

ΚΕΙΜΕΝΟ

- Ο πρώτος τρόπος απεικόνισης πληροφορίας (και βασικός ως σήμερα).
- Αδυναμία πρώτων υπολογιστών να χειριστούν άλλη μορφή πληροφορίας.
- Πρόβλημα με καθιερωμένα πρότυπα (π.χ. ASCII), π.χ. σε πολυγλωσσικές εφαρμογές.
- Κείμενο: βασικός φορέας πληροφορίας (απαραίτητη ολοκλήρωση σε εφαρμογή πολυμέσων).

ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΕΙΜΕΝΟΥ

- Αρχικές χρήσεις: συγγραφή προγραμμάτων, εισαγωγή δεδομένων.
- Απλές κωδικοποιήσεις μέχρι λεπτομερείς γλώσσες περιγραφής.
- ASCII (American Standard Code for Information Interchange): κωδικοποίηση χαρακτήρων των 7 bits.
- Χαρακτήρες ASCII: λατινικοί χαρακτήρες και ειδικά σύμβολα.
- Υπόλοιποι 128 χαρακτήρες του byte: αναπαράσταση μη-ASCII χαρακτήρων.

ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΕΙΜΕΝΟΥ

- ISO character sets: αντεπεξέρχονται στην αδυναμία του ASCII να υποστηρίξει άλλες γλώσσες.
- *Παράδειγμα:* ISO Latin (ISO 1987b) είναι υπερσύνολο ASCII για τις περισσότερες ευρωπαϊκές γλώσσες.
- Κωδικοποιήσεις ISO χρησιμοποιούνται στο X Windows System, στο PostScript, κλπ.

ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΕΙΜΕΝΟΥ

- Κείμενο: μορφή και νοηματικό περιεχόμενο.
- Mark-up κείμενο: μίξη κανονικού κειμένου με οδηγίες παρουσίασης.
- Παραδείγματα γλωσσών mark-up: SGML (Standardized General Markup Language) (πρότυπο ISO), T_EX (μέσω του συστήματος L^AT_EX).

ΑΝΑΠΑΡΑΣΤΑΣΗ ΚΕΙΜΕΝΟΥ

- Σε mark-up κείμενο η δομική πληροφορία είναι δυσδιάκριτη.
- Υπάρχουν πρότυπα για δομημένη αναπαράσταση κειμένου: μέρη, κεφάλαια, παράγραφοι κλπ. (π.χ. δενδρική δομή).
- Γραμμικό κείμενο vs hypertext: κείμενο χωρίζεται σε κόμβους που συνδέονται σε ένα γράφο (επεξηγήσεις, υποσημειώσεις).
- Πρόγραμμα παρουσίασης υπερκειμένου (hypertext browser).
- Ευρύτατη διάδοση της HTML μέσω του WWW.

ΧΕΙΡΙΣΜΟΣ ΚΕΙΜΕΝΟΥ

- Πράξεις χαρακτήρων: ορίσματα είναι απλοί χαρακτήρες (χρήση: προγραμματιστές).
- Πράξεις συμβολοσειρών: ορίσματα είναι ακολουθίες χαρακτήρων (π.χ. λεξικογραφική σύγκριση, συνένωση, αναζήτηση χαρακτήρων κλπ.)
- Editing κειμένου: εισαγωγή-διαγραφή χαρακτήρων, αποθήκευση κειμένου σαν αρχείο (γίνεται μέσω των editors σε κάθε υπολογιστική πλατφόρμα).
- Άλλες δυνατότητες editors: επιλογή, αποκοπή, αντιγραφή, επικόλληση, αρίθμηση λιστών, κλπ.

ΧΕΙΡΙΣΜΟΣ ΚΕΙΜΕΝΟΥ

- Φορμάρισμα κειμένου: επιβολή χαρακτηριστικών εμφάνισης
- Γίνεται μέσω interactive εφαρμογών: επεξεργαστές κειμένου (επίσης δυνατότητα μίξης κειμένου με εικόνες, διαγράμματα, ήχους, κλπ.: πολυμεσικό κείμενο).
- Έξοδος εφαρμογής: bitmap εικόνα του κειμένου ή περιγραφή σε γλώσσα περιγραφής σελίδας (Page Description Language): π.χ. PostScript (Adobe).
- Βασικό στοιχείο: γραμματοσειρές (fonts) και οικογένεια fonts (typeface).

ΧΕΙΡΙΣΜΟΣ ΚΕΙΜΕΝΟΥ

- Fonts του ίδιου typeface διαφέρουν σε: μέγεθος, bold, italics, underlined, συνδυασμός.
- Ψηφιογραφικές (bitmap fonts): επεξεργασία, εκτύπωση γρήγορη, πρόβλημα σε αποθηκευτικό χώρο και ποιότητα μετασχηματισμού.
- Διανυσματικές: μαθηματική περιγραφή, αυξημένος χρόνος εκτύπωσης (π.χ. PostScript Type I στην PostScript και TrueType στα MS Windows).

ΧΕΙΡΙΣΜΟΣ ΚΕΙΜΕΝΟΥ

- Αναζήτηση προτύπων: απλές λέξεις ή λέξεις με χαρακτήρες ειδικής σημασίας (wildcards) αναζητούνται στο κείμενο.
- Περιγραφή προτύπων με κανονικές εκφράσεις (regular expressions).
- Αναζήτηση σε μεγάλους όγκους κειμένου: ενδιαφέρει ανεύρεση κειμένων γύρω από ένα θέμα.
- Συμπίεση: συνήθως αντικατάσταση προτύπων, κωδικοποίηση Huffman.

ΧΕΙΡΙΣΜΟΣ ΚΕΙΜΕΝΟΥ

- Κρυπτογράφηση: απαραίτητη λόγω διάδοσης δικτύων, ηλεκτρονικού ταχυδρομείου, ηλεκτρονικού εμπορίου.
- DES (Data Encryption Standard): αναπτύχθηκε από την IBM για την αμερικανική κυβέρνηση.
- Αλγόριθμοι με δημόσιο κλειδί (public key): για κρυπτογράφηση υπάρχει το δημόσιο κλειδί (γνωστοποιείται στους αποστολείς), ενώ για αποκρυπτογράφηση χρειάζεται και το ιδιωτικό κλειδί (το έχει μόνο ο παραλήπτης).