πραγματικό χρόνο, από πολλές θέσεις ή πλατφόρμες
- Ψηφιακές υπογραφές για την πιστοποίηση της αυθεντικότητας
- Προστασία και διαχείριση δικαιωμάτων χρήσης που επιτρέπουν στο δημιουργό να διατηρήσει τον έλεγχο του εγγράφου και των σχετικών δικαιωμάτων
- Δυνατότητα πρόσβασης περιεχομένου σε άτομα με ειδικές ανάγκες
- Εξαγωγή και αξιοποίηση του περιεχομένου σε άλλες μορφές αρχείων και εφαρμογές λογισμικού

### **6.2. Προβλέψεις**

Από πολλές απόψεις, το PDF είναι ένα ψηφιακό υποκατάστατο για το χαρτί-αλλά καλύτερο. Όπως τα έντυπα, έτσι και τα αρχεία PDF μπορούν να απεικονίσουν τόσο το περιεχόμενο όσο και τη σχεδίαση με πλήρη πιστότητα. Αλλά το PDF είναι επίσης ένα έξυπνο σχήμα που υποστηρίζει την ανταλλαγή στοιχείων XML, καθώς και δυναμικές φόρμες, ψηφιακές υπογραφές, ηλεκτρονικά σχόλια, και άλλα. Αυτό το επίπεδο των ψηφιακών υπηρεσιών επιτρέπει την αυτοματοποίηση των

διαδικασιών που σχετίζονται με τη διαχείριση των εγγράφων, καθώς επίσης ενισχύει την ενσωμάτωση τους σε άλλα συστήματα αυτοματισμού. Επειδή αυτές οι υπηρεσίες περιλαμβάνονται μέσα στο αρχείο PDF, ταξιδεύουν με το έγγραφο οπουδήποτε αυτό πηγαίνει.

Το λογισμικό Adobe Acrobat και το Adobe PDF εξελίσσονται γρήγορα σε εφαρμογή παραγωγής και σε προτιμώμενη μορφή αρχείου, αντίστοιχα, για μια σειρά χρήσεων. Η ανοικτή αρχιτεκτονική της τεχνολογίας PDF επιτρέπει στις επιχειρήσεις και τους οργανισμούς να την ενσωματώσουν με άλλες εφαρμογές, για να αντιμετωπίσουν εξειδικευμένα επιχειρησιακά σενάρια. Οι ενδιαφερόμενοι φορείς πρέπει να αναλύσουν τις εσωτερικές διαδικασίες τους και να αποτιμήσουν τις ανάγκες τους, για να δουν ποια πλεονεκτήματα μπορούν να αποκομίσουν από την υιοθέτηση της τεχνολογίας PDF.

## ΠΑΡΑΠΟΜΠΕΣ

Μια αναζήτηση μέσω ίου Google, σε μια προσπάθεια να εκτιμήσουμε τον αριθμό των διαθέσιμων αρχείων PDF στο Διαδίκτυο, επέστρεψε 27.800.000 αποτελέσματα. Περιορίζοντας την έρευνα στον Ελλαδικό χώρο (κατάληξη .gr), η μηχανή επέστρεψε 56.100 αποτελέσματα. Μια παρόμοια αναζήτηση, για αρχεία της μορφής DOC (Microsoft Word Document), επέστρεψε 3.730.000 και 35.600 αποτελέσματα, αντίστοιχα. Για να επαναλάβετε, μπορείτε να χρησιμοποιήσετε τις ακόλουθες διατυπώσεις:

[http://www.google.com/search?as\\_a=pdf&as\\_filetype=pdf](http://www.google.com/search?as_a=pdf&as_filetype=pdf)

[http://www.google.com/search?as\\_a=pdf&as\\_filetype=pdf&as\\_sitesearch=.gr](http://www.google.com/search?as_a=pdf&as_filetype=pdf&as_sitesearch=.gr)

[http://www.google.com/search?as\\_a=doc&as\\_filetype=doc](http://www.google.com/search?as_a=doc&as_filetype=doc)

[http://www.google.com/search?as\\_a=doc&as\\_filetype=doc&as\\_sitesearch=.gr](http://www.google.com/search?as_a=doc&as_filetype=doc&as_sitesearch=.gr)

<sup>2</sup>Ο ρυθμός εξάπλωσης της εφαρμογής Adobe Reader, του δωρεάν λογισμικού ανάγνωσης και εκτύπωσης αρχείων PDF, αποτελεί μια ουσιαστική ένδειξη της καθολικής αποδοχής του PDF. Σύμφωνα με τα στοιχεία της Adobe, έως το Μάρτιο του 2000 είχαν διανεμηθεί 100.000.000 αντίγραφα της εφαρμογής, το Μάιο του 2001 ο αριθμός τους ανερχόταν σε 200.000.000, ενώ σήμερα η Adobe θεωρεί πως έχει ξεπεράσει κατά πολύ τα 500.000.000 διανεμημένα αντίτυπα. (Adobe, 2000: 2001: 2003d)

<sup>3</sup>Το PDF είναι σχεδιασμένο ώστε να είναι ευέλικτο. Όχι μόνο νέα χαρακτηριστικά μπορούν να προστεθούν, αλλά εφαρμογές βασισμένες στις προηγούμενες εκδόσεις της προδιαγραφής μπορούν να συμπεριφερθούν ικανοποιητικά όταν αντιμετωπίζουν νεότερα χαρακτηριστικά που δεν καταλαμβάνουν. Ιδανικά, αυτό σημαίνει ότι κάθε εφαρμογή Acrobat οφείλει να είναι σε θέση να απεικονίζει τα περιεχόμενα ενός εγγράφου PDF, ακόμα και αν το αρχείο δημιουργήθηκε αρκετό καιρό πριν, ή μετά, από την ίδια την εφαρμογή. (Adobe, 2003d)

<sup>4</sup>Πληροφορίες σχετικά με το φορέα, τις δραστηριότητες του, και τις συμμετέχουσες εταιρείες είναι διαθέσιμες στο Διαδίκτυο στη διεύθυνση <http://www.npres.org/> .

<sup>5</sup>Πληροφορίες σχετικά με το φορέα, είναι διαθέσιμες στο Διαδίκτυο στη διεύθυνση <http://www.aiim.org/>

<sup>6</sup>Βλ. <http://www.adobe.com/products/acrobat/> .

<sup>7</sup>Τα προϊόντα που έχει αναπτύξει η Adobe γύρω από την τεχνολογία PDF στοχεύουν σε όλες τις κατηγορίες της αγοράς. Συμπεριλαμβάνουν την εφαρμογή Adobe Reader-για ανάγνωση των εγγράφων-που διανέμεται δωρεάν, το Adobe Acrobat 6.0 Professional-για παραγωγή και επεξεργασία των εγγράφων-που απευθύνεται στους επαγγελματίες και τις μικρές επιχειρήσεις, αλλά και προϊόντα με υψηλό κόστος που απευθύνονται αποκλειστικά στις μεγάλες επιχειρήσεις και τους κυβερνητικούς οργανισμούς. Ενδεικτικά, αναφέρουμε τον Adobe Document Server, που προσφέρει αυτοματοποιημένη παραγωγή εγγράφων και διαχείριση ψηφιακών δικαιωμάτων πρόσβασης. Για περισσότερες πληροφορίες σχετικά με λύσεις που απευθύνονται στα υψηλά κλιμάκια της αγοράς, μπορείτε να ανατρέξετε στη διεύθυνση <<http://www.adobe.com/products/server/>> .

<sup>8</sup>Βλ. <<http://partners.adobe.com/asn/tech/pdf/specifications.isp>> .

Εως τη στιγμή της δημοσίευσης, το βιβλίο δεν έχει εκδοθεί. Ωστόσο, είναι διαθέσιμο στο Διαδίκτυο. (Adobe, 2003d)

<sup>10</sup>Είναι χαρακτηριστικό πως, αν και η Adobe κατέχει τα πνευματικά δικαιώματα (© COPYRIGHT) της προδιαγραφής του Portable Document Format (PDF), έχει αποφύγει έως τώρα να καταθέσει την ονομασία ως ένα εμπορικό σήμα TRADEMARK). (Adobe, 2003a)

<sup>11</sup>Ενδεικτικά, αναφέρουμε το ελεύθερο λογισμικό μετατροπής αρχείων PostScript σε PDF, με την ονομασία Ghostscript. Περισσότερες πληροφορίες είναι διαθέσιμες στη διεύθυνση: <<http://www.cs.wisc.edu/~ghost/>> ,

<sup>12</sup>Δύο άριστες πηγές πληροφόρησης σχετικά με το Adobe Acrobat και τρίτους κατασκευαστές λογισμικού μπορούν να βρεθούν στις διευθύνσεις <http://www.pdfzone.com> και <http://www.pdfplanet.com>,

<sup>13</sup>Λίγο έως πολύ, τα λειτουργικά συστήματα και οι περισσότερες εφαρμογές σήμερα, προσφέρουν σε έμβροτο βαθμό την υποστήριξη χρήσης μεταδεδομένων, που συνήθως εκφράζεται μέσα από το πλαίσιο διαλόγου των ιδιοτήτων του κάθε αρχείου.

<sup>14</sup>Η σχετική προδιαγραφή είναι ελεύθερα διαθέσιμη από το World Wide Web Consortium (W3C), στη διεύθυνση <<http://www.w3.org/XMIL/>>.

<sup>15</sup>Το ακροατήριο στο οποίο κυρίως απευθύνεται το παρόν έγγραφο είναι οι υπεύθυνοι για την ανάπτυξη εφαρμογών που θα παράγουν, θα επεξεργάζονται, ή θα διαχειρίζονται αρχεία που περιέχουν μεταδεδομένα της μορφής XMP. Η σχετική προδιαγραφή και άλλες πηγές είναι ελεύθερα διαθέσιμα από την Adobe, στη διεύθυνση <<http://partners.adobe.com/asn/tech/xmp/>>.

<sup>16</sup>Η προδιαγραφή του Resource Description Framework (RDF) είναι ελεύθερα διαθέσιμη από το World Wide Web Consortium (W3C), στη διεύθυνση <<http://www.w3.org/RDF/>>.

Για περισσότερες πληροφορίες Βλ. στη βιβλιογραφία. (Ράπτης, 1999)

<sup>17</sup>Αν και τα δύο πλαίσια μοιράζονται τα περισσότερα κοινά στοιχεία, η ουσιαστική τους διαφορά συνοψίζεται στο εύρος των πιθανών χρήσεων. Το RDF επικεντρώνεται σχεδόν αποκλειστικά σε χρήσεις που αφορούν το διαδίκτυο, ενώ η XMP αποσκοπεί να αποτελέσει μια γενικευμένη πλατφόρμα ανταλλαγής μεταδεδομένων μεταξύ διαφορετικών συστημάτων και εφαρμογών.

<sup>18</sup>Το πλαίσιο αναφοράς Dublin Core προτείνει ένα προκαθορισμένο σύνολο δέκα πέντε στοιχείων-όπως Τίτλος, Δημιουργός, θέμα, Περιγραφή, κ.α.-με σκοπό την κοινή καταγραφή των μεταδεδομένων στα διάφορα συστήματα και τις εφαρμογές. Είναι το αποτέλεσμα μιας πρωτοβουλίας που ξεκίνησε το 1995, με σκο-

πό την τυποποίηση της περιγραφής των διακινούμενων πληροφοριών. Η προδιαγραφή και άλλες πηγές, μπορούν να αναζητηθούν στη διεύθυνση <<http://dublincore.org/>> .

Για περισσότερες πληροφορίες Βλ. στη βιβλιογραφία. (Ράπτης, 1999)

<sup>19</sup>Ο αναγκώστης που ενδιαφέρεται να ενσωματώσει απομακρυσμένες υπηρεσίες αναζήτησης στο περιεχόμενο των εγγράφων PDF, μπορεί να ανατρέξει στη διεύθυνση <<http://www.searchtools.com/info/Ddf.html>>

<sup>20</sup>Το Google αποτελεί σήμερα, ίσως την πλέον διαδεδομένη μηχανή αναζήτησης περιεχομένου στο Διαδίκτυο. Η υπηρεσία είναι διαθέσιμη στη διεύθυνση <http://www.google.com/> .

<sup>21</sup>Οι υποστηριζόμενες μορφές αρχείων που ανιχνεύει και ταξινομεί η μηχανή αναζήτησης του Google, παρτίθενται στη διεύθυνση <<http://www.google.com.gr/corporate/facts.html>> .

<sup>22</sup>Για να μάθετε περισσότερα σχετικά με τα λογικά ερωτήματα, τη σύνταξη τους, και τις παραμέτρους που μπορείτε να καθορίσετε στις αναζητήσεις σας, μπορείτε να ανατρέξετε σε οποιαδήποτε πηγή αναφέρεται σε αυτή τη μεθοδολογία αναζήτησης.

<sup>23</sup>Ένα πολύ χρήσιμο χαρακτηριστικό, σχετικά με τη δημιουργία καταλόγων, είναι πως αν και απαιτείται η αγορά του Adobe Acrobat για τη δημιουργία τους, μπορούν κατόπιν να διατίθενται ελεύθερα και να χρησιμοποιούνται μέσω του δωρεάν Adobe Reader.

<sup>24</sup>Για περισσότερες πληροφορίες σχετικά με την προδιαγραφή του Filter, μπορείτε να επισκεφθείτε τη σχετική διεύθυνση της Microsoft.

Για περισσότερες πληροφορίες Βλ. στη βιβλιογραφία. (Microsoft, Αγν.)

<sup>25</sup>Πολλές φορές θεωρούμε αυτονόητο ότι ο παραλήπτης της πληροφορίας διαθέτει το κατάλληλο λογισμικό. Αυτό δε συμβαίνει πάντοτε. Είναι πολύ σημαντικό να παραδεχτούμε ότι η πρόσβαση σε κάποιο έγγραφο, οφείλει να μην επιβαρύνει το χρήστη με το κόστος αγοράς του απαραίτητου λογισμικού ανάγνωσης. Δυστυχώς, αυτή δεν είναι μια πρακτική που ακολουθείται συχνά.

## **BIBΛΙΟΓΡΑΦΙΑ**

1. ADOBE SYSTEMS, INC., 1993a. *Portable Document Format reference manual*. Reading, MA: Addison-Wesley.
- 2.\_\_\_\_\_, 1993b. *PostScript Language Reference Manual*, 3rd ed. Reading, MA: Addison-Wesley. <<<http://partners.adobe.com/asn/developer/pdfs/tn/PIJMpdf>>> [πρόσβαση 31/10/2003]
- 3.\_\_\_\_\_, 2000. *PDF Reference*, 2nd ed. Reading, MA: Addison-Wesley. <[http://partners.adobe.com/asn/acrobat/docs/File\\_Format\\_Soecifications/PPFReferencel3.odf](http://partners.adobe.com/asn/acrobat/docs/File_Format_Soecifications/PPFReferencel3.odf)> [πρόσβαση 31/10/2003]
- 4.\_\_\_\_\_, 2001. *PDF Reference*, 3rd ed. Reading, MA: Addison-Wesley. «[http://partners.adobe.com/asn/acrobat/docs/File\\_Format\\_Specifications/PDFReferencel3.odf](http://partners.adobe.com/asn/acrobat/docs/File_Format_Specifications/PDFReferencel3.odf)» [πρόσβαση 31/10/2003]

5. \_\_\_\_\_, 2003a. *Adobe Trademark Database For General Distribution*.  
<<http://www.adobe.com/misc/Ddfs/USGenExtTMdb022803.pdf>>  
[πρόσβαση 31/10/2003]
6. \_\_\_\_\_, 2003b. *PDF as a Standard for Archiving*.  
<<http://www.adobe.com/products/acrobat/pdfs/pdfarchiving.pdf>>  
[πρόσβαση 31/10/2003]
7. \_\_\_\_\_, 2003c. *PDF as a Standard for Managing and Archiving Documents and Forms*.  
<<http://www.adobe.com/products/acrobat/pdfs/docforms.pdf>>  
[πρόσβαση 31/10/2003]
8. \_\_\_\_\_, 2003d. *PDF Reference*, 4th ed.  
<[http://partners.adobe.com/asn/acrobat/sdk/public/docs/PDFReference15\\_v6.pdf](http://partners.adobe.com/asn/acrobat/sdk/public/docs/PDFReference15_v6.pdf)>  
[πρόσβαση 31/10/2003]
9. \_\_\_\_\_, Αγν. *Searching the Contents of PDF Files on a Web Site*.  
<<http://www.adobe.com/support/techdocs/12b42.htm>> [πρόσβαση 31/10/2003]
10. ASSOCIATION FOR INFORMATION AND IMAGE MANAGEMENT (AIIM), 2002. *AIIM international and NPES partner to standardize use of pdf for document archive and preservation*.  
<[http://www.aiim.org/documents/standards/press\\_release-pdfa-augQ2-2.pdf](http://www.aiim.org/documents/standards/press_release-pdfa-augQ2-2.pdf)>  
[πρόσβαση 31/10/2003]
11. DEEGAN, M., 2001. Managing documentary collections through digitization. *10ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών*. 15-17/10, Θεσσαλονίκη.  
<<http://www.lib.uom.gr/palciO/greek/Deegan.doc>> [πρόσβαση 31-10-2003]
12. INTERNATIONAL HERALD TRIBUNE (IHT), Αγν. *Publishing Partnerships*.  
<<http://www.iht.com/global.html>> [πρόσβαση 31-10-2003]
13. INTERNATIONAL ORGANIZATION FOR STANDARDIZATION (ISO), 2003. *ISO/WD 19005, Document management-Electronic document file format for long-term preservation-Use of PDF (PDF/A)*. <[www.aiim.org/documents/standards/ISO\\_19005\\_\(E\).pdf](http://www.aiim.org/documents/standards/ISO_19005_(E).pdf)>  
<[http://www.aiim.org/documents/standards/ISO\\_19005\\_\(E\).pdf](http://www.aiim.org/documents/standards/ISO_19005_(E).pdf)>  
[πρόσβαση 31/10/2003]
14. MICROSOFT CORPORATION, Αγν. *Filter*.  
<[http://msdn.microsoft.com/librarv/en-us/indexsrv/html/ixrefint\\_9sfm.asp](http://msdn.microsoft.com/librarv/en-us/indexsrv/html/ixrefint_9sfm.asp)>  
[πρόσβαση 31-10-2003]
15. SINHA, V., 2003. For e-gov, feds turn to reliable tool. *Government Computer News*. 28/04. <<http://www.adobe.com/government/pdfs/rn097.pdf>>  
[πρόσβαση 31/10/2003]
16. WARNOCK, J., 1991. *The Camelot Project*.  
<[http://www.planetpdf.com/planetpdf/pdfs/warnock\\_camelot.pdf](http://www.planetpdf.com/planetpdf/pdfs/warnock_camelot.pdf)>  
[πρόσβαση 31/10/2003]
17. ΡΑΪΤΗΣ, Π., 1999. Πλαίσιο περιγραφής πηγών και μεταπληροφορίες στο Διαδίκτιο. *8ο Πανελλήνιο Συνέδριο Ακαδημαϊκών Βιβλιοθηκών*. 20-22/10, Ρόδος.