Pocket

Editoria digitale

Letizia Sechi

Linguaggi, strumenti, produzione e distribuzione dei libri digitali

Editoria Digitale

Letizia Sechi

Prima edizione nella collana “Pocket” marzo 2010 - ISBN 978-88-503-2917-5

Copyright © 2010 Apogeo S.r.l Socio Unico Giangiacomo Feltrinelli Editore S.r.l.

Edizione elettronica luglio 2010

ISBN: 978-88-503-1097-5

La presente Opera è rilasciata gratuitamente sotto una licenza Creative Commons 3.0 BY-NC-SA.
Se decidi di distribuire quest'Opera devi farlo nel rispetto della licenza Creative Commons, indicando sempre e chiaramente l'Autore (Letizia Sechi) e l'Editore (Apogeo).
L'Editore si riserva ogni diritto di commercializzazione.
L'edizione cartacea è in vendita nelle migliori librerie.

Nomi e marchi citati nel testo sono generalmente depositati o registrati dalle rispettive case produttrici.

PREFAZIONE

«I libri stanno vivendo il cambiamento più profondo e strutturale dall'era di Gutenberg», scrive John Mekinson (CEO di Penguin) sul Wall Street Journal. «Gli editori reagiranno con ansioso entusiasmo. La pirateria è una delle fonti di ansia, ma siamo vulnerabili a tutti gli effetti del digitale, a partire dal crollo dei prezzi e della protezione del copyright. La soluzione? Leggere il contesto in prospettiva e anticipare il cambiamento».

Anticipare il cambiamento significa anche ripensare il processo produttivo del libro, guardando al libro cartaceo come a una delle possibili versioni del libro stesso, alla quale affiancare la versione digitale nelle sue numerose declinazioni.

In questo libro Letizia Sechi ci guida alla scoperta dei formati e delle tecnologie che si nascondono dietro la parola ebook, fornendo un utile strumento introduttivo per i professionisti dell'editoria e i semplici appassionati che vogliano capire come trasformare i contenuti editoriali già presenti a catalogo o realizzare nuove opere editoriali, sfruttando la duttilità dei nuovi standard e sperimentando con le possibilità offerte dalla tecnologia.

Marco Ghezzi

INTRODUZIONE

Sono molti anni che si parla di libri digitali, ma mai come adesso i tempi sono sembrati maturi per una loro effettiva diffusione. Il mondo editoriale è in fermento, i dispositivi di lettura si moltiplicano con velocità sorprendente, l'interesse e la curiosità dei lettori sono vivissimi.

Questo libro nasce in questo clima ricco di prospettive. Quando l'ebook smette di essere un'idea di difficile realizzazione, una speculazione su quanto può essere "oltre la carta", un oggetto dai contorni indefiniti,

allora, si sente il bisogno di lasciare per un momento da parte le teorie e di entrare nel vivo della sua produzione, di capire cosa bisogna fare concretamente per pubblicare libri fatti di bit.

Non è un caso che proprio un editore come Apogeo abbia deciso di occuparsi del lato pratico di questo argomento, e – tra l'altro – non è neanche una novità: nel 2001 usciva un interessante volumetto ormai fuori catalogo, *e-book*, firmato da Virginio Sala, che cercava di capire dove i libri digitali potessero portare l'editoria e la lettura, senza la pretesa di dare risposte definitive (aspetto in comune con il libro che tenete fra le mani), ma condividendo alcune riflessioni dettate dall'esperienza di un addetto ai lavori.

Sono onorata di poter proseguire in qualche modo quella riflessione, affrontando nuovi aspetti di un tema che – a distanza di quasi dieci anni – ha vissuto così tante evoluzioni e sembra ora raggiungere una certa maturità, almeno rispetto ad allora. Mi sono lasciata contagiare dalla curiosità e dalla ricerca di *e-book*, cercando di sviluppare un percorso che potesse mettere in luce tutti gli aspetti, tecnici e non, sull'attuale panorama dei libri digitali.

Briciole di pane

Due parole sul percorso che affronteremo in queste pagine per orientarci tra tutti gli aspetti che riguardano l'editoria digitale.

Nel primo capitolo passeremo rapidamente in rassegna tutti i discorsi a cui siamo in qualche modo abituati quando si parla di ebook: il libro inteso come tecnologia e la sua tipica organizzazione dei contenuti, ipertesti e associazioni mentali in rapporto a gerarchia e linearità; il Web e i nuovi strumenti per la gestione della conoscenza e la diffusione delle informazioni. Tutti elementi che hanno modificato radicalmente il nostro modo di rapportarci con la conoscenza, il testo e l'informazione in genere.

Nel secondo capitolo inizieremo a entrare nel vivo delle competenze tecniche necessarie per affrontare in modo pratico la produzione di libri digitali. Parliamo dei linguaggi per la codifica del testo, concentrando l'attenzione su quelli che offrono più strumenti per gli ebook.

Il terzo capitolo è invece una panoramica sui formati per i libri digitali. L'attenzione in questo caso è rivolta all'ePub, lo standard aperto basato su XML pensato apposta per gli ebook: ne analizzeremo le principali caratteristiche tecniche in maniera da capire quali siano i metodi migliori per sfruttarne al meglio le opportunità.

Nel quarto capitolo proponiamo alcune possibilità per avviare la produzione sistematica dei libri digitali. Non si tratta certamente delle uniche vie, ma sono quelle che permettono di controllare meglio la struttura del libro e di sfruttare nel modo più efficace le opportunità dei libri digitali. In particolare, la ricerca di un unico metodo che permetta di ottenere dalla stessa lavorazione sia i file per la stampa sia quelli per gli ebook costituisce un punto di partenza interessante

per ottimizzare gli attuali processi di produzione del libro, gestirli meglio e renderli più flessibili.

Nel quinto capitolo ragioneremo sulle numerose implicazioni che comporta la produzione sistematica di ebook dal punto di vista editoriale: l'impatto sulle figure professionali coinvolte, la collocazione dei libri digitali nelle linee editoriali, le tecniche di distribuzione. Riguardo quest'ultimo aspetto osserveremo da vicino le esperienze di editori già attivi nella realizzazione di ebook, e delle nuove figure che si fanno largo nella distribuzione, come i librai e gli stessi autori, questi ultimi attraverso meccanismi di *self-publishing*.

Il sesto capitolo propone uno sguardo sui dispositivi in circolazione. Difficilissimo da

tenere aggiornato per via delle continue novità immesse sul mercato, ci permette di introdurre alcuni aspetti tecnici che occorre tenere presenti nel progettare la produzione di libri digitali.

Il settimo capitolo è un breve spaccato sul copyright e sulla gestione dei diritti digitali: non intende esaurire gli aspetti legali legati a DRM e diritto d'autore, ma piuttosto riflettere sull'opportunità di certi sistemi di protezione dei file in un mondo pervaso da strumenti sempre più in grado di copiare e diffondere i dati alla velocità di un clic.

Una breve appendice, infine, su un software gratuito e open source – Calibre – capace di gestire una biblioteca digitale personale e di convertire i file in un numero veramente ampio di formati, per chi è

interessato a uno strumento agile e veloce per organizzare i propri libri digitali.

RINGRAZIAMENTI

Questo libro porta il mio nome sulla copertina, ma è frutto di un grandioso lavoro di squadra. Per questo sono orgogliosa di poter dire grazie a ognuna delle persone che ha contribuito alle idee “contenute” in queste pagine. A Fabio per la disponibilità a discutere di ogni aspetto anche a tarda ora; a Matteo per la pacatezza e la serenità con cui raggiunge gli obiettivi; a Patrizia per la pazienza con cui ha risposto a infinite domande; a Enrico per l’esperienza preziosa; a Gaia e Matteo per il sostegno psicologico; ad Aurelia per i rinforzi al cioccolato; a Stefano,

Alberto e Francesca per la curiosità e l’entusiasmo; e a Marco – *last but not least* – per averci creduto: banale ma mai scontato.

Prima di partire: dalla carta, al Web, all'ebook

Che ci si riferisca al prodotto, alle modalità di distribuzione, ai formati in uso, o al flusso di lavoro sul contenuto, quello che è certo è che l'editoria digitale ruota attorno ai libri digitali. Ma di cosa si tratta in effetti? Per quali aspetti – oltre che per la loro "immaterialità" – si differenziano da quelli "fisici"? E quali abitudini culturali sono in grado di influenzare?

Per provare a rispondere a queste domande può essere utile fare un passo indietro e chiarire cosa sia in realtà il libro stesso. Si tratta infatti di un oggetto ormai talmente quotidiano da apparire come un mezzo del tutto naturale per l'espressione delle idee.

I libri sono una tecnologia

In realtà nel libro cartaceo non c'è mai stato niente di naturale: è una tecnologia, sotto ogni punto di vista. Nessuna fase della sua produzione, con l'unica eccezione del concepimento dell'idea da parte dell'autore, vive lontana dall'artificio tecnologico. La carta su cui si scrive, la penna, la macchina da scrivere un tempo, il computer oggi, tutti

questi oggetti sono tecnologie. L'intero processo produttivo del libro si svolge in ogni sua fase – scrittura, redazione, impaginazione, stampa – grazie all'aiuto delle macchine. Per poter inquadrare nel modo giusto il libro digitale bisogna ricominciare a considerare anche il suo corrispettivo materiale come una tecnologia, e non ritenerlo "naturale" in virtù della maggiore familiarità che abbiamo acquisito nei suoi confronti.

Parlare del libro come oggetto ci permette inoltre di osservare altri due aspetti dovuti alla sua fisicità: le idee di *consumo* e di *possesso* dell'informazione.

Terminologia *Utilizziamo il termine informazione per indicare in senso lato tutti i contenuti testuali che possono essere pubblicati su carta o su altri supporti, includendo in questa espressione*

qualsiasi genere di "contenuto", dall'intrattenimento alle notizie e alla cultura in senso più ampio.

Si dà per scontato che l'informazione venga *consumata*, né più né meno di una risorsa esauribile che va ricreata per poter essere ancora utilizzata e diffusa. In realtà non è così che succede: l'informazione, una volta consumata, non scompare, ma anzi prolifera e si diffonde in mille modi, dal semplice passaparola al più raffinato e moderno mezzo di comunicazione che ci può venire in mente. L'informazione viene piuttosto *trasferita*, e qualunque sia il mezzo attraverso il quale questo "spostamento" si compie – una pagina web o di carta, un'immagine, un video e così via – esso rimane tale: un tramite, uno strumento per stabilire un collegamento tra chi ha concepito l'idea da comunicare o chi

ha trovato l'informazione da diffondere e chi viene a conoscenza di quell'idea o notizia, sia che questa venga letta, guardata o ascoltata. Che la trasmissione delle idee avvenga in modo analogico o digitale non fa alcuna differenza: si tratta comunque di mezzi distinti in modo netto dal loro contenuto. Quest'ultimo, proprio grazie alla sua assenza di fisicità, è in grado di adattarsi alle forme più varie di diffusione.

Anche per questo motivo, parlare di *possesso* dell'informazione è fuorviante: si dovrebbe parlare di *accesso* all'informazione. L'idea del possesso deriva dal fatto che, considerato il legame di identità profondo che unisce il libro fisico al suo contenuto, il possedere l'oggetto lascia pensare di possedere le idee in esso contenute. In realtà, possedere intere biblioteche

non è di alcuna utilità se i libri non si leggono: in quel caso non si possiede alcuna informazione ma solo una grande quantità di carta rilegata, ossia *il mezzo* di informazione. La differenza viene fatta dalla possibilità di accedere ai contenuti, il che non comporta necessariamente il possedere fisicamente i mezzi.

Dopo queste considerazioni si può forse osservare il libro con il giusto distacco e concordare sul fatto che nell'oggetto non si identifica il contenuto che esso veicola. Il libro è uno dei metodi più efficaci per fissare e divulgare il pensiero umano, ma non è certo l'unico. Del resto, da quando la stampa a caratteri mobili ha permesso che la lettura diventasse un fenomeno diffuso, dando un'enorme spinta all'alfabetizzazione di massa, l'idea che "libro" sia sinonimo di

“contenuto” ha avuto più di mezzo millennio per stratificarsi nel nostro pensiero.

Riferimenti *A proposito del libro inteso come “contenitore” di idee si può leggere l’introduzione a* Content *di Cory Doctorow (giornalista e attivista in favore della liberazione del copyright), a cura di John Perry Barlow. Il libro, in inglese e rilasciato gratuitamente sotto licenza Creative Commons, può essere scaricato all’indirizzo http://craphound.com/content/download/. Esiste anche una traduzione italiana, http://www.apogeonline.com/libri/9788850310852/scheda.*

L’organizzazione del contenuto nel libro: gerarchia e linearità

Non solo: il libro ha rafforzato ulteriormente l’abitudine a strutturare il contenuto in maniera gerarchica e lineare, rendendo questa prassi un modello imprescindibile – almeno all’apparenza – per l’organizzazione e la comunicazione delle idee.

Terminologia *La* gerarchia *è un sistema per ordinare e organizzare oggetti e concetti in modo graduato. Sulla base della gerarchia vengono costruite le* tassonomie, *termine che può indicare sia la classificazione gerarchica, sia la classificazione in senso lato.*

*L'immagine che rappresenta la tassonomia, secondo la matematica, è l'*albero*: a capo della struttura si trova una radice le cui proprietà si applicano a tutti gli elementi sottostanti.*

Terminologia *Con* linearità *si indica il modo di esprimere le idee secondo un processo graduale, senza salti logici, di tempo o di spazio. Gli elementi che caratterizzano il pensiero lineare sono i rapporti di causaeffetto, il principio di non contraddizione, la gerarchia e l'ordine logico, temporale e spaziale, appunto. La struttura lineare di un libro è resa evidente dall'indice, che elenca in sequenza progressiva i capitoli e la loro suddivisione gerarchica in paragrafi (Figura 1.1).*

Pensiamo come scriviamo: associazioni mentali come ipertesti

Il pensiero umano non si sviluppa in maniera naturale attraverso questo tipo di strutture: si tratta di conquiste raggiunte in secoli di formalizzazione dell'organizzazione della conoscenza, con lo scopo di sfruttare al meglio le nuove possibilità offerte prima dalla scrittura e poi dalla stampa. Il pensiero umano, in altre parole, non segue spontaneamente strutture gerarchiche nel suo sviluppo, ma procede per *associazione*.

L'associazione è una forma di pensiero più naturale rispetto a quella razionale-lineare. A differenza di quest'ultima è in grado di tollerare l'incoerenza, la contraddizione,

l'illogicità e la sovrapposizione arbitraria delle idee, applicando a tutto questo le capacità organizzative acquisite nel corso del tempo. Il pensiero associativo non si limita ad aggiungere qualcosa a ciò che c'era prima, ma può trasformare l'insieme in qualcosa di nuovo, attraverso un vero e proprio processo creativo.

Prefazione xi

I. Il testo pensato I

1.1 Storia

Evoluzione del concetto di testo 2

1.1.1 Premessa 2

1.1.2 Testo e scrittura 4

1.1.3 Il testo orale 8

1.1.4 Il mondo come testo 11

1.1.5 Il testo come intenzione comunicativa 12

1.1.6 Il testo come struttura 16

1.1.7 Il testo come oggetto 17

1.1.8 Il testo senza materia 21

1.1.9 La nuova interattività 22

1.2 Progettazione

Scrivere per pensare 26

1.2.1 Mettere ordine nei pensieri 26

1.2.2 Pianificare il testo 31

1.2.3 Scrivere le idee 33

Figura 1.1 L’indice rappresenta la struttura gerarchica e lineare di un libro.

L’idea dell’ipertesto come spazio per il contenuto svincolato dai confini fisici del libro e dalle sue necessità strutturali piuttosto rigide ha affascinato a lungo teorici e tecnologi, che hanno visto in esso possibilità prima impensabili per la fruizione del testo. La svolta è data dall’utilizzo dell’ipertesto attraverso l’informatica, un’attività piuttosto recente rispetto a tutte le versioni cartacee già ampiamente diffuse (indici analitici, mappe concettuali e così via). L’ipertesto digitale mette in discussione il testo stesso, che non può essere più definito come una sorta di entità localizzata all’interno del libro “fisico”.

Riferimenti *Non ci dilungheremo qui in un discorso sulle radici degli*

*ipertesti. Per chi volesse approfondire questo aspetto, suggeriamo di partire da Ted Nelson e il suo progetto Xanadu del 1960 (*Literary Machines 90.1*, Muzzio, Padova, 1992; sito ufficiale http://www.xanadu.net/) oppure dall'articolo scritto nel 1945 per l'Atlantic Monthly da Vannevar Bush, "As we may think" (http://www.theatlantic.com/doc/194507/bush, in inglese).*

Ancora, l'ipertesto ci porta a riflettere sulla ridefinizione dello *spazio della scrittura* determinato dalla tecnologia dell'informazione che si sta utilizzando.

Riferimenti *Per approfondire il discorso sul concetto di spazio della scrittura una lettura interessante può essere il saggio di Jay David Bolter,* Lo

spazio dello scrivere, *Vita e Pensiero, Milano, 2002*.

Ogni forma di scrittura è *spaziale*, dato che siamo in grado di vedere e comprendere solo quei segni scritti che abbiano un'estensione nello spazio, e ogni tecnologia di scrittura delimita il proprio, stabilendo anche modelli specifici di creazione, produzione, trasmissione e diffusione del sapere. Partendo dalle tavolette d'argilla, passando per papiri, codex in pergamena, carta, fino ad arrivare allo schermo del computer, il testo ha gradualmente perso parte della sua fisicità, trasformandosi in un'entità immateriale composta di 0 e 1 resa nuovamente comprensibile grazie a un hardware e a un software specifici. Attraverso il processo di digitalizzazione della cultura possiamo muoverci con un clic nella Rete di informazioni

sempre più vasta e sempre meno materiale. L'ipertesto ha ridefinito le modalità di produzione e trasmissione dei contenuti nella nostra epoca: il Web ne è la manifestazione più evidente.

Riferimenti *Negli anni Ottanta la casa editrice Edizioni E. Elle (poi diventata EL) cominciò a pubblicare i romanzi della serie Lupo solitario, avviando così una collana di libri-gioco, conosciuta come* Librogame. *I libri sono strutturati attraverso blocchi di testo numerati al termine dei quali il lettore può trovarsi a compiere una scelta secondo l'idea che si è fatto dello sviluppo della storia: a ogni opzione corrisponde il numero della sezione successiva. Una struttura in qualche modo ipertestuale, quindi!*

Il Web rivoluziona (anche) l’editoria?

A partire dal 1993 il World Wide Web ha diffuso due strumenti rivoluzionari per l’accesso all’informazione: gli ipertesti e i motori di ricerca.

Il Web è un’immensa ragnatela di ipertesti. Il contenuto del Web è navigabile grazie alla sua organizzazione in siti, blocchi di contenuto suddivisi in entità più piccole definite *pagine*, con un evidente richiamo alla struttura fisica del libro. Ma le pagine web sono collegate tra loro in maniera non sequenziale, attraverso dei *link* che rimandano a un’altra porzione della pagina o a un riferimento esterno al sito che si sta navigando in

quel momento. La base del Web è in gran parte costituita da testi e grafica, ma grazie alla diffusione delle connessioni a banda larga anche contenuti audio e video sono diventati quotidiani, trasformando il Web in un mezzo di comunicazione in grado di sfruttare tutte le tecnologie dell'informazione che fino a un paio di decenni fa convivevano in luoghi distinti e in qualche modo distanti. La possibilità di trasformare in digitale qualunque forma la comunicazione assuma (testo, audio, video, immagini e tutte le loro possibili combinazioni) ha permesso la diffusione di nuovi modelli di informazione e intrattenimento ormai irrinunciabili.

L'abitudine alla fruizione dell'informazione non solo attraverso il testo, e soprattutto non solo attraverso il testo stampato, fa sì

che le esigenze dei destinatari stessi cambino: la multimedialità, la possibilità di stabilire percorsi personali all'interno dei contenuti e la scomparsa dell'obbligo di seguire schemi precostituiti, l'immediatezza degli aggiornamenti, la possibilità di conservare le pagine più importanti non solo sul computer ma anche attraverso sistemi di archiviazione online spesso condivisibili, hanno portato nuovi bisogni informativi che obbligano i mezzi di informazione "tradizionali" (radio, televisione, giornali, libri) se non ad adeguarsi quantomeno a reinventarsi, per tenere testa all'enorme varietà di strade aperte dalla Rete per la diffusione dell'informazione.

Nuovi editori: i motori di ricerca

La quantità di informazioni che iniziarono ad accumularsi nel Web fin dal primissimo periodo della sua diffusione ha fatto sì che ci si ponesse subito il problema di come ritrovarle. A pochi anni della nascita del Web, nel 1995, si fa strada il primo motore di ricerca, Altavista, che determinava i risultati in base alle occorrenze della parola chiave inserita dall'utente rinvenute nelle varie pagine web.

Riferimenti *Il metodo utilizzato da Altavista (http://it.altavista.com/) finì ben presto vittima dello spam: molti siti web iniziarono infatti a inserire parole chiave non pertinenti al loro contenuto per scalare la classifica dei risultati, falsandoli e rendendoli così*

meno rilevanti. Dal 2004, in seguito a scelte strategiche poco fortunate, Altavista è stato acquisito da Yahoo!.

I motori di ricerca sono i nuovi editori del Web. Se il ruolo di un editore è quello di selezionare e proporre in forma strutturata e organica i contenuti, allora possiamo affermare che in Rete sono i motori di ricerca ad assolvere questo ruolo, partendo direttamente dalle richieste del lettore. Inserendo una chiave di ricerca si restringe l'area di interesse. Il motore si occupa – con metodi differenti a seconda di quello che stiamo utilizzando – di raccogliere per noi tutte le risorse attinenti disponibili in Rete, di segnalarci in qualche modo quelle che ritiene più attendibili e in qualche caso di suggerircene altre potenzialmente interessanti: di fatto si comporta come un editore che

sceglie per noi il contenuto che ritiene migliore su un determinato argomento, gli dà una struttura e ci propone in qualche caso contenuti affini (per esempio tramite una bibliografia). Da questo punto di vista Google è il più grande editore del mondo, considerata la quantità di pagine che è in grado di indicizzare e grazie alla velocità con cui restituisce i contenuti più interessanti.

Riferimenti *Data la mole di informazioni pubblicata sul Web, oltre ai motori di ricerca generalisti, che presentano comunque un ventaglio sempre più ampio di possibilità di raffinare la ricerca, ne esistono numerosi tematici, ossia dedicati alla ricerca solo in un determinato campo: ricerche scientifiche, blog, copertine di CD e così via. Esistono anche dei metamotori,*

come MetaCrawler o Ixquick, che riorganizzano i principali risultati di ricerca ottenuti sui motori generalisti.

Google è stato il primo motore di ricerca a rivoluzionare il sistema di indicizzazione, inventando il criterio del *PageRank*: la rilevanza di un sito non è più definita dalla ricorrenza di una parola chiave al suo interno, ma da quanti altri siti rimandano a quella stessa pagina. Più link rimandano a un sito, più quel sito viene considerato autorevole da Google, facendolo salire nella classifica dei risultati, che vengono elencati in ordine di pertinenza.

Riferimenti *“L’algoritmo di Google si basa sull’idea che le persone sono abili nel comprendere le cose e i computer sono efficienti nel contarle”, afferma Cory Doctorow nel saggio “Wikipedia.*

Una genuina guida galattica per auto-stoppisti, senza editori”, contenuto in Content, *già citato in questo capitolo.*

Quali che siano le necessità nel reperire informazioni, il concetto fondamentale rimane la rivoluzione profonda che i motori di ricerca hanno portato nella maniera di accedere ai contenuti. In meno di vent’anni il Web è diventato il punto di riferimento per ogni necessità di informazione, la base da cui partire per farsi almeno un’idea su un qualunque argomento, magari da approfondire in seguito, e su altri strumenti, ma con accanto il continuo supporto della Rete. La possibilità di digitare una parola e ottenere in pochi secondi milioni di risultati in ordine di rilevanza comporta la possibilità di portare sugli schermi dei nostri computer un livello almeno epidermico di conoscenza su

ogni materia. In questo nuovo modello di fruizione della cultura ciò che fa la differenza è ancora una volta la possibilità di *accesso* all'informazione, e non il possesso.

Riferimenti *In realtà, il possesso non della conoscenza stessa ma dei mezzi per accedervi – le nuove tecnologie – costituisce uno spartiacque fondamentale tra coloro che possono permettersi l'accesso alle informazioni del Web e tutti quelli che ne restano tagliati fuori. Il fenomeno del* digital divide, *il divario digitale tra chi può accedere a tecnologie come Internet e personal computer e chi no, è un fenomeno molto studiato, e le cause comunemente individuate sono principalmente di ragione economica, di*

assenza di un livello di istruzione adeguato e di infrastrutture.

Tutto ciò influisce sui libri per almeno due aspetti.

In primo luogo colpisce direttamente i libri di consultazione intesi come *reference*, termine che indica quei libri da leggere non in maniera lineare ma come *riferimenti*, appunto, su una determinata materia (dizionari, enciclopedie e così via). Questo settore editoriale è passato dall'allegato multimediale su supporti ottici alla diffusione dei contenuti via Web con siti dedicati, prevalentemente in modo gratuito.

La gratuità e la disponibilità di contenuti sono la chiave di volta del Web come mezzo di comunicazione. Realtà come Wikipedia (http://it.wikipedia.org), «l'enciclopedia

multilingue collaborativa, online e gratuita», come si legge nel lemma su Wikipedia in Wikipedia stessa, hanno segnato con il loro successo un punto di non ritorno.

Riferimenti *Un effetto di questo successo è rappresentato dalla decisione di Microsoft di abbandonare la pubblicazione di Encarta, l'enciclopedia multimediale avviata nel 1993 e disponibile online in abbonamento. Come si legge nella dichiarazione di Microsoft «[...] il settore delle enciclopedie tradizionali e dei materiali di reference è cambiato. Le persone oggi ricercano e consumano le informazioni in modi considerevolmente diversi rispetto al passato». Il testo della dichiarazione può essere consultato per intero all'indirizzo*

http://encarta.msn.com/guide_page_FAQ/FAQ.html, *in inglese*.

Si pensi anche all'operatore di ricerca che – su Google – permette di rintracciare tutte le definizioni di un termine premettendo alla chiave di ricerca `define:`. I risultati della ricerca mostrati nella Figura 1.2 sono definizioni del termine "Libro"; `define:` consente in pochi secondi di fare una ricerca mirata su tutte le fonti online che forniscono una definizione del termine di nostro interesse, possibilità con cui un libro fisico non può competere.

Figura 1.2 Le definizioni del termine “Libro” trovate con Google grazie all’operatore define:.

In secondo luogo, attraverso i motori di ricerca è possibile ritrovare le informazioni nelle pagine dei libri stessi: un esempio è Google Books (Google Libri, nella versione italiana), progetto controverso avviato nel 2004 con il nome di Google Print per estendere la ricerca al testo dei libri

digitalizzati da Google. Del progetto e dei suoi punti critici parleremo ancora nel Capitolo 7; ciò che qui interessa osservare è che un simile strumento di ricerca crea un'opportunità in più per mettere in contatto il potenziale lettore con il libro, offrendo la possibilità di consultarlo per valutarne il contenuto e di sapere dove si trova l'informazione che sta cercando senza costringerlo a sfogliarlo fisicamente.

Ci sono altri aspetti del Web che influenzano il modo di interagire con l'informazione. Il Web 2.0 è ormai una realtà talmente affermata da non poter essere certo considerata una novità, ma qui vale la pena ricordarne alcuni elementi che hanno contribuito a diffondere alcune abitudini fondamentali per il futuro dei libri digitali.

Riferimenti *Il documento del 30 settembre 2005 a cui si può far risalire la diffusione della definizione* Web 2.0 *può essere letto per intero e in inglese all'indirizzo http://oreilly.com/web2/archive/what-is-web-20.html; una versione italiana è invece disponibile su http://www.awaredesign.eu/articles/14-Cos-Web-2-0. Scritto da Tim O'Reilly, editore ed esperto di nuove tecnologie, l'articolo intende chiarire alcuni punti sul valore del termine e tracciare una linea di confine con il Web della New Economy (definito* Web 1.0*), ritenuta una fase del mercato sostanzialmente conclusa.*

La trasformazione del Web 1.0 in 2.0 viene identificata in una serie di tecnologie che conoscevano o avevano in parte già

conosciuto un grande sviluppo: il peer-to-peer; la programmazione “leggera” consentita da linguaggi come XML; la diffusione dell’RSS e più in generale la trasformazione del software da prodotto a servizio, di cui la ricerca su Google è il principale esempio.

Folksonomie

Uno dei maggiori problemi del Web 1.0 era l’organizzazione dell’informazione all’interno dei siti. Applicando alla Rete quello che era ormai da secoli il modello di indicizzazione tipico della scrittura e del libro fisico, anche per la gestione dei contenuti online si erano affermate strutture gerarchiche.

Riferimenti *Un metodo alternativo ai motori di ricerca per il ritrovamento delle informazioni sul Web è rappresentato dalle web directory, ossia*

realtà caratterizzate da una struttura ad albero (Yahoo! È l'esempio principale, Figura 1.3). Una web directory indicizza solo l'home page dei siti che vengono inseriti, classificati e gestiti manualmente dagli addetti al servizio.

Questo tipo di struttura ripropone sul Web gli stessi limiti del libro: la necessità di collocare una risorsa in un unico punto della struttura e la conoscenza precisa dei metodi di ricerca per individuarla. Finché non si è pensato di sfruttare l'intelligenza collettiva e lasciare l'indicizzazione agli utenti stessi: ciascuno può etichettare le informazioni come preferisce, attraverso dei *tag* (etichette, appunto) che ne descrivano il contenuto. Questo modello di classificazione è ciò che si definisce *folksonomy*, folksonomia.

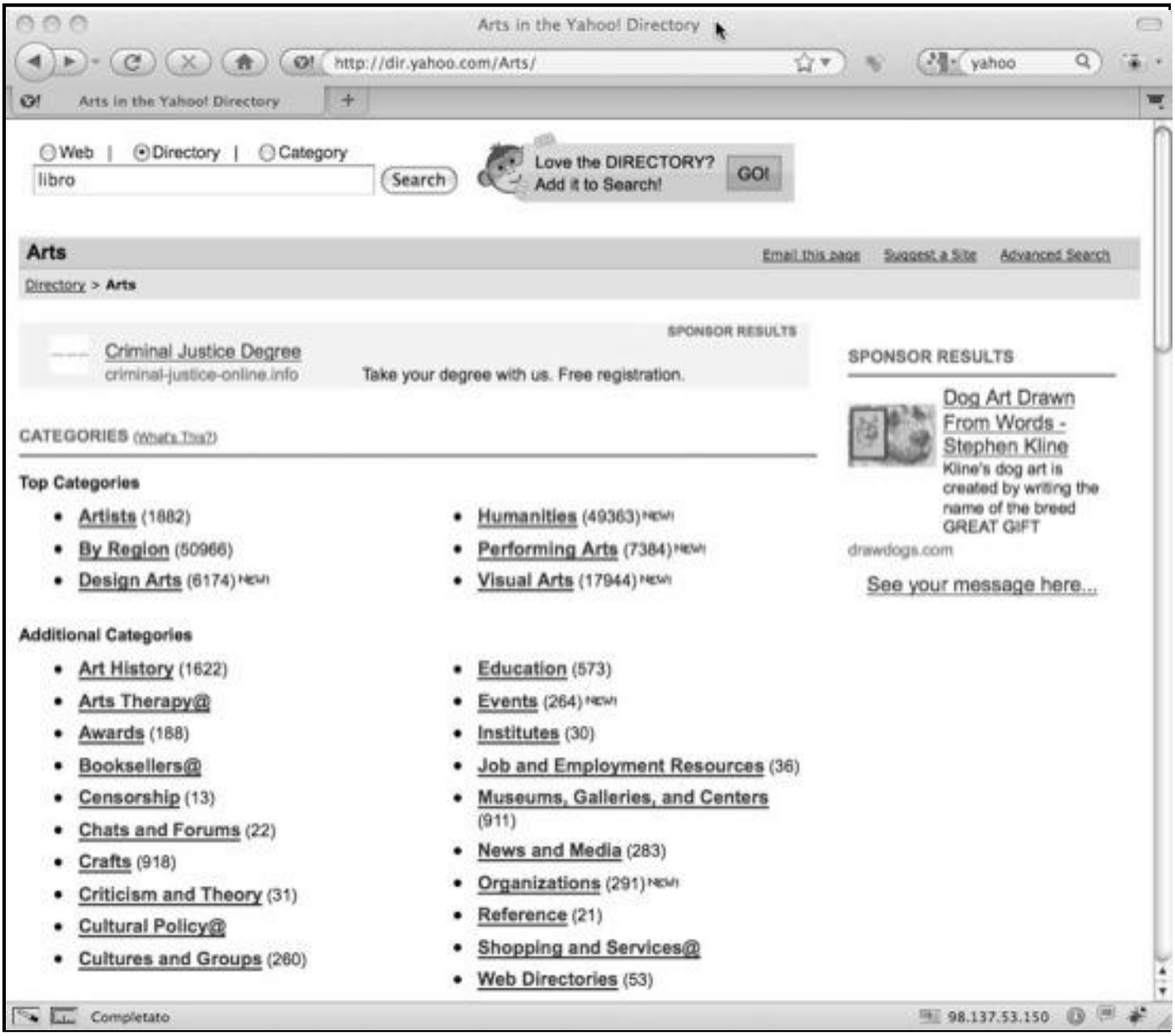

Figura 1.3 La ricerca per categorie su Yahoo!, attraverso la quale il motore propone una serie di risorse relative ai diversi argomenti contenuti nella directory.

Terminologia Folksonomia *è un neologismo derivato dall'unione delle*

parole folk *e* tassonomia*: indica una classificazione creata dagli stessi utenti in base a criteri individuali. Si riferisce al metodo utilizzato da gruppi di persone che organizzano in categorie le informazioni disponibili sul Web in modo spontaneo e collaborativo. Le folksonomie funzionano meglio quando è un gran numero di utenti a descrivere la stessa informazione: in questo modo essa acquista rilevanza e può essere rintracciata attraverso un numero maggiore di parole chiave, tra cui prevalgono quelle più pertinenti.*

Tale sistema di catalogazione “dal basso” ci conduce all’altro aspetto essenziale del Web 2.0: l’emergere del contenuto generato dagli utenti. A partire dai blog, fino ad arrivare all’esplosione dei social network, la Rete è

diventata una fitta trama di conversazioni incrociate, un punto di riferimento per una vera e propria vita sociale, anche se virtuale.

Terminologia *I* blog *(contrazione di web-log, traccia su rete), nati negli Stati Uniti nel 1997 e diffusi in Italia dal 2001, sono, semplificando al massimo, delle pagine personali in forma di diario, con una disposizione dei contenuti prevalentemente cronologica. La fortuna dei blog è dovuta anche alla diffusione della tecnologia dei feed RSS, di cui si parlerà più avanti in questo capitolo. Per rendersi conto della varietà di tipologie di blog esistenti si può consultare Wikipedia all'indirizzo http://it.wikipedia.org/wiki/Blog#Tipologie_di_blog.*

Terminologia *I* social network *(reti sociali), si diffondono negli Stati Uniti dal 2003 e mirano a riprodurre la rete di relazioni sociali del mondo reale su Internet. Friendster è stato uno dei primi ad affermarsi tra quelli di carattere generale; oggi tra i più popolari figura Facebook, con centinaia di milioni di iscritti in tutto il mondo. Ne esistono però anche di tematici, per esempio LinkedIn per il mondo del lavoro e Dopplr per i viaggiatori.*

Non a caso, una delle più note applicazioni del Web 2.0 è una sorta di motore di ricerca con l'indice determinato dagli utenti: Del.icio.us. Nato nel 2003, ha come scopo la semplificazione della gestione dei siti preferiti da un utente (o dei segnalibri, a seconda delle etichette utilizzate dai diversi browser

per indicare la stessa funzione), permettendo di creare un profilo personale nel quale salvare e catalogare attraverso dei tag che ne descrivano il contenuto i link verso qualsiasi altro sito.

L'innovazione non è tanto nel servizio al singolo utente – qualunque browser web permette oggi di salvare e organizzare i siti preferiti e spesso anche di sincronizzarli con un profilo online per averli a disposizione su computer diversi – quanto nella possibilità di condividere l'archivio di indirizzi Internet con gli altri, onorando l'aspetto "social" del Web 2.0.

Il vantaggio nell'utilizzo della folksonomia invece della tassonomia sta nella multidimensionalità del nuovo metodo organizzativo. Le tassonomie, infatti, costringono a

collocare l’informazione in un punto della classificazione unico e determinato, a meno di non creare dei doppioni. Se l’immagine ideale per rappresentare la tassonomia è l’albero, quella per la folksonomia è la nuvola (in inglese *cloud*), dove la scritta con il carattere di dimensione più grande indica una maggiore ricorrenza per quel determinato tag (Figura 1.4).

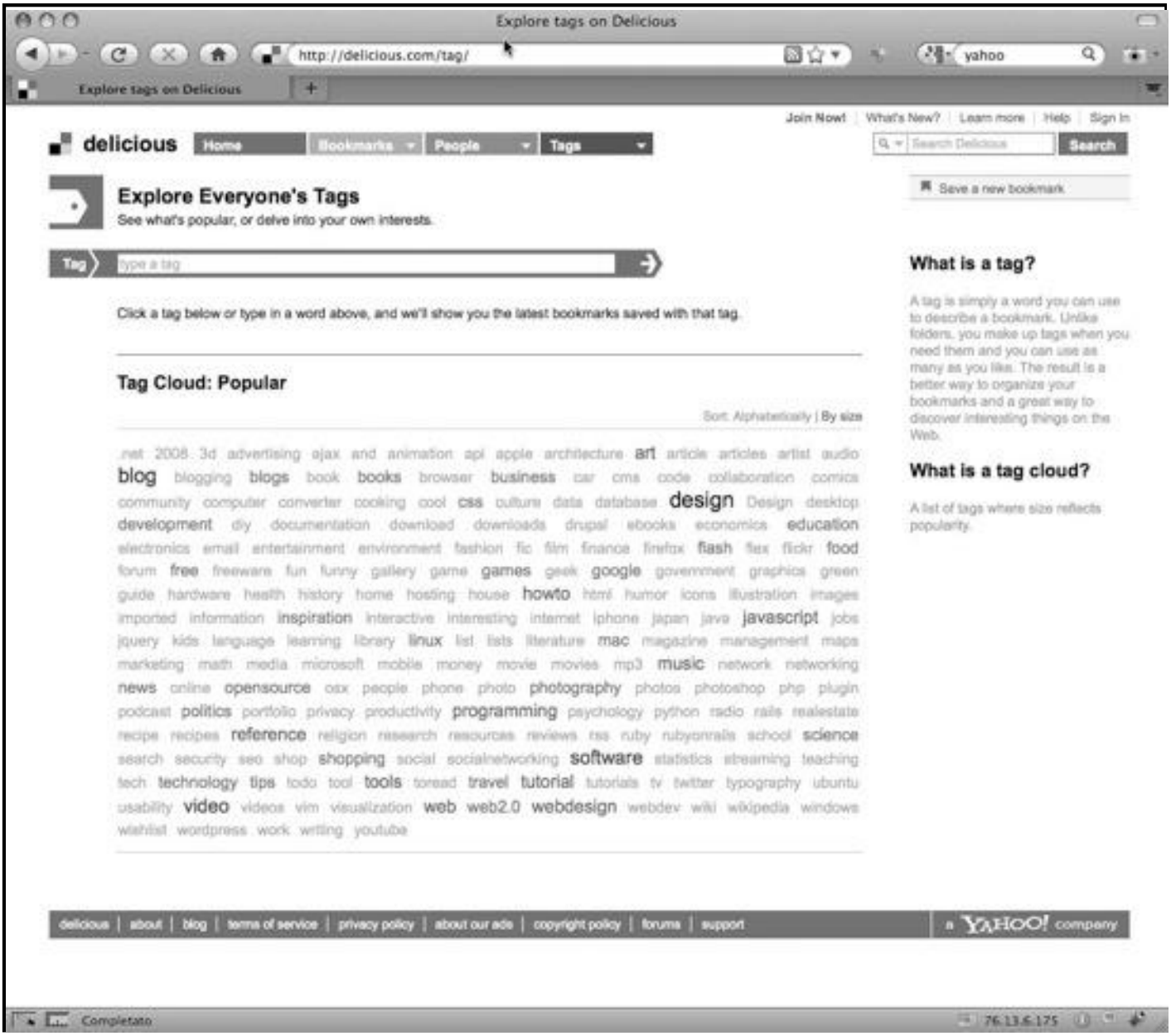

Figura 1.4 La nuvola dei tag più popolari sul sito Del.icio.us.

Rispetto a un normale motore di ricerca, Del.icio.us offre alcuni vantaggi: è più facile procedere nella ricerca per affinità, dato che

un tag è associato ad altre parole simili; ogni sito ha un indice di popolarità dato dal numero di persone che ne hanno salvato l'indirizzo; si può navigare tra i link degli utenti che permettono l'accesso al proprio profilo, aspetto che si rivela molto utile specialmente quando quell'utente ha compiuto ricerche simili alle nostre e potrebbe già essersi imbattuto in pagine di nostro interesse salvando su Del.icio.us i risultati più validi.

Terminologia *La* serendipità *è la sensazione che si prova quando si scopre qualcosa di imprevisto mentre si era impegnati nella ricerca di qualcos'altro. Oltre a questo, il termine indica anche l'elemento tipico della ricerca scientifica, quando scoperte importanti vengono fatte mentre si ricerca altro. Il termine si adatta molto*

bene anche alle nuove consuetudini di ricerca sul Web, nelle quali una parola chiave può condurre a risultati inaspettati ma altrettanto interessanti.

RSS

Un altro strumento che ha cambiato il modo di avere a che fare con l'informazione sono i feed RSS. Questi consentono non solo di collegarsi a una pagina sfruttando il link, ma anche di "abbonarsi" a essa, inviando un avviso ogni volta che la pagina in questione viene modificata.

Riferimenti *Il* link, *ossia il collegamento ipertestuale, è la base della struttura della Rete. Quando vengono aggiunti nuovi contenuti gli utenti, prima o poi, li scoprono e, se li trovano interessanti, creano un collegamento*

che punta a essi, formando delle connessioni tanto più forti quante più volte lo stesso contenuto viene "linkato", come avviene per le sinapsi nel cervello. La rete delle connessioni sul Web cresce come risultato dell'attività dell'intelligenza collettiva. Di fatto, alla base del Web c'è il "diritto a linkare".

RSS è un formato per la distribuzione dei contenuti web che contiene in forma strutturata informazioni sulla pagina a cui è legato e i contenuti della pagina stessa.

Terminologia *Alla base di RSS sta* XML, *acronimo di* eXstensible Markup Language *(linguaggio di marcatura estensibile). "Estensibile" indica la possibilità di creare dei marcatori personalizzati, ossia di creare gli elementi sintattici che descrivono "marcando"*

porzioni di documento. Viene utilizzato per creare nuovi linguaggi in grado di gestire documenti strutturati, e per questo è detto "metalinguaggio". Le potenzialità di XML verranno analizzate meglio nel Capitolo 2.

In altre parole, pubblicando le notizie in questo formato si definisce una struttura adatta a contenerle, che comprenderà vari campi, per esempio autore, titolo, testo della notizia, riassunto e così via. Man mano che nuove notizie vengono pubblicate la struttura predefinita si aggiorna con i nuovi contenuti, che vengono inviati, raccolti e visualizzati in maniera uniforme da qualunque lettore RSS, anche se provenienti da fonti molto diverse.

Riferimenti *La tecnologia RSS nasce nel 1997 dal confluire della* Really

Simple Syndication *di Dave Winer, utilizzata per pubblicare gli aggiornamenti dei blog, e dal* Rich Site Summary *di Netscape, che permetteva la creazione di pagine personali con flussi di dati aggiornati regolarmente. Questa tecnologia è stata portata avanti dalla stessa società di Winer dopo che Netscape aveva perso interesse a riguardo.*

Per tenere sotto controllo tutte le fonti di informazione di nostro interesse – blog, quotidiani, azioni di borsa, meteo e così via – dovremmo collegarci a esse almeno una volta al giorno senza la certezza di trovarle aggiornate o, viceversa, col rischio che lo siano con una frequenza maggiore. I feed RSS ci consentono di sottoscrivere un abbonamento ai siti: tramite un *aggregatore* possiamo raccogliere tutti gli indirizzi dei feed

che ci interessano lasciando a lui il compito di verificare – a intervalli di tempo da noi stabiliti – la presenza di aggiornamenti, scaricandoli e permettendoci di visualizzarli velocemente, conservarli e approfondirli aprendo la pagina da cui provengono, o semplicemente dare loro un'occhiata e poi buttarli via.

In un sito che offre la possibilità di sottoscrivere un feed RSS viene visualizzata un'icona generalmente arancione, che raffigura delle onde radio, spesso presente anche nella barra degli indirizzi del browser (Figura 1.5).

Figura 1.5 Nella parte destra della barra degli indirizzi del browser potete osservare l'icona per abbonarsi al feed della pagina aperta.

Facendo clic sull'icona si accede al feed: salvando (o meglio sottoscrivendo) il suo URL sarà possibile ricevere gli aggiornamenti del sito che verranno gestiti dall'aggregatore secondo le preferenze dell'utente.

Riferimenti *La ricezione dei feed, a differenza per esempio delle newsletter, non comporta per l'autore del contenuto la necessità di inviare gli aggiornamenti ai propri iscritti, ma è richiesta esplicitamente dagli utenti.*

Riferimenti *Esistono due tipi di aggregatori: Desktop e Web. I primi sono dei software da installare sul computer – tra cui anche i principali client di posta elettronica – che permettono di consultare le notizie anche offline ma non offrono strumenti di pubblicazione. I secondi, invece, (come Bloglines e*

Google Reader) offrono la possibilità di pubblicare le notizie provenienti dai feed a cui l'utente si abbona e di condividerle con altri.

I feed RSS sono diventati uno dei servizi web più diffusi grazie alla loro semplicità sia d'uso sia strutturale, dovuta a quel tipo di programmazione "leggera" permessa dall'XML di cui si accennava riguardo alle innovazioni del Web 2.0. Un altro motivo del loro successo è la scarsità di vincoli posti per il loro utilizzo: sono stati progettati per permettere all'utente di vedere il contenuto desiderato nel momento in cui lui ha deciso di vederlo, senza alcuna imposizione (o possibilità di controllo) da parte dell'autore di quel contenuto. La propensione a rendere disponibili i dati piuttosto che a controllare

cosa succede quando arrivano al destinatario è un punto di forza di questa tecnologia.

La rinuncia al controllo totale dell'uso che viene fatto dell'informazione dopo che questa è stata rilasciata è una tendenza che si riflette anche in licenze come Creative Commons (CC), in cui la frase "alcuni diritti riservati" si contrappone a "tutti i diritti riservati" tipica del copyright. Le licenze CC non sono in alcun modo collegate a RSS, ma il loro intento di regolamentare una diversa gestione dei diritti sui contenuti è ancora più interessante, proprio perché rispecchia l'inclinazione a incoraggiare una diffusione sempre più libera e flessibile delle informazioni.

Riferimenti *Creative Commons (http://www.creativecommons.it/) è un'organizzazione non-profit che si*

occupa di rendere possibile l'utilizzo e la diffusione delle opere di ingegno nel rispetto delle norme vigenti. Con l'idea di espandere la portata delle opere di creatività, nel 2002 ha rilasciato il primo pacchetto di licenze, studiate per la diffusione delle opere in Rete e grazie alle quali si può scegliere tra diversi gradi di rinuncia ai diritti d'autore in favore di una più libera condivisione delle informazioni. Parleremo del rapporto tra diritto d'autore e opere digitali nel Capitolo 7.

Il Web non è solo per i computer

Tim O'Reilly rileva nel Web 2.0 un altro aspetto innovativo: l'utilizzo del Web non è più limitato ai soli computer.

L’idea alla base di questa affermazione è che i software sono tanto più utili quanto più sono in grado di funzionare e sincronizzarsi con più dispositivi. Un modello da questo punto di vista è iTunes, il software gratuito di Apple per la gestione dei file musicali, che consente sia di organizzare localmente la propria musica, sia di accedere all’iTunes Store e di acquistare tramite Internet brani musicali in assoluta legalità. Grazie a questa possibilità iTunes non può essere considerato semplicemente un’applicazione (ciò che sarebbe considerandone il solo utilizzo in locale), ma si configura come un vero e proprio servizio.

Riferimenti *La prima versione di iTunes risale al gennaio del 2001. Nel novembre dello stesso anno viene implementato il supporto per l’iPod, il*

lettore musicale portatile prodotto dalla stessa Apple. A partire dalla versione 4.0, nel 2003, viene introdotto l'iTunes Music Store (oggi iTunes Store), attraverso il quale è possibile acquistare e scaricare file musicali e diversi altri contenuti tramite Internet.

Parlare di iTunes ci permette fare diverse considerazioni, spostando per un momento il discorso dai libri alla musica. In questo campo, infatti, la rivoluzione digitale è avvenuta da tempo ed è ormai innegabilmente parte della nostra quotidianità. A partire dall'affermazione dell'MP3 come formato di compressione per la musica digitale e dalla seguente diffusione di lettori portatili capaci di riprodurli (i primi nel 1998), il modo di produrre, promuovere,

distribuire e ascoltare la musica è cambiato radicalmente.

L’MP3 e i relativi lettori hanno reso possibile portare con sé intere collezioni di dischi e ascoltare le canzoni preferite ovunque ci si trovi con ingombri minimi: i CD musicali non sono scomparsi e non sono neanche diventati una nicchia nel mercato discografico, ma hanno dovuto adattarsi alle nuove esigenze di ascolto e fruizione della musica.

Consideriamo alcuni elementi.

- L’innovazione e la concorrenza nella produzione dell’hardware, con il conseguente abbassamento dei costi per il consumatore e il miglioramento continuo nelle prestazioni tecniche dei dispositivi, sono

state una delle spinte più forti alla diffusione dell'abitudine al consumo di musica digitale.

- La pirateria musicale ha sconvolto il mercato discografico tra la fine degli anni Novanta e i primi anni del Duemila. Con lo sviluppo del mercato della musica digitale le perdite dovute a questo fenomeno sono state in parte arginate grazie a una politica di prezzi vantaggiosa e a nuove modalità di distribuzione più flessibili e immediate, offrendo di fatto all'utente un nuovo servizio.
- Secondo l'NPD Group, il mercato della musica digitale negli Stati Uniti è pari al 35%. All'interno di questa percentuale, iTunes detiene – al 2009 – il 25% delle vendite, seguito da Amozon (8%) e lasciando la percentuale rimanente ad attori minori (iTunes è passato dal 14% del 2007 al 21% del 2008,

confermando la crescita nel 2009). Il rimanente 65% è costituito dalle vendite di CD, che restano ancora in una posizione dominante.

Queste osservazioni non ci servono di certo per tracciare un modello di sviluppo dell'editoria digitale. Piuttosto, sono utili per riflettere sulle innovazioni che hanno stravolto un settore dell'industria dell'intrattenimento, e possono offrire degli spunti per ragionamenti paralleli sulle possibilità aperte dai libri digitali nel mercato editoriale. Vediamone alcuni.

- Se uno degli stimoli fondamentali per la musica digitale è stato lo sviluppo di dispositivi sempre più "performanti" e accessibili dal punto di vista economico, nel settore degli ereader la tecnologia sta guadagnando

terreno. L'apripista in questo senso è stato Kindle di Amazon; vedremo nel dettaglio le evoluzioni e le caratteristiche dei dispositivi di lettura nel Capitolo 6.

- Se la diffusione massiccia della pirateria musicale (intesa sia dal punto di vista delle copie pirata sia dei download illegali) ha rappresentato l'apice della crisi e allo stesso tempo un punto di svolta obbligato per l'industria discografica, l'editoria combatte da tempo sul fronte delle copie illegali e delle digitalizzazioni dei libri che vengono poi messe in Rete e scaricate illegalmente. Tuttavia, mentre la pirateria interessa senza differenze ogni genere musicale e in quell'ambito la produzione di copie illegali è poco laboriosa e rapida, per i libri non è lo stesso: la scansione è una procedura laboriosa e difficilmente automatizzabile e la qualità

ottenuta non è paragonabile a quella degli MP3 per immediatezza e precisione.

- Il mercato dei libri digitali, specialmente negli Stati Uniti, ha visto aumentare in modo costante la sua dimensione e lascia intravedere ulteriori margini di crescita.
- I dispositivi attualmente in commercio per la riproduzione di file musicali di rado si limitano a questa sola funzione. Sempre più spesso integrano schermi a colori per la visualizzazione di foto, video, giochi e altre applicazioni, così come del resto i telefoni cellulari di ultima generazione possono riprodurre MP3, scattare fotografie e catturare video. Si è diffusa l'abitudine, insomma, ad avere a portata di mano dispositivi multifunzione, in grado di rispondere alle esigenze più diverse di intrattenimento ovunque ci

si trovi, con poco ingombro. L'esempio di riferimento è il telefonino multifunzione di Apple: l'iPhone dispone di rete Wi-Fi, lettore MP3, fotocamera integrata, navigatore satellitare e migliaia di applicazioni per ampliare e creare sinergie tra le funzioni del dispositivo. Tra queste molte sono dedicate alla lettura dei libri digitali: ne vedremo alcune nel Capitolo 6.

- Gli ereader sono ancora dispositivi molto costosi rispetto alla quantità di funzioni che riescono a svolgere. Il loro punto forte è la tecnologia E-Ink, che permette un minore affaticamento della vista durante la lettura, e la sempre più diffusa possibilità di connettersi senza fili alle risorse messe a disposizione sul Web per scaricare libri digitali sia a pagamento sia gratuitamente. A questo vanno aggiunti una serie di strumenti pensati

specificamente per migliorare l’esperienza di lettura, sfruttando tutti vantaggi che la digitalizzazione del libro offre: come spesso accade a una minore varietà di funzioni corrisponde un maggiore sviluppo di quelle presenti.

In effetti, uno dei più grossi ostacoli per la lettura di libri digitali è rappresentato dall’affaticamento per la vista che deriva dalla luminosità dei monitor (o dalle dimensioni, per certi dispositivi portatili): per questo l’E-Ink è la soluzione tecnologica migliore. A questo proposito, sono interessanti alcune considerazioni di Cory Doctorow proprio riguardo alle difficoltà della lettura su schermo:

“Non mi piace leggere da cima a fondo lo schermo di un computer”: è un luogo comune

---

> del mondo dell'ebook. Significa "non leggo interi romanzi sullo schermo di un computer" (o cellulari, o palmari, o dispositivi per la lettura di ebook), e, spesso, le persone che affermano ciò, sono le stesse che trascorrono ogni secondo della giornata leggendo interamente lo schermo del computer. È come guardare qualcuno infilarsi un'intera tavoletta di cioccolato in bocca mentre ti dice quanto lo odia [...] Questo non significa che [il computer] non sia un *mezzo di intrattenimento*, anzi: qualsiasi cosa io faccia sul computer mi diverte un sacco. E in più, è quasi tutto basato sul testo. Sostanzialmente, ciò che faccio sul computer è una sorta di lettura piacevole. Ma è fondamentalmente un tipo diverso di piacere, più dispersivo e frammentario.
> --Cory Doctorow, "Voi adorate leggere dallo schermo del computer", Content.

Quindi, ciò che rende complesso leggere dallo schermo di un computer non è tanto l'affaticamento della vista (di fatto leggiamo già a lungo dallo schermo sopportando bene

la fatica che ne deriva), quanto piuttosto l'elevato numero di stimoli che un computer ci propone, soprattutto se connesso alla Rete, interrompendo continuamente la concentrazione sulla lettura. La criticità, dunque, è più forte sulle opere lunghe come i romanzi, per esempio. Per opere di questo tipo i dispositivi dovrebbero poter essere maneggiati con la stessa comodità di un libro e percepiti con la stessa naturalezza in modo da favorire l'"immersione" nella lettura, senza che l'immedesimazione e l'illusione create dal racconto vengano interrotte continuamente da altri stimoli.

La lettura digitale non sostituirà quella su carta. Un forte lettore non smetterà di comprare i libri stampati: semmai comprerà *anche* quelli digitali. Le due cose si affiancano e non si sovrappongono, perché hanno

diverse funzioni e luoghi. Ci sono generi, come si osservava a proposito dei libri di reference, che vivono sconvolgimenti maggiori nel passaggio al digitale. Ma anche le guide, i manuali, sia che seguano uno sviluppo del discorso lineare sia che abbiano una struttura a blocchi di contenuto abbastanza indipendenti, hanno davanti nuove prospettive nel diventare estremamente portatili occupando solo l'ingombro del dispositivo che li contiene.

I lettori che per primi potrebbero riconoscere i vantaggi dei libri digitali sono gli studenti e i professionisti, ossia coloro i quali hanno più di altri necessità di consultare frequentemente un ampio numero di libri, effettuare delle ricerche al loro interno, prendere appunti, inserire vari segnalibri, rileggere più volte un determinato passaggio e

metterlo in relazione con altre parti del libro e così via. Potersi spostare portando con sé tutti i testi necessari ai propri studi o per il proprio lavoro con tutti i vantaggi del formato digitale rappresenta senza dubbio una grande opportunità.

La situazione non è certo priva di aspetti critici. La digitalizzazione di questo tipo di libri, in cui spesso più che in altri si è meno disposti a chiudere un occhio su imperfezioni ed errori formali che inevitabilmente possono verificarsi, rischia di incontrare ulteriori ostacoli nel processo di conversione dai file per la stampa a quelli per il digitale, a causa dell'introduzione di errori che spesso non vengono corretti. Questo non è un problema che dipende né dal dispositivo né dal formato: è un problema di ordine economico. Come osserva Jeffrey Zeldman in

Kindling, i costi del libro stampato riescono a coprire anche delle fasi di verifica della correttezza del testo che un libro digitale non è, per ora, in grado di sostenere allo stesso modo.

Riferimenti *Jeffrey Zeldman, tra i maggiori esperti internazionali sugli Standard Web, è docente e autore specializzato nell'ambito del web design. Potete leggere il post citato sul suo sito all'indirizzo http://www.zeldman.com/2009/08/24/kindling/, in inglese.*

Finché il mercato dei libri digitali non diventerà abbastanza forte da sostenere economicamente fasi più accurate di revisione del testo convertito, il processo di conversione comporterà sempre un certo margine

di errore che non sarà possibile correggere. Finché non si potrà seguire un processo di produzione pensato specificamente per i libri digitali, eliminando il problema della conversione, quando si tratta di un libro che ha bisogno di estrema precisione in ogni dettaglio per poter essere fruito correttamente, sarà meglio affidarsi all'edizione cartacea.

Questo non vuol dire che pensare i libri digitali sia ancora un'idea buona solo per un romanzo di fantascienza. La tecnologia diventerà sempre più raffinata e i processi di produzione si adatteranno e svilupperanno di conseguenza, assecondando le richieste del mercato. La situazione attuale presenta già situazioni e iniziative degne di nota sotto diversi aspetti: sarà su questi che ci soffermeremo a partire dal prossimo capitolo.

La Bibbia di Lutero non ebbe successo sugli stessi binari che vedevano protagoniste Bibbie copiate a mano da monaci; principalmente perché erano brutte, non erano scritte nel latino ecclesiastico, non c'era nessuno che le leggesse in pubblico e le interpretasse, non rappresentava il frutto del lavoro devoto, con la D maiuscola, di qualcuno che ha dedicato la propria vita a Dio. La cosa che ha portato al successo la Bibbia di Lutero è stata la sua duttilità: era più popolare perché più diffusa, tutti i fattori di successo di un nuovo mezzo di comunicazione impallidiscono accanto alla sua rapida diffusione.

--Cory Doctorow, "Discorso sui DRM", Content

[...] nessuna delle cose che ha fatto della Bibbia di Lutero un successo ha fatto anche il successo delle Bibbie scritte a mano. Allo stesso modo, le ragioni per cui amare gli ebook hanno ben poco a che fare con le ragioni per cui amare i libri cartacei.

--Cory Doctorow, "Ebook, ovvero né e né book", Content

LINGUAGGI PER TESTI DIGITALI

Dalla codifica del testo ai linguaggi per definirne la struttura, passando per i principali formati per i libri digitali e infine analizzando i software attraverso i quali è possibile leggerli, produrli e manipolarli, in questo capitolo e nei due successivi cercheremo di descrivere un quadro preciso degli strumenti per gli ebook, cercando di chiarire sia i vantaggi sia i punti critici delle procedure per l'elaborazione digitale del testo.

Terminologia Digitale *è un aggettivo che in campo informatico definisce tutto ciò che viene rappresentato attraverso i numeri o che agisce manipolandoli. Il suo contrario è* analogico *e indica qualcosa che non è numerabile. Il procedimento di conversione dall'analogico al digitale è detto* digitalizzazione.

Cominciamo quindi dalle basi per creare un testo che sia comprensibile dal computer, ossia un testo codificato.

Il testo digitale: codifica e markup

Il testo è un oggetto complesso, che contiene delle informazioni strutturate su più

livelli: dalle sequenze di caratteri che, combinandosi, compongono le parole fino ad arrivare alle strutture linguistiche che determinano legami di significato in senso più ampio. I computer non sono in grado di interpretare direttamente questi diversi gradi di complessità: ciò che possono fare è interpretare delle sequenze di codice binario. Per far sì che riescano a cogliere i vari tipi di informazione contenuti nella struttura dobbiamo perciò predisporre il testo in modo che possano comprenderlo: dobbiamo, cioè, *codificarlo*.

L'operazione di codifica del testo opera su due livelli. Possiamo avere una codifica di livello zero, in cui ci si occupa di rappresentare ciascun carattere che compone il testo in codice binario, oppure una codifica di alto livello, in cui si rappresenta l'organizzazione

strutturale del testo. Per l'editoria digitale è importante questo secondo tipo di codifica.

La rappresentazione del testo tramite la codifica si basa su determinati linguaggi informatici, chiamati appunto *linguaggi di codifica*: essi permettono – oltre alla rappresentazione – anche il controllo di uno o più livelli della struttura testuale.

Utilizzare un linguaggio di codifica di alto livello significa in primo luogo individuare, selezionare e classificare gli elementi rilevanti del testo a seconda del livello strutturale che si intende codificare e in seguito stabilire uno schema che tracci le relazioni tra gli elementi così stabiliti.

I linguaggi che consentono di compiere questo tipo di operazioni vengono chiamati

linguaggi di marcatura (in inglese *markup language*).

Terminologia *Il termine* marcatura *deriva dall'abitudine tipografica di marcare con varie annotazioni quelle parti del testo che avevano bisogno di essere evidenziate o corrette, per segnalarle al compositore o al dattilografo.*

In generale, un linguaggio di marcatura descrive tramite convenzioni standardizzate i meccanismi di rappresentazione del testo a livello strutturale, semantico o di presentazione. I suoi elementi fondamentali sono:

- un insieme di marcatori o *tag*, che individuano dei "blocchi" testuali a cui verrà assegnata una determinata interpretazione;

- una grammatica, che regola l'uso dei tag;
- una semantica, che definisce la funzione della marcatura.

La marcatura viene inserita direttamente all'interno del testo, a cui viene applicata attraverso delle apposite istruzioni. Ogni istruzione utilizza quindi dei caratteri speciali che la delimitano e permettono all'elaboratore di distinguere il testo dal markup (Figura 2.1). Vedremo meglio in cosa consista un tag quando parleremo dei linguaggi in senso più specifico.

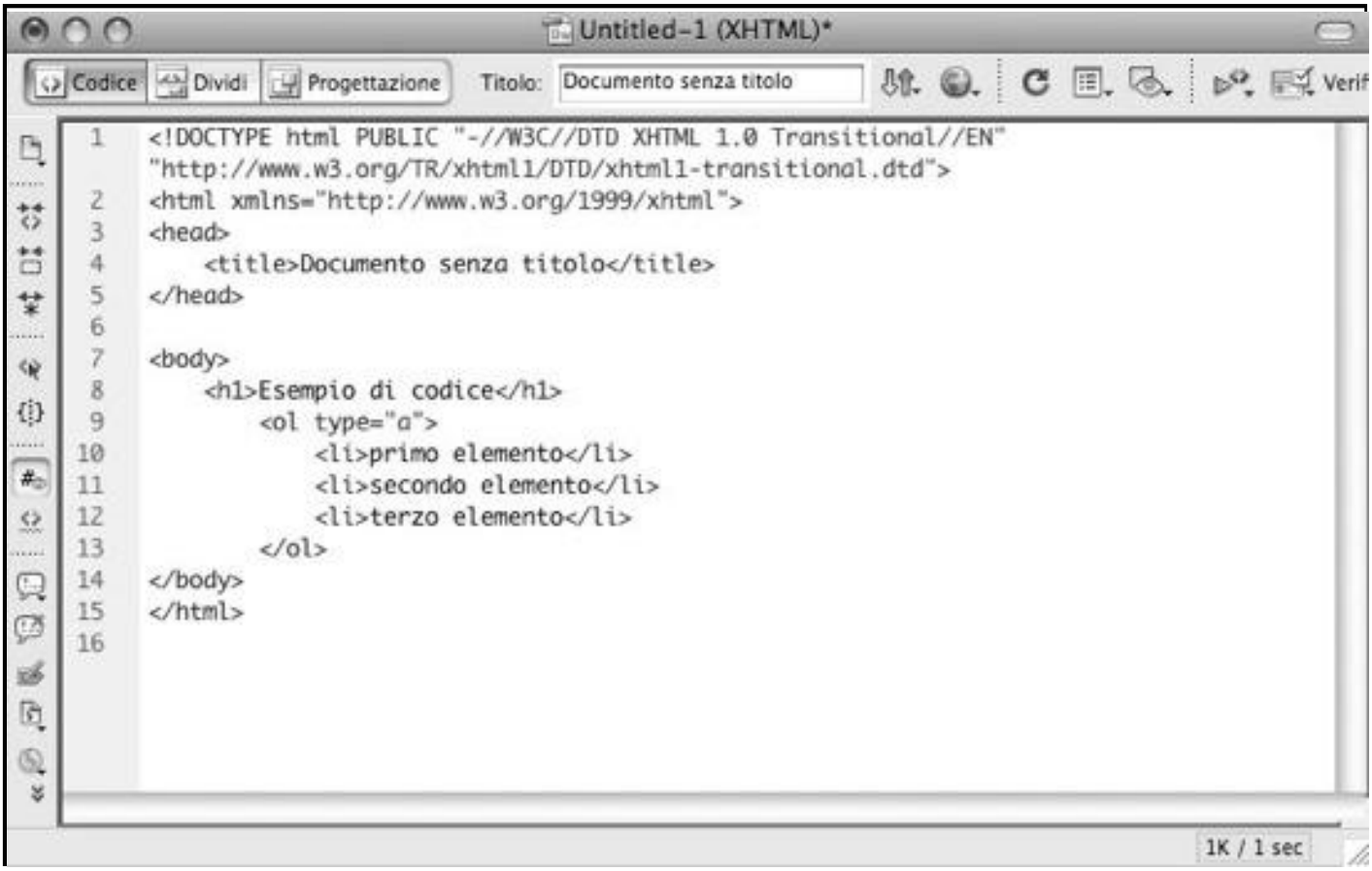

Figura 2.1 Ecco un esempio di come appare il testo delimitato dalla marcatura. I tag sono racchiusi tra i caratteri < e >.

Un linguaggio di marcatura, inoltre, deve rispondere a una serie di requisiti che ne misurano il valore per verificare se e in che misura esso sia adeguato alle esigenze della elaborazione e trasmissione digitale dei dati.

Questi requisiti possono essere suddivisi in tre aree:

- potenza espressiva;
- portabilità e riusabilità;
- standardizzazione e apertura.

La *potenza espressiva* è la caratteristica che permette al linguaggio di essere applicato al numero più ampio possibile di tipologie e generi testuali. Essa consente inoltre di rappresentare il maggior numero di livelli strutturali e di esprimere le caratteristiche del testo secondo diverse prospettive (per esempio dal punto di vista della metrica o della grammatica). Infine, permette di associare metadati descrittivi e gestionali ai documenti digitali.

Terminologia *I* metadati *sono, letteralmente, dei dati sui dati, ossia informazioni utilizzate per descrivere insiemi di dati. Possono essere di diverse tipologie: quelli* descrittivi, *per esempio, servono per l'identificazione e il recupero del documento, e descrivono la fonte alla quale sono legati; quelli* gestionali, *invece, racchiudono informazioni sulle modalità di archiviazione e manutenzione.*

La *portabilità* fa sì che il documento possa essere leggibile su qualunque tipo di piattaforma e dispositivo informatico senza problemi di compatibilità: in altre parole, il documento è indipendente sia dall'hardware sia dal software. La *riusabilità*, invece, riguarda la possibilità di accedere al documento

anche a distanza di tempo dalla sua creazione.

Infine, la *standardizzazione* richiede che il linguaggio sia utilizzato universalmente dalla comunità di utenti interessati, oltre che *aperto* e di pubblico dominio.

Terminologia *Il termine* aperto *significa che il formato (o il software) non è legato a un produttore o a un ambito specifico, rendendo di conseguenza il suo utilizzo libero da vincoli di proprietà.*

Esistono due categorie di linguaggi di markup: linguaggi procedurali (o specifici) e linguaggi descrittivi (o generici o analitici).

Linguaggi procedurali

I linguaggi procedurali specificano al computer quali operazioni deve compiere sul documento per presentarlo in una determinata maniera, ossia specificano la struttura tipografica e compositiva della pagina, le spaziature, l'interlinea e le caratteristiche dei font. La loro funzione principale è la resa grafica del documento: non forniscono alcuna informazione sulla struttura o sul contenuto del testo.

Terminologia *Il termine* procedurale *si riferisce al fatto che il linguaggio indica al computer le procedure di trattamento a cui deve sottoporre il testo al momento della presentazione.*

Due esempi di linguaggi procedurali sono RTF e LaTeX.

RTF

RTF, acronimo di *Rich Text Format* (formato di testo ricco) è un formato aperto realizzato da Microsoft con lo scopo di semplificare lo scambio di documenti fra diverse applicazioni. In RTF la marcatura indica che tipo di formattazione applicare al testo semplice: allineamento, impaginazione, tipo di carattere e così via.

Terminologia *Per* testo semplice *in questo caso si intende quello definito sfruttando soltanto set di caratteri molto semplici e diffusi, come per esempio l'ASCII (*American Standard Code for Information Interchange*), il primo standard per la rappresentazione*

binaria di caratteri, nonché il nucleo comune per numerosi set successivi.

Il testo così marcato è definito “ricco” perché passa dal puro testo (dalla sola codifica in ASCII) all’arricchirsi – appunto – di nuove informazioni riguardanti la formattazione.

RTF non è stato pensato per essere letto o codificato da un operatore umano, ma per essere utilizzato tramite un programma: quasi tutti i word processor sono in grado di importare ed esportare dei documenti in formato RTF.

Terminologia *Un* word processor, *o elaboratore di testo elettronico, è un software che permette di elaborare, formattare e salvare il testo inserito dall’utente. Può offrire funzioni più o meno complesse; tra i più noti*

ricordiamo Word (Microsoft), Pages (Apple) e Writer (OpenOffice).

Possiamo comunque dare un'occhiata all'aspetto di una porzione di testo marcata attraverso RTF:

```
{\rtf
Ciao! \par
{\b Questo testo è in grassetto.}
\par
}
```

che in un word processor in grado di interpretare i comandi del formato RTF apparirà così:

Ciao!
Questo testo è in grassetto.

La barra rovesciata `\` (o *backslash*) indica l'inizio di un codice di comando RTF. `\par`

segnala l'inizio di un nuovo paragrafo; `\b` attiva il testo in grassetto (*bold*, in inglese). Le parentesi graffe `{` e `}` definiscono un gruppo: la porzione di testo in grassetto nell'esempio è racchiusa in questo modo e distinta dal resto del paragrafo. `\rtf` è il primo comando che costituisce il gruppo-radice entro cui è racchiuso tutto il testo. I codici di comando devono essere separati da uno spazio dal testo a cui vengono applicati: questo spazio non comparirà nella visualizzazione del testo, ma è necessario per la corretta interpretazione del linguaggio da parte della macchina.

Attenzione *Aprire un documento con il comando* `\rtf` *e lasciare lo spazio tra il tag e il testo anche se questo non verrà visualizzato significa rispettare la sintassi del linguaggio che si sta*

utilizzando: sono operazioni fondamentali affinché il computer interpreti correttamente le istruzioni che gli diamo.

LaTeX

LaTeX è un linguaggio di markup distribuito con una licenza di software libero: questo ha permesso la sua disponibilità per qualsiasi tipo di sistema operativo.

Riferimenti *Pronunciato /'l?:t?k/, LaTeX è un linguaggio di alto livello derivato da TeX, un programma di tipografia digitale adatto alla scrittura di documenti matematici e scientifici, ideato nel 1978 da Donald Knuth per scrivere il libro* The Art of Computer Programming. *La pronuncia del nome deriva dalla radice della parola greca* τέχνη, *arte, tecnica, tecnologia:*

l'ultima lettera presente nel nome è una "chi" greca maiuscola (X) la cui grafia si confonde con quella della latina "ics".

Viene utilizzato soprattutto per la preparazione di testi scientifici e nelle scienze dell'informazione, ed è molto apprezzato in ambito accademico. Questo è dovuto alle sue enormi potenzialità riguardo alla gestione dell'impaginazione delle formule matematiche, alla gestione dei riferimenti bibliografici, delle note, delle tabelle e delle figure, oltre che dell'organizzazione generale delle pagine.

Riferimenti *La gestione delle liste di riferimenti bibliografici viene fatta in LaTeX tramite BibTeX, attraverso il quale l'informazione bibliografica viene separata dalla modalità di presentazione. Si tratta di un file*

testuale distinto, con estensione `.bib`, *nel quale gli elementi bibliografici possono venire inseriti con un ricchissimo livello di dettaglio. Lo stile per il documento* `.bib` *deve essere specificato a parte dal documento LaTeX.*

LaTeX è un linguaggio molto complesso, che in questa sede non è fondamentale analizzare più a fondo. Esistono numerosissime comunità online a riguardo: si può cominciare l'esplorazione dal sito ufficiale: http://www.latex-project.org/.

Linguaggi dichiarativi

Se i linguaggi procedurali si caratterizzano per la loro attenzione specifica sulla formattazione del testo, i linguaggi dichiarativi si concentrano sul contenuto, ossia sul

tentativo di rappresentarne la struttura astratta. I simboli di codifica (come i codici di comando di RTF) non sono impiegati per definire informazioni di formattazione o di impaginazione, ma diventano indicatori di struttura. Il *generic markup* (la marcatura generica) si basa su un insieme di marcatori – o tag – che, inseriti direttamente nel testo, indicano non le caratteristiche di presentazione di quel blocco di testo ma la funzione che esso assolve. In altre parole, a ciascun blocco viene assegnato un tag che lo definisce come titolo, paragrafo, citazione, nota e così via, senza occuparsi di stabilire l'aspetto di quella porzione di testo.

Non occupandosi della resa grafica del documento, il linguaggio dichiarativo è indipendente dal dispositivo che lo visualizzerà. Per quanto infatti possano essere precise e

dettagliate le istruzioni riguardo la presentazione del documento indicate dai linguaggi procedurali, queste dipenderanno sempre, almeno in una certa misura, dalle caratteristiche della macchina che dovrà interpretarle.

Riferimenti *Se per esempio il linguaggio procedurale indica come font del documento un carattere che non è presente nel computer che dovrà interpretare le sue istruzioni, la macchina ne sceglierà uno predefinito per la resa del testo, perdendo così l'informazione indicata dal linguaggio e modificandone di fatto l'aspetto in maniera sostanziale.*

In genere la resa grafica di un documento trattato con un linguaggio di markup

dichiarativo è affidata a un secondo documento testuale chiamato *foglio stile*. Questo, su modello dei linguaggi procedurali, associa ai tag delle istruzioni di resa grafica: soluzione particolarmente vantaggiosa in quanto consente – grazie alla possibilità di collegare più fogli stile a uno stesso documento – di associare al testo numerose soluzioni di presentazione, anche molto diverse tra loro, che rendono possibile soddisfare le esigenze di visualizzazione più diverse a partire dallo stesso documento sorgente.

Riferimenti *Approfondiremo questo aspetto nel paragrafo dedicato agli standard correlati a XML.*

I punti di forza dei linguaggi dichiarativi sono immediatamente visibili quando si pensa alle possibilità offerte dal trattamento

informatico dei testi. Tanto per fare un esempio, un'applicazione per la ricerca e l'estrazione delle informazioni potrebbe permetterci di conoscere le occorrenze di un termine esclusivamente nei titoli di un documento, oppure quante volte lo stesso termine viene pronunciato da un determinato personaggio.

SGML

SGML, acronimo di *Standard Generalized Markup Language*, fu sviluppato per consentire lo scambio di documenti – specie in progetti governativi, industriali e legali – che potessero essere interpretati dai computer per diverse decadi.

Riferimenti *Le origini di SGML risalgono agli anni Sessanta, quando il markup più diffuso era quello*

procedurale, che comportava non pochi problemi nello scambio di documenti tra macchine e applicazioni diverse, a causa della diversa resa grafica delle istruzioni. Per questo si pensò di elaborare un linguaggio che mantenesse distinte le informazioni di presentazione da quelle di struttura.

Oltre alla semplificazione della condivisione dei documenti, un altro aspetto che si rendeva sempre più necessario per la diffusione delle informazioni era la definizione di uno standard per la trasmissione e l'archiviazione dei documenti stessi. SGML divenne uno standard ufficiale nel 1986: si trattava del primo sistema di codifica dichiarativa astratto e generalizzato.

In realtà, SGML è un *metalinguaggio* più che un vero e proprio linguaggio di

marcatura. Esso infatti non definisce direttamente i tag per la struttura logica del testo, ma fornisce piuttosto una serie di regole di sintassi attraverso le quali è possibile definire le norme da applicare per la marcatura di specifici documenti.

Riferimenti *Due delle applicazioni più note di SGML sono HTML e XML, che vedremo nei paragrafi successivi.*

Potremmo definirlo come una super-grammatica con cui definire le grammatiche. È, in altre parole, un linguaggio per creare altri linguaggi, ognuno dei quali costituisce un'*applicazione* SGML, definita così in quanto applicazione dello standard. SGML è una sorta di capostipite dei linguaggi di marcatura in cui i tag sono sempre delimitati dalle parentesi angolari < e >, e sono

caratterizzati da un nome (per esempio capitolo, titolo e così via) ed eventualmente accompagnati da attributi espressi come coppie `NomeAttributo=”valore”`.

Tra i concetti teorici su cui si basa l’impostazione del linguaggio troviamo quello di *tipo di documento*. Il tipo di documento si occupa della descrizione delle caratteristiche di un gruppo di documenti dalla struttura omogenea, identificandone la struttura, appunto, gli elementi che la compongono e le relazioni sia di dipendenza sia di ricorrenza che si instaurano tra questi. L’immagine che rappresenta questa struttura è quella del grafico ad albero.

Il tipo di documento viene definito attraverso un DTD (*Document Type Definition*), strumento attraverso il quale si crea

un’applicazione SGML. Tramite un DTD si determinano le regole per la presenza, l’ordine e la posizione delle etichette di marcatura e dei loro attributi all’interno di un documento.

> **Attenzione** *Mentre un documento SGML deve essere obbligatoriamente associato a un DTD, XML, come vedremo, permette anche documenti ben formati: questi presentano una struttura sufficientemente regolare e comprensibile da poter essere utilizzata anche in assenza di DTD.*

In altre parole, un DTD elenca e definisce tutti gli oggetti necessari all’elaborazione di un certo tipo di documento. Vediamoli nel dettaglio.

- *Elementi*: i marcatori e il relativo contenuto.

- *Content model* (modello di contenuto): indica per ciascun elemento quali altri elementi, quando e in che numero; possono comparire al suo interno. Indica inoltre gli attributi che si possono associare a un elemento.

  Suggerimento *Gli attributi consentono di aggiungere informazioni, utili per descrivere più accuratamente la struttura dei dati rappresentata. Possono anche specificare funzioni e caratteristiche non propriamente strutturali, ossia metainformazioni. Possono essere paragonati a degli aggettivi.*

- Entità: permettono di creare collegamenti a dati esterni al documento; possono anche indicare dei caratteri che non sono presenti

nella tabella codici del documento o che potrebbero non essere rappresentati correttamente nel passaggio da un sistema all'altro.

Attenzione *Nel secondo caso, le entità sono utilizzate per caratteri che non fanno parte della tavola ASCII.*

Sintatticamente, la definizione di un elemento all'interno di un DTD può essere divisa in due parti. Nella prima si riporta il suo identificatore generico (in pratica il nome del marcatore), mentre nella seconda si descrive la struttura del suo contenuto:

La definizione degli attributi si effettua per ogni elemento. La struttura della dichiarazione di un attributo è la seguente:

```
<!ATTLIST nome_elemento

della specificazione
>
```

Dopo la parola chiave `ATTLIST` si inserisce il nome dell'elemento di riferimento; segue la dichiarazione vera e propria degli attributi (uno per riga), l'indicazione del contenuto e l'obbligo/possibilità della loro specificazione nell'elemento.

Vediamo un frammento di DTD:

```
<!ELEMENT capitolo (titolo,
paragrafo+)>
<!ATTLIST capitolo
n CDATA #IMPLIED
id ID #REQUIRED
>
<!ELEMENT titolo (#PCDATA)>
```

Abbiamo alcune dichiarazioni, dette *markup declaration*. Ciascuna è aperta dal simbolo `<!` cui segue una parola chiave che identifica il tipo di dichiarazione: `ELEMENT` per gli elementi, `ATTLIST` per gli attributi. Segue l'identificatore dell'elemento, ossia il nome del tag.

Tra parentesi tonde troviamo il modello di contenuto dell'elemento, che contiene l'elenco degli elementi figli che possono essere contenuti dall'elemento padre. Se gli elementi figli sono separati da una virgola, significa che hanno l'obbligo di presentarsi nell'ordine definito dalla dichiarazione di markup: nel nostro esempio l'elemento `titolo` deve sempre precedere l'elemento `paragrafo`.

Il simbolo + presente nella dichiarazione di elemento è un indicatore di ricorrenza: specifica quante volte un elemento può essere presente dentro un altro. Nel nostro esempio, l'elemento capitolo può contenere uno o più paragrafi.

Riferimenti *Esistono tre tipi di indicatori di ricorrenza:* + *indica che l'elemento ricorre 1 o più volte;* * *0 o più volte;* ? *1 o 0 volte.*

La dichiarazione degli attributi prevede, dopo il simbolo di apertura e il nome dell'elemento a cui gli attributi sono associati, la dichiarazione del nome di ogni attributo seguito dal tipo di valori ammissibili. (La parola chiave CDATA indica che l'attributo può avere come valori stringhe di testo non definibili a priori: in pratica l'attributo può

avere qualsiasi valore testuale.) Quindi segue l'obbligatorietà o l'opzionalità dell'attributo. Questa viene espressa con le parole chiave `#IMPLIED`, `#REQUIRED` e `#FIXED`.

Riferimenti #REQUIRED *indica che l'attributo è obbligatorio,* #IMPLIED *facoltativo, mentre* #FIXED *specifica che l'attributo ha un valore fisso che deve essere opportunamente dichiarato. La parola chiave* #FIXED *stabilisce una dichiarazione di costante per il valore di un attributo. Se il valore è diverso da quello dichiarato, il documento non è valido.*

Linguaggi derivati da SGML, come XML e HTML, sono oggi più diffusi rispetto al pieno utilizzo di SGML stesso: ce ne occuperemo nei paragrafi seguenti.

HTML

HTML è l'acronimo di *HyperText Markup Language*, ossia linguaggio di marcatura per ipertesti. Fu ideato da Tim Berners-Lee nel 1990 con lo scopo di sviluppare un sistema di pubblicazione e ritrovamento dell'informazione in grado di tenere in contatto la comunità internazionale dei fisici.

Riferimenti *Per concretizzare la sua idea, Berners-Lee ideò un linguaggio per creare ipertesti consultabili via Rete – HTML, appunto – e un protocollo di trasferimento, HTTP,* HyperText Transfer Protocol. *Definì inoltre un nuovo metodo per l'indirizzamento dei documenti in Internet, URL (*Universal Resource Locator*). L'applicazione derivata dall'insieme di queste tecnologie è il World Wide Web:*

la sua prima concezione era incentrata sul testo attraverso un browser con un'interfaccia interamente testuale per il reperimento e la navigazione tra informazioni messe in relazione tra loro secondo associazioni non sequenziali. Sulle origini del Web si può leggere il libro di Tim Berners-Lee L'architettura del nuovo Web, Feltrinelli, Milano 2001.

HTML è un linguaggio di pubblico dominio e il suo standard è definito dal *World Wide Web Consortium* (W3C); la versione attualmente in uso è la 4.01. È concepito per definire la struttura logica di un documento e non il suo aspetto: per questo motivo anche dispositivi senza grandi capacità grafiche possono elaborarne il contenuto, anche se questo viene visualizzato in modi

differenti. Attualmente i documenti HTML sono in grado di incorporare altre tecnologie che ne espandono le potenzialità, arricchendole di colori, immagini, animazioni, streaming audio video e così via.

Riferimenti *Fondato nel 1994 da Tim Berners-Lee in collaborazione con il Cern, il World Wide Web Consortium (http://www.w3.org/) ha come scopo principale la creazione di standard e linee guida per il Web. Sul sito si possono trovare pagine di riferimento su SGML, HTML, XML e molto altro ancora. All'indirizzo http://www.w3.org/TR/html4/intro/sgmltut.html, per esempio, si possono approfondire le relazioni tra SGML e HTML.*

I documenti ipertestuali scritti in HTML hanno come estensione `.html` o `.htm`. Anche qui l'*elemento* costituisce la base della sintassi del linguaggio: si occupa di formattare i dati rendendo possibile la loro interpretazione al dispositivo che dovrà rappresentarli. Per esempio, `<h1>` indica l'apertura di un elemento che definisce un titolo di primo livello. Il tag di chiusura si differenzia da quello di apertura per la presenza della barra `/` (slash):

```
<h1>Questo è il testo del titolo</h1>
```

Il testo compreso tra `<h1>` e `</h1>` è indicato in questo modo come titolo di primo livello.

Riferimenti *Altri tag non hanno bisogno del tag di chiusura poiché richiedono un'applicazione puntuale,*

come accade per `<img>`*, che serve per inserire un'immagine in un determinato punto della pagina; per questo sono definiti tag a chiusura implicita.*

I documenti HTML si aprono con l'indicazione della definizione di tipo di documento (DTD) che segnala le specifiche utilizzate e indica così implicitamente quali elementi, attributi ed entità possono essere presenti. Il dispositivo è così in grado di identificare le regole di interpretazione e visualizzazione appropriate:

Nell'esempio la versione di riferimento è la 4.01. In seguito a questa dichiarazione il documento si sviluppa attraverso una struttura ad albero:

```
<html>
 <head>
  <title>Titolo della pagina</title>
 </head>
 <body>
  <h1>Corpo della pagina</h1>

 </body>
</html>
```

Dopo la dichiarazione del DTD l'elemento più esterno e che comprende l'intero documento è `<html> </html>`, al cui interno sono sempre previste due sezioni distinte e in sequenza: l'intestazione (o *header*), compresa tra i tag `<head> </head>`, che contiene informazioni di controllo, e il corpo della pagina (`<body> </body>`), che racchiude il contenuto vero e proprio della pagina. La Figura 2.2 mostra l'output del codice sopra riportato così come appare in un browser.

Figura 2.2 Il listato precedente visualizzato in un browser.

Le ulteriori sottosezioni all'interno dell'intestazione o del corpo della pagina non hanno obblighi particolari di ordine e posizionamento, eccetto la correttezza negli annidamenti: ogni sottosezione deve essere chiusa prima che inizi la successiva.

Attenzione *HTML è un linguaggio* case insensitive*: significa che non fa*

differenza se i tag sono scritti in lettere maiuscole o minuscole.

Ogni tag può avere una serie di attributi che possono assumere diversi valori. La struttura per aggiungere uno o più attributi a un tag è la seguente:

Nell'esempio sottostante vediamo che l'elemento `<ol>` (*ordered list*, elenco ordinato) ha per attributo `type="a"`: significa che la lista avrà per indice le lettere dell'alfabeto in minuscolo al posto dei numeri (Figura 2.3).

```
<h1>Esempio di codice</hi>
 <ol type="a">
  <li>primo elemento</li>
  <li>secondo elemento</li>
  <li>terzo elemento</li>
</ol>
```

Attraverso gli attributi si possono specificare anche caratteristiche stilistiche, come altezza e larghezza degli elementi, colore, margini, allineamento del testo e così via. Tuttavia, è preferibile raccogliere questo tipo di informazioni in un documento distinto, il *foglio stile*, che si occupa di definire l'aspetto grafico della struttura HTML in maniera più efficace e leggera. Parleremo dei fogli stile nei paragrafi successivi.

Figura 2.3 Ecco come compare in un browser l'esempio di codice: un elenco numerato con lettere per indice.

Riferimenti *Non è possibile in questa sede approfondire oltre le regole del linguaggio HTML. Per chi volesse saperne di più un ottimo punto di partenza è il sito www.html.it, ricco di tutorial, articoli e guide adatte anche chi si avvicina per la prima volta all'argomento.*

XML

XML (*eXtensible Markup Language*) è un metalinguaggio di marcatura derivato da SGML, del quale definisce una versione semplificata per creare in modo semplice nuovi linguaggi di markup adatti al Web. *eXtensible* (estensibile) si riferisce alla possibilità di creare dei tag personalizzati.

Anche se XML nasce con lo scopo di ottenere una maggiore flessibilità per la gestione dei documenti sul Web, le sue enormi potenzialità lo hanno reso adatto all'utilizzo nei contesti più vari – dalla definizione della struttura di documenti, allo scambio delle informazioni tra diversi sistemi, alla rappresentazione di immagini e alla definizione di formati di dati – permettendogli di trovare un numero molto maggiore di applicazioni rispetto a quelle pensate in principio. Ecco alcuni dei vantaggi derivati dal suo utilizzo.

- È indipendente dal tipo di piattaforma hardware o software su cui viene utilizzato, oltre che dai dispositivi di archiviazione e visualizzazione: in altre parole, permette l'interoperabilità del dato.

Terminologia *L'*interoperabilità *è la possibilità di scambiare dati attraverso sistemi informativi diversi con semplicità e affidabilità.*

- Consente la descrizione di qualsiasi tipo di documento o struttura testuale indipendentemente dalle finalità applicative.
- È un formato di pubblico dominio.
- Sono disponibili numerose applicazioni e librerie open source basate su diversi linguaggi di programmazione per la manipolazione di dati in formato XML.

In qualità di metalinguaggio, XML è in grado di creare delle regole sintattiche che definiscono formalmente un linguaggio di markup, che in questo caso viene detto *applicazione XML*. Ognuna di queste applicazioni eredita da XML determinate

caratteristiche di sintassi e allo stesso tempo definisce alcune regole specifiche.

Ogni applicazione XML si basa, come SGML, sulla definizione di tipo di documento che descrive le caratteristiche di un insieme di documenti strutturalmente omogenei. Il riferimento al tipo di documento deve essere espresso implicitamente mediante il markup in ogni documento XML, mentre è facoltativa la sua espressione esplicita attraverso un DTD.

Struttura di XML

La struttura fondante di XML si sviluppa su tre basi:

- elementi;
- attributi;
- documenti validi e ben formati.

Gli *elementi* sono i componenti strutturali del documento. A ogni elemento possono venire associati uno o più *attributi* che ne specificano ulteriori caratteristiche o proprietà non strutturali.

Un documento XML *ben formato* dichiara implicitamente la propria struttura nel markup e rispetta i vincoli di buona formazione indicati nelle specifiche del linguaggio. Perché un documento sia ben formato non è necessario che sia associato a un DTD. Nel momento in cui, invece, il DTD è dichiarato esplicitamente e il documento risponde alle sue norme sintattiche, questo è definito, oltre che ben formato, anche *valido*.

Attenzione *A differenza di HTML, XML è* case sensitive, *ossia è sensibile*

all'uso delle maiuscole o delle minuscole nella scrittura dei tag.

Come in tutti i linguaggi di marcatura, anche in XML i tag servono la struttura e il valore semantico del testo. A differenza di HTML la sintassi è però molto più rigida. Ciò che è fondamentale è che il documento sia prima di tutto ben formato, cioè che risponda ad alcuni principi di base della sintassi XML.

- Deve essere presente un unico elemento radice che contenga tutti gli altri elementi del documento. Possono trovarsi al di fuori solamente i commenti e le direttive di elaborazione, come per esempio la dichiarazione della versione di XML utilizzata.

Suggerimento *I commenti sono delle sezioni di testo delimitate dalle*

sequenze di caratteri `<!--` *e* `-->` *che vengono ignorati dal programma che interpreta il documento e quindi non vengono visualizzati. Sono utili perché permettono di inserire osservazioni e annotazioni nella composizione del documento e consentono di renderlo più comprensibile anche a chi non l'ha creato.*

- Ogni elemento deve avere un tag di chiusura, o la sua forma abbreviata in caso sia vuoto.
- La nidificazione dev'essere corretta, ossia i tag aperti per primi devono essere chiusi per ultimi.
- XML è case sensitive, pertanto i nomi di tag e attributi devono coincidere in relazione a quest'aspetto.

- I valori degli attributi devono essere sempre racchiusi tra apici singoli o doppi.

Il non osservare queste regole comporta che il documento non sia ben formato e quindi che non venga interpretato correttamente dall'applicazione che lo esegue.

Concretamente, un file XML è un file di testo contenente una serie di elementi, attributi e testo strutturati secondo una precisa gerarchia. L'organizzazione degli elementi, dunque, segue una struttura ad albero che prevede l'esistenza di un elemento *root* (radice) che contiene tutto il contenuto del documento. Vediamola con un esempio concreto:

```
<?xml version="1.0" ?>
 <capitolo>
```

```
   <paragrafo>
    Blocco di testo del primo
    paragrafo
    <immagine file="immagine1.jpg"/>
   </paragrafo>
   <paragrafo>

   <nota>
    Esempio di nota
   </nota>
   </paragrafo>
</capitolo>
```

Nella prima riga del documento troviamo la dichiarazione obbligatoria del tipo di documento e della versione del linguaggio utilizzata: `<?xml version="1.0" ?>`. Quello che segue è il contenuto del documento vero e proprio, i cui tag sono delimitati, come in HTML e SGML, dai caratteri `<` e `>`. I nomi dei tag sono definibili liberamente:

la loro organizzazione gerarchica è impostata attraverso la nidificazione.

Non intendiamo dare conto in queste poche pagine di tutte le caratteristiche e norme di XML. I pochi concetti fondamentali che abbiamo espresso dovrebbero consentirci di proseguire nella lettura senza troppe difficoltà: ciò che importa in questo momento è l'avere ben chiare le possibilità di questo linguaggio di marcatura. Vediamo adesso come espanderle attraverso l'utilizzo di altri linguaggi.

Standard correlati a XML

Come già osservato, una delle maggiori potenzialità di XML sta nella possibilità di delegare ad altre tecnologie la definizione di quegli aspetti che non vengono espressi

dalla marcatura, come per esempio la formattazione.

Attraverso la definizione di DTD e XML Schema si possono creare grammatiche mirate a determinare e verificare la correttezza formale del documento; tramite CSS e XSL possiamo intervenire sulla presentazione dei dati o trasformarli in formati differenti; con XLink e XPointer si possono creare collegamenti tra diversi documenti XML, mentre con XQuery e XQL possiamo estrarre informazioni in base a determinati criteri. Nei paragrafi seguenti vedremo le principali caratteristiche delle tecnologie associate a XML.

DTD e XML Schema

Il DTD è – come abbiamo visto per SGML – un metodo per definire la grammatica di

un linguaggio di marcatura, definendo i tipi di elemento che possono essere contenuti dal documento, gli attributi che possono essere a essi associati e le relazioni gerarchiche che intercorrono tra loro. Il suo funzionamento per XML è lo stesso che abbiamo descritto per SGML.

Il W3C, inoltre, ha validato XML Schema come unico linguaggio ufficiale di descrizione del contenuto di un file XML. Analogamente a un DTD, esso permette di stabilire la grammatica del linguaggio di markup, ma la differenza fondamentale sta nel fatto che mentre un DTD utilizza una sintassi specifica, XML Schema è scritto in XML: in questo modo è in grado di stabilire le regole con una flessibilità ancora maggiore.

CSS e XSL

A differenza di HTML, XML non integra alcun tipo di informazione riguardo alla presentazione: in altre parole non è possibile in alcun modo dare istruzioni sulla visualizzazione del documento attraverso il linguaggio di marcatura. Per questo si utilizzano diversi tipi di fogli stile, ossia dei documenti collegati a XML che contengono le informazioni necessarie per presentare il contenuto nella maniera desiderata. Allo stesso documento possono essere associati più fogli stile, distinti – per esempio – in base al dispositivo che dovrà visualizzarlo.

I CSS (acronimo di *Cascading Style Sheet*, foglio stile a cascata) permettono di separare le informazioni di visualizzazione dal codice che descrive la struttura astratta del

documento, rendendo quest'ultima più leggibile e flessibile.

Un foglio stile a cascata si basa su delle *regole* che indicano gli elementi del documento a cui applicare una determinata serie di informazioni di stile. La struttura di una regola è la seguente:

```
selettore {
proprietà1 : valore1;
proprietà2 : valore1, valore2;
}
```

Il selettore indica a quale elemento della pagina vanno applicate le regole racchiuse tra le parentesi graffe. Le proprietà sono una serie di caratteristiche che possono essere modificate secondo le indicazioni espresse nel valore indicato accanto:

```
body {
font-size: 12px;
}
```

Nell’esempio sopra, si stabilisce che il testo contenuto nell’elemento body abbia una dimensione di 12 pixel.

Riferimenti *Naturalmente questo non può essere che un accenno alla complessità dei CSS. Per approfondire l’argomento, oltre al gran numero di risorse web, si può valutare la lettura di* CSS*, di Gianluca Troiani, pubblicato da Apogeo sempre nella collana Pocket.*

Essendo pensati specificamente per il Web i CSS mancano di alcune caratteristiche necessarie ad altri ambiti di presentazione, come per esempio la stampa. Per questo il W3C ha definito un insieme di specifiche in

grado di gestire in modo più flessibile la presentazione e la trasformazione di documenti XML: XSL.

Acronimo di *eXtensible Stylesheet Language*, XSL è in realtà una vera e propria famiglia di specifiche, con tre componenti non dipendenti l'una dall'altra e per questo ancora più flessibili.

- *XSL Transformation*, ossia XSLT, un linguaggio per la trasformazione di XML che controlla le operazioni che permettono di rendere presentabili i dati.
- *XML Path Language* (detto XPath), utilizzato da XSLT per individuare gli attributi e gli elementi del documento XML sui quali applicare le operazioni di trasformazione.

- *XSL Formatting Objects*, o XSL-FO, un vocabolario per definire la formattazione di un documento.

L'indipendenza dei linguaggi tra loro fa sì che se uno di essi non è adatto per raggiungere gli obiettivi prefissati per il documento possiamo sostituirlo con un altro senza influire sugli altri elementi.

In altre parole, ciò che XSL consente di fare attraverso le sue componenti è di selezionare i marcatori di un documento XML e di intervenire sulla loro presentazione ed eventuale trasformazione.

Riferimenti *Grazie a XSLT possiamo trasformare un documento XML (detto documento di origine) in un altro documento (detto documento risultante) che può essere ancora in XML o in un altro*

formato, come RTF, XHTML, PDF e così via.

Applicazioni rilevanti di XML sui testi

Esistono diversi tipi di applicazioni XML progettate per rispondere alle esigenze più varie. Per quanto riguarda la gestione di testi complessi ne prenderemo in esame quattro: ognuna definisce il suo insieme di marcatori e l'insieme di regole sul come usarli.

XHTML

L'acronimo XHTML significa *eXstensible HyperText Markup Language*, linguaggio di marcatura di ipertesti estensibile. In pratica si tratta di una versione di HTML conforme allo standard XML. Questo significa principalmente che la sintassi è più restrittiva: se in

HTML sono tollerate indicazioni di layout e resa grafica all'interno della struttura logica della pagina, in XHTML queste sono gestite e imposte tramite i CSS.

Riferimenti *Anche XHTML è uno standard ufficiale del W3C. È nato nel 2000 e può essere considerato come un ponte tra HTML 4.01 e XML 1.0, come vedremo tra poco.*

XHTML risponde alla necessità di avere un linguaggio con una sintassi più definita rispetto a HTML, in modo da poter creare pagine web estensibili e portabili su qualsiasi piattaforma, caratteristica propria di XML, come abbiamo visto.

Le potenzialità di XHTML vengono accentuate, come accennavamo, dall'utilizzo combinato con i fogli stile, che permettono di

mantenere separata la struttura logica dei dati dalla sua presentazione.

La differenza principale rispetto a HTML è che i tag devono essere *ben strutturati*: devono cioè obbedire ad alcune regole che ne assicurano la coerenza reciproca. Non possono esistere tag a chiusura implicita, mentre è diffusa la forma di chiusura abbreviata: `<img />`, per esempio. XHTML è, inoltre, *case sensitive*; in più i tag devono essere scritti in lettere minuscole. Tutti i valori degli attributi devono essere obbligatoriamente racchiusi tra virgolette, aspetto non sempre necessario né in HTML né in SGML.

Esistono tre differenti DTD per XHTML, che corrispondono ad altrettanti DTD di HTML 4.01, dato che si tratta di una riformulazione di questo linguaggio.

- XHTML 1.0 Transitional: è un DTD particolarmente tollerante nei confronti di alcuni elementi HTML deprecati e accetta anche un certo margine di non conformità sintattica della struttura.
- XHTML 1.0 Strict: è il DTD più severo. Non tollera tag HTML deprecati né le difformità sintattiche; esegue controlli rigorosi anche rispetto agli attributi dei tag.
- XHTML 1.0 Frameset: ormai in disuso, si basava sulla pratica ora deprecata dal W3C di suddividere la finestra visualizzata dal browser in sottofinestre (*frame*, appunto).

Terminologia *Un elemento si dice* deprecato *nel momento in cui viene indicato come obsoleto e ne viene quindi sconsigliato l'utilizzo in favore di una nuova prassi.*

Attualmente è in uso la versione XHTML 1.1. Si tratta di una riformulazione dell'XHTML Strict, e fa riferimento a un solo DTD.

Riferimenti *Esistono versioni di XHTML dedicate a particolari dispositivi, come la Basic (per i dispositivi non in grado di interpretare il set completo di comandi e tag) e la Mobile Profile (pensata per i cellulari e sviluppata a partire dalla versione Basic con l'integrazione delle funzionalità proprie dei terminali telefonici).*

DocBook

DocBook è un linguaggio di marcatura semantico basato su XML nato inizialmente per semplificare la stesura di documenti tecnici relativi a computer e software, ma

presto utilizzato anche per altri argomenti. Il numero di elementi semantici che DocBook può individuare è vastissimo ed estremamente dettagliato. I tag sono divisi in tre macrocategorie.

- Strutturali, ossia la categoria che individua il contenitore – libro, articolo, capitolo, appendice e così via – all'interno del quale si annidano gli altri elementi o, in altre parole, gli elementi che costituiscono la struttura del documento.
- Elementi di blocco, come titoli, paragrafi, liste, che possono essere contenuti nel livello superiore (strutturale).
- Elementi in linea, ossia dei marcatori che sono in grado di individuare una porzione più piccola all'interno di un elemento di

blocco, dandole valore semantico (tutti i tipi di enfasi, per esempio).

> **Terminologia** *Gli elementi di* blocco *costituiscono una porzione di testo "finita", un blocco, appunto: unità di contenuto concluse (nel testo, infatti, sono evidenziate dall'andata a capo). Gli elementi* in linea*, invece, possono essere contenuti all'interno di un blocco, e servono per enfatizzarne una parte. Questa distinzione ci tornerà molto utile nel Capitolo 4, a proposito di stili di paragrafo e di carattere.*

In questo modo possiamo dire che un *libro* contiene dei *capitoli* che a loro volta contengono alcuni *paragrafi* al cui interno possono essere indicati elementi rilevanti (in *enfasi*). Vediamo un esempio di codice:

```
 <book>
  <title>DocBook</title>
  <part>

    <chapter>
      <title>Capitolo 1</title>

      <sect1>



        sezione</para>
      </sect1>
    </chapter>
  </part>
<appendix>
  <title>Titolo
  dell’Appendice</title>
  <para>Contenuto
  dell’Appendice</para>
</appendix>
</book>
```

La marcatura è piuttosto leggibile: dopo la specifica del DTD utilizzato (`DocBook`), si apre l'elemento radice `<book>` (libro) con il suo `<title>` (titolo). L'elemento radice racchiude tutta la struttura del libro: parti (`<part>`), capitoli (`<chapter>`), sezioni (`<sect1>`, dove il numero sta per il livello della sezione nella struttura) e paragrafi (`<para>`). In fondo all'esempio, dopo la chiusura della parte (`</part>`) vediamo l'apertura di un nuovo elemento `<appendix>` (appendice), che si trova allo stesso livello strutturale di una parte all'interno del libro. Se dovessimo visualizzare la struttura ad albero del codice d'esempio avremmo una figura come la 2.4.

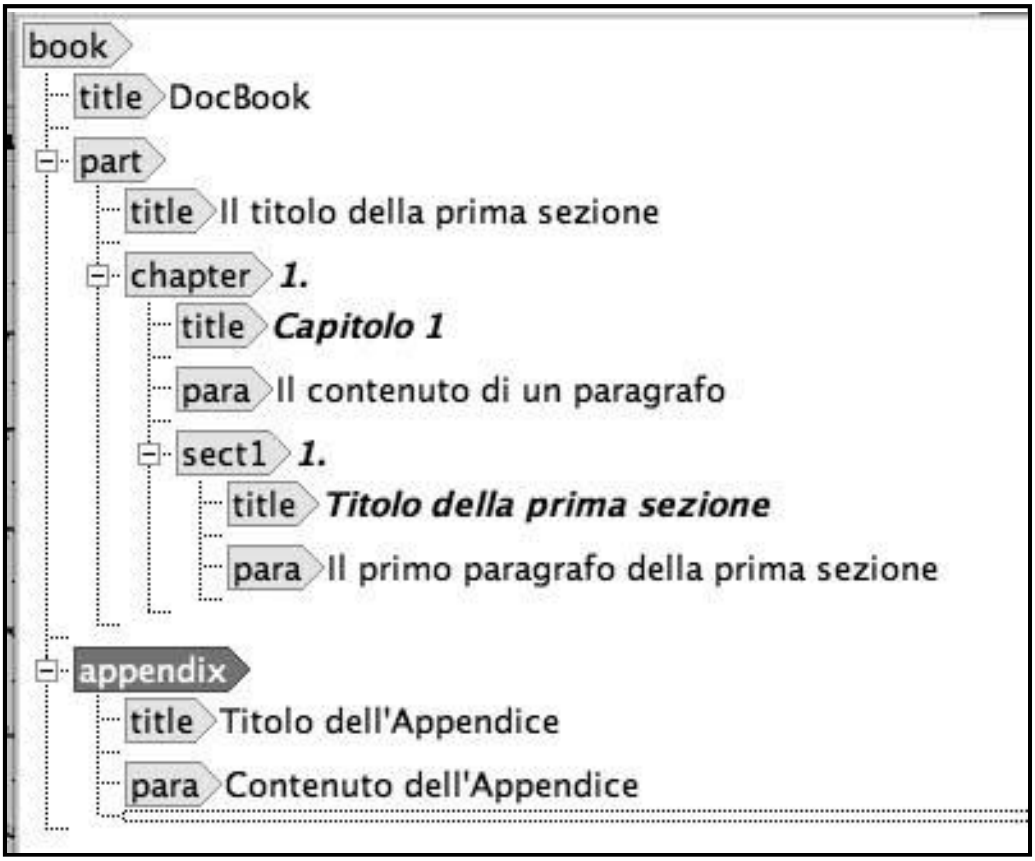

Figura 2.4 La struttura ad albero del listato precedente.

DocBook definisce un numero preciso e limitato – per quanto ampio – di marcatori e stabilisce le regole per usarli correttamente, indicando quali attributi si possono associare a un dato marcatore, quali elementi possono essere contenuti in un determinato elemento e così via. Tali regole devono

essere osservate se si vuole che il documento sia riconosciuto come un documento DocBook: il rischio è che il documento non possa essere interpretato e utilizzato. Infine, esattamente come per XML, i marcatori indicano solamente la struttura logica e semantica, mai quella grafica.

Text Encoding Initiative

La TEI, acronimo di *Text Encoding Initiative*, è uno standard internazionale nato con l'obiettivo di permettere di rappresentare fedelmente ogni genere di testo in formato digitale, con attenzione particolare alle esigenze di biblioteche, musei, case editrici e studiosi in genere. Nello specifico, si tratta di uno standard pensato per i testi letterari del passato, nell'ottica della loro conservazione e preservazione più a lungo termine

possibile. Il vantaggio dell'utilizzo dello standard TEI è quello di produrre documenti portabili, facilmente archiviabili e gestibili attraverso strumenti informatici.

Riferimenti *Il Consorzio TEI è composto da istituzioni internazionali che operano in ambito linguistico e letterario. Emana delle direttive che garantiscono la validità scientifica della codifica dei testi effettuata con lo standard. Il progetto è iniziato nella seconda metà degli anni Ottanta, grazie alla collaborazione tra l'Association for Computers and the Humanities, l'Association for Computational Linguistics e l'Association for Literary and Linguistic Computing, a cui si sono uniti numerosi gruppi di studiosi provenienti da diversi paesi.*

Gli sviluppatori della TEI hanno scelto inizialmente SGML per la codifica dello standard e la normalizzazione dei formati di memorizzazione dei documenti testuali; in seguito sono passati a XML per arrivare poi alla definizione di un DTD studiato appositamente per questo tipo di documenti. Le sue caratteristiche sono state pubblicate nel documento *Guidelines for Electronic Text Encoding and Interchange*, consultabile all'indirizzo http://www.tei-c.org/release/doc/tei-p4-doc/html/. Una lista dei progetti che utilizzano il TEI è disponibile invece all'indirizzo http://www.tei-c.org/Activities/Projects/.

Encoded Archival Description

L'*Encoded Archival Description* (EAD) è uno standard XML creato per semplificare il

reperimento delle informazioni negli archivi, sostené to dalla Library of Congress e dalla Society of American Archivists. L'obiettivo del progetto, nato nel 1993 presso l'Università di Berkeley, era quello di creare uno standard che descrivesse le collezioni degli archivi, dei musei e delle biblioteche, con un occhio di riguardo alle raccolte di manoscritti: gli elenchi generati con lo standard dovevano essere interpretabili dalle macchine e semplici da ricercare, conservare e scambiare. Le linee guida per l'utilizzo dello standard sono state scritte dal Research Libraries Group e stabiliscono elementi e attributi utilizzabili per codificare i documenti in maniera conforme.

Riferimenti *Il Research Libraries Group ha anche ideato un sistema per indicizzare tutte le raccolte che*

> *utilizzano lo standard EAD in maniera da poter offrire un'interfaccia di ricerca da cui poter accedere ad alcune centinaia di collezioni di diverse istituzioni in una sola volta. Il gruppo è anche in grado di individuare le divergenze dallo standard dei documenti, attraverso il sistema automatico per l'analisi qualitativa chiamato* RLG Report Card.

Il DTD dell'Encoded Archival Description specifica quali elementi devono essere utilizzati per descrivere una collezione di manoscritti, oltre al modo in cui tali elementi devono essere utilizzati. Come nel caso della TEI, anche l'EAD è partita da SGML per passare poi a XML, ed è in grado di descrivere una collezione nella sua interezza così come di sgranare il suo

contenuto nel dettaglio, creando vari livelli di indici per gli inventari. Inoltre, l'EAD può contenere altri standard, come il Dublin Core, aumentando notevolmente la flessibilità e l'interoperabilità dei dati.

Terminologia *Il* Dublin Core *è un sistema di metadati ideato per catalogare, descrivere e identificare qualsiasi risorsa accessibile tramite una rete informatica. Il nucleo principale di metadati comprende informazioni come titolo, autore, data di pubblicazione, editore, soggetto, tipo di documento, formato e così via: proprio questa semplicità ha fatto sì che il Dublin Core si affermasse come standard trasversale per la catalogazione di risorse molto diverse fra loro, grazie*

anche all'assenza di eccessiva specificità.

La codifica EAD inserisce all'inizio di ogni documento una sezione detta eadheader, in cui sono contenuti elementi come il titolo del documento e informazioni dettagliate necessarie per il reperimento e la catalogazione del documento: chi l'ha creato e quando, la storia delle sue revisioni, la lingua in cui è scritto e così via. Gli elementi in esso contenuti possono essere messi in relazione con quelli di altri standard, come il già citato Dublin Core; per esempio, l'elemento "autore" dell'EAD può essere messo in relazione con l'"autore" del Dublin Core, favorendo così l'interoperabilità del documento. Lo stesso principio regola le altre sezioni del documento, che si occupano di descriverne il contenuto o la struttura

dell’intera collezione: i loro elementi possono essere messi in relazione con altri standard.

L’Encoded Archival Description è uno standard molto utilizzato dalle istituzioni e collezioni statunitensi. Un elenco può essere consultato all’indirizzo http://www.archivists.org/saagroups/ead/implementors.html.

FORMATI PER I LIBRI DIGITALI

A seconda di come si è deciso di codificare il documento possiamo consultarne e manipolarne il contenuto in maniere diverse. Ognuno dei linguaggi di codifica analizzato nel capitolo precedente determina un particolare formato del file a cui viene applicato: tuttavia questo non è ancora sufficiente per poter definire un documento come un libro digitale.

In questo capitolo ci occuperemo proprio dei formati progettati specificamente per gli

ebook e dei software dedicati alla loro lettura e fruizione.

Testo o immagini?

Una prima divisione tra i formati per libri digitali è quella tra formati testuali e formati immagine. Tra i primi, possiamo distinguere ulteriormente tra quelli ideati appositamente per gli ebook e quelli che invece non lo sono. Di questi ultimi, in qualche caso si è già parlato (RTF, LaTeX, HTML); per tutti gli altri casi non si tratta di formati abbastanza specifici da rientrare nel nostro discorso.

Riferimenti *Oltre a quelli di cui si è già detto, tra i formati non dedicati troviamo il DOC, generato*

dall'elaboratore di testo Microsoft Word e il PostScript, ideato da Adobe come un vero e proprio linguaggio di programmazione adatto a comunicare con le stampanti.

Nelle prossime pagine ci soffermeremo brevemente sui formati immagine, per poi entrare nel vivo dei formati testuali per ebook.

Formati immagine

I libri digitali in formato immagine sono meno frequenti da trovare rispetto ai loro corrispettivi testuali. Questo è dovuto a una serie di svantaggi, come le limitazioni di intervento sul testo (a volte anche nelle comuni operazioni di copia/incolla e ricerca) e le dimensioni eccessive che tendono ad assumere.

Riferimenti *In realtà esistono metodi che permettono di convertire un'immagine contenente del testo in un documento manipolabile con un qualunque elaboratore di testo: si tratta dei metodi di riconoscimento ottico dei caratteri (*Optical Character Recognition*). Con i caratteri stampati l'attendibilità e l'accuratezza del riconoscimento sono ormai pressoché totali, mentre con la scrittura a mano si hanno ancora numerose difficoltà.*

I libri digitali in formato immagine, in genere, prevedono la creazione di un file per ogni pagina di carta, e si prestano alla riproduzione di libri d'arte e fumetti: di quei documenti, insomma, in cui la resa dell'immagine è predominante rispetto al testo.

DjVu

Il formato immagine più interessante per i libri digitali è DjVu (pronunciato come il francese *déjà vu*). Si tratta di un formato sviluppato sul finire degli anni Novanta con l'obiettivo di realizzare raccolte digitali di documenti di alta qualità e a colori. L'alta qualità è la caratteristica più complessa da ottenere, perché per garantire la massima leggibilità delle immagini bisogna creare documenti con dimensioni tali da comportare tempi di accesso non accettabili, specialmente sul Web.

DjVu indica allo stesso tempo una tecnologia di compressione per le immagini, un formato file e una piattaforma per la distribuzione dei contenuti tramite la Rete. Grazie alla separazione in livelli delle parti

che compongono l'immagine, compressi ciascuno diversamente a seconda delle esigenze richieste dal contenuto del livello stesso, DjVu può garantire un'elevata risoluzione del livello che contiene il testo e allo stesso tempo ottenere un documento dalle dimensioni complessive molto basse e adatte a una pagina web o alle prestazioni di un computer di vecchia generazione.

Per visualizzare i file DjVu occorre scaricare e installare un plug-in per il proprio browser web, disponibile per la maggior parte dei software e per tutti i sistemi operativi.

Riferimenti *Un plug-in è un'estensione in grado di aggiungere alcune funzionalità a un programma. Il plug-in che consente di interpretare il formato DjVu può essere scaricato*

gratuitamente all'indirizzo http://www.opendoc.it/plugin/.

Questa estensione permette di utilizzare una serie di funzioni evolute per la consultazione e la gestione del documento. Zoom, rotazione, salvataggio, esportazione, stampa, separazione del testo dallo sfondo e anche la ricerca su tutto il testo, se l'immagine è stata trattata con un sistema di riconoscimento ottico dei caratteri.

Riferimenti *Per dare una misura della velocità di caricamento di un file DjVu e della sua perfetta fruibilità anche con macchine datate, basti pensare che il formato è in grado di visualizzare con solo 2 MB di RAM immagini che in condizioni normali ne richiederebbero 25. La visualizzazione progressiva del documento, inoltre, fa*

sì che l'utente possa accedere subito a una versione di qualità più bassa, che migliora progressivamente con il caricamento successivo di blocchi di dati, in tempi accettabili anche con connessioni lente.

DjVu è un formato aperto e dal 2006 rientra ufficialmente tra gli standard per i progetti di digitalizzazione dei documenti. È molto utilizzato da biblioteche ed emeroteche: in Italia è possibile consultare, per esempio, i documenti dell'emeroteca Braidense, all'indirizzo http://emeroteca.braidense.it/.

Formati per ebook

In questo paragrafo analizzeremo tutti quei formati dedicati esplicitamente alla gestione e presentazione dei libri digitali. La loro particolarità è quella di comprendere, oltre alla codifica del singolo documento, tutta una serie di strumenti per la gestione organica e strutturale di più documenti in relazione tra loro, come in effetti accade nel caso di un libro, considerando l'insieme di capitoli e altri apparati. Questi formati, inoltre, mettono a disposizione dell'utente alcuni strumenti di supporto alla lettura, come i segnalibri, la possibilità di inserire commenti e annotazioni, di navigare nel documento attraverso collegamenti ipertestuali e così via.

Una distinzione interessante da farsi a questo proposito è quella tra formati *aperti* e formati *proprietari*. Questi ultimi, infatti, comportano una serie di limitazioni nell'utilizzo e nello scambio dei file, come anche nella scelta del software capace di interpretarli.

Terminologia *Per proprietario si intende generalmente un software le cui specifiche tecniche non siano di pubblico dominio o non siano accessibili gratuitamente. La definizione comprende, oltre al software, anche il formato che esso produce. Esempi di formati proprietari sono DOC (testo), MP3 (musica) e AVI (video).*

La capacità di supportare sistemi di gestione dei diritti digitali, o DRM, verrà qui

analizzata in maniera sintetica: per riferimenti più approfonditi alle questioni legate al copyright rimandiamo al Capitolo 7.

Formato OeBPS

Open eBook Publication Structure (OeBPS) è un formato per i libri digitali basato su XML. Si tratta di un vero e proprio standard rilasciato nel 1999 dall'Open eBook Forum (adesso *International Digital Publishing Forum*, IDPF), ente al quale partecipano diverse aziende, centri di ricerca e università.

Attenzione *OeBPS è un formato in disuso: al suo posto si è affermato ePub, di cui parleremo tra poco. Considerato che molti formati attualmente in uso derivano da questo alcuni elementi della loro struttura, è*

interessante farvi comunque un accenno.

Riferimenti *L'IDPF è un'organizzazione che si occupa dello sviluppo, della produzione e del mantenimento di formati standard per l'editoria elettronica. Il suo sito è http://www.id-pf.org/.*

L'obiettivo era quello di avere delle regole comuni per la realizzazione dei libri digitali, in grado di fornire funzioni avanzate di impaginazione e catalogazione, disponibili per il maggior numero di piattaforme hardware e software. A questo scopo, OeBPS è un formato aperto e non proprietario. Derivando da XML, può supportare i fogli stile e i metadati dello standard Dublin Core.

In una pubblicazione in OeBPS il contenuto del libro è codificato in file XHTML,

legati tra loro per mezzo di un file XML, detto *package file*, e salvato con l'estensione .opf. Le specifiche stabilite dall'Open eBook Forum definiscono il vocabolario e la sintassi con cui realizzare il package file, il documento che – in altre parole – costituisce la radice dell'albero dei documenti che insieme compongono e descrivono il libro digitale.

Il package file è composto da sei elementi fondamentali.

- *Package identity*: l'identificatore univoco della pubblicazione.
- *Metadata*: contiene le metainformazioni sulla pubblicazione (autore, titolo dell'opera, curatori e così via). L'insieme di metadati a cui ricorre OeBPS è il Dublin Core.

- *Manifest*: contiene la lista dei file che costituiscono la pubblicazione, quindi tutti i file con il contenuto, eventuali immagini, fogli stile e così via.
- *Spine*: è l'elemento che stabilisce l'ordine di lettura predefinito dei file che compongono la pubblicazione.
- *Tours*: elemento facoltativo con cui è possibile stabilire percorsi di lettura alternativi.
- : elemento facoltativo che permette di raccogliere alcuni apparati di supporto al contenuto vero e proprio del libro (indice, bibliografia, copertina e così via).

Attenzione *L'elemento Tours è deprecato nell'attuale formato ePub: ciò significa che l'uso è fortemente sconsigliato poiché l'elemento è considerato*

obsoleto e con il tempo non verrà più supportato.

OeBPS era stato pensato per essere utilizzato come una sorta di formato “sorgente” del libro digitale, più che per la diretta fruizione da parte dell’utente finale. L’idea era infatti quella di poterlo impiegare come base per la generazione di file in altri formati.

Lo standard non prevedeva alcuna indicazione riguardo all’utilizzo di sistemi per la gestione dei contenuti digitali (DRM). Ci fu una collaborazione con l’*Electronic Book Exchange* (EBX, un’organizzazione nata con l’obiettivo di elaborare standard per la protezione dei diritti d’autore nell’ambito dell’editoria elettronica) nel tentativo di contribuire alla formazione di una tecnologia di DRM standard e aperta.

Formato LIT

LIT è il formato sviluppato da Microsoft per i libri elettronici, e prende il nome dall'estensione del file (`.lit`). Il software in grado di leggere i file `.lit` è Microsoft Reader, un'applicazione gratuita disponibile solo per Windows e sviluppata a partire dal 2000. Il formato deriva dall'OeBPS, a cui aggiunge sistemi di protezione del contenuto di vario livello.

In particolare, il formato LIT impedisce la modifica non autorizzata del contenuto del libro digitale. Il sistema di gestione dei diritti digitali applicato è il DAS (*Digital Asset Server*), basato – analogamente a Content Server di Adobe – su tecniche crittografiche

sofisticate. Il DAS opera principalmente su tre livelli di protezione.

- *Sealed*, “sigillato”: è il livello di protezione di base attraverso il quale il contenuto viene criptato in modo da assicurarne l’integrità. Non impedisce la copia del file.
- *Inscribed*, “iscritto”: a questo livello di protezione si aggiungono alle caratteristiche del *Sealed* i dati sensibili dell’acquirente sulla copertina del file. I file così protetti possono essere letti su ogni reader, ma permettono di risalire all’acquirente del file in ogni momento.
- *Owner exclusive*, “proprietà esclusiva”: il terzo livello di protezione prevede delle chiavi di protezione complesse che consentono di associare univocamente il documento a un determinato lettore, a sua volta

legato tramite una procedura di attivazione a un dato computer. Il file protetto in questo modo non può essere copiato. Si tratta di un tipo di protezione fortemente limitante della libertà dell'acquirente, che corre il rischio di non poter più accedere ai propri documenti digitali in seguito a modifiche elementari sul suo computer.

Una delle carenze maggiori di questo formato consiste nell'impossibilità di essere utilizzato al di fuori dei sistemi operativi Windows, il che ne limita la diffusione.

Microsoft Reader

La particolarità di questo software (arrivato alla versione 2.4.1) consiste nell'utilizzo della tecnologia ClearType, in grado di migliorare in maniera notevole la risoluzione sugli schermi LCD e di offrire

una visualizzazione simile alla stampa. Il ClearType è una tecnica di anti-aliasing piuttosto sofisticata che migliora la leggibilità del testo sullo schermo. In pratica, questa tecnologia influisce sulla griglia di pixel che compongono l'immagine sullo schermo del computer, intervenendo per smorzare l'effetto "a gradini" che assumono alcuni elementi grafici come le linee curve o diagonali, interessando di conseguenza in maniera forte la visualizzazione dei caratteri. L'anti-aliasing è un algoritmo che può contrastare questo effetto, sfumando la calettatura di testo e immagini. Questo è possibile grazie all'aggiunta ai bordi dell'immagine di pixel di minore intensità cromatica rispetto all'immagine stessa, dando così l'impressione di un arrotondamento.

Terminologia *Microsoft ClearType è una tecnologia brevettata, e rientra tra quelle dette di sub-pixel font rendering, finalizzate al miglioramento della leggibilità dei testi. Il nome di queste tecnologie deriva dal fatto che gli interventi non avvengono più direttamente sui pixel ma a un livello ancora più profondo, agendo sulla scomposizione del pixel stesso in elementi più piccoli (sub-pixel). Adobe ha brevettato CoolType con lo stesso scopo.*

Essendo basato su OeBPS, anche il LIT ha origine da XML. Per sfruttare al meglio questa caratteristica, Microsoft Reader è dotato di un motore di lettura dinamico, capace di visualizzare la pagina dell'ebook a seconda delle caratteristiche del dispositivo di visualizzazione, senza dover ricorrere allo

scorrimento orizzontale o verticale della pagina. Le dimensioni del testo, inoltre, possono essere aumentate o diminuite a seconda delle necessità dell'utente. Il reader permette la ricerca nel testo, l'evidenziazione di parole e frasi e l'aggiunta delle note; non è possibile invece, per questioni di copyright, la stampa del documento.

Formato Mobipocket

Anche il formato Mobipocket si basa sullo standard OeBPS, utilizzando documenti in XHTML e permettendo anche l'inclusione di JavaScript e frame. Supporta nativamente anche SQL, un linguaggio di interrogazione per basi di dati, consentendo così l'utilizzo del formato in sinergia con i database.

Mobipocket Reader Desktop

Il software in grado di visualizzare, gestire e modificare i file `.mobi` e `.prc` è l'omonimo Mobipocket Reader. Il software (gratuito) è disponibile solo per Windows e un gran numero di cellulari e smartphone (Windows, Symbian, Blackberry e Palm). Per Linux e Mac si possono utilizzare software alternativi come Okular e FBReader: di entrambi i software parleremo nel paragrafo dedicato ai PDF.

Terminologia *L'estensione .prc è tipica dei Palm, i dispositivi per cui Mobipocket era nato. Tutti i file per Palm hanno estensione* `.prc`*, e per questo è stata creata l'estensione* `.mobi`*, in modo da evitare ambiguità.*

Il Mobipocket Reader è stato sviluppato nel 2000 da una compagnia francese, inizialmente come lettore universale per PDA. La compagnia (Mobipocket.com) è stata acquisita da Amazon nel 2005.

Riferimenti *Come vedremo nel paragrafo successivo, il formato supportato dall'ereader di Amazon, l'Amazon Kindle (AZW), si basa appunto su una versione leggermente modificata di Mobipocket.*

A seconda del dispositivo sul quale viene utilizzato, il software mette a disposizione dell'utente diversi strumenti che vanno dalla gestione della biblioteca digitale, la rotazione del documento, i segnalibri, fino alla possibilità di creare collegamenti ipertestuali tra documenti diversi,

sottolineatura, commenti e disegni. Nel trasferimento di un file verso un dispositivo con meno funzionalità quelle non supportate verranno ignorate senza comunque tralasciare o alterare l'informazione che contengono.

Nella versione Desktop del software (Figura 3.1) i formati che si possono importare sono diversi, tra cui HTML, PDF e TXT. Attraverso questa applicazione l'utente può creare documenti in formato Mobipocket e aggiungere commenti, segnalibri e altre funzionalità che verranno visualizzate a seconda di quelle supportate dal dispositivo portatile in cui il documento verrà letto.

Figura 3.1 La schermata iniziale di Mobipocket Reader Desktop su Windows. In primo piano la copertina del libro in lettura.

Tra le operazioni che il lettore può effettuare sul testo c'è la possibilità di aggiungere

pagine bianche in qualunque parte del libro e inserire annotazioni e disegni a mano libera. Evidenziazioni, segnalibri, correzioni, note e disegni possono essere applicati, organizzati e richiamati con semplicità. Il programma comprende un dizionario integrato ed è in grado di adattare il testo al dispositivo di lettura che lo sta visualizzando.

Dalla versione Desktop per Windows ci si può anche collegare a diversi cataloghi di ebook online, dai quali scaricare nuovi libri digitali gratuitamente o a pagamento.

Formato Amazon Kindle

Amazon Kindle indica in realtà sia un formato sia un ereader. Il software, di cui ci occupiamo in questo capitolo, è stato sviluppato da Amazon stessa, è supportato

dall’ereader omonimo e da un’applicazione gratuita per iPhone (Kindle for iPhone, che può sincronizzarsi con la libreria del dispositivo).

Il formato AZW di Kindle è sostanzialmente il formato Mobipocket con uno schema differente per la generazione del numero seriale, in cui utilizza il simbolo del dollaro ($) al posto dell’asterisco. Inoltre i file sono protetti da uno speciale sistema DRM.

Il discorso sul formato Amazon Kindle non è tanto rilevante dal punto di vista tecnico, quanto come spunto per introdurre vari aspetti riguardo l’editoria digitale in genere (copyright, modelli di business e così via). Avremo modo di parlarne più diffusamente nei capitoli successivi.

Formato eReader

eReader è un termine destinato a creare confusione. In questo libro, infatti, lo usiamo per riferirci ai dispositivi di lettura, nel Capitolo 6. Allo stesso tempo, però, indica sia un formato di codifica per i libri digitali sia un software di lettura e modifica degli stessi.

Il formato eReader si basa sul *Palm Markup Language* (PML): il linguaggio non si occupa della struttura semantica del testo ma della sua formattazione, e per questo può essere considerato simile a un linguaggio di marcatura procedurale (anche se molto semplificato rispetto a RTF o LaTeX). eReader prevede l'integrazione di un dizionario, in modo che una parola evidenziata dal lettore

possa essere istantaneamente ricercata nel dizionario stesso.

Merita attenzione anche perché nel luglio 2009, Barnes & Noble – la più grande catena di librerie degli Stati Uniti – ha annunciato che il formato eReader sarà quello che utilizzerà per distribuire i suoi libri digitali.

Riferimenti *Si parlerà delle strategie di Barners & Noble riguardo ai libri digitali nel Capitolo 5.*

Il software eReader

eReader è un programma gratuito per la visualizzazione di libri digitali in formato *Palm Digital Media* (.pdb). Ne esistono versioni per diversi dispositivi: iPhone, PalmOS, Symbian, BlackBerry, Windows

Mobile, Pocket PC/Smartphone, Windows e Mac OS (Figura 3.2). Il programma mostra il testo una pagina alla volta e supporta collegamenti ipertestuali e immagini.

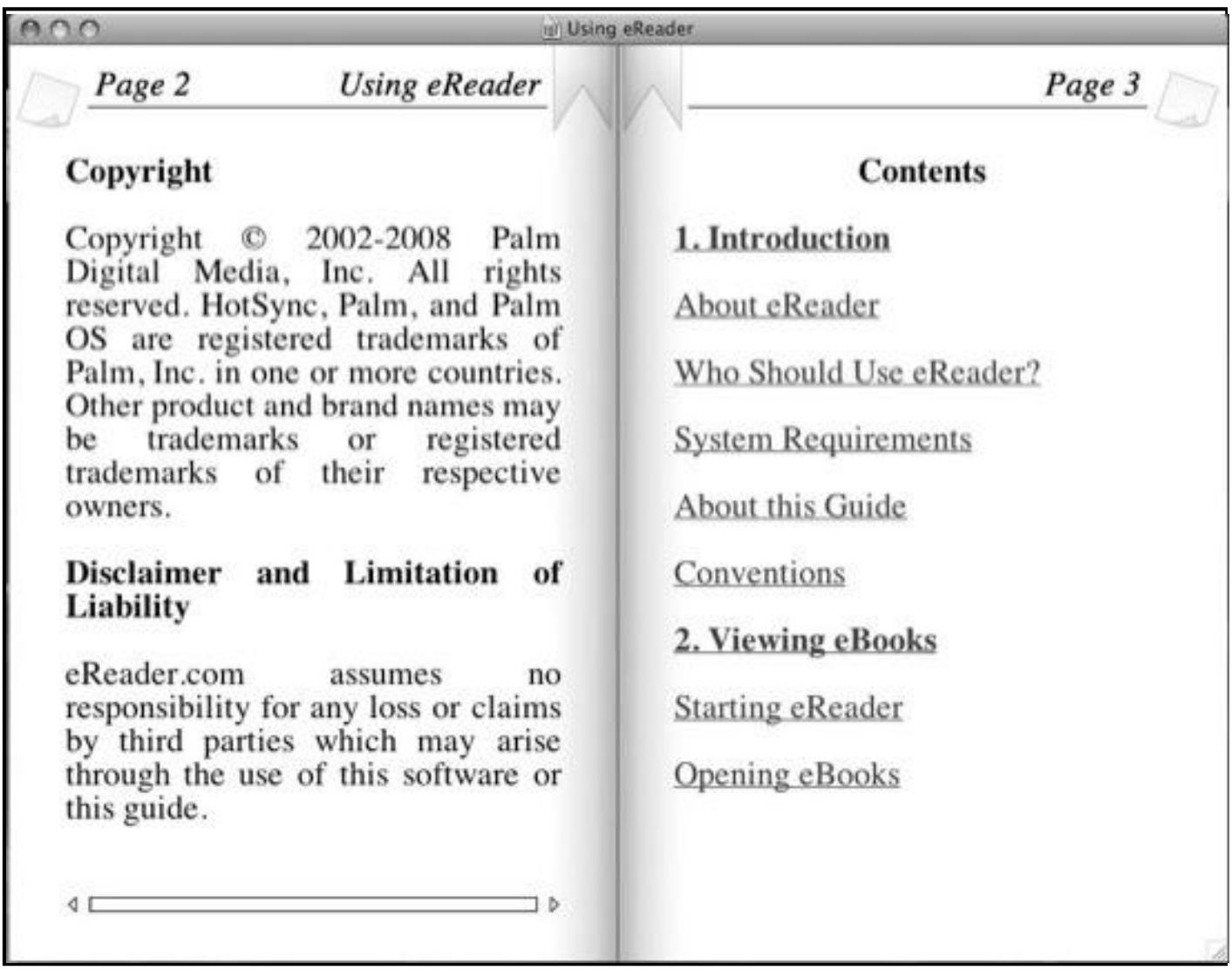

Figura 3.2 Ecco come appare il contenuto di un ebook visualizzato con eReader su un sistema Mac OS X.

Su siti come http://www.ereader.com e https://www.fictionwise.com/home.html è presente un'ampia selezione di ebook in formato eReader, disponibili per l'acquisto e il download. Questi titoli sono criptati: la chiave è composta dal nome completo dell'acquirente e dal suo numero di carta di credito.

Il software supporta funzionalità come segnalibri e note a piè di pagina, che possono essere sfruttate anche per inserire commenti e successivamente esportate.

All'indirizzo http://www.ereader.com/ereader/software/browse.htm sono disponibili anche altri due software: Drop-Book, gratuito, permette di convertire file di testo nel formato .pdb; eBook Studio, a pagamento, permette di formattare i propri

documenti per eReader e di gestirne la struttura di navigazione per creare ebook professionali e adatti al numero più ampio possibile di dispositivi.

Formato PDF

Acronimo di *Portable Document Format*, PDF è uno dei primi formati multipiattaforma per la distribuzione dei documenti elettronici, in seguito adottato anche per la diffusione di ebook. È un prodotto di Adobe, nato da una trasformazione del PostScript: il PDF, infatti, è un linguaggio di descrizione della pagina ma anche un linguaggio di programmazione procedurale, attraverso il quale si possono dare indicazioni a un dispositivo di output su come rappresentare accuratamente la pagina descritta. Il PDF, soprattutto, eredita dal PostScript la

fondamentale caratteristica di essere indipendente dai vari dispositivi di uscita, che hanno l'unico obbligo di supportare il linguaggio.

Il PDF è riuscito nell'obiettivo di rendere possibile la condivisione e distribuzione dei documenti con testi e grafica formattati su piattaforme differenti senza alterazioni nel layout originale. Installando su un computer Acrobat Reader, il software gratuito e multipiattaforma per la lettura dei PDF, si riescono a visualizzare i file codificati in tale formato a prescindere dalla piattaforma hardware e software con cui il documento originale è stato creato. Nel 2008 il formato è diventato uno standard internazionale per la distribuzione dei documenti elettronici.

Riferimenti *Per la creazione dei PDF si utilizza il software Acrobat, prodotto*

dalla stessa Adobe. Esistono sul mercato anche soluzioni alternative, in certi casi anche gratuite. Ne vedremo alcune nel capitolo seguente.

Poiché il PDF si occupa della descrizione della pagina dal punto di vista grafico e non da quello testuale, il suo utilizzo come formato per gli ebook presenta alcuni limiti. Innanzitutto occorre creare il file PDF tenendo in considerazione il dispositivo che lo visualizzerà: infatti le pagine del documento potranno solamente essere ingrandite o rimpicciolite per adeguarsi allo schermo sul quale verranno visualizzate. Questo inconveniente non si verifica con i documenti digitali codificati utilizzando i linguaggi di marcatura, che sono svincolati dal concetto di pagine e in grado di adattare il contenuto

del documento allo schermo del dispositivo, in maniera “fluida”.

Proprio per risolvere questo problema, a partire dalla versione 5 di Acrobat i PDF possono contenere dei tag che consentono di riformattare il flusso del testo per permettergli di adattarsi ai dispositivi più diversi. Tuttavia, considerata la natura prevalentemente grafica del formato, esso non possiede la stessa ricchezza di metainformazioni tipica dei documenti prodotti con codifiche mirate sul contenuto, e di conseguenza non ha la stessa versatilità e adattabilità alle esigenze più diverse.

Riguardo ai sistemi di gestione dei diritti digitali, Adobe propone una gamma piuttosto variegata di strumenti professionali per proteggere i contenuti in varie forme.

I software per visualizzare, modificare e creare i PDF sono piuttosto numerosi e in grado di accontentare le esigenze più diverse. Considerato che il formato costituisce uno standard di fatto per la diffusione dei documenti digitali da molto più tempo rispetto alla sua approvazione ufficiale, la quantità di programmi capaci di interpretarlo e manipolarlo è conseguentemente ampia. Di seguito ne prenderemo in esame alcuni.

Acrobat Reader

Come accennato, il software Adobe in grado di leggere i PDF è Acrobat Reader, attualmente arrivato alla versione 9. Il software permette di visualizzare, stampare ed effettuare ricerche all’interno dei file PDF. Se usato in sinergia con Acrobat (che a

differenza del Reader è a pagamento), le sue funzionalità aumentano, rendendo possibile aggiungere commenti ai documenti, compilare e salvare le modifiche ai moduli che possono essere contenuti al loro interno e aggiungere firme digitali.

Digital Editions

Digital Editions è un software gratuito e multipiattaforma sviluppato da Adobe appositamente per il download, la gestione e la lettura dei libri digitali. In particolare, l'applicazione può compiere tutte queste operazioni anche su ebook coperti da DRM, previa l'attivazione di un account Adobe, che permette di autorizzare il computer alla gestione e visualizzazione di file protetti dal diritto d'autore.

Riferimenti *Digital Editions è in grado non solo di gestire libri digitali, ma ogni tipo di documento elettronico, come quotidiani, riviste o altre pubblicazioni. Può essere scaricato per Windows e Mac all'indirizzo http://www.adobe.com/it/products/digitaleditions/#fp.*

I formati supportati sono PDF, ePub e i contenuti in Adobe Flash.

Digital Editions consente lo scambio di documenti coperti da DRM tra computer autenticati attraverso l'account Adobe.

Per gestire i file protetti con tecnologie DRM, Digital Editions utilizza lo schema proprietario ADEPT (*Adobe Digital Editions Protection Technology*), che viene supportato anche da alcuni ereader. ADEPT

permette all’utente di vedere il contenuto protetto su un massimo di sei computer autorizzati.

L’interfaccia di lettura rende possibile l’utilizzo di segnalibri, annotazioni e la navigazione del documento tramite il sommario, quando presente. I PDF possono essere visualizzati a pagina singola, su due pagine o attraverso modalità di visualizzazione personalizzate; gli ePub invece, essendo in un formato “fluido” non hanno vincoli di pagina e il testo può essere ridimensionato con maggiore facilità.

La biblioteca digitale può essere gestita ordinando i documenti per titolo, autore o editore, e visualizzando le copertine di libri. Si possono creare veri e propri scaffali

personalizzati per organizzare i contenuti in modo ancora più specifico (Figura 3.3).

Figura 3.3 Una collezione di libri digitali visualizzata su Adobe Digital Editions.

L'attivazione del software tramite un account Adobe consente, oltre che di visualizzare ebook coperti da DRM su più computer,

anche di sincronizzare la biblioteca digitale con diversi tipi di dispositivi, tra cui i Sony Reader, di cui parleremo nel Capitolo 6. È inoltre possibile “prendere in prestito” ebook da biblioteche pubbliche che dispongono di pubblicazioni digitali con una procedura del tutto simile a quella di acquisto e download, senza naturalmente effettuare la procedura di pagamento.

Riferimenti *La New York Public Library, per esempio, offre una scelta vastissima di ebook per il prestito, previa autenticazione con il numero della propria tessera della biblioteca. Le collezioni possono essere consultate all’indirizzo http://ebooks.nypl.org: non si tratta solamente di libri digitali ma anche di collezioni di musica, video e audiolibri.*

Il programma è stato sviluppato con la tecnologia Adobe Flash. Se da un lato questo costituisce un punto di forza (è multipiattaforma, si aggiorna automaticamente, è semplice da installare), dall'altro ne costituisce anche un punto critico. A causa della sua interfaccia Flash, infatti, l'applicazione non ha barre dei menu, non permette l'accesso ai menu di scelta rapida poiché inibisce il clic destro del mouse e richiede la versione più aggiornata di Flash per funzionare.

Foxit

Foxit Reader è un visualizzatore di PDF gratuito disponibile per Windows e Linux. A differenza di altri lettori consente di disegnare, evidenziare e commentare i file, oltre che di stampare e salvare il documento così annotato. Inoltre è possibile convertire il

PDF in un semplice file di testo. Foxit, la casa produttrice del software, dispone di un completo insieme di strumenti (a pagamento) per la creazione e manipolazione dei PDF. Inoltre produce un ereader, eSlick (http://www.foxitsoftware.com/ebook/).

Okular

Okular è un visualizzatore di documenti gratuito per Linux. Oltre ai PDF, è in grado di interpretare molti altri formati, come per esempio DjVu ed ePub. A seconda dei formati supporta diverse funzionalità, come l'esportazione del testo, collegamenti ipertestuali, annotazioni, ricerche nel testo, DRM e così via.

Anteprima

Anteprima è un'applicazione preinstallata su Mac OS capace di visualizzare immagini e

PDF (Figura 3.4). Per questi ultimi dispone di alcuni strumenti di modifica in genere disponibili solo in software professionali come per esempio l'estrazione di una singola pagina da un documento e strumenti avanzati di commento e annotazione sul file. Anteprima permette anche di criptare un documento grazie all'applicazione di una password e di convertire il PDF in un formato immagine.

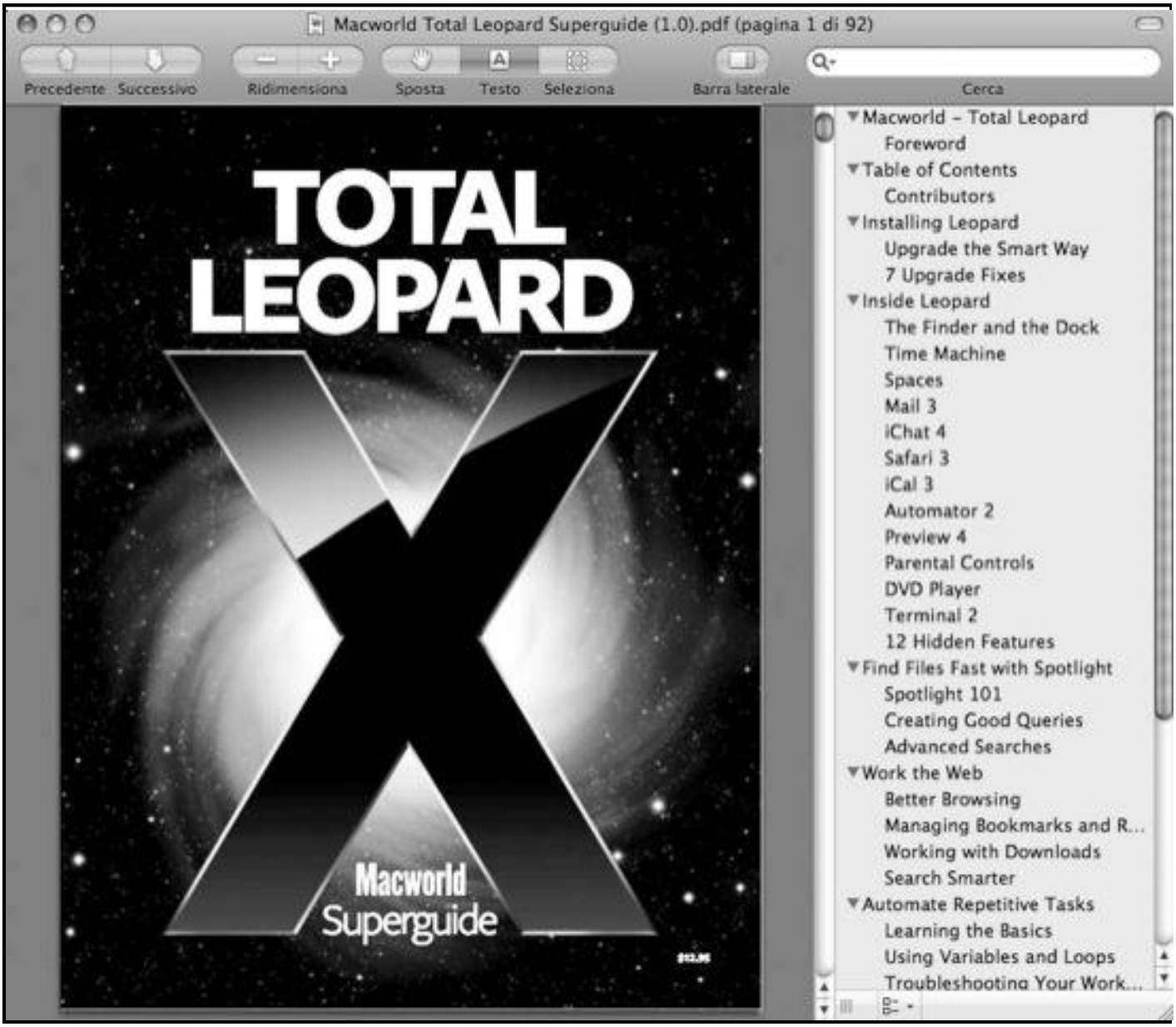

Figura 3.4 La visualizzazione di un PDF su Anteprima.
A destra l’indice dei contenuti.

Formato ePub

Uno dei maggiori problemi del libro digitale, come si può intuire anche solo dai formati presentati finora in questo capitolo (che per altro non esauriscono la lista ma si limitano alla descrizione di quelli principali), è l'assenza di uno standard univoco e riconosciuto da tutti che possa affermarsi uniformemente per software e dispositivi di lettura. La creazione dell'ePub da parte dell'IDPF nel 2007 ha contribuito a dare una svolta a questa situazione: si tratta di un formato basato su XML e composto da tre specifiche aperte ideate per favorire la compatibilità potenziale con diversi dispositivi. Quindi non solo i dispositivi dedicati alla

lettura degli ebook, ma tutti i dispositivi in grado di visualizzarli, anche quelli multifunzione come gli smartphone.

ePub è per certi versi una versione aggiornata e sostitutiva dell'*Open eBook* (OeBPS). L'estensione del file è .epub, e la sua caratteristica principale è quella di permettere al testo di adattarsi in modo fluido al dispositivo di visualizzazione. Come dicevamo, il formato è costituito di tre specifiche:

- l'*Open Publication Structure* (OPS), che descrive come marcare i contenuti del file;
- l'*Open Packaging Format* (OPF), che descrive in XML la struttura del file .ePub;

- l'*OEBPS Container Format* (OCF), che descrive il metodo per costruire l'archivio ZIP che conterrà la pubblicazione.

Figura 3.5 Alcuni logo di pubblico dominio per il formato ePub.

In pratica, ePub rappresenta il testo attraverso file codificati in XHTML e ne gestisce la struttura attraverso XML. L'insieme dei file è archiviato e compresso in un unico file `.zip`, rinominato poi in `.epub`.

Elenchiamo alcune informazioni fondamentali riguardo le caratteristiche supportate e richieste dal formato ePub.

- La versione di XHTML supportata è quella attualmente in uso, la 1.1 (di cui si è parlato

nel Capitolo 2), con alcune limitazioni: non sono supportati i form, le mappe di immagine e gli script.

- Nei fogli stile CSS esistono limitazioni; il posizionamento assoluto è altamente sconsigliato.
- I font utilizzati devono essere OpenType (ovvero in un formato multipiattaforma); devono essere indicati nel foglio stile attraverso il comando `@font-face`.
- I formati immagine supportati sono JPG, PNG, GIF e SVG.

Nei paragrafi seguenti analizzeremo nel dettaglio le tre specifiche alla base di ePub, descrivendo per prima cosa il contenitore (OCF), quindi la struttura del suo contenuto

(OPF) per arrivare infine al contenuto vero e proprio (OPS).

OCF: definire il contenitore

Lo standard *OEBPS Container Format* (OCF) definisce una tecnologia di contenitore per contenuti digitali versatile e non finalizzata unicamente alle pubblicazioni elettroniche. Nel caso degli ePub, è in grado di archiviare e rendere interpretabili correttamente i file che compongono il libro digitale e i documenti che ne descrivono la resa su diversi software e dispositivi.

Nella Figura 3.6 vediamo com'è strutturato il contenuto di un file ePub.

Attenzione *Lo schema proposto è un modello, e in quanto tale descrive la struttura ideale di un ePub: non*

implica l'obbligo di posizionare esattamente in questo modo tutti i documenti dell'archivio. Specificheremo di volta in volta quali elementi del modello sono obbligatori e quali invece sono facoltativi.

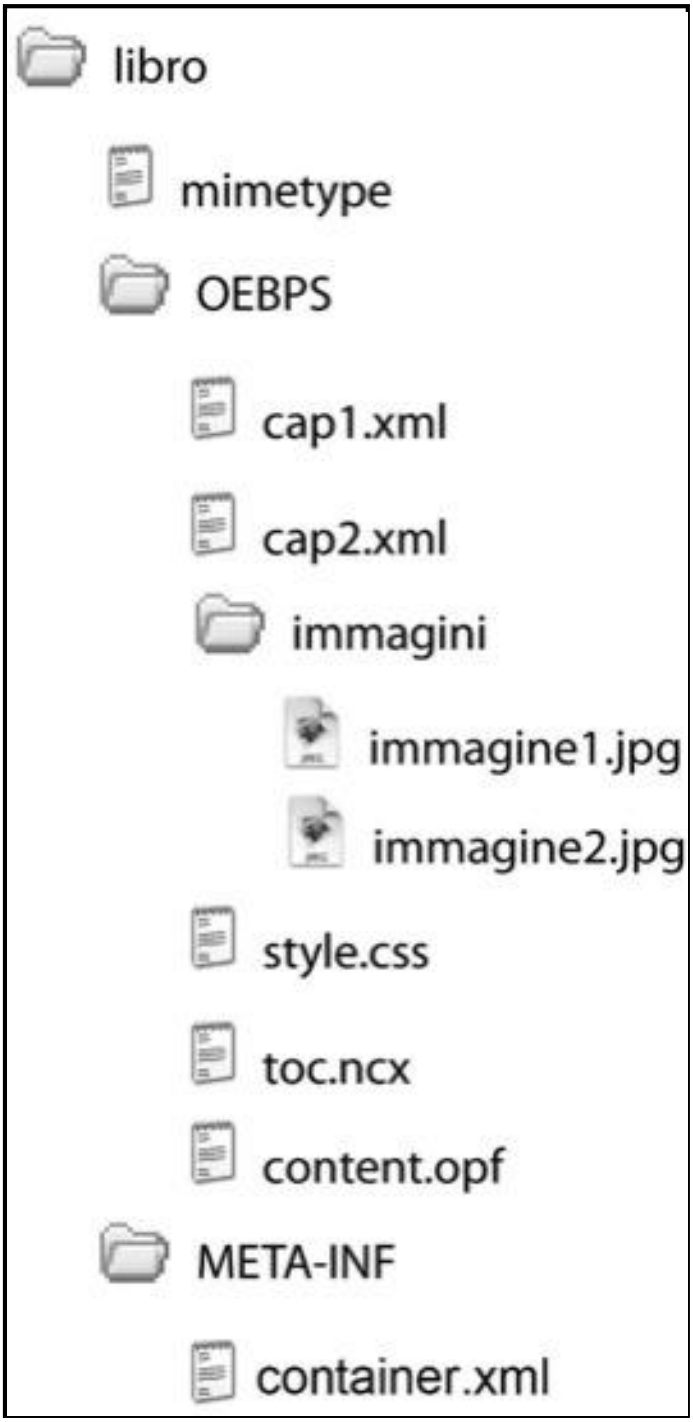

Figura 3.6 La struttura fondamentale di un ePub.

Il modello nello schema considera il file `libro.epub`: modificando l'estensione

`.epub` in `.zip` potremo aprire la cartella proprio come facciamo normalmente con un archivio compresso.

Riferimenti *Lo ZIP è un formato di compressione dati molto diffuso, col vantaggio di non comportare alcuna perdita di informazione per i file compressi. Per scompattarli occorre un software apposito: ne esistono diversi, molti dei quali gratuiti, come per esempio ZipGenius, disponibile all'indirizzo http://www.zipgenius.it/.*

All'interno troviamo due cartelle e alcuni file singoli. Delle due cartelle, quella denominata *META-INF* deve essere inclusa nell'archivio nel primo livello della directory. In questa cartella devono essere presenti i file che descrivono il contenuto, i metadati,

le firme digitali, le informazioni sull'eventuale cifratura del contenuto o sulla presenza di eventuali DRM e così via. Non tutti questi elementi sono obbligatori, come vedremo più avanti. Questo può significare, in alcuni casi, che certi software o dispositivi non supportino in pieno tutti i file opzionali previsti dallo standard, o che non li gestiscano alla stessa maniera: utilizzarli non significa compromettere la corretta interpretazione del documento, ma si potrebbe determinare la perdita di una parte dell'informazione a livello strutturale.

Riferimenti *Gli esempi di codice da questo punto in avanti sono tratti (e dove necessario adattati alle esigenze del discorso) dai documenti IDPF che stabiliscono le specifiche dei tre standard che compongono l'ePub. Tutti i*

*documenti possono essere consultati all'indirizzo http://www.idpf.org/specs.htm. Una guida più generale ma altrettanto completa è invece l'*Epub Format Construction Guide, *consultabile all'indirizzo http://www.hxa.name/articles/content/epub-guide_hxa7241_2007.html. Entrambi i documenti sono in inglese.*

Allo stesso livello della cartella *META-INF* è necessario collocare il file mimetype. Si tratta di un file in ASCII contenente la seguente stringa:

```
application/epub+zip
```

Lo scopo di questo file è quello di permettere alle applicazioni di determinare di che tipo di file si tratta. In genere questo è possibile grazie all'estensione del file stesso,

che consente di individuare rapidamente il tipo di file senza doverne esaminare il contenuto. L'archivio, quindi, dovrebbe utilizzare l'estensione `.epub` per permettere l'identificazione corretta del file da parte delle applicazioni. Tuttavia, non sempre l'identificazione per mezzo dell'estensione è attendibile: per renderla più precisa possibile è necessario inserire il file `mimetype` nella cartella in modo che le applicazioni possano verificarne l'esistenza per determinare se si tratti – nel nostro caso specifico – di un OCF Container.

Riferimenti *Con la prima parte della stringa il mimetype definisce il tipo di media che dovrà interpretare il documento a cui si riferisce (in questo caso si tratta di un'applicazione, ma potrebbe anche essere un file audio,*

un'immagine e così via), mentre con la seconda parte ne specifica il formato.

Esaminiamo adesso il contenuto della cartella *META-INF*: i file elencati di seguito, lo ricordiamo, fanno parte dello standard OCF (*OEBPS Container Format*). Le specifiche IDPF per questo standard sono disponibili all'indirizzo http://www.idpf.org/ocf/ocf1.0/download/ocf10.htm.

Container

Il file `container.xml`, deve contenere la dichiarazione del media-type, il path per individuare il file OPF (di cui ci occuperemo più avanti); la versione OeBPS della pubblicazione elettronica e tutte le altre eventuali possibili codifiche incluse nell'archivio. È l'unico file la cui presenza all'interno della

cartella *META-INF* è obbligatoria. Vediamo il codice di un file `container.xml`:

```
<?xml version="1.0"?>

  <rootfiles>


  </rootfiles>
</container>
```

Dopo la dichiarazione del tipo di documento della prima riga si apre l'elemento `<container>`, che indica attraverso i valori dei suoi attributi la specifica dello standard a cui si fa riferimento.

Riferimenti *L'attributo* `xmlns` *sta per XML Namespace: identifica i nomi dei tag e degli attributi che si possono*

usare all'interno di un documento XML.

`<rootfiles>` contiene al suo interno tanti elementi `<rootfile>` quanti sono i file che descrivono i contenuti all'interno dell'archivio.

Riferimenti *Capiremo meglio questo concetto analizzando il file OPF, tra qualche paragrafo. Per ora è sufficiente sapere che il file* `My Crazy Life.opf` *nell'esempio costituisce una sorta di elenco dettagliato dei contenuti che compongono l'ePub.*

L'elemento `<rootfile>` è specificato dagli attributi `full-path` (*path* in inglese significa "percorso") e `media-type`. Se l'archivio contenesse anche altri documenti in formati differenti, questi dovrebbero essere elencati

ciascuno in un elemento `<rootfile>`. Se per esempio all'interno dell'archivio ci fosse anche un PDF bisognerebbe aggiungere questo elemento, con i valori degli attributi modificati di conseguenza:

```
media-type="application/pdf" />
```

Manifest

La presenza del file `manifest.xml` nella cartella *META-INF* non è obbligatoria. Tuttavia può risultare utile, per esempio, nel momento in cui il contenuto è composto da più file in XHTML messi in relazione da collegamenti ipertestuali. Per far sì che tali file correlati vengano considerati come un unico documento si può utilizzare il file `manifest`, che permette di riunire in un set un insieme

di file e di interpretarlo così come una singola unità di contenuto.

Dopo la dichiarazione del tipo di documento, la struttura del file manifest è la seguente:

```
<package>
  <metadata>
  </metadata>
  <index>
  </index>
</package>
```

L'elemento `<package>` contiene due elementi, `<metadata>` e `<index>`, che descrivono rispettivamente i metadati relativi all'unità di contenuto che si intende isolare e un indice che determini un ordine di navigazione per il software o il dispositivo che visualizzerà l'ePub.

Metadata

Si tratta di un file (`metadata.xml`) destinato a contenere informazioni generali relative alla pubblicazione digitale. L'ordine dei metadati non è obbligato, ma alcuni devono necessariamente essere presenti, come `title`, `language` e `identifier`: quest'ultimo deve essere univoco, e a questo scopo è consigliabile utilizzare il codice ISBN.

Altri metadati opzionali sono: `creator`, `contributor`, `publisher`, `subject`, `description`, `date`, `type`, `format`, `source`, `rights` e altri.

Firme digitali

Il file `signatures.xml` opzionalmente incluso in META-INF contiene eventuali firme digitali sull'archivio e il suo contenuto. Il suo

elemento radice è `<signatures>` che contiene uno o più elementi `<signature>`. Le firme possono essere applicate all'intera pubblicazione digitale o ad alcune sue parti e possono riferirsi a qualunque file dell'archivio eccetto al file `signatures.xml` stesso.

Encryption

Il file opzionale `encryption.xml` contiene le informazioni per decifrate eventuali file cifrati all'interno della pubblicazione. Il suo elemento radice è `<encryption>` e contiene i due elementi `<EncryptedKey>` ed `<EncryptedData>`, per individuare rispettivamente la chiave per decifrare i dati e i dati eventualmente cifrati. L'esistenza di dati cifrati determina l'obbligatorietà della presenza del file `encryption.xml` all'interno della cartella *META-INF*.

Attenzione *Nessuno dei file contenuti nella cartella* META-INF, *né il file mimetype, né l'eventuale cartella* OEBPS *possono essere cifrati.*

Gestione dei diritti

Il file `rights.xml` è opzionale ed è destinato a contenere gli eventuali DRM applicati alla pubblicazione. In presenza di DRM il file è in grado di gestire la comunicazione e gli scambi tra l'utente e il detentore dei diritti applicati o eventuali intermediari.

OPF: nominare ed elencare il contenuto

Lo standard *Open Packaging Format* (OPF) definisce i meccanismi attraverso i quali le varie componenti di una pubblicazione OPS sono connesse tra loro e allo

stesso tempo si occupa di dare una struttura semantica alla pubblicazione digitale.

Riferimenti *Prenderemo in esame lo standard OPS tra poche pagine.*

Le azioni dell'OPF sono le seguenti.

- Descrivere ed elencare tutti i contenuti della pubblicazione elettronica: immagini, file XML e strutture di navigazione.
- Occuparsi della presenza e della gestione dei metadati.
- Specificare l'ordine lineare per la lettura della pubblicazione.

La divisione fra OPF e OPS permette di tenere separate le metodologie di descrizione e strutturazione del "pacchetto" dei contenuti in quanto tale dalla descrizione e strutturazione del contenuto nello specifico:

questo significa fondamentalmente che con lo standard OPF si interviene sulla macrostruttura del libro digitale, mentre OPS si occupa delle sue porzioni di contenuto, come possono essere per esempio i capitoli.

Riferimenti *Come abbiamo già visto, invece, OCF definisce il meccanismo attraverso il quale tutte le componenti della pubblicazione elettronica possono essere racchiuse insieme in modo da creare un singolo file semplice da diffondere e archiviare. Usando una metafora, è come se l'ePub fosse composto da tre scatole cinesi: la più piccola è l'OPS contenuta nell'OPF contenuta a sua volta nell'OCF.*

Per essere conforme allo standard un ePub deve includere un solo file XML OPF, il

quale specifica i documenti OPS (ossia il contenuto vero e proprio del libro digitale), le immagini e gli altri oggetti della pubblicazione e le relazioni che intercorrono tra loro.

> **Attenzione** *Il file OPS deve essere presente nella cartella* OEBPS, *a sua volta collocata allo stesso livello della cartella* META-INF.

Nonostante si tratti di un file XML, l'estensione del file OPF dev'essere `.opf`, così da renderlo immediatamente identificabile nell'insieme di file che compongono l'ePub. Vediamo ora nel dettaglio gli elementi principali di un file OPF.

- *Package*: l'identificatore univoco della pubblicazione.

- *Metadata*: i metadati sul libro (titolo, autore, editore e così via).
- *Manifest*: l'elenco di tutti i file che compongono il libro, comprese le immagini, i fogli stile e così via.
- *Spine*: l'ordine lineare di lettura dei documenti.
- *Guide*: i riferimenti agli apparati del libro, come la prefazione, la bibliografia, la postfazione e così via.
- *Tours*: una serie di sequenze di lettura alternative allo schema lineare dello *Spine*. Questo elemento è deprecato: l'utilizzo è altamente sconsigliato.

Osserviamo ora il codice elementare di un documento OPF, che in quest'esempio chiameremo `content.opf`:

```
<?xml version="1.0"?>

  <metadata>...</metadata>
  <manifest>...</manifest>
  <spine>...</spine>
  <guide>...</guide>
</package>
```

L'elemento radice del file OPF è `<package>`, all'interno del quale si annidano tutti gli altri. Nell'elemento `<package>` viene specificato il valore dell'attributo `unique-identifier` che, come abbiamo già osservato, è obbligatorio. È assolutamente necessario che la chiave sia univoca, perché i software di lettura e i dispositivi non possono trovarsi davanti a due documenti diversi con lo stesso identificatore.

Metadata

L'elemento `<metadata>` fornisce informazioni sulla pubblicazione nel suo insieme, ed è un elemento obbligatorio. È compatibile con lo standard Dublin Core e può contenerne direttamente gli elementi. Ecco un esempio dell'elemento `<metadata>` all'interno di `content.opf`:

```
xmlns:opf="http://www.idpf.org/2007/o
pf">
  <title>Tale of Two Cities</title>

  <language>en</language>

  123456789X
  </identifier>

</metadata>
```

È necessario includere almeno i metadati espressi dagli elementi `<title>`, `<identifier>` e `<language>`.

Riferimenti *Gli elementi* `<title>`, `<identifier>` *e* `<language>` *sono tratti direttamente dallo standard del Dublin Core. Può capitare, però, di aver bisogno di aggiungere ai metadati informazioni non previste da questo standard: nell'esempio è il caso del prezzo. In questo caso si specifica un nome per l'informazione attraverso l'attributo* `name` *e il suo valore attraverso l'attributo content.*

Il Dublin Core non distingue ruoli specifici ma solo le figure di `<creator>` e `<contributor>`. Per questo si ricorre all'attributo

`role`, attraverso il quale si può specificare di volta in volta se l'informazione riguardi l'autore, l'illustratore, il curatore e così via.

Riferimenti *Per un'analisi dettagliata di tutti gli elementi che possono essere contenuti dall'elemento* `<metadata>` *rimandiamo alla relativa sezione del documento delle specifiche dello standard, consultabile all'indirizzo* *http://www.idpf.org/2007/opf/OPF_2.0_final_spec.html*.

Manifest

L'elemento `<manifest>` deve elencare la lista di tutti i file che compongono la pubblicazione. Ogni suo oggetto (`<item>`) descrive un documento, un'immagine, un foglio stile o qualunque altro file sia parte

del libro digitale. Ogni `<item>` deve avere alcuni attributi: un id per l'identificazione univoca della risorsa; un `href` (un collegamento) che punti al file in oggetto e un `media-type` che specifichi il formato del file. Non è rilevante l'ordine in cui gli elementi vengono elencati. Vediamo il codice:

```
<manifest>

    media-type="application/xhtml+xml"
    />
  <item id="c1" href="chapter-1.html"
    media-type="application/xhtml+xml"
    />
  <item id="c2" href="chapter-2.html"
    media-type=application/xhtml+xml"
    />
  <item id="toc" href="contents.xml"
    media-type="application/xhtml+xml"
    />
  <item id="oview" href="arch.png"
```

media-type=”image/png” />
</manifest>

Il primo oggetto è, in questo caso, il documento che contiene l'introduzione della pubblicazione; l'ultimo è un file immagine. Ogni oggetto può essere elencato una sola volta nel `<manifest>`.

Spine

Subito dopo l'elemento `<manifest>` deve trovarsi l'elemento `<spine>`, che può contenere annidati uno o più elementi `<itemref>`. Ciascuno di essi si riferisce a un documento già elencato nel `<manifest>`: in questo caso l'ordine in cui gli oggetti sono elencati è rilevante, perché il compito dell'elemento `<spine>` è proprio quello di stabilire l'ordine lineare della lettura della pubblicazione digitale.

È possibile specificare attraverso gli attributi dell'oggetto se quell'elemento abbia un'importanza primaria o secondaria nell'ordine di presentazione dei contenuti da parte del dispositivo. Questa scelta, a seconda delle funzionalità supportate dal software o dal dispositivo di lettura, può influire sulla presentazione dei contenuti in maniera più o meno significativa: in certi casi, per esempio, il software potrebbe presentare i contenuti elencati come secondari in finestre diverse da quella principale.

Per default tutti gli oggetti sono considerati primari; se si stabiliscono degli oggetti secondari, è necessario ricordarsi che deve esserci almeno un oggetto primario. Il sistema di lettura, infatti, per presentare i contenuti, cerca prima di tutto un `<itemref>` primario

per poter avviare la visualizzazione della pubblicazione, poi visualizza gli altri documenti secondo l'ordine stabilito all'interno dell'elemento `<spine>`:

```
<manifest>
 <item id="intro"   href="intro.html"
  media-type="application/xhtml+xml"
  />
 <item id="c1"  href="chap1.html"
  media-type="application/xhtml+xml"
  />
 <item id="c1-answerkey"
  href="chap1-answerkey.html"
  media-type="application/xhtml+xml"
  />
 <item id="c2"  href="chap2.dtb"

 <item id="ncx" href="toc.ncx"

</manifest>
<spine toc="ncx">
 <itemref idref="intro"    />
```

```
<itemref idref="c1"    />
```

```
  <itemref idref="c1-answerkey"
  <itemref idref="c2"     />
</spine>
```

Nel codice possiamo vedere che gli `<item>` elencati nel `<manifest>` vengono richiamati nell'elemento `<spine>` tramite l'attributo id che diventa il valore dell'attributo `idref`. Per l'elemento in cui viene specificato il valore "`no`" per l'attributo `linear` avremo un trattamento differente da parte del sistema di lettura, che lo separerà dal flusso principale del documento.

toc.ncx

Prima di proseguire apriamo una breve parentesi per accennare al file `toc.ncx`, chiamato in causa nell'esempio precedente. L'acronimo "toc" sta per *table of content*, letteralmente "tavola dei contenuti", e

rappresenta l'esposizione sistematica dei contenuti della pubblicazione, come una sorta di indice. Il formato NCX (*Navigation Center eXtended*) è supportato dal file `content.opf` per semplificare la navigazione e migliorare l'accessibilità del contenuto:

```
  <?xml version="1.0"?>



version="2005-1">
  <head>
  </head>
  <docTitle>
    <text>Titolo del libro</text>
  </docTitle>
  <navMap>

      <navLabel>
        <text>Frontespizio</text>
```

```
      </navLabel>
      <content
      src="Frontespizio.html"/>
    </navPoint>

      <navLabel>
        <text>Indice</text>
      </navLabel>
      <content src="Indice.html"/>
    </navPoint>

      <navLabel>
        <text>Introduzione</text>
      </navLabel>
      <content src="Introduzione.html"/>
    </navPoint>
  </navMap>
</ncx>
```

Attenzione *Come il file .ops, anche il file .ncx deve essere presente nella cartella OEBPS.*

Nell'esempio di codice, dopo l'apertura del tag `<ncx>`, la sezione più interessante inizia col marcatore `<navMap>` che apre la mappa della navigazione dell'ePub. L'elemento `<navPoint>` viene ripetuto tante volte quanti sono i documenti da elencare nell'indice del documento, e prevede gli attributi `id` (l'identificatore univoco) e `playOrder` (che stabilisce attraverso un numero l'ordine di visualizzazione del documento specificato). L'elemento `<navLabel>` contiene l'etichetta che verrà visualizzata per ogni voce dell'indice; infine in `<content>`, attraverso l'attributo `src`, viene specificato l'indirizzo del file da visualizzare.

Attenzione *In questo esempio il tag <head> è vuoto, ma potrebbe contenere i metadati sul documento.*

Un documento NCX associato alle applicazioni di XML permette di navigarne più agevolmente la struttura descrivendone la gerarchia, rendendo possibile saltare da un elemento a un altro esattamente come accade per un indice, che consente di individuare velocemente parti, capitoli, paragrafi e così via. La sua visualizzazione è simile a quella della struttura ad albero familiare a chiunque utilizzi un PC, con la possibilità di espandere gli elementi principali per rivelare quelli contenuti al loro interno. Il file `toc.ncx` permette anche di elencare elementi come note a piè di pagina, figure, tabelle e così via in elenchi separati e non gerarchici, ma accessibili comunque in maniera rapida ed efficace. Non si tratta di un file richiesto dallo standard, ma di un file

opzionale: i sistemi di lettura, tuttavia, hanno l'obbligo di supportarlo.

Guide

L'elemento `<guide>` è opzionale. Al suo interno possono essere annidati uno o più elementi `<reference>`, simili agli `<item>` già descritti: `<guide>` identifica alcune parti fondamentali per la struttura della pubblicazione e permette al sistema di lettura di provvedere a un accesso conveniente ai loro contenuti. Vediamo un esempio:

```
<guide>
 <reference type="loi" title="List
  Of Illustrations"
  href="toc.html#figures" />
 <reference type="other.intro"
  title="Introduction"
```

```
  href="intro.html" />
</guide>
```

Gli elementi `<reference>` si riferiscono agli apparati testuali come l'introduzione, la lista delle illustrazioni, la prefazione e così via. I sistemi di lettura non sono tuttavia obbligati a supportare l'elemento `<guide>`: pertanto non necessariamente, anche includendolo nel codice, otterremo un accesso distinto per tali apparati.

Riferimenti *A differenza degli* `<item>`, *gli elementi* `<reference>` *richiedono anche l'attributo* `type`, *il cui valore indica la funzione che l'apparato assolve all'interno della pubblicazione. La lista dei valori possibili è consultabile nel documento delle specifiche*

dello standard OPF, già citato in questo capitolo.

Tours (deprecato)

Non ci dilungheremo su quest'elemento, visto che il suo utilizzo è stato deprecato. Permetteva di creare veri e propri percorsi di lettura alternativi all'interno del testo, stabiliti in anticipo dall'autore della pubblicazione digitale. Rendeva possibile, per esempio, in un libro di ricette organizzate per portate (antipasti, primi, secondi e così via), stabilire un ordine di lettura impostato su tutte le ricette in cui veniva utilizzato l'ingrediente "arancia", proponendo così un'alternativa alla struttura preordinata del libro. Nel caso di un manuale tecnico, invece, si sarebbero potuti prevedere più percorsi di lettura in base, per esempio, al grado di esperienza del lettore – principiante o esperto

–, rendendo così l’esperienza di lettura più personale e interessante.

OPS: descrivere il contenuto del libro digitale

L’*Open Publication Structure* (OPS) è lo standard che descrive il contenuto vero e proprio della pubblicazione digitale. Il suo obiettivo è di fornire delle linee guida che consentano di ottenere documenti leggibili sul più ampio numero possibile di sistemi di lettura, sia software sia hardware, e di assicurare la massima uniformità di presentazione a prescindere dalla piattaforma utilizzata.

Un documento viene considerato conforme allo standard OPS se utilizza XHTML secondo le specifiche stabilite dal

documento stilato dall'IDPF e disponibile all'indirizzo http://www.idpf.org/2007/ops/OPS_2.0_final_spec.html. Ciò significa che dev'essere un documento XML ben formato e valido secondo lo schema allegato alla documentazione IDPF.

Allo stesso tempo, un sistema di lettura è conforme allo standard se interpreta il documento soddisfacendo le seguenti richieste:

- processa i file XML come richiesto dalle specifiche XML 1.1;
- riconosce tutti i marcatori descritti nelle specifiche IDPF e nei linguaggi XHTML 1.1 e CSS 2;
- non mostra immagini o oggetti con media type non supportati, a meno che non sia specificato diversamente.

L’OPS (come il suo predecessore OeBPS) stabilisce un vocabolario XHTML 1.1 di base, per così dire, che i sistemi di lettura devono supportare. Allo stesso modo, elenca le regole dei fogli stile CSS supportate dallo standard. Elencare in questa sede tutti gli elementi e le norme conformi o meno allo standard sarebbe – oltre che impossibile – poco interessante: rimandiamo per questo alla documentazione ufficiale.

Nella Figura 3.7 vediamo invece l’aspetto del contenuto di un ePub vero e proprio.

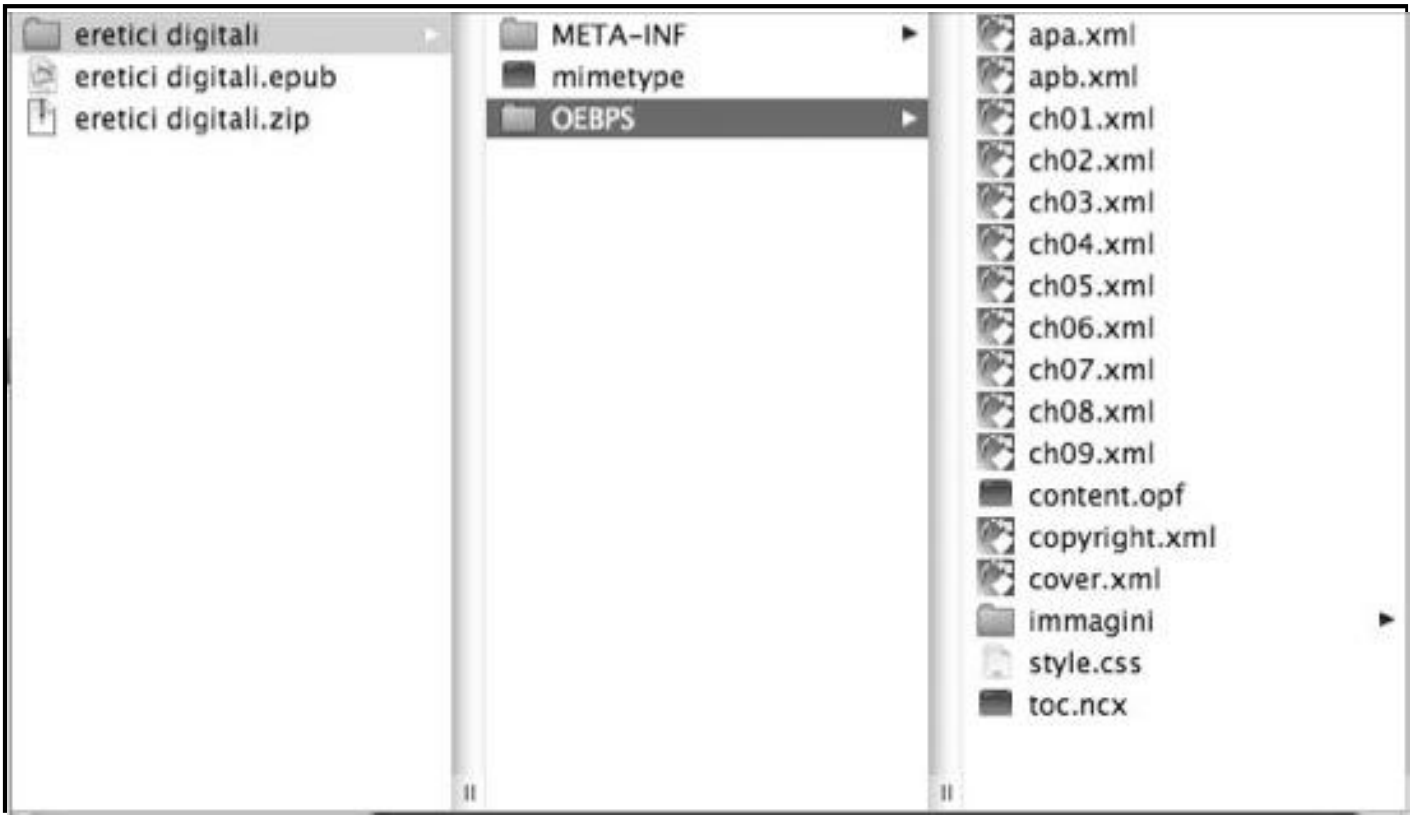

Figura 3.7 I file dell'ePub "Eretici digitali": nella colonna a destra della schermata vediamo il contenuto della cartella OEBPS.

L'ePub *eretici digitali* è stato rinominato come archivio `.zip` e poi compattato in una cartella omonima. Al suo interno troviamo la cartella *META-INF*, del cui contenuto abbiamo già parlato, il file `mime-type` e la cartella *OEBPS*, di cui possiamo osservare il contenuto nella parte destra della figura. Si

tratta dei file XML dei capitoli e degli altri apparati del libro; della copertina; del foglio stile CSS; dell'indice (`toc.ncx`) e della cartella con tutte le immagini: in altre parole si tratta del libro vero e proprio.

Software per la lettura degli ePub

L'ePub, come abbiamo più volte ribadito, fa dell'estrema flessibilità e compatibilità con il numero più ampio possibile di sistemi di lettura I suoi punti di forza, svincolandosi al massimo dalla dipendenza da hardware e software specifici. Per questo indicare un unico software di lettura in grado di interpretarlo è piuttosto difficile: ci soffermeremo quindi su alcuni programmi che ci

forniscono spunti ulteriori per lo sviluppo del discorso nei capitoli successivi.

Stanza

Stanza è un software gratuito per la lettura di pubblicazioni digitali sviluppato da Lexcycle. I formati supportati oltre a ePub sono eReader, LIT, Amazon Kindle, Mobipocket e PalmDoc, ma anche HTML, PDF, DOC e RTF. Il software esiste in due versioni: per iPhone/iPod touch e Desktop (Mac e PC).

Riferimenti *Nell'aprile del 2009 Lexcycle è stata acquisita da Amazon. Il sito in cui potete trovare il software è http://www.lexcycle.com.*

La versione del software per iPhone e iPod touch permette di gestire una libreria digitale e di sfruttare un'interfaccia di lettura

personalizzabile. I contenuti scaricati possono essere sincronizzati con Stanza Desktop: questa versione del programma dispone di diversi layout di lettura ed è in grado di esportare libri digitali in un vasto numero di formati supportati da dispositivi diversi come il Mobipocket o l'Amazon Kindle.

Stanza offre la possibilità di collegarsi a un gran numero di librerie online che propongono sia libri a pagamento sia una vastissima scelta di titoli gratuiti (si pensi per esempio alla libreria del Progetto Gutenberg, http://www.gutenberg.org).

Suggerimento *Per esempio, Stanza può collegarsi al catalogo di Feedbooks, un sito (http://www.feedbooks.com/) sul quale sono presenti numerosissimi*

titoli liberi dal diritto d'autore e quindi scaricabili gratuitamente e titoli di autori esordienti (gratuiti o a pagamento). Feedbook offre una scelta ampia di formati per il download dei libri disponibili (ePub, Mobipocket, Kindle, PDF). Nel caso del PDF offre anche degli strumenti per personalizzare il file a piacimento o con le caratteristiche più adatte per alcuni dispositivi di lettura (iLiad, Cybook Gen 3 e altri).

L'aspetto interessante di Stanza è che si tratta dell'applicazione per la lettura di libri digitali in assoluto più scaricata per iPhone. Chiariremo meglio perché questo aspetto sia così importante nei prossimi capitoli, per ora basti sapere che si tratta del software che per molti versi ha aperto la strada a delle

serie opportunità di lettura su dispositivi multifunzione, cioè non studiati specificamente per questa attività.

FBReader

FBReader (http://www.fdbreader.org) è un software gratuito, open source e multipiattaforma per la lettura dei libri digitali: è in grado di funzionare infatti sia su Desktop che su dispositivi mobili, in particolare su quelli con sistema operativo Android. Supporta una grande varietà di formati quali ePub, Palm, RTF e Mobipocket non protetti da DRM.

Tra le sue funzioni c'è la possibilità di generare automaticamente gli indici, il supporto diretto di vari formati di archiviazione (ZIP, per esempio) e un'ottima interpretazione di note a piè pagina e collegamenti ipertestuali.

Attenzione *La generazione automatica degli indici crea qualche problema con il formato ePub: di questo, infatti, non viene supportato il file* `toc.ncx`*, rendendo così la navigazione del documento più difficoltosa.*

Il vantaggio del software sta nella sua disponibilità per dispositivi portatili come l'iLiad dell'IREX, un ereader che vede così aumentare notevolmente le sue potenzialità (ne parleremo più diffusamente nel Capitolo 6), oltre che per alcuni modelli di cellulare.

Bookworm

Bookworm è una piattaforma online e gratuita per gli ePub. Il software è open source: è stato sviluppato da Threepress e dall'inizio del 2009 l'O'Reilly Lab ha acquisito il progetto, mantenendolo open source e

continuando a svilupparlo costantemente (`http://bookworm.oreilly.com/`).

Bookworm permette di creare la propria biblioteca di ePub online e di leggere libri su qualsiasi dispositivo dotato di browser web. Se si possiede un ereader in grado di interpretare gli ePub, la biblioteca online può essere sincronizzata con quella del dispositivo. Bookworm può inoltre esportare gli ebook direttamente su Stanza per iPhone.

La maggior parte del contenuto di un ePub, come abbiamo visto, è basato su XHTML e quindi è un formato capace di interagire con un browser web in modo naturale. Grazie a questo, Bookworm può supportare in maniera semplice e diretta anche layout complessi, comprendenti tabelle e liste, fondamentali per libri tecnici e scientifici, per

esempio. I fogli stile CSS inclusi nel documento vengono riprodotti senza difficoltà: tuttavia vengono leggermente modificati per evitare che interferiscano con i fogli stile nativi di Bookworm.

Riferimenti *Bookworm include anche dei particolari fogli stile pensati per la pagina stampata: il risultato è visibile utilizzando l'anteprima di stampa del browser.*

Queste modifiche limitano l'intervento del CSS dell'ePub all'area di contenuto del libro. Gli script, invece, non sono supportati per motivi di sicurezza: il marcatore `<script>`, se presente, viene rimosso automaticamente con tutto il suo contenuto.

Il file `.ncx` viene gestito senza difficoltà: i punti annidati nel sommario si

espanderanno man mano che si procede nella lettura del documento. L'ePub si aprirà automaticamente sul primo punto del file `.ncx` con l'attributo `linear` impostato su `yes`.

Essendo un'applicazione basata sul browser, il supporto delle immagini è valido per tutti i formati supportati dall'ePub.

Bookworm si pone tra gli obiettivi il riuscire a visualizzare anche ePub non completamente conformi alle specifiche. Ci sono alcuni casi in cui il file verrà comunque rifiutato perché è impossibile interpretarlo correttamente; in caso di gravi errori di formattazione, invece, il documento passerà a un servizio di controllo, che procederà a identificare il problema del file con un buon grado di attendibilità e restituirà all'utente un

report con l'elenco dei problemi riscontrati nel file caricato.

Riferimenti *L'epubcheck service è disponibile all'indirizzo http://www.thre-epress.org/document/epub-validate/*.

Se l'esecuzione di Bookworm sul browser presenta senza dubbio un insieme di vantaggi, non sono da trascurare le limitazioni. Alcuni ebook con contenuti molto lunghi (fumetti a pagina singola o romanzi senza interruzioni nel testo) non sono ideali per una lettura sul Web. Tendenzialmente è preferibile suddividere testi così lunghi in più file XHTML: questo è fondamentale, per esempio, per alcuni dispositivi mobili, che spesso hanno problemi con file di grandi

dimensioni e possono avere impostate delle limitazioni riguardo al peso del file.

Attenzione *Bookworm non è un sistema di lettura completamente conforme. I problemi e i bug sono elencati all'indirizzo http://code.google.com/p/threepress/issues/list. L'applicazione, inoltre, non supporta gli ebook protetti da DRM.*

Conclusioni

Dare conto di tutti i formati di codifica che possono essere utilizzati per i libri digitali e, per ciascuno, di tutti i lettori in grado di interpretarli è un'impresa assai ardua in queste poche pagine. Presentando i formati più diffusi e quelli dalle caratteristiche più

interessanti per l'editoria digitale speriamo di aver fornito un panorama sufficientemente completo che possa lasciar intuire quali e quante siano le possibilità offerte dalle varie codifiche. Sui software di lettura, inoltre, si sono potuti fare solo pochi accenni: entrare nel merito dell'utilizzo puntuale di ciascuno sarebbe stato tutto sommato poco utile, e si è scelto di tralasciare altre applicazioni altrettanto interessanti per questioni di spazio.

Riferimenti *Se desiderate avere una panoramica più ampia di formati e soft-ware per i libri digitali un ottimo punto di partenza può essere il lemma* Comparison of e-book formats *nella versione inglese di Wikipedia, all'indirizzo* *http://en.wikipedia.org/wiki/Comparison_of_e-book_formats*,

oppure la voce E-book formats *(sempre in inglese) all'indirizzo http://wiki.mobileread.com/wiki/E-book_formats, sulla Wiki del sito mobileread.com.*

REALIZZARE UN LIBRO DIGITALE

I metodi per realizzare libri digitali sono diversi a seconda degli obiettivi che si intendono raggiungere: creare una collezione personale, convertire in altri formati documenti già esistenti o intraprendere la pubblicazione sistematica di libri elettronici. Come si può intuire, quest'ultima è una delle operazioni più complesse, soprattutto per la pianificazione che richiede riguardo risorse, procedure e cambiamenti nel ciclo di produzione del libro.

Dal punto di vista di un editore, il lavoro di preparazione dei file necessari per il libro digitale presenta una serie di interrogativi che vanno al di là delle questioni tecniche e spostano la discussione su temi nevralgici riguardo il processo di lavorazione e l'identità editoriale stessa. Si tratta di aspetti tutt'altro che marginali: dedicheremo il capitolo successivo alla loro analisi.

Nei paragrafi seguenti, invece, descriveremo un iter per la produzione del libro digitale, cercando di tenere conto delle esigenze più diverse. La procedura illustrata non è di certo l'unica possibile; tuttavia ci sembra una maniera efficace per sfruttare nel modo più flessibile le opportunità offerte dal digitale. I formati che otterremo al termine della procedura saranno ePub, mobi e PDF.

Nel descrivere la procedura di lavorazione faremo spesso riferimento al libro che state leggendo per alcuni esempi pratici. Trattandosi di un manuale, presenta una serie di difficoltà che per la maggior parte dei libri – e per la maggior parte degli editori – sono straordinarie: non tutti devono gestire tabelle e immagini, e ancora meno si trovano di fronte alle complessità date dai listati di codice. Tuttavia riteniamo interessante accennare a una gamma più ampia di difficoltà, per offrire un panorama di possibilità più variegato possibile.

Partiamo dal principio: il manoscritto

Come abbiamo già osservato, l'attuale processo di produzione del libro stampato si allontana dal mondo digitale solo al momento della stampa vera e propria: le altre fasi di produzione, inclusa la stesura da parte dell'autore, si svolgono attraverso un computer e sono di fatto già digitali.

La questione allora è: come sfruttare al meglio le fasi di produzione precedenti alla stampa in modo da poter ricavare oltre al libro fisico anche quello digitale? In realtà, anche il prodotto finale delle lavorazioni (stesura del testo, redazione, impaginazione) è di per sé un libro digitale: i file di stampa

che si consegnano al tipografo non sono nient'altro che la versione in PDF del libro da stampare.

Tuttavia, nell'ottica del libro digitale per come l'abbiamo descritto finora, abbiamo visto che esistono formati più flessibili che permettono di sfruttare al meglio le caratteristiche e le opportunità offerte dalla lettura su supporti differenti rispetto alla carta: per questo dovremmo riuscire a modificare le fasi produttive di quel tanto che serve a produrre, insieme ai PDF per la stampa, anche gli altri formati più adatti all'ebook.

A questo scopo occorre ripensare il processo di lavorazione del libro a partire (idealmente) dalla stesura da parte dell'autore. In realtà non si tratta di grandi rivoluzioni, quanto piuttosto di sfruttare in

maniera più consapevole e completa strumenti di cui siamo già ampiamente in possesso e con i quali siamo abituati a lavorare da tempo, come per esempio gli elaboratori di testo.

Il lavoro di un autore, salvo rari casi, si avvantaggia notevolmente delle opportunità offerte dai word processor. La semplificazione di tutte le operazioni di modifica e intervento su ciò che si sta scrivendo, oltre alle possibilità di organizzarne semplicemente la struttura, costituiscono senza dubbio un’evoluzione importante nel modo in cui si approccia la scrittura.

La nostra attenzione si focalizzerà proprio sull’organizzazione della struttura del lavoro dell’autore, un aspetto che spesso viene trascurato o perché non si è in possesso delle

sufficienti competenze per gestirla correttamente, o perché semplicemente non vi si dedica la dovuta attenzione.

Il metodo più semplice per produrre un documento con una struttura coerente e ben organizzata è quello di definire un *foglio stile*, ossia un documento che faccia da modello stabilendo a priori la gerarchia del contenuto e un'efficace valorizzazione della sua semantica. Nel prossimo paragrafo vedremo perché il foglio stile è un elemento fondamentale per la produzione di un libro digitale e come crearne uno utilizzando due tra i più comuni elaboratori di testo: Microsoft Word e Writer di OpenOffice.

Il foglio stile

La difficoltà principale nella definizione del foglio stile sta nel fatto che a partire da questo modello si dovranno ricavare sia i file necessari per la stampa sia quelli per il libro digitale. Questa duplice funzione comporta la necessità di una straordinaria attenzione verso l'equilibrio tra esigenze così differenti delle due versioni del libro. Su tutte basti pensare all'idea di *pagina*, fondamentale per un libro cartaceo, inesistente nel caso di numerosi formati di file per ebook.

Attenzione *In realtà, nella definizione di un foglio stile da parte di un editore intervengono anche ragionamenti più complessi legati all'identità delle collane, di cui parleremo nel capitolo seguente.*

Il vantaggio nella definizione di un foglio stile ben strutturato sta nella possibilità di ottenere un documento di testo che può essere salvato e condiviso in formati differenti dal DOC, a cui in genere si fa riferimento, come per esempio XML. Word e Writer sono entrambi in grado di salvare il documento in formato XML, grazie all'applicazione corretta degli *stili* che determinano la struttura del documento.

Come abbiamo visto nei capitoli precedenti, XML è un linguaggio versatile e flessibile, alla base di molti dei formati più diffusi per i libri digitali. Ottenere dal documento di testo prodotto dall'autore un contenuto con una marcatura XML già ben formata e valida offre la possibilità di utilizzare quel contenuto per le successive

lavorazioni con molta semplicità, qualunque sia il formato di destinazione a cui si aspira.

Attenzione *Stiamo dando per scontato che l'autore abbia le competenze necessarie per mettere in pratica quanto diremo in seguito. Non sempre è così: per quanto gli elaboratori di testo siano "pane quotidiano" per i professionisti della scrittura, le pratiche che descriveremo per l'utilizzo di questi software potrebbero sembrare delle "inutili complicazioni" che interferiscono col processo creativo. In realtà si sottovaluta il vantaggio di lavorare su un documento con una struttura ben definita, entro la quale inserire il contenuto. Inoltre, il compito di attribuire i corretti valori semantici ai*

vari elementi del testo dovrebbe essere strettamente autoriale.

Il foglio stile deve quindi definire una serie di caratteristiche del documento che si intende creare. Ricordiamo ancora una volta che deve servirci sia per la stampa sia per il digitale, pertanto dovrà determinare:

- le dimensioni della pagina, i margini, il tipo di carattere, la sua dimensione, l'interlinea e tutti gli altri elementi che concorrono a una collocazione quanto più verosimile dei contenuti all'interno del documento rispetto a ciò che si otterrà con la stampa;

Suggerimento *Produrre un modello che dia all'autore una percezione verosimile dell'ingombro del suo manoscritto nella pagina permette sia a lui sia all'editore di lavorare con un'idea*

più precisa dell'effettivo sviluppo del lavoro.

- titoli e paragrafi e la loro corretta collocazione tra i livelli della struttura; tanti stili differenti quanti sono i tipi di paragrafo che si intendono utilizzare nel testo;
- gli stili di carattere, da applicare alle porzioni di testo da isolare all'interno di un paragrafo, per esempio corsivo e grassetto;
- la collocazione e la gestione di elementi complessi come immagini e tabelle.

Ognuno di questi aspetti presenta problematiche differenti e richiede decisioni che non permettono troppe generalizzazioni. In particolare occorre riflettere di volta in volta su quali decisioni siano legate solamente alla presentazione del testo e quali invece abbiano un preciso valore semantico: le prime

comportano scelte grafiche o sono determinate dalle necessità della stampa, le seconde assumono un ruolo ancora più centrale per il libro digitale, che grazie alle sue peculiarità consente di dare maggior rilievo ad alcuni aspetti.

Vediamo nella pratica come affrontare questi temi.

Immaginiamo di dover definire il foglio stile di questo libro, lavorando su Word 2003 e Writer di OpenOffice 3.0, una combinazione software abbastanza diffusa da essere familiare alla maggior parte dei lettori. Il foglio stile definito farà da *modello* per la stesura di un singolo capitolo: il libro non sarà contenuto in un unico documento ma in tanti documenti quanti sono i capitoli e gli

apparati che lo compongono. In seguito saranno chiari i motivi di questa scelta.

Attenzione *Per ragioni di spazio, ci soffermeremo sugli aspetti grafici legati alla stampa nella definizione del foglio stile solo nei casi in cui la loro presenza comporti delle scelte complesse in relazione alla creazione del libro digitale; in tutti gli altri li daremo per scontati.*

Una questione di stile: paragrafo e carattere

Dopo aver definito le misure della pagina, i margini e il tipo di carattere da utilizzare, il cuore della creazione del modello sta nella definizione degli *stili*. Occorre subito fare una distinzione tra *stili di paragrafo* e *stili di carattere*: i primi si applicano a blocchi di

testo quali titoli e paragrafi, appunto, mentre i secondi definiscono ed enfatizzano piccole porzioni all'interno dei paragrafi stessi. Per esempio, queste righe hanno applicato lo stile di paragrafo A-Body, e al loro interno lo stile di carattere A-EnfasiCorsivo sui termini che si è ritenuto importante *mettere in rilievo*.

Attenzione *I nomi degli stili che abbiamo appena indicato non sono quelli predefiniti di Word ma quelli che si è stabilito di utilizzare nella definizione del foglio stile della collana Pocket di Apogeo:* A-Body *corrisponde, per esempio, allo stile* Normale *o* Predefinito. *Nel seguito della spiegazione, per semplicità, ci rifaremo agli stili predefiniti proposti da Word e Writer.*

Gli stili che stiamo definendo hanno due funzioni fondamentali (Figura 4.1).

- Semantica: ci dice che ruolo ha quella parte di testo nella struttura del documento (titolo e di che livello, paragrafo e che tipo di paragrafo e così via).
- Grafica: attraverso le caratteristiche della formattazione ci dà un'idea verosimile di ciò che si otterrà a libro stampato.

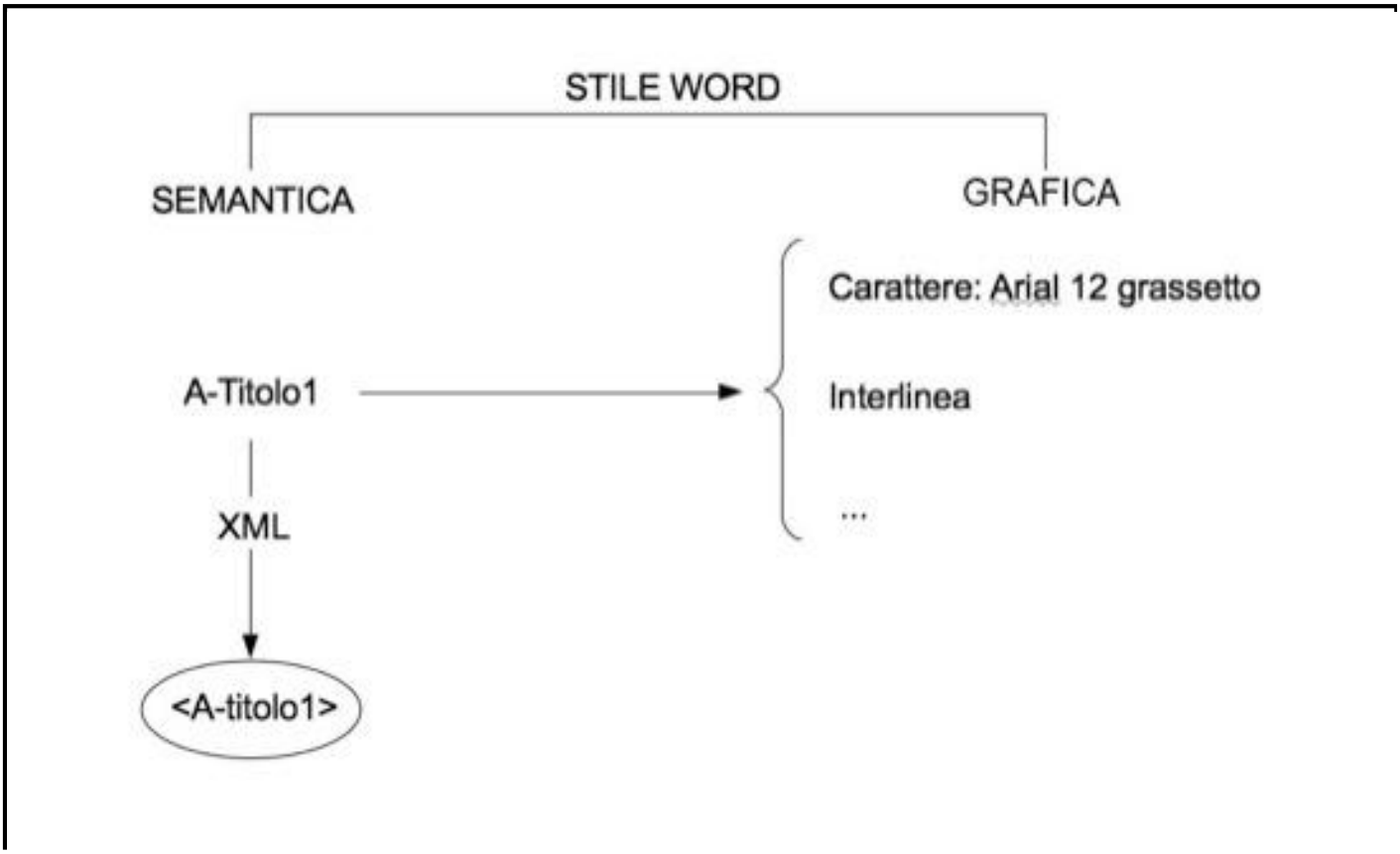

Figura 4.1 Lo schema mostra visivamente la separazione tra valore semantico e grafico. La marcatura in XML deriva direttamente dal valore semantico che individuiamo per quel determinato stile.

In altre parole, ai fini del libro digitale le caratteristiche grafiche stabilite all'interno degli stili sono completamente indifferenti, al punto che potrebbero anche non essere definite affatto. Ciò che conta per la corretta produzione di un documento ricco dal punto

di vista semantico è che a ogni contenuto sia applicato lo stile che ne indica la funzione all'interno del testo. Vale a dire che applicare lo stile A-Titolo1 ci porterà a ottenere nell'XML un codice simile al seguente:

```
<A-Titolo1>Esempio titolo</A-Titolo1>
```

Come arrivare a ottenere un codice simile a partire da un documento Word ben strutturato sarà l'argomento delle pagine seguenti.

Creare il modello su Word

Creare un foglio stile significa creare un file modello che costituisca la base per tutte le fasi di scrittura e lavoro sul testo. I file modello di Microsoft Word hanno l'estensione `.dot`. Il vantaggio di lavorare su questo formato è costituito dalla possibilità di personalizzare in maniera molto flessibile anche

l'area di lavoro, proponendo così solo gli strumenti necessari per la corretta strutturazione del documento, prevenendo errori di formattazione e semplificando il lavoro sia dell'autore – che con meno strumenti a disposizione avrà meno dubbi sul da farsi – sia del redattore – che dovrà intervenire meno per rimediare a tali errori e potrà dedicarsi maggiormente al contenuto.

Suggerimento *Per salvare e quindi creare un file* `.dot` *occorre selezionare il formato relativo dalla tendina* Formato *nella finestra* Salva con nome.

È utile iniziare a definire gli stili a partire da quelli di paragrafo. Dal menu *Formato*, selezionando la voce *Stili*, si visualizzeranno tutti gli stili predefiniti per i documenti Word. Si possono utilizzare gli stili esistenti,

modificarli secondo le proprie esigenze o crearne di nuovi. Se decideremo di creare degli stili completamente nuovi, il nome che attribuiremo loro sarà fondamentale per la produzione del documento XML.

Suggerimento *Nella scelta del nome raccomandiamo di privilegiare chiarezza, brevità e facilità di individuazione. Nel caso dei fogli stile delle collane di Apogeo, per esempio, si è scelto di aggiungere il prefisso "A-" ai nomi degli stili, in maniera da rintracciarli immediatamente senza ambiguità.*

Nel pensare gli stili di paragrafo rilevanti nella definizione del foglio stile di questo libro abbiamo individuato i seguenti elementi.

- Numero e Titolo di capitolo: compongono l'intestazione di un capitolo e ricorrono una sola volta all'interno del documento.

  Attenzione *La spaziatura che potete osservare in questo libro tra il titolo di capitolo e il paragrafo che lo segue è un perfetto esempio di scelta grafica legata alla stampa ma totalmente priva di valore semantico e indifferente ai fini della creazione del libro digitale.*

- Titoli di paragrafo: i titoli di paragrafo determinano la struttura del documento, generando la classica struttura ad albero che siamo abituati a ritrovare nell'indice del libro attraverso l'indentazione dei titoli. A seconda della collocazione nella gerarchia del testo, a ogni titolo viene associato un livello: in questo libro esistono titoli di primo,

secondo e terzo livello, annidati come si può vedere nella Figura 4.2.

- Paragrafo: la maggior parte del corpo del testo ha applicato questo stile di paragrafo. È lo stile di base, per così dire, a partire dal quale si determinano armonicamente tutti gli altri.
- Note: le note sono una caratteristica della collana Pocket. La loro resa grafica – caratterizzata da un'icona e dal corsivo – è una scelta strettamente editoriale; tuttavia hanno un evidente valore semantico, in quanto ognuno dei quattro tipi di note corrisponde a un determinato contenuto, spiegato nelle pagine introduttive del libro. La loro identificazione corretta attraverso uno stile specifico permette di tener traccia di questo

valore e di utilizzarlo nel modo più efficace all’interno dell’ebook.

- Codice: gli esempi di codice hanno uno stile grafico molto differente da quello del paragrafo normale, e devono rispettare le regole del linguaggio in cui sono scritti. È una necessità tipica dei manuali di informatica, ma costituisce un ulteriore grado di complessità nella strutturazione del foglio stile.

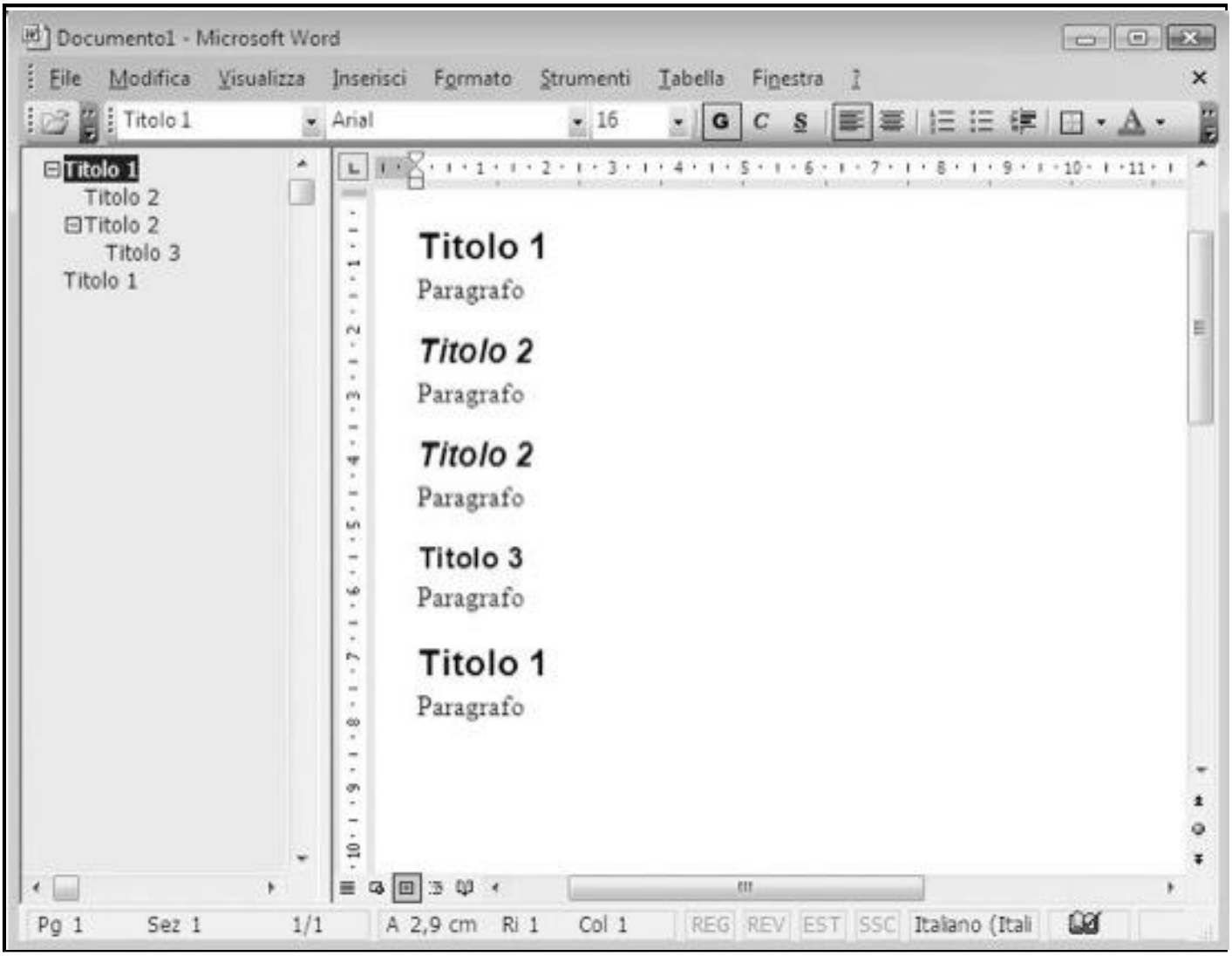

Figura 4.2 Visualizzando la mappa del documento su Word, si apre a sinistra un riquadro con l'albero dei contenuti del documento in cui abbiamo correttamente applicato gli stili, e dal quale potremo spostarci nella struttura con semplicità.

Se capire i principi alla base della definizione degli stili di paragrafo è piuttosto

semplice, riguardo agli stili di carattere occorre fare più attenzione. Oltre ai più diffusi (corsivo e grassetto), in questo libro abbiamo visto spesso applicato lo stile `codice`, per parti di linguaggio informatico presenti all'interno di un paragrafo normale. Riguardo agli stili di carattere possiamo fare due osservazioni di ordine differente.

- L'applicazione della sola *formattazione* corsivo o grassetto a un termine non implica l'applicazione di uno stile di carattere. Questa distinzione è fondamentale: l'applicazione della formattazione senza la definizione dello stile appropriato (indicata in genere come applicazione *in linea* della formattazione) non aggiunge alcun valore semantico al termine ma si limita a intervenire sulla presentazione: per questo non

lascia traccia in una successiva esportazione del documento in un linguaggio semantico come XML.

- Da un punto di vista strettamente grafico, gli stili di carattere applicati nella collana Pocket per frammenti di codice e link sono identici. Tuttavia, è evidente che il valore semantico costituisce la differenza: l'applicazione dell'uno o dell'altro – indifferente ai fini del libro cartaceo proprio per l'aspetto identico – è assolutamente rilevante ai fini del libro digitale, che può valorizzare in maniera ben diversa il collegamento ipertestuale rendendolo attivo durante la lettura dell'ebook.

Vediamo nelle Figure 4.3 e 4.4 gli strumenti per la creazione degli stili in Word e in Writer.

Figura 4.3 Il riquadro Stili e formattazione in Word.

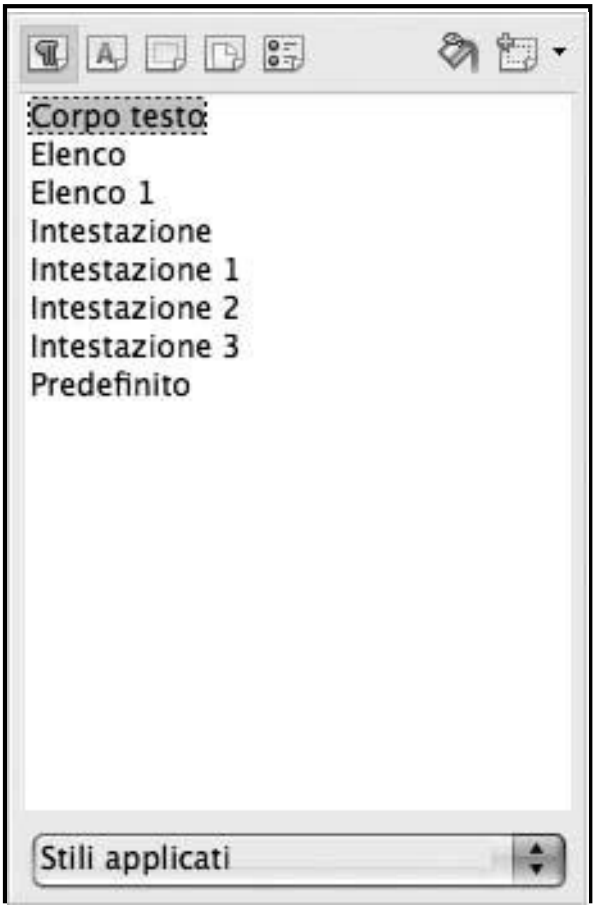

Figura 4.4 Il riquadro Stili e formattazione in Writer.

Nonostante le differenze di interfaccia, in entrambi gli elaboratori di testo troviamo le stesse convenzioni per indicare i diversi raggruppamenti di stili. Quelle che maggiormente ci interessano sono le seguenti.

- L'icona per lo stile di paragrafo: su Word visibile a destra dello stile *Normal* ¶, su

Writer è il primo pulsante da sinistra nella barra degli strumenti

- L'icona per lo stile di carattere: su Word visibile a destra dello stile *Enfasi* , su Writer è il secondo pulsante da sinistra nella barra degli strumenti
- L'icona per gli stili di elenco (puntato o numerato): su Word visibile a destra dello stile 1 1/1.1/1.1.1 – che mostra così anche la struttura della numerazione dell'elenco –, su Writer è il quinto pulsante da sinistra nella barra degli strumenti

Attenzione *La principale differenza di interfaccia tra Writer e Word è che quest'ultimo visualizza in un'unica scheda tutti gli stili disponibili per il documento, mentre Writer li suddivide*

in schede (nell'ordine: paragrafo, carattere, cornice, pagina e lista).

A partire da queste due finestre possiamo scegliere tra gli stili già definiti che vogliamo utilizzare, eventualmente modificarli o crearne di nuovi. Ricapitoliamo quanto ci occorre per quanto riguarda gli stili di paragrafo.

- Intestazione capitolo:
  - numero capitolo;
  - titolo capitolo.
- Titoli:
  - titolo di primo livello;
  - titolo di secondo livello;
  - titolo di terzo livello.
- Paragrafo:

- paragrafo base;
- nota – attenzione;
- nota – suggerimenti;
- nota – terminologia;
- nota – riferimenti;
- codice.

Riferimenti *Questa struttura così dettagliata potrebbe essere ridotta, nel caso di un romanzo, per esempio, alle sole intestazioni di capitolo e a un paio di stili di paragrafo.*

Per quanto riguarda invece gli stili di carattere:

- corsivo;
- grassetto;

- codice;
- link.

Attenzione *Uno stile di paragrafo e uno di carattere che in questo esempio non sono previsti sono quelli per le citazioni, siano esse introduttive, interne al paragrafo o contenute in paragrafi distinti. Se presenti, si dovrebbero prevedere appositi stili di paragrafo e di carattere.*

Si tratta di un elenco piuttosto semplice: possiamo trovare tra gli stili predefiniti ciò che fa al caso nostro per i titoli, il paragrafo base, per il codice e per tutti e quattro gli stili di carattere. Nella Tabella 4.1 vediamo le corrispondenze.

Iniziare a lavorare con gli stili corretti è dunque molto semplice: restano da definire quelli per l'intestazione del capitolo e le note, caratteristica della collana di cui questo libro fa parte.

Vediamo quindi come creare un nuovo *stile di paragrafo*, *di carattere* e *di elenco* in Word.

Stili di paragrafo		
Stile necessario	Predefinito Word	Predefinito Writer
Numero capitolo	Assente	Assente
Titolo capitolo	Assente	Assente
Titolo primo livello	Titolo 1	Intestazione 1

Stili di paragrafo		
Titolo secondo livello	Titolo 2	Intestazione 2
Titolo terzo livello	Titolo 3	Intestazione 3
Paragrafo base	Normale	Predefinito
Codice	Preformattato HTML	Testo preformattato
Nota Attenzione	Assente	Assente
Nota Suggerimenti	Assente	Assente
Nota Terminologia	Assente	Assente
Nota Riferimenti	Assente	Assente

Tabella 4.1 Gli stili di paragrafo e di carattere in Word e in Writer.

Stili di carattere		
Corsivo	Enfasi (corsivo)	Enfasi
Grassetto	Enfasi (grassetto)	Enfasi forte
Codice	Codice HTML	Testo sorgente
Link	Collegamento ipertestuale	Collegamento

Nel riquadro *Stili e formattazione* fate clic sul pulsante Nuovo stile.

1. Nella finestra*Nuovo stile* occorre:
2. inserire il nome scelto per il nuovo stile che si sta creando nel campo *Nome*;

1. selezionare il tipo di stile (paragrafo, carattere, tabella o elenco), ciascuno contrassegnato da un'icona differente;
2. stabilire se basare il nuovo stile su uno stile già impostato;
3. impostare lo stile da applicare al paragrafo successivo.

Attenzione *Se si decide di basare il nuovo stile su uno stile già impostato, occorre tener presente che qualunque modifica allo stile di base verrà applicata a cascata su tutti gli stili da esso derivati.*

Le altre informazioni servono per modificare la resa grafica dello stile, e non le prenderemo qui in esame. Ci soffermeremo invece sul menu a tendina in basso a sinistra

della finestra, *Formato*: da lì occorre selezionare la voce *Paragrafo* per visualizzare l'omonima finestra (Figura 4.5). Da qui, nella scheda *Rientri e spaziatura*, possiamo selezionare dal menu a discesa *Livello struttura* la posizione nell'albero dei contenuti del testo a cui applicheremo lo stile che stiamo definendo. *Corpo del testo* è il livello dei paragrafi di base, *Livello 1* rappresenta i titoli di primo livello, *Livello 2* quelli di secondo e così via.

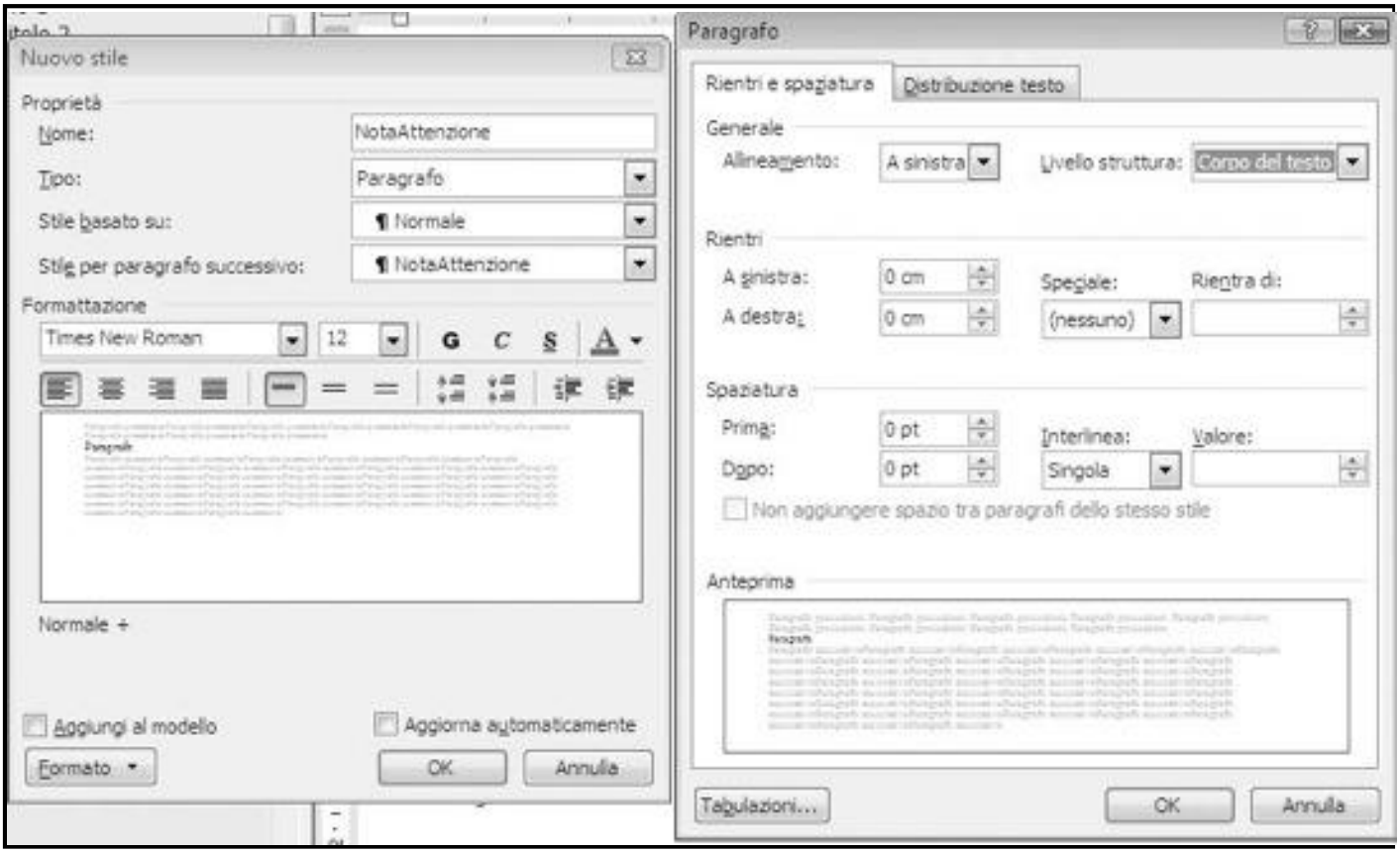

Figura 4.5 Le finestre Nuovo stile e Paragrafo.

Per creare uno *stile di carattere* occorre seguire una procedura simile, scegliendo come tipo di stile *Carattere* e ricordandosi di selezionare *Carattere predefinito paragrafo* come base dello stile (Figura 4.6): in questo modo il testo a cui verrà applicato lo stile di carattere sarà contenuto

correttamente nella struttura del paragrafo in cui è inserito.

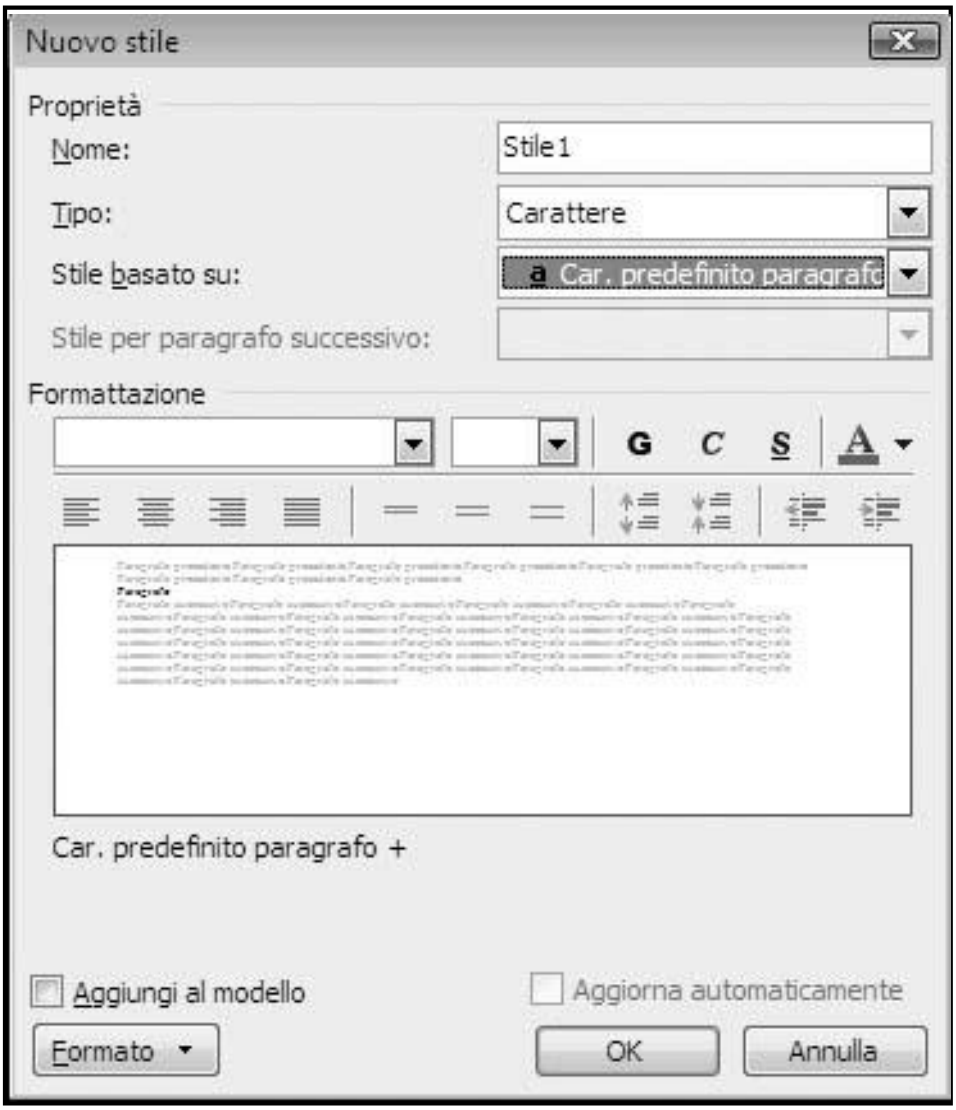

Figura 4.6 La finestra Nuovo stile per la creazione di uno stile di carattere.

Nella definizione di un nuovo *stile di elenco*, invece, occorrerà selezionare come

tipo di stile *Elenco* e in seguito definire le caratteristiche necessarie al tipo di lista che si vuole creare. In questo caso l'attenzione va soprattutto ai diversi livelli di struttura dell'elenco stesso: occorre definirli con attenzione in maniera che vengano esportati nel modo giusto in XML. Per esempio:

1. questa è una voce di primo livello, e il suo stile è definito selezionando la voce *1° livello* nel menu a tendina *Applica formattazione a* (Figura 4.7);

1. questa è una voce di secondo livello, il cui stile è definito selezionando la voce *2°livello* nel menu a tendina *Applica formattazione a*.

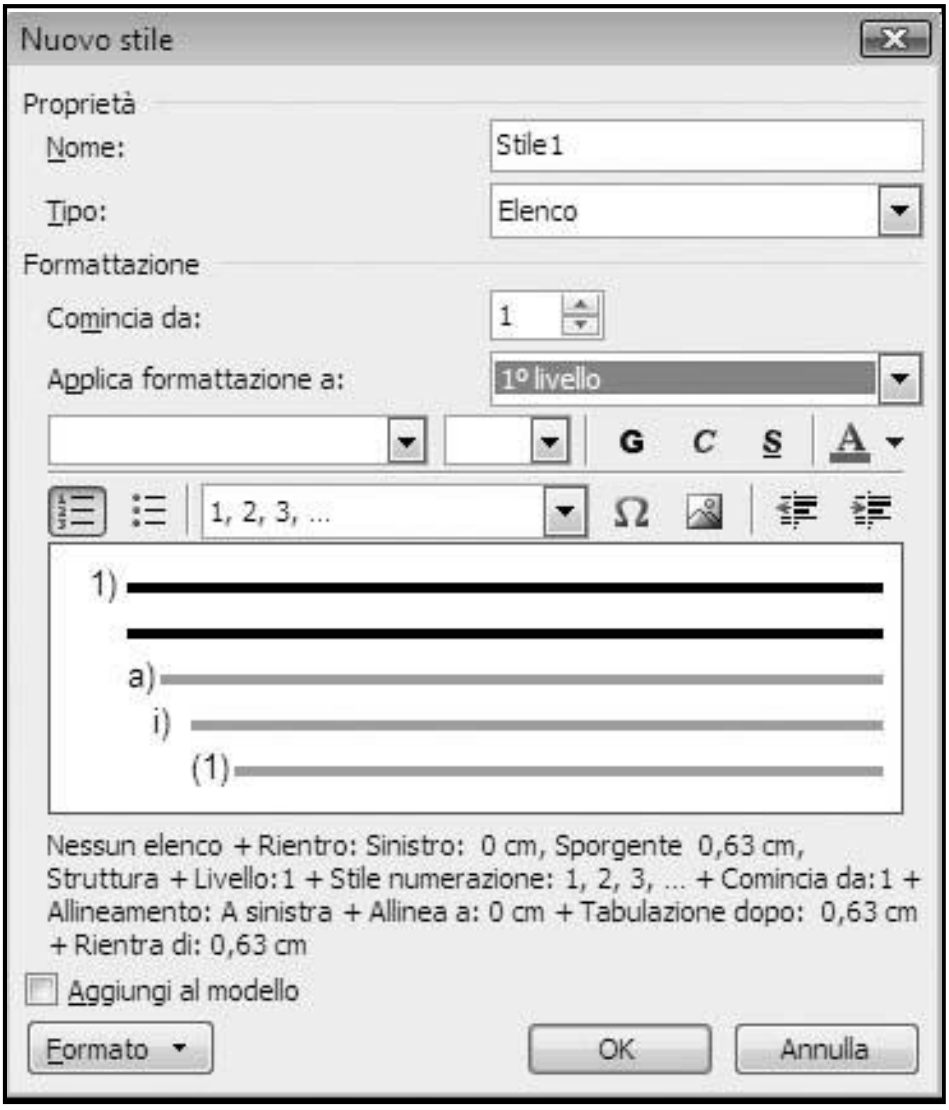

Figura 4.7 La finestra Nuovo stile per la creazione di uno stile di elenco.

L'applicazione del corretto livello di struttura della voce di elenco si ottiene facendo clic sull'icona *Aumenta rientro* presente nella barra degli strumenti *Formattazione*

nella sezione dedicata agli elenchi puntati e numerati.

> **Attenzione** *L'aspetto grafico delle voci dell'elenco (tipo di numerazione, indentazione e così via) non ha alcun valore nell'esportazione in XML. Ciò che conta è che la struttura sia applicata correttamente attraverso i comandi* Aumenta rientro *e* Riduci rientro *nella barra degli strumenti* Formattazione *oppure da* Visualizza > Struttura.

Le scelte fatte per la formattazione dei diversi livelli di struttura dello stile elenco che stiamo creando concorrono tutte a definire un solo stile elenco (Figura 4.7). In altre parole, lo stile A-ElencoNumerato avrà al suo interno definite le caratteristiche per la voce

1° livello, *2° livello* e così via, racchiuse tutte all'interno di un unico stile.

La visualizzazione del documento attraverso la sua struttura (*Visualizza > Struttura*) permette di gestirne l'organizzazione in maniera molto accurata e diretta, oltre che di tenerne sotto controllo lo sviluppo (Figura 4.8). Questo non deve comunque far perdere di vista l'applicazione corretta degli stili definiti da foglio stile, ma va utilizzata piuttosto come un supporto per far sì che il documento lo rispetti quanto più possibile.

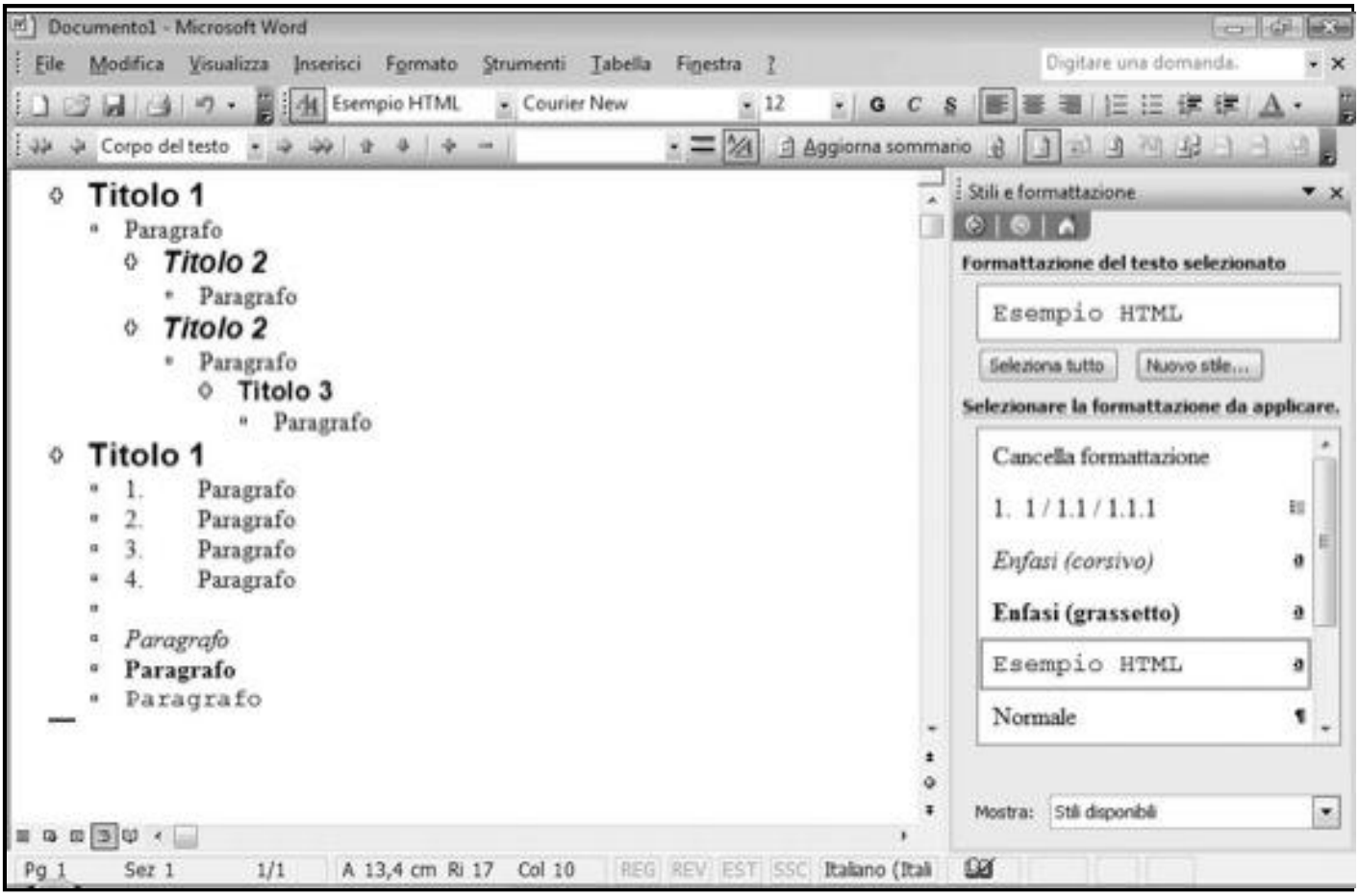

Figura 4.8 La modalità Visualizza > Struttura in Word.

Ora che la procedura per creare nuovi stili ci è chiara e che abbiamo a disposizione gli strumenti per compilare il modello, possiamo riempire un paio di pagine del file .dot con del testo di esempio che mostri come e dove applicare gli stili predefiniti, in

maniera simile a quanto mostrato nella Figura 4.9.

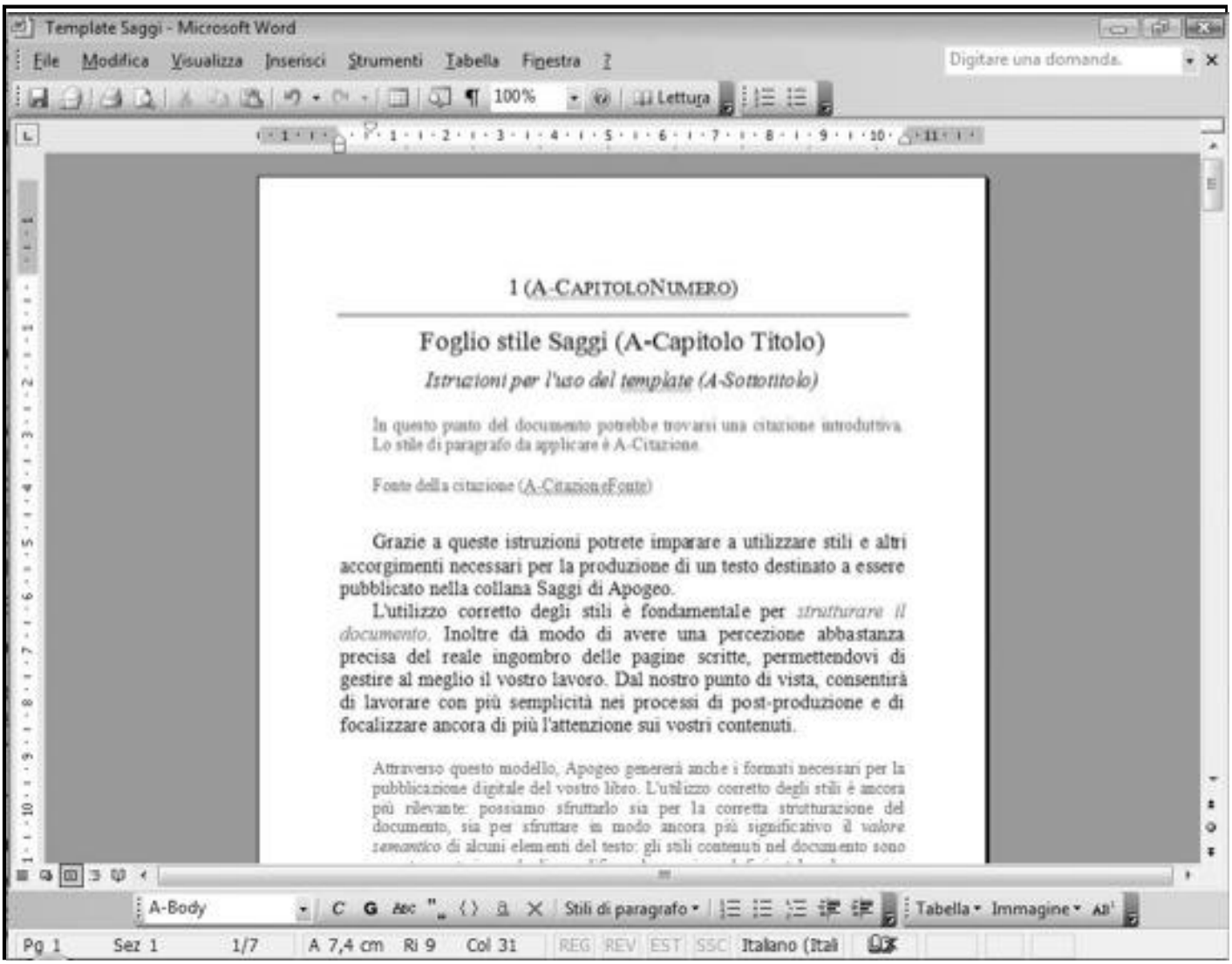

Figura 4.9 Nel creare un modello di documento è possibile – e consigliabile – anche definire l'area di lavoro con barre degli strumenti personalizzate.

Suggerimento *L'inserimento del testo segnaposto nel modello non è certo obbligatorio, ma può risultare utile per rendere più semplice l'utilizzo del foglio stile, perché fornisce un esempio immediato della sua applicazione.*

Dato che la struttura del foglio stile è ormai definita – almeno nei suoi elementi di base – possiamo iniziare a ragionare sulla sua esportazione in XML, o più precisamente in una sua particolare applicazione pensata esplicitamente per la descrizione di contenuti editoriali: DocBook.

Attenzione *Un modello creato in Word con estensione* `.dot` *è utilizzabile anche su OpenOffice. Da OpenOffice non si possono però creare file* `.dot`*:*

l'estensione per i modelli di Writer è `.ott`.

Prima trasformazione: DocBook

Nel Capitolo 2 abbiamo elencato alcune delle applicazioni più rilevanti di XML, tra cui anche DocBook. Ricordiamone brevemente le principali caratteristiche.

- DocBook definisce un insieme ampio ma finito di marcatori e di regole per il loro utilizzo corretto.
- I marcatori definiscono sempre e soltanto la struttura semantica del documento, mai la resa estetica, delegata ai fogli stile CSS.

- Un documento DocBook (o più in generale XML) può essere trasformato in un altro formato attraverso l'utilizzo dei fogli di trasformazione XSL.

DocBook propone un vocabolario di marcatori ideale per la descrizione del contenuto di un libro ed è un'applicazione di XML non proprietaria molto versatile, grazie al gran numero di fogli di trasformazione disponibili in Rete e adatti a ottenere i più diffusi formati per i documenti digitali.

Riferimenti *Un'ampia sezione di marcatori di DocBook è pensata per contenuti di tipo tecnico e informatico: questo non significa che sia possibile trattare solo testi di questo tipo.*

Da Writer di OpenOffice, se abbiamo creato il modello utilizzando gli stili predefiniti

proposti dall'elaboratore di testo, abbiamo l'opportunità di effettuare il salvataggio direttamente in DocBook, accedendo al menu *File > Salva* con nome e selezionando nel menu a tendina *Tipo file* la voce *DocBook (.xml)*.

Tuttavia, come abbiamo visto, difficilmente gli stili predefiniti avranno soddisfatto tutte le esigenze del foglio stile che abbiamo progettato, e con molta probabilità avremo preferito definirne di nuovi: in questo caso il salvataggio in DocBook così com'è di default non sarà utilizzabile. Come procedere per sfruttare comunque questa opportunità offerta dal software? Vediamo a questo scopo di spiegare più nel dettaglio la struttura di un documento codificato in DocBook, in maniera da familiarizzare con i suoi aspetti

principali e poter così seguire più facilmente le successive fasi del lavoro.

Le basi di DocBook

DocBook è un'applicazione XML studiata appositamente per la composizione di documenti tecnici, ma versatile anche per la marcatura di altri tipi di contenuti. Osserviamo il seguente esempio di codice:

```
<?xml version="1.0" encoding="UTF-
8"?>

  version="5.0">
   <title>Esempio di libro</title>
   <chapter>
    <title>Capitolo 1</title>
     <sect1>
      <title>Titolo section 1</title>
       <para>Questo è il testo del
       primo
```

```
        <sect2>
         <title>Titolo section
         2</title>
          <para>Qui vediamo
          l'utilizzo
           di un tag utile per


         </sect2>
     </sect1>
   </chapter>
</book>
```

L'elemento radice (`<book>`) contiene la dichiarazione del namespace XML. Al suo interno annideremo il contenuto dell'intero libro. A seguire, l'elemento `<title>` contiene il titolo del libro che stiamo creando.

Suggerimento `<title>` *può ricorrere in un gran numero di elementi a vari livelli della struttura del docu-*

mento, come possiamo vedere dalla sua

presenza anche all'interno dell'elemento `<chapter>`.

`<chapter>` indica l'apertura di un nuovo capitolo. All'interno dei capitoli il testo è suddiviso in sezioni (`<sect1>`, `<sect2>` e così via) e paragrafi (`<para>`). Le sezioni (`sect1` sta per *section 1*) indicano il livello di un blocco di testo nella struttura di documento: in altre parole è come se ogni Titolo1 che applichiamo nel nostro documento aprisse una section 1, un Titolo2 una section 2 e così via. Gli elementi `<para>`, invece, marcano il testo dei singoli paragrafi. Infine, con l'elemento `<emphasis>` si individua una porzione di testo che deve assumere una particolare rilevanza all'interno di un paragrafo.

Attenzione *Osservate ancora una volta che visto che stiamo utilizzando un linguaggio di marcatura tutti gli elementi si limitano a segnalare il valore semantico e strutturale del contenuto che indicano, e non forniscono alcuna informazione riguardo alla sua resa grafica o stilistica.* `<emphasis>` *individua solo una porzione di testo rilevante all'interno di un paragrafo, ma non fornisce alcuna indicazione su come questa debba venire rappresentata.*

Anche con questi pochi elementi possiamo provare a accennare l'aspetto del capitolo che state leggendo se fosse scritto in DocBook:

```
<book>
  <chapter>




    <sect1>

        <!--Altri elementi para-->
      <sect2>
        <title>Il foglio
        stile</title>
        <!--Altri elementi para-->
      </sect2>
    </sect1>
  </chapter>
</book>
```

In corrispondenza del primo titolo di primo livello si apre un elemento `<sect1>`, all'interno del quale vanno annidate tutte le sezioni di livello inferiore, come accade alla

`<sect2>` poco oltre. Ogni sezione ha il proprio elemento `<title>`.

Naturalmente l'enorme gamma di opportunità offerte da DocBook non può essere esaurita in questo paragrafo. Una volta osservato come la logica della marcatura sia del tutto simile a quella di altri linguaggi come XML e HTML, però, ci sarà più semplice rintracciare gli elementi che fanno al caso nostro per la codifica del documento che intendiamo avere in questo formato.

Prima di proseguire indichiamo alcuni riferimenti fondamentali per lavorare con DocBook, tutti disponibili liberamente sul Web.

- All'indirizzo http://www.docbook.org trovate il manuale *DocBook: The Definite Guide*, scritto da Norman Walsh e Leonard

Muellner per O' Reilly. Il sito non si limita a diffondere il contenuto della guida, ma costituisce una vera e propria risorsa di riferimento imprescindibile per chiunque intenda lavorare con DocBook.

- All'indirizzo http://docbook.sourceforge.net/ potete trovare un gran numero di risorse open source riguardo DocBook, ma soprattutto fogli di trasformazione XSL già sviluppati per passare da DocBook ad altri formati.

Esistono inoltre degli editor visuali per scrivere direttamente in DocBook: alcuni di questi software potrebbero essere interessanti da sperimentare, qualunque sia il metodo scelto per creare libri digitali.

- Serna (http://www.syntext.com/products/serna): è un editor XML open source e

multipiattaforma, di cui è fornita anche una versione completamente gratuita. Tra i vari standard XML supportati troviamo anche DocBook: lo consigliamo per l'immediatezza e la semplicità di utilizzo.

- <oXygen/> (http://www.oxygenxml.com): è un editor professionale per XML, con strumenti per la scrittura di documenti in DocBook molto avanzati, per i quali prevede anche delle funzioni di conversione ad altri formati. È un software a pagamento.

Riferimenti *Per un elenco più completo di editor XML in grado di lavorare anche con DocBook consigliamo di consultare la pagina http://wiki.docbook.org/topic/DocBookAuthoringTools.*

Creare un salvataggio personalizzato in DocBook

La possibilità di salvare un documento direttamente in DocBook è data in OpenOffice da un file contenuto nella cartella *xslt > docbook* che si trova tra i file installati insieme all'applicazione.

> **Suggerimento** *A seconda del sistema operativo che state utilizzando l'indirizzo per la cartella* xslt > docbook *potrebbe essere differente.*

Il file in questione è `sofftodocbookheadings.xsl`, ed è il foglio di trasformazione in grado salvare il documento strutturato con l'elaboratore di testo direttamente in DocBook. Osservando questo documento con attenzione possiamo individuare le

sezioni in cui viene specificato come trasformare gli stili applicati nel documento di Writer in marcatori DocBook. Nel paragrafo seguente vedremo alcune delle caratteristiche fondamentali dei fogli di trasformazione in maniera da riuscire a orientarci meglio nel codice del documento e per poter apportare alcune semplici modifiche necessarie per ottenere un salvataggio in DocBook dal nostro foglio stile personalizzato.

XSLT: a metà tra programmazione e marcatura

Attraverso l'*eXstensible Stylesheet Language for Trasformation* (XSLT) è possibile definire una serie di regole che – tramite l'interpretazione da parte di un processore – consentono di trasformare il contenuto di un documento XML in un altro formato.

Terminologia *Il* processore *è un programma in grado di applicare XSLT a XML, verificando la correttezza della sintassi di entrambi e permettendone la trasformazione.*

Alla base del linguaggio sta la corrispondenza tra la regola di trasformazione e l'elemento a cui essa viene applicata: in altre parole XSLT individua i nodi del documento interessati e applica loro la trasformazione. In questo paragrafo prenderemo in esame gli elementi principali del linguaggio analizzando le sezioni più importanti del documento `sofftodocbookheadings.xsl` che utilizzeremo come base per la nostra trasformazione personalizzata.

Teniamo presente che la trasformazione in DocBook di Writer parte dall'XML prodotto

automaticamente dall'editor stesso, l'OpenDocument. Per questo, l'XSLT che utilizzeremo sarà pensato appositamente per l'XML prodotto da Writer, con i tag da esso stabiliti. Questo passaggio è più chiaro se osserviamo l'esempio seguente, in cui viene mostrato frammento di codice OpenDocument (l'XML prodotto da Writer) e la relativa trasformazione in DocBook grazie all' XSLT che stiamo per utilizzare.

> **Attenzione** *L'esempio seguente si trova nelle prime righe del documento: è commentato, e costituisce una sorta di mini-tutorial alle trasformazioni successive.*

```
Example of OpenDocument content:
 <office:body>
  <text:h text:style-name="Heading 1"
   text:level="1">Heading 1</text:h>
```

```
  Content</text:p>
 <text:h text:style-name="Heading 2"
  text:level="2">Heading 2</text:h>

  Content</text:p>
 <office:body>
```

Example of DocBook content:

```
 <article lang="en-US">
  <sect1>
   <title>Heading 1</title>
    <para>Heading 1 Content</para>
     <sect2>
      <title>Heading 2</title>
       <para>Heading 2 Content</para>
     </sect2>
  </sect1>
 </article>
```

Il primo listato mostra la codifica OpenDocument di OpenOffice, la quale,

oltre a individuare elementi come titoli (`<text:h>`) e paragrafi (`<text:p>`), associa loro tramite l'attributo `text:style-name` il nome dello stile che quella porzione di testo ha applicata, oltre al livello che quel testo ha nella struttura attraverso l'attributo `text:level`.

Il secondo listato mostra come la stessa struttura viene trasformata in DocBook.

La prima cosa che possiamo osservare è che nell'esempio basato su DocBook la struttura del documento è cambiata: gli elementi hanno la possibilità di contenere al proprio interno altri elementi (`<sect1>` contiene `<title>`, `<para>` e `<sect2>`, per esempio), mentre nell'Open-Document questo non avviene.

Osserviamo adesso il valore dell'attributo `text:style-name`: nel file OpenDocument troviamo tra i doppi apici il nome degli stili che nel documento sono stati applicati a quella porzione di testo: `Heading 1`, `Heading 2` e `Text body`, che corrispondono rispettivamente a titoli di primo e secondo livello e al corpo del testo.

Se visualizzassimo un documento realizzato con il nostro foglio stile creato su Word attraverso l'OpenDocument di Writer, troveremmo come valori dell'attributo `text:style-name` il nome dello stile che abbiamo definito nel nostro modello per quella porzione di testo. Per esempio, nel caso di questo paragrafo, troveremmo `text:style-name="A-body"`.

Una volta compreso questo principio, dobbiamo essere in grado di individuare nel foglio di trasformazione XSL quelle sezioni che agiscono sugli stili applicati per trasformarli in marcatori DocBook. Alla luce di queste nuove osservazioni, e ai fini di realizzare un documento capace di sfruttare appieno le numerose opportunità offerte dalla marcatura, consigliamo di ideare il foglio stile Word con un occhio di attenzione alle possibilità semantiche di DocBook: la conoscenza del linguaggio, infatti, ci permetterà di avere già chiara la corrispondenza tra stili e marcatori, e di utilizzare i due strumenti in sinergia e al pieno del loro potenziale.

Riferimenti *Per avere un'idea ancora più precisa riguardo quanto suggerito, consigliamo di consultare una tabella riassuntiva proposta da O'Reilly*

all'indirizzo http://oreilly.com/oreilly/author/stylesheet.html#typography_and_font_conventions: potrete vedere come per ogni elemento del documento venga proposta una specifica resa grafica e il relativo marcatore DocBook. Naturalmente, data la natura dei libri pubblicati da O'Reilly, si tratta spesso di stili che rispecchiano esigenze tecniche; tuttavia risulta utile per osservare quanto può essere sfaccettata la struttura di un documento prodotto dalla combinazione di un foglio stile e DocBook.

Iniziamo adesso a entrare nel vivo del nostro foglio di trasformazione `softtodocbookheadings.xsl`. L'elemento radice di un

foglio di trasformazione XSL è `<xsl:stylesheet>`.

Subito dopo, grazie all'elemento `<xsl:output>` e al valore del suo attributo `method`, indichiamo che il documento risultante dalla trasformazione sarà un documento XML:

Suggerimento *L'attributo* `indent`, *con valore positivo, stabilisce che l'output sarà indentato rispettando la struttura gerarchica del documento.*

Come abbiamo già osservato, nel passaggio dall'OpenDocument al DocBook occorre riordinare la struttura annidando nel modo corretto gli elementi. La parte di codice che segue il mini-tutorial di cui abbiamo parlato

si occupa proprio di compiere questa trasformazione: non ci occorre modificarla, per questo non la vedremo nel dettaglio.

Attenzione *Si tratta di una trasformazione complessa che si occupa innanzitutto di individuare i titoli contenuti del documento, il loro livello e gli elementi che devono essere annidati al loro interno per generare correttamente la struttura delle sezioni su DocBook. Una volta individuati, si occupa di associare a ciascuno la sezione giusta in base al livello e di annidarvi correttamente all'interno gli elementi figli.*

La parte che ci interessa, in cui possiamo modificare le regole che riguardano la trasformazione degli stili applicati nei corretti

elementi DocBook, inizia intorno alla riga 655 del documento, dove troviamo la prima regola, relativa ai titoli di primo livello:

```
  <xsl:element name="title">
    <xsl:apply-templates/>
  </xsl:element>
</xsl:template>
```

Terminologia *I fogli di trasformazione sono composti da* regole, *cioè porzioni di codice che descrivono la trasformazione da applicare ai vari nodi del documento di origine.*

L'elemento `<xsl:template>` crea una regola che produce un modello: ogni volta che si verificano le circostanze espresse genera l'output che abbiamo stabilito. L'attributo match individua, grazie al suo valore, il

sul quale la regola deve intervenire. In questo caso, quindi, l'elemento dell'OpenDocument su cui si deve intervenire è `<text:p>`, e più precisamente *ogni elemento* `<text:p>` *che ha nell'attributo* `text:style-name` *il valore* `Section1 Title`.

> **Attenzione** *In questo caso dobbiamo ricordarci che il foglio di trasformazione è già intervenuto per modificare il valore* `Heading 1` *dell'attributo* `text:style-name` *in modo da generare e annidare correttamente le sezioni richieste da DocBook. Per questo, là dove ci aspetteremmo di trovare* `Heading 1`, *ossia il nome dello stile applicato ai titoli di primo livello, nella regola di trasformazione per creare il*

tag corretto del titolo troviamo invece `Section1 Title`.

L'elemento `<xsl:element>` indica che una volta rintracciato il nodo in cui si deve intervenire nel documento di origine, nel documento risultante si apre un marcatore `<title>`, il cui nome è stabilito dal valore dell'attributo `name`. Con l'elemento `<xsl:apply-templates/>` il processore può applicare eventuali altre trasformazioni necessarie al contenuto dell'elemento processato (per esempio l'elaborazione di un tag `<emphasis>` o di una nota). Quindi la regola si chiude con i marcatori `</xsl:element>` e `</xsl:template>`.

Per definire le nuove regole necessarie per la trasformazione degli stili personalizzati che abbiamo definito nel nostro foglio stile,

possiamo procedere sulla falsa riga di questo esempio. È consigliabile posizionarle in questa sezione del documento, in maniera da mantenere ordine nel foglio di trasformazione, e sarebbe buona norma inserire anche dei commenti appropriati, in maniera da potersi orientare nel codice anche a distanza di tempo dalla sua realizzazione. Vediamo come procedere per creare una regola per uno stile personalizzato.

Consideriamo per esempio lo stile creato per il paragrafo base, nel caso di questo libro A-Body; il marcatore corrispondente in DocBook è `<para>`. Pertanto, la regola che andremo a costruire sarà la seguente:

```
<xsl:element name="para">
  <xsl:apply-templates/>
```

```
  </xsl:element>
</xsl:template>
```

Come abbiamo visto nell'esempio precedente, attraverso il valore dell'attributo match dell'elemento `<xsl:template>`, ci posizioniamo nel nodo che ha per valore dell'attributo `text:style-name` il valore `A-Body`; successivamente apriamo con `<xsl:element>` un elemento (il nostro obiettivo per la trasformazione in DocBook), applichiamo le modifiche con `<xsl:apply-templates/>`, chiudiamo il marcatore e infine la regola.

La sezione che riguarda gli stili di carattere comincia alla riga 981. In questo caso, poiché gli stili di carattere sono contenuti nell'OpenDocument invariabilmente nell'elemento `<text:span>`, abbiamo

bisogno di individuare tale elemento e leggere di volta in volta il valore dell'attributo `text:style-name` per attribuire a ciascuno stile il marcatore DocBook appropriato. A questo scopo ci serviremo dell'elemento `<xsl:choose>` per scegliere il valore dell'attributo, e dell'elemento `<xsl:when>` per stabilire cosa fare a seconda del valore trovato. Vediamo un esempio:

```
<xsl:template match="text:span">
 <xsl:choose>

    <xsl:element name="emphasis">
     <xsl:value-of select="."/>
    </xsl:element>
   </xsl:when>
   <xsl:when

    <xsl:element name="code">
```

```
      <xsl:apply-templates/>
     </xsl:element>
    </xsl:when>
  </xsl:choose>
</xsl:template>
```

Dopo aver individuato l'elemento `<text:span>`, con `<xsl:choose>` scegliamo cosa fare a seconda del valore (`<xsl:when>`) che ritroviamo nell'attributo `text:style-name`. *Quando* l'esito della ricerca del valore `Emphasis` è positivo, *allora* apriamo l'elemento Doc-Book `<emphasis>`, inseriamo al suo interno l'esatto contenuto ritrovato (`select="."`) con l'elemento `<xsl:value-of>`, chiudiamo `<emphasis>` e la condizione appena stabilita (`</xsl:when>`).

Suggerimento *Emphasis è il nome attribuito allo stile di carattere corsivo tra gli stili predefiniti sia di Word sia di Writer. In DocBook,* `<emphasis>` *individua un termine enfatizzato in maniera generica: il tipo di enfasi viene di volta in volta definito dal valore dell'attributo* `role`*. Per esempio, il grassetto è indicato da* `<emphasis role="strong">`.

La seconda parte dell'esempio ci mostra la regola per uno stile di carattere personalizzato. Quando l'esito della ricerca dello stile `A-CodiceCarattere` dà esito positivo, allora si apre l'elemento DocBook `<code>`, si applicano le modifiche con `<xsl:apply-templates/>`, si chiude l'elemento e la condizione. Una volta stabilite tutte le

trasformazioni da effettuare sugli stili di carattere si chiude anche l'elemento `<xsl:choose>` e infine la regola stessa.

Queste indicazioni non pretendono certamente di esaurire tutte le questioni poste da una trasformazione accurata dell'XML prodotto dal foglio stile. In queste poche pagine ci è sembrato utile fornire una serie di indicazioni di massima per orientarsi nell'XSLT e proporre spunti di approfondimento per creare le proprie regole di trasformazione personalizzate a seconda delle esigenze dei libri digitali che si devono produrre.

Per concludere questa parte non ci rimane che affrontare la gestione di elementi complessi come tabelle e immagini, che comportano diverse difficoltà sia dal punto

di vista del foglio stile sia della successiva trasformazione in DocBook. Facciamo un passo indietro, dunque, e vediamo per ciascuna le procedure più efficaci per integrarle nel nostro modello di documento e in seguito come riuscire a trasformarle correttamente nel linguaggio di marcatura.

Elementi complessi: tabelle, immagini ed elenchi

Tabelle e immagini presentano difficoltà di gestione su più livelli. Per prevederle nel foglio stile occorre tenere costantemente presente non solo le esigenze del modello ma allo stesso tempo anche il risultato da ottenere sia nell'XML sia per la stampa. Se questo discorso vale per tutti gli altri elementi che compongono il testo, per tabelle e

immagini le complicazioni sono maggiori e in certi casi di non semplice soluzione. Vediamo di accennare ad alcuni metodi possibili per trattare questi oggetti.

Tabelle

Inserire una tabella nel nostro foglio stile in Word non è di per sé una procedura troppo complicata. Dal menu *Tabella* nella barra dei menu selezioniamo la voce *Inserisci > Tabella* e inseriamo le caratteristiche desiderate nei campi della finestra *Inserisci tabella* così richiamata (Figura 4.10).

Figura 4.10 La finestra Inserisci tabella.

Tuttavia, una tabella inserita in questo modo non soddisferà quasi sicuramente le nostre esigenze, specialmente dal punto di vista del foglio stile. Possiamo vedere che alla voce *Stile tabella*, di default, il valore selezionato è *Griglia tabella*. Per modificare tale valore in maniera che risponda alle esigenze del foglio stile dobbiamo prima

definire uno stile personalizzato per la tabella, così come abbiamo fatto per gli stili di paragrafo e carattere.

Sempre dal riquadro *Stili e formattazione* facciamo clic sul pulsante *Nuovo stile* e dal menu a discesa *Tipo* selezioniamo la voce *Tabella* (Figura 4.11).

Figura 4.11 La finestra Nuovo stile per creare uno stile di tabella.

Suggerimento *Ricordiamoci di specificare sempre un nome appropriato per il nuovo stile che stiamo creando!*

A questo punto dobbiamo fare attenzione al menu a discesa *Applica formattazione a*, all'interno del quale possiamo scegliere le parti della tabella a cui applicare le definizioni di stile che ci interessano. Tutte le caratteristiche che definiremo per righe, celle, colonne o per l'intera tabella saranno poi racchiuse e salvate nello stile indicato dal nome che abbiamo scelto.

Suggerimento *Definire uno stile di tabella predefinito ci permetterà inoltre di evitare che l'autore del testo si trovi incerto sul come formattare la tabella che intende inserire, evitando che scelga altri stili che non intendiamo utilizzare, ma lasciandolo comunque libero di definire il numero di righe e colonne.*

A questo punto torniamo alla finestra *Inserisci tabella* (Figura 4.10) e qui facciamo clic sul pulsante *Formattazione automatica*. Selezioniamo dall'area *Stili tabella* lo stile che abbiamo appena creato e facciamo clic su *OK*. Nella finestra *Inserisci tabella* possiamo adesso vedere che al posto del valore predefinito di *Stile tabella* compare il nome che abbiamo scelto per lo stile di tabella creato: non ci rimane che spuntare la casella *Imposta come predefinito per le nuove tabelle* così che la scelta ricada sempre sullo stile che abbiamo deciso di applicare.

Probabilmente vorremo che le tabelle abbiano una didascalia: consigliamo a questo proposito di utilizzare l'inserimento automatico di Word che permette di automatizzare la numerazione ed è ben riconosciuto nel salvataggio in DocBook.

Per inserire una didascalia ci sono diversi modi, e conducono tutti alla finestra *Didascalia* (Figura 4.12), nella quale dobbiamo selezionare un'etichetta appropriata ("tabella", "figura" o un'etichetta personalizzata facendo clic sul pulsante *Nuova etichetta*) e scegliere le caratteristiche che preferiamo per la numerazione e il posizionamento.

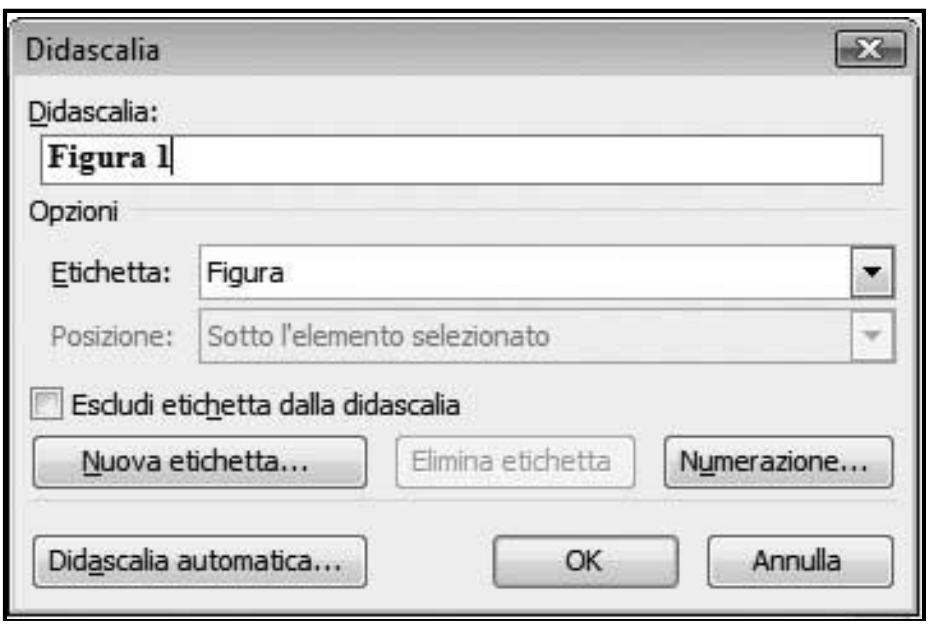

Figura 4.12 La finestra Didascalia.

Suggerimento *Per inserire una didascalia possiamo fare clic sulla voce*

Inserisci > Didascalia *nella barra dei menu, oppure sulla voce* Didascalia *che compare nel menu contestuale con un clic destro sull'oggetto a cui vogliamo assegnarla.*

Suggerimento *È possibile anche definire l'inserimento di didascalie automatiche (evitando quindi di doverle inserire una per una ogni volta che creiamo un elemento che ne richiede la presenza) facendo clic sul pulsante* Didascalia automatica*: può essere utile se sappiamo già che nel nostro documento, per esempio, non ci saranno mai tabelle senza didascalia.*

Una volta creato lo stile per le tabelle, torniamo al foglio di trasformazione, e osserviamo in che modo gestisce le tabelle. La sezione che ci interessa inizia alla riga 697

del codice, e data la sua lunghezza non la commenteremo per intero. In breve, ci occorre sapere che – una volta individuata la tabella grazie all'elemento `<table:table>` che la marca – la trasformazione si occupa di aprire i marcatori necessari per il DocBook, di contare celle e righe per definirne la struttura e di individuare l'eventuale presenza di un'intestazione. Si occupa inoltre di rintracciare gli stili di paragrafo, di carattere o di elenco presenti all'interno della tabella e di trasformarli correttamente in base alle regole per essi definite.

Attenzione *Se abbiamo deciso di inserire una didascalia per la tabella, dobbiamo aggiungere una regola in grado di individuarla e trasformarla correttamente in DocBook: l'elemento DocBook che dobbiamo creare è*

`<caption>`. *La complessità della regola non ci permette di descriverla in poche righe.*

Questo breve accenno alle tabelle lascia purtroppo molte difficoltà in sospeso. Una di queste, per esempio, è la gestione di tabelle orientate orizzontalmente nella pagina: un classico caso in cui le necessità della resa grafica e della collocazione degli oggetti nelle dimensioni della pagina fisica creano problemi vincolanti per la definizione del foglio stile (che deve continuare a utilizzare una pagina a dimensioni fisse) e delle sue successive trasformazioni.

Immagini

La gestione delle immagini è forse l'aspetto più complesso di tutto il processo. Nelle varie trasformazioni di formato, infatti, tenere

traccia della corretta posizione delle figure e conseguentemente esportarle nel modo giusto richiede molta attenzione e revisioni accurate difficilmente automatizzabili. Vediamo quindi, a partire dal foglio stile, qual è il metodo più efficace per ottenere più informazioni possibili e ridurre il numero di errori e ambiguità.

In un documento prodotto con un elaboratore di testo i metodi per inserire un'immagine sono principalmente due:

- inserirla direttamente nel documento;
- collegarla al documento.

Il primo caso non ci è utile per una varietà di motivi:

- nel caso di documenti lunghi e con molte immagini, il peso del file aumenta eccessivamente;
- quasi sempre le immagini devono subire ulteriori lavorazioni, sia per la stampa sia per i formati digitali: averle inserite nel documento rende più complicate tali modifiche;
- l'immagine inserita direttamente nel file non viene esportata correttamente in DocBook, perdendo ogni riferimento necessario per la produzione di un libro digitale (indirizzo dell'immagine, formato e così via).

Per questo, la maniera più efficace di trattare le immagini è quella di collegarle al documento: sarà così possibile localizzarle tramite il loro indirizzo e farne l'uso più opportuno a seconda del formato che si intende ottenere.

Per collegare un'immagine al documento su cui si sta lavorando occorre fare clic sulla voce *Inserisci* > *Immagine* > *Da file* nella barra dei menu. Nella finestra successiva, dopo aver selezionato il file da inserire, fare clic sul pulsante *Inserisc*i e scegliere la voce *Collega al file*, come si vede nella Figura 4.13.

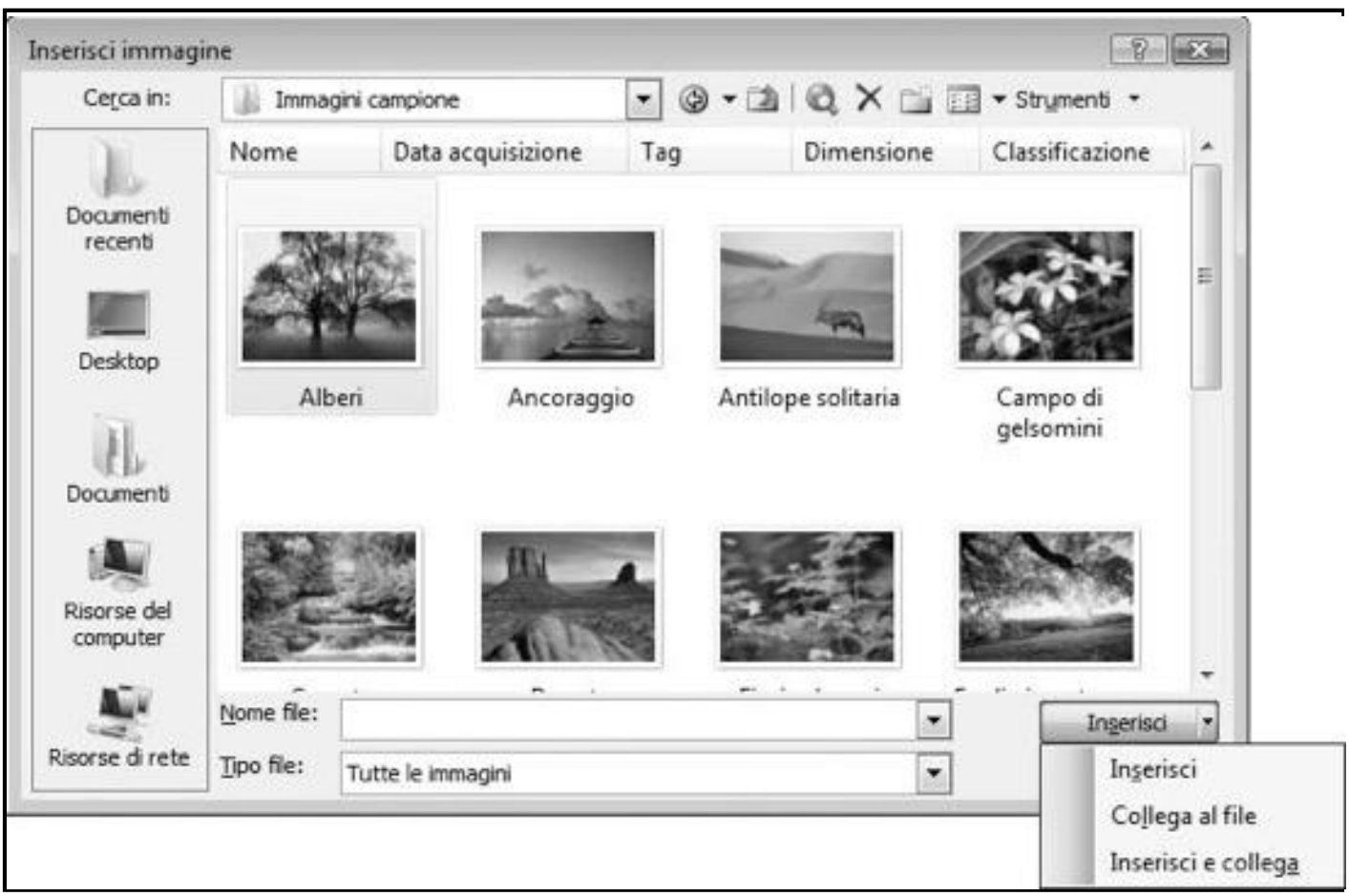

Figura 4.13 In basso a destra: il menu del pulsante Inserisci che permette di collegare l'immagine al documento.

Inserendo l'immagine in questo modo, otterremo nell'XML il percorso dell'immagine inserito nell'elemento `<draw:image>` come valore dell'attributo `fileref`. Se invece l'immagine fosse semplicemente inserita nel documento, avremmo nell'XML l'informazione della presenza dell'immagine in una modalità inutile ai fini di qualunque trasformazione, poiché perderebbe ogni riferimento riguardo alla sua posizione.

Suggerimento *È importante creare una cartella in cui salvare tutte le immagini collegate al documento su cui si sta lavorando.*

La procedura di inserimento delle immagini è quindi semplice, ma in tutto il percorso di lavorazione sarà sempre necessario prevedere dei controlli mirati, in particolare per verificare che passando da un formato all'altro l'immagine (e il relativo elemento XML) conservi la sua esatta collocazione.

DocBook è in grado di gestire le immagini in maniera molto precisa e dettagliata, strutturando al massimo le informazioni che esse veicolano e permettendo un alto grado di definizione di ogni aspetto riguardante l'immagine. Vediamo alcune possibilità offerte dai suoi marcatori per le immagini.

Riferimenti *Potete trovare l'esempio seguente all'indirizzo http://www.docbook.org/tdg5/en/html/imageobject.html, insieme a molte altre*

informazioni riguardo tutti i marcatori utilizzati e DocBook in generale.

```
<title>Example imageobject</title>
 <mediaobject>
  <imageobject>

    format=”EPS”/>
  </imageobject>
  <imageobject>

    format=”PNG”/>
  </imageobject>
  <textobject>
   <phrase>The Eiffel Tower</phrase>
  </textobject>
  <caption>



  </caption>
```

```
  </mediaobject>
</article>
```

Nell'esempio l'elemento radice è `<article>`; segue il `<title>` dell'articolo.

Riferimenti *Oltre ai libri, identificati dall'elemento* `<book>`*, DocBook è pensato anche per diversi tipi di pubblicazioni, come appunto gli articoli (*`<article>`*).*

L'elemento `<mediaobject>` individua ogni tipo di oggetto multimediale inserito nel documento (audio, video, immagini...). Al suo interno troviamo `<imageobject>`, ossia il contenitore per tutte le informazioni relative all'immagine che intendiamo marcare. L'elemento `<imagedata>` contiene le informazioni che puntano al file immagine esterno al documento: leggiamo infatti

nell'attributo `fileref` l'indirizzo dell'immagine che va collocata in questo punto del documento.

> **Suggerimento** `<imagedata>` *è l'elemento che dovrà contenere nell'attributo* `fileref` *l'indirizzo che otteniamo tramite il collegamento di un'immagine al documento scritto con il foglio stile.*

L'elemento `<textobject>` contiene delle metainformazioni relative all'immagine, non visualizzate nel documento. L'elemento `<caption>`, infine, contiene il testo della didascalia dell'oggetto.

Elenchi

Un breve accenno, infine, ad alcuni aspetti a cui prestare attenzione nella definizione degli stili per gli elenchi puntati e numerati.

Le difficoltà proposte dagli elenchi derivano dalla loro struttura in livelli e sono perciò simili a quelle dei titoli: le voci possono contenere al loro interno delle sottovoci, e l'annidamento crea delle complessità di trasformazione.

La sezione del foglio di trasformazione che riguarda gli elenchi non richiede particolari interventi, a patto che nel foglio stile la struttura dell'elenco sia individuata in modo corretto dagli stili che abbiamo definito, e soprattutto che sia stata applicata nel modo giusto: in caso contrario non saremo in grado di rintracciare le informazioni necessarie per ricostruirla in DocBook.

Suggerimento *Se vogliamo dare un'occhiata alla parte di codice che riguarda la gestione degli elenchi nel foglio di trasformazione dobbiamo*

osservare il documento a partire dalla riga 819, dove si individua l'elemento `<text:ordered-list>`*, corrispondente all'elenco numerato. Poco oltre (riga 835) troviamo invece l'elenco puntato, individuato nell'elemento* `<text:unordered-list>`*.*

Aggiungere il foglio di trasformazione personalizzato in OpenOffice

Vi consigliamo di fare una copia di `sofftodocbookheadings.xsl` e di rinominarlo per distinguere il foglio di trasformazione con le vostre modifiche senza alterare quello nativo di OpenOffice. Sarete così liberi di testare e sperimentare tutte le regole che vi sembrano necessarie per la trasformazione personalizzata senza

rischiare di perdere il documento di partenza.

Mentre create le nuove regole avrete certamente la necessità di testarne i risultati. Per poterlo fare, e anche per poter trasformare definitivamente i vostri documenti in DocBook, dovete aggiungere il vostro foglio di trasformazione a quelli presenti in OpenOffice. Vediamo come.

- Selezionate la voce*Impostazioni filtro XML* dal menu *Strumenti* e nella relativa finestra fate clic sul pulsante *Nuovo*.
- Nella scheda*Generale* della finestra successiva, inserite il nome del vostro foglio di trasformazione nel campo *Nome del filtro*, selezionate *OpenOffice.org Writer (.sxw)* dal menu *Applicazione* e il nome che volete

che compaia per il salvataggio nel menu *Tipo file* nel campo *Nome per tipo di file*.

Attenzione *Quest'ultimo parametro verrà utilizzato e proposto da OpenOffice nell'elenco dei formati di salvataggio possibili, consentendovi così di scegliere di salvare un documento nel vostro formato DocBook personalizzato.*

- Nella scheda *Trasformazione* aggiungete il foglio di trasformazione personalizzato alla voce *XSLT per esportazione* e fate clic su *OK*.

A questo punto potrete vedere nella finestra *Impostazione filtri XML* il nome che avete scelto per la vostra trasformazione personalizzata. Da qui, con un clic sul pulsante *Test XSLTs* potete testare l'output

restituito dalla trasformazione selezionata sul documento sul quale state lavorando.

Il risultato

Dopo aver salvato il documento prodotto con l'elaboratore di testo nel formato DocBook modificato *ad hoc*, sarà necessario aprirlo con un software a vostra scelta in grado di evidenziare la sintassi XML, in maniera da poter effettuare i controlli necessari sulla struttura del documento con semplicità e immediatezza. Se le fasi di test nella stesura delle regole sono state accurate e se il documento rispetta le caratteristiche definite nel foglio stile, avrete pochi interventi da effettuare e il documento DocBook sarà pronto per essere utilizzato per tutte le trasformazioni successive.

È bene compiere le verifiche sulla struttura con la massima precisione in questa fase: se infatti il documento in DocBook è corretto, le successive trasformazioni non daranno problemi. Se ci accorgeremo di eventuali errori nella struttura dopo la seconda trasformazione dovremo comunque riprendere il processo a partire dalla generazione del DocBook: per questo è preferibile essere il più precisi possibile a questo punto della lavorazione.

Seconda trasformazione: verso il formato digitale

Abbiamo concluso la prima parte del nostro percorso: il contenuto, prodotto attraverso il nostro modello .dot è stato salvato in

DocBook ed è quindi pronto per essere trasformato ulteriormente in tutti i formati che ci occorre ottenere sia per la stampa sia per il digitale. Per fare questo dovremo proseguire in maniera analoga a quanto già fatto per il salvataggio in DocBook personalizzato e utilizzare dei fogli di trasformazione appositi a seconda del formato di destinazione.

Questo tipo di trasformazione, tuttavia, è molto più complessa rispetto a quella vista finora. Fortunatamente, la grande versatilità e diffusione di DocBook fanno sì che esistano liberamente disponibili in Rete un gran numero di fogli di trasformazione già pronti da utilizzare per ottenere dai propri documenti in XML i formati di file di cui si ha bisogno.

Una risorsa molto utile a questo scopo è il sito http://docbook.sourceforge.net/, nel quale è possibile trovare fogli di trasformazione verso tutti i principali formati, tra cui ePub, mobi e HTML.

Un discorso differente va invece fatto per il PDF. Per ottenere la trasformazione verso questo formato, infatti, XSLT non sarà sufficiente: dovremo utilizzare l'XSL-FO, a cui abbiamo già accennato nel Capitolo 2.

Riferimenti *XSL-FO è un linguaggio di markup che consente di formattare i dati contenuti in un documento XML per la conversione in formati particolari, tra cui il PDF. XSL-FO è parte di XSL, che indica le tre distinte raccomandazioni del W3C riguardo la formattazione e trasformazione dei dati XML: XSLT per la trasformazione,*

XSL-FO per la formattazione e XPath per la navigazione nel documento XML.

Ma perché XSLT non è sufficiente per trasformare DocBook in un documento in formato PDF? XSLT può individuare dei nodi nell'XML e *trasformarli* nel linguaggio del formato di destinazione: ePub si basa su XML, il mobi su XHTML e per questo sono compatibili con questo tipo di trasformazione. Tuttavia, in questo processo non è prevista alcuna informazione riguardo alla formattazione del documento, alla sua resa grafica: infatti sarà necessario creare a parte un foglio stile CSS per stabilire come visualizzare il formato ottenuto, sia nel caso dell'ePub sia del mobi, che essendo basati su linguaggi di marcatura non hanno bisogno di informazioni grafiche. Il PDF, invece, non

può essere creato senza stabilire tali informazioni a priori: ecco perché è necessario utilizzare XSL-FO. Attraverso questo linguaggio, infatti, è possibile individuare i nodi dell'XML su cui intervenire (XPath), trasformarli (XSLT) e formattarli (XSL-FO) secondo le proprie esigenze.

Terminologia *Ricordiamo che XSL-FO è l'acronimo di* eXtensible Stylesheet Language Formatting Objects.

XSL-FO prevede elementi e attributi per la definizione di informazioni tipografiche e di impaginazione, con alla base il concetto di area: un documento XSL-FO definisce una serie di aree tipografiche in cui inserire i contenuti secondo stili di rappresentazione personalizzabili. Quando il parser interpreta

“l’albero delle aree”, restituisce in output quanto si è determinato.

Approfondire le basi del linguaggio XSL-FO, così come le trasformazioni necessarie per ottenere da DocBook ePub e mobi, è un argomento che per difficoltà e spazio non può rientrare in queste pagine. Per questo segnaliamo una serie di risorse che potrebbero essere utili a chi intendesse cominciare a lavorare con questi linguaggi.

In Rete:

- http://www.w3schools.com/xslfo/default.asp: il tutorial su XSL-FO su w3schools.com (sullo stesso sito anche tutorial su XML, XSLT e molti altri linguaggi citati in queste pagine);

- http://www.tizag.com/xmlTutorial/: guide di base su XML e linguaggi connessi, come XPath e XSLT;
- http://wiki.docbook.org/topic/DocBookStylesheets: sul wiki ufficiale di docbook.org, tutte le risorse relative ai fogli di trasformazione XSL.

Su carta:

- Michiel van Otegem,*XSLT Guida Completa*, Apogeo, 2002;
- Doug Tidwell,*XSLT*, O'Reilly Media, 2001;
- Dave Pawson, *XSL-FO Making XML Look Good in Print*, O'Reilly Media, 2002.

Altri metodi

Come abbiamo premesso all'inizio di questo capitolo, la via descritta fin qui e che abbiamo indicato come nostra preferita non è certamente l'unica percorribile per ottenere i formati adatti per il libro digitale. Vediamo alcuni altri metodi che prevedono l'utilizzo di software differenti e percorsi diversi di conversione dei formati.

Adobe InDesign

InDesign è l'applicazione di Adobe dedicata all'impaginazione: grazie a essa si può curare la corretta messa in pagina di testi, immagini, tabelle e tutto ciò che costituisce un testo in maniera molto precisa e

professionale, con un grande controllo di tutti gli elementi necessari per la stampa e non solo.

Riferimenti *Nel momento in cui si scrivono queste pagine, l'ultima versione disponibile di InDesign è la CS4. Si tratta di un software a pagamento; per testarlo è possibile scaricare la prova gratuita valida per trenta giorni direttamente dal sito di Adobe.*

Fin dalla versione CS2, infatti, InDesign è in grado di esportare i contenuti in XML: questo perché tale formato può essere trasformato e riutilizzato facilmente per la produzione e diffusione di testi su vari supporti e piattaforme. Nelle successive versioni del software questa capacità è stata perfezionata e i miglioramenti continuano a

maturare in maniera parallela rispetto alle evoluzioni dei formati per i libri digitali.

L'esperienza di Adobe nel realizzare un reader efficace come Digital Editions (di cui abbiamo parlato nel capitolo precedente) ha fatto sì che già dalla versione CS3 fosse presente una voce nel menu *File* per esportare il documento in uso nel formato compatibile con il reader, *Esporta per Digital Editions*: in questo modo è possibile ottenere in maniera diretta e automatica un ePub del documento a cui si sta lavorando.

InDesign creerà un singolo file .epub contenente il documento esportato in XHTML e includendo una miniatura JPEG della prima pagina utilizzata da Digital Editions come copertina del file. L'ePub prodotto con InDesign in maniera automatica

permette un certo numero di personalizzazioni, accessibili dalla finestra di dialogo *Opzioni di esportazione Digital Editions*.

- *Generali*: consente di aggiungere al documento i metadati relativi, tra cui le informazioni sull'autore, l'editore e così via. È possibile anche specificare quale formattazione del testo vada inclusa nel foglio stile CSS dell'ePub, utilizzando come base gli stili applicati (anche in linea).
- *Immagini*: da questa sezione si può gestire il comportamento delle immagini durante l'esportazione. Possono essere incluse nell'ePub nel formato originale, ottimizzate definendone le caratteristiche o conservandone la formattazione applicata nel documento InDesign e convertite automaticamente in GIF o JPEG.

- *Contenuto*: permette di specificare alcune impostazioni relative alla visualizzazione su Digital Editions in modo specifico, come per esempio la gestione e la generazione del sommario.

Malgrado questa ricca premessa, la generazione automatica dell'ePub da parte dell'applicazione potrebbe non soddisfare del tutto le nostre esigenze. Vediamo quindi come ottenere un documento XML ben formato utilizzando l'impaginazione di InDesign.

Produrre un documento XML con InDesign

Per produrre un documento XML ben formato dal nostro impaginato, occorre impostare prima di tutto un foglio stile, simile a quello descritto per Word. In InDesign

dovremo anche occuparci di creare i tag che utilizzeremo per marcare il documento (vedremo tra poco come).

Attenzione *L'impaginazione del documento dal punto di vista grafico in InDesign può essere decisamente molto più accurata che in Word. Tuttavia, se intendiamo produrre XML, dobbiamo tener presente che il foglio stile che stiamo definendo dovrà essere versatile anche dal punto di vista del linguaggio di marcatura.*

Consideriamo un documento nel quale abbiamo creato gli stessi stili già descritti per lavorare in Word. Anche in questo caso, il nome che abbiamo attribuito loro sarà fondamentale: su InDesign serve a definire il

collegamento tra lo stile e il tag che deve essere applicato al testo.

Come in Word, anche in InDesign gli stili del testo si dividono principalmente in stili di carattere e di paragrafo (Figura 4.14). Tuttavia, a differenza di Word (che è in grado di gestire la struttura gerarchica del testo attraverso l'assegnazione del livello allo stile), InDesign non consente alcun intervento di questo tipo: per questo qualsiasi struttura XML generata avrà bisogno di essere ulteriormente lavorata per stabilire la corretta gerarchia tra gli stili di paragrafo assegnati. Lo stile in InDesign, insomma, definisce solamente una serie di caratteristiche grafiche; i valori semantici vanno applicati a posteriori.

Figura 4.14 I pannelli per gli stili di carattere e di paragrafo in Adobe InDesign CS3.

Per questo dobbiamo creare i tag che vogliamo compaiano nel nostro documento XML e associarli ai relativi stili permettendo così di marcare il testo correttamente. Dal menu *Finestra* selezioniamo la voce *Tag*, facendo comparire il pannello omonimo, dal quale possiamo creare e gestire i marcatori per la generazione dell'XML.

La creazione di un tag è molto semplice: nel pannello *Tag* occorre fare clic sul pulsante *Nuovo tag* e attribuirgli il nome desiderato, tenendo presenti le osservazioni già fatte riguardo ai nomi degli stili in Word nelle pagine precedenti. Dovremo creare un tag per ciascuno stile che desideriamo venga marcato nel documento XML. Alla fine il nostro pannello avrà un aspetto simile a quello nella Figura 4.15.

Figura 4.15 Il pannello per la gestione dei tag.

Suggerimento *InDesign attribuisce automaticamente a ogni tag un colore, che può essere modificato in base alle nostre esigenze. Questo permette poi di visualizzare immediatamente nel documento l'applicazione della marcatura.*

Una volta definiti tutti i tag necessari possiamo procedere con la loro mappatura sugli stili già presenti nel documento. Vediamo come.

- Aprite il menu del pannello Tag e selezionate la voce Mappa stili su tag (Figura 4.16).

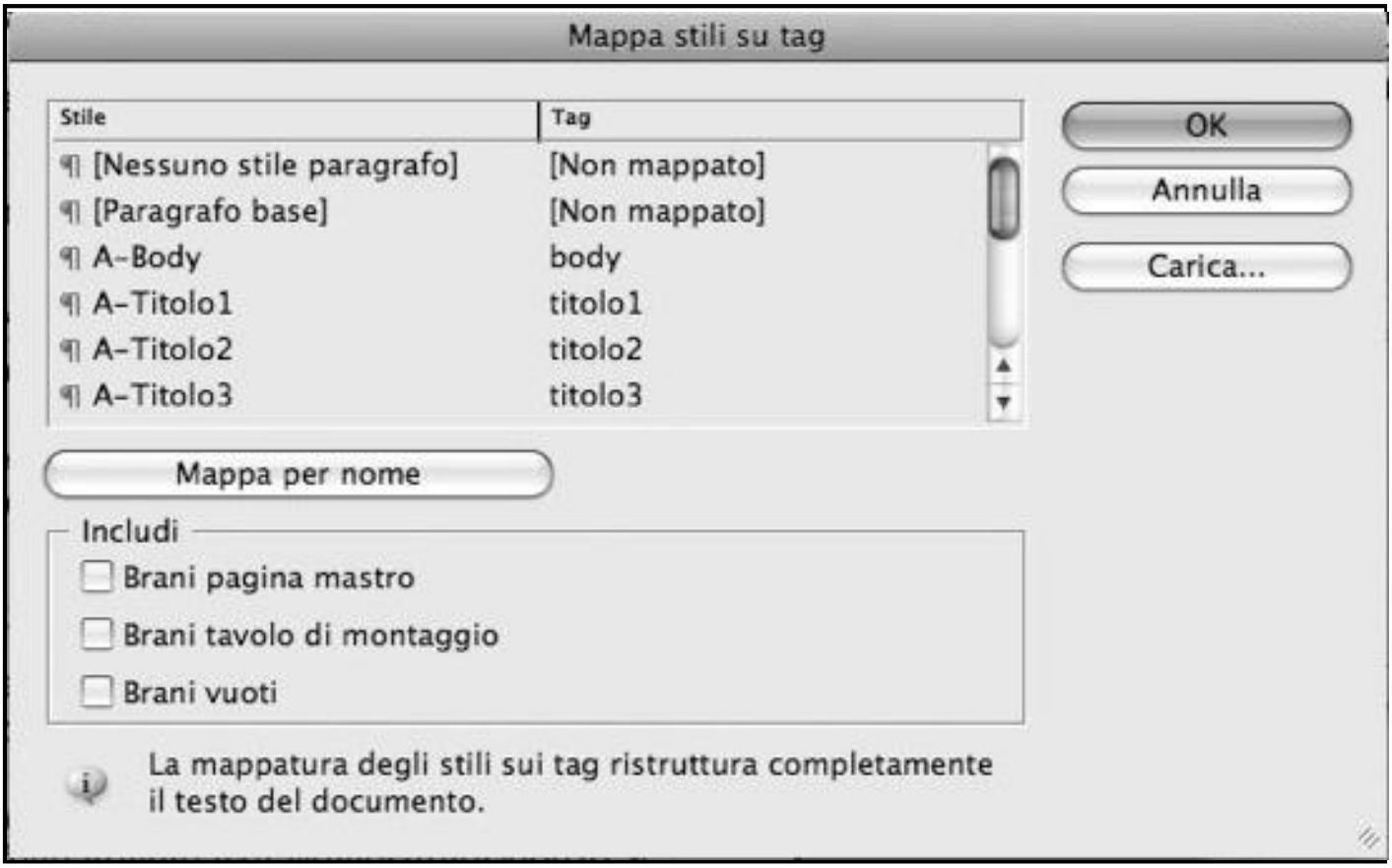

Figura 4.16 La finestra Mappa stili su tag con le relative corrispondenze assegnate.

- Nella finestra*Mappa stili su tag* selezionate per ciascuno stile che intendete esportare il tag che avete creato per marcarlo, facendo clic sulla voce che vi interessa nella colonna *Tag*.

- Una volta stabilite tutte le relazioni necessarie fate clic su *OK* per applicare la marcatura stabilita a tutto il documento.

  Suggerimento *Se i nomi degli stili e quelli dei tag coincidono, potete utilizzare il comando* Mappa per nome *e InDesign traccerà le relazioni corrette. Non sempre è però possibile mantenere questa coincidenza nella definizione del foglio stile.*

Per verificare che la marcatura del documento sia a questo punto andata a buon fine, avete a disposizione diversi strumenti. Dal menu *Visualizza > Struttura* selezionate la voce *Mostra struttura*: sulla sinistra dell'area di lavoro compare un pannello contenente l'albero del documento

determinato dalla nostra mappatura degli stili sui tag.

Come mostrato nella Figura 4.17, la struttura è racchiusa in un elemento *Root*, creato automaticamente da InDesign come radice del documento, contenente un elemento *Brano* applicato anch'esso automaticamente. Al suo interno troviamo i tag che abbiamo mappato in base agli stili applicati al testo. Possiamo inoltre visualizzare nel documento stesso l'applicazione della marcatura attraverso la voce *Mostra marcatori tag* dal menu *Visualizza > Struttura*: in questo modo il testo mappato su un determinato tag attraverso l'applicazione dello stile apparirà racchiuso tra parentesi del colore scelto per il tag relativo.

Figura 4.17 La mappatura degli stili sui tag è visibile sia nell'albero della struttura nel riquadro a sinistra, sia attraverso le parentesi quadre del colore al tag associato corrispondente.

Un'altra possibile verifica è data dalla visualizzazione con lo strumento Editor brani, accessibile sia dal menu *Visualizza > Editor* brani sia dal menu *Modifica > Modifica in*

Editor brani. La marcatura del documento è resa evidente dalla presenza dei tag che delimitano il testo a cui sono applicati (Figura 4.18).

Figura 4.18 La visualizzazione attraverso Editor brani.

Quando siamo soddisfatti della struttura ottenuta dobbiamo fare clic sul menu *File > Esporta* e selezionare il formato XML dal menu *Formato*. Nella finestra di esportazione (Figura 4.19) possiamo stabilire se applicare un foglio di trasformazione XSLT già pronto, per esempio per ottenere un DocBook, e decidere dall'apposita

sezione come dovranno essere gestite le immagini.

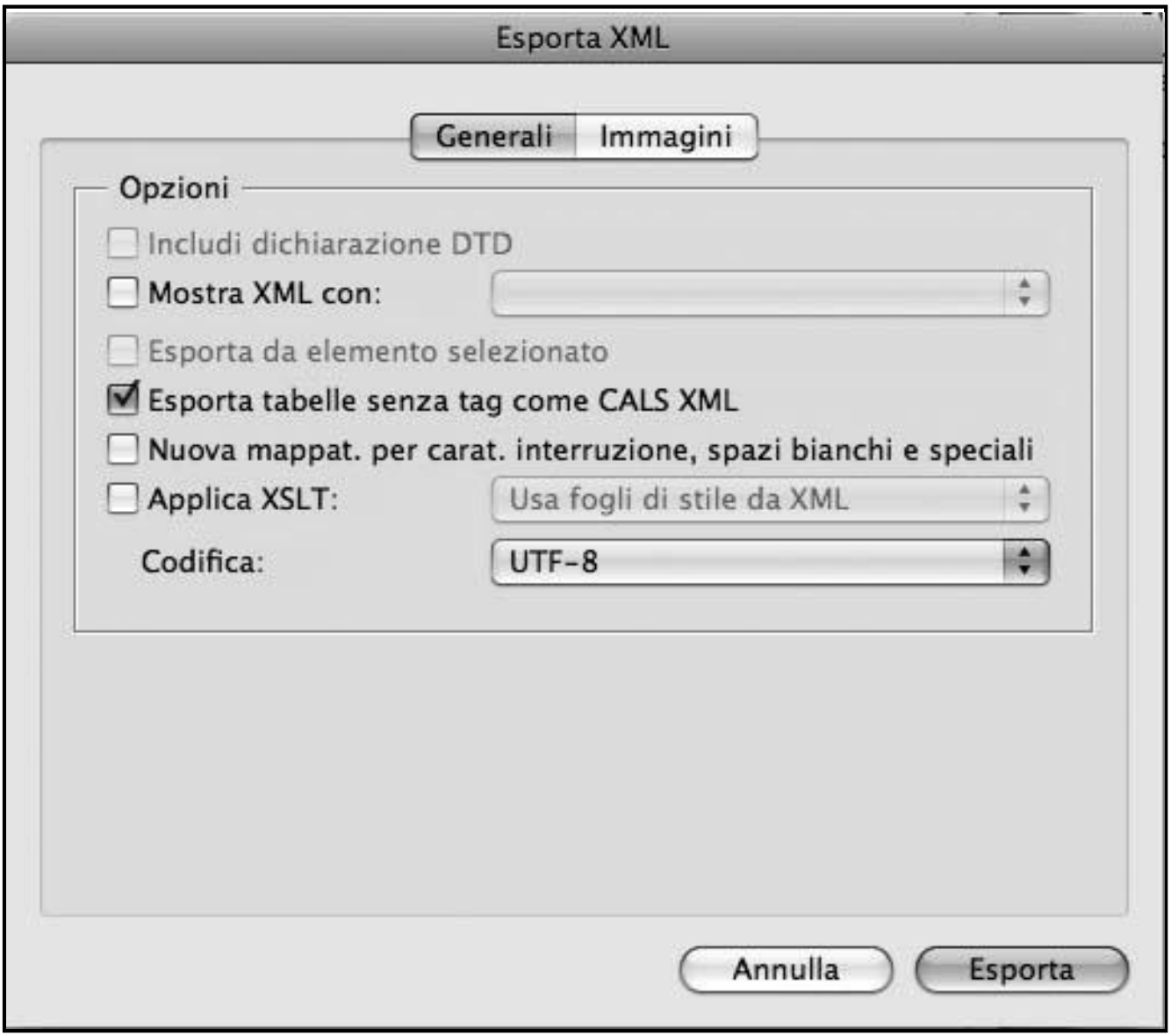

Figura 4.19 La finestra Esporta XML.

Considerazioni di fine percorso

In questo capitolo abbiamo cercato di proporre suggestioni e spunti da cui partire per scegliere la strada più appropriata per produrre libri digitali. Ciò che è importante comprendere è che qualunque sia il metodo che si sceglie di adottare, rimangono alcuni punti fermi.

- DocBook è la via più efficace per ottenere da un’unica codifica più formati da distribuire in maniere differenti.
- Per poter essere sfruttato al meglio DocBook richiede di prestare massima attenzione ai contenuti, dedicando più cura alla loro

strutturazione e alla valorizzazione della loro semantica.

In qualunque modo si decida di ottenere la versione XML dei documenti, è necessario modificare il metodo di lavoro in maniera da riuscire a concentrare l'attenzione su questi aspetti e lavorare sui libri con una diversa consapevolezza, a prescindere dalle tecnologie che si intendono impiegare.

È implicita nel discorso la necessità di acquisire le competenze tecniche necessarie almeno a supervisionare il processo, se non a gestirlo per intero. Ferma restando la possibilità di affidarsi a dei fornitori esterni che si occupino dell'utilizzo delle tecnologie specifiche al nostro posto, rimane comunque il bisogno di essere in grado di gestire i file e la loro struttura con il giusto grado di

consapevolezza e preparazione: la valorizzazione della semantica del testo e la gestione della sua struttura – al di là delle conoscenze tecniche che richiedono – sono compiti fondamentalmente editoriali e tali dovrebbero rimanere.

FACCIAMO ORDINE!

Veniamo adesso agli aspetti legati alle complessità non tecniche che i libri digitali pongono nel momento in cui un editore intende sistematizzarne la produzione. Il discorso non potrà che essere molto generale: ogni casa editrice ha le proprie necessità determinate dal settore in cui si colloca, dall'ampiezza del catalogo, dal numero delle novità che immette sul mercato, dal tipo di libri che produce... L'infinita varietà di casi non ci impedisce comunque di individuare alcuni spunti di riflessione che si dovrebbero

generalmente tenere presenti al momento di avviare la produzione sistematica di ebook.

La pubblicazione di libri digitali, inoltre, non riguarda solo aspetti tecnici e organizzativi, ma comporta anche precise scelte culturali. Occorre ragionare sulle diverse esigenze del digitale rispetto a quelle consolidate del libro cartaceo, e pianificare al meglio le strategie distributive.

Nei paragrafi seguenti affronteremo questi temi senza la pretesa di fornire delle soluzioni, ma cercando di mettere in luce degli aspetti su cui è necessario riflettere nel momento in cui si intendono inserire i libri digitali nella linea editoriale.

Sistematizzare la produzione

Gli aspetti coinvolti nell'avvio della produzione sistematica di ebook sono i più vari, da quello economico a quello organizzativo: se infatti la produzione di libri digitali non grava su voci di bilancio come distribuzione e spese di magazzino, richiede comunque un certo margine di investimento; dal punto di vista dell'organizzazione del flusso di lavoro, invece, potrebbe comportare alcuni cambiamenti la cui importanza varia a seconda del metodo scelto per produrre gli ebook e delle risorse (anche umane) che si intendono impiegare a questo scopo. Gran parte di queste considerazioni

dipende dalla struttura della casa editrice, dalle sue dimensioni, possibilità economiche e interesse nell'investire nei libri digitali.

Come abbiamo visto, i libri digitali possono essere ottenuti modificando in una certa misura la prassi di lavorazione del libro, scegliendo secondo le proprie esigenze su quali momenti della produzione impattare maggiormente. Vediamo alcuni pro e contro dell'introdurre certi cambiamenti nel processo produttivo.

Impatto maggiore sull'autore

Si potrebbe scegliere di lasciare all'autore il compito di lavorare sul suo manoscritto seguendo le norme stabilite nel foglio stile che l'editore ha creato per la produzione dei suoi libri. In questo modo arriverebbe

all’editore un manoscritto idealmente già pronto per le successive fasi redazionali, con una struttura ben definita e lavorabile sia come bozza per il libro cartaceo sia per quello digitale.

Tuttavia, l’utilizzo corretto del foglio stile richiede alcune competenze che non tutti gli autori possiedono. La potenza degli elaboratori di testo non viene impiegata che al minimo e in maniera superficiale: proprio a causa di ciò l’autore potrebbe persino vivere il foglio stile come un’imposizione da parte dell’editore che lo distoglie dal contenuto e non gli permette di scrivere con sufficiente libertà.

È chiaro che partire da un manoscritto già ben strutturato sarebbe ideale per le successive lavorazioni, ma occorre ancora del

tempo perché la mentalità degli autori sia sufficientemente matura per affrontare la necessità di un certo grado di competenza tecnologica e per viverla non come un ostacolo o un'inutile complicazione, ma anzi come uno strumento per semplificare il proprio lavoro.

Naturalmente non tutti gli autori sono uguali e ciascun editore saprà individuare tra i propri i più disponibili a cimentarsi con una nuova prassi lavorativa, ai quali potrà proporre con la dovuta preparazione e guida il proprio foglio stile. Quest'ultimo, inoltre, sarà tanto più complesso quanto più strutturato è il contenuto del libro: anche questo è un elemento da tenere in considerazione prima di partire con l'ideazione e l'utilizzo del file `.dot`.

Suggerimento *Per esempio, se si parla di un classico romanzo, il foglio stile, come potete immaginare, non presenterà molto più che dei titoli di capitolo e forse qualche stile di carattere. Un saggio invece, con tabelle, immagini e vari livelli di titoli, comporta un grado di complessità differente.*

Ci si potrebbe anche trovare davanti a un'altra difficoltà: la diffidenza nei confronti del formato digitale stesso. La convinzione comune che qualsiasi cosa transiti per uno schermo possa essere piratata con facilità non è di per sé errata: è un aspetto che prenderemo in esame nel Capitolo 7.

Diversa e più complessa, invece, è la diffidenza riguardo al libro digitale stesso, alla sua validità culturale e al prestigio che ne deriva

rispetto all'autorevolezza del libro stampato. Un'argomentazione interessante per l'autore, a questo proposito, potrebbe essere la stessa scelta dell'editore di pubblicare libri digitali: si tratta di un'iniziativa con un preciso valore culturale, che richiede investimenti e risorse, e dovrebbe risultare convincente e autorevole di per sé.

Tornando al lato pratico del dover affidare un maggior carico di responsabilità all'autore nella gestione e nella valorizzazione semantica del testo, dobbiamo considerare che il grado di alfabetizzazione informatica degli autori non sempre è tale da rendere affidabile – per il momento – la loro gestione autonoma del testo attraverso il foglio stile, e richiederà un intervento da parte di altre figure professionali.

Tuttavia sarà utile iniziare comunque a rendere partecipi anche gli autori delle nuove tecniche che si adotteranno per la produzione dei libri: guidandoli con gli strumenti più adatti per comprendere la nuova importanza che il loro ruolo potrebbe assumere nei confronti del testo si potrà col tempo contare su una loro partecipazione efficace e attiva alla strutturazione del contenuto.

Impatto maggiore sul redattore

Il redattore sembra essere la figura più indicata a seguire e coordinare il nuovo flusso di lavorazione del libro: già abituato a occuparsi della struttura dei contenuti e più pratico nell'utilizzare gli elaboratori di testo, è anche il professionista che più di tutti è in grado di realizzare gli strumenti necessari

per mettere in atto il nuovo processo di produzione del libro.

Nessuno meglio del redattore sa infatti ricondurre a un modello tutte le necessità stilistiche e semantiche dei libri a cui lavora; creare le guide e i supporti necessari per l'utilizzo dei nuovi strumenti destinati agli altri attori della produzione del libro (autori, traduttori, impaginatori); supervisionare la correttezza del risultato prodotto attraverso l'utilizzo di questi stessi strumenti.

Tuttavia, anche il redattore dovrà impadronirsi di nuove competenze per riuscire a gestire al meglio il nuovo flusso di lavoro. Come abbiamo visto, per realizzare un foglio stile veramente efficace non è sufficiente determinare degli stili, ma occorrerà riflettere attentamente sulla struttura semantica del

documento tenendo conto delle necessità che la trasformazione in DocBook comporta.

Non solo: così come la conoscenza piena degli elementi fondamentali per la stampa è già un requisito in possesso del redattore, analogamente questi dovrà diventare un conoscitore delle possibilità e delle necessità dei formati digitali che intende ottenere attraverso la lavorazione del libro secondo la nuova procedura.

Fatte queste premesse, dovrà poi essere in grado di verificare l'esattezza del risultato della prima trasformazione dal foglio stile in DocBook: come abbiamo visto, è in quella fase che ogni errore trascurato può diventare un problema per le successive trasformazioni verso i formati definitivi.

Anche in questo caso, ogni editore potrà gestire come meglio crede le risorse professionali in suo possesso. In alcuni casi sarà utile produrre internamente i fogli stile e poi affidarsi ad altri professionisti per gestire le successive trasformazioni: seguendo questa strada, il redattore potrebbe trovarsi a lavorare sempre sui formati finiti, senza dover acquisire troppe competenze tecniche sui linguaggi necessari per la gestione concreta del processo di trasformazione.

In altri casi, invece, si potrà ricorrere a un compromesso: far sviluppare esternamente il foglio di trasformazione necessario per passare dal foglio stile al DocBook e poi richiedere al redattore le competenze per verificare la struttura una volta riavuto il documento in questo formato. In questo modo, la verifica sui formati definitivi

potrebbe essere più agile, considerato che partendo da un DocBook corretto le successive fasi di trasformazione richiedono interventi minimi.

Al redattore, comunque, spetterà il compito di compilare alcune delle parti istituzionali del libro, come per esempio quelle che permettono di generare il colophon e che contengono tutte le informazioni legali e bibliografiche della pubblicazione. Queste infatti non possono essere generate completamente in automatico né tramite il foglio stile: il redattore quindi dovrà essere capace di orientarsi nella marcatura e di compilare correttamente l’insieme di metadati sul libro necessari alla creazione di queste pagine e di altre simili.

Concludendo, la scelta migliore sarebbe considerare come necessaria l'acquisizione di competenze sul linguaggio di marcatura che diventa il perno del nuovo flusso editoriale: solo così il redattore potrà controllare pienamente il flusso di lavorazione e supervisionare in maniera efficace il risultato prodotto.

Come nel caso dell'autore, potrebbero sorgere delle resistenze verso questa nuova esigenza professionale: il linguaggio di marcatura potrebbe essere considerato come un qualcosa di troppo tecnico, che esula dalle competenze "umanistiche" di un redattore, già così indaffarato tra correzioni di bozze e riscontri. In realtà, DocBook è sì uno strumento tecnico che richiede un certo livello di competenza informatica, ma il suo funzionamento è così fortemente basato sulla

semantica e sulla definizione dei contenuti da rendere evidente come il suo utilizzo da parte dei professionisti del testo sia in qualche modo una naturale deriva del loro metodo di lavoro. I computer sono da tempo uno strumento imprescindibile del lavoro redazionale: ora forse è tempo di sfruttarne le possibilità con maggiore consapevolezza e senza timori o pregiudizi.

Impatto maggiore sull'impaginatore

Quanto detto a proposito dell'autore e del redattore lascia intendere che la strada da seguire sia necessariamente quella descritta a partire dalla generazione di un foglio stile con gli elaboratori di testo. Eppure abbiamo visto che ci sono altri metodi per ottenere lo

stesso risultato: la produzione di un foglio stile in grado di generare un XML ben formato da cui partire per le successive trasformazioni è possibile anche da altre applicazioni che richiedono l'intervento di altre figure professionali, come i grafici e gli impaginatori.

La definizione di un foglio stile è una pratica ben nota per queste figure che, tuttavia, come richiesto dal loro ruolo, sono esperte riguardo la resa grafica dei documenti più che sulla struttura e la semantica del testo, compiti che – come abbiamo visto nel paragrafo precedente – spettano per formazione e ruolo al redattore. Questo rende difficile per un impaginatore creare un foglio stile che tenga conto delle esigenze strettamente semantiche imposte dal linguaggio di marcatura, e d'altra parte limita l'autonomia del

redattore perché questi non sempre è in possesso né del software necessario per l'impaginazione né delle competenze per utilizzarlo.

Il processo che abbiamo descritto a partire dalla produzione di un foglio stile con un elaboratore di testo, in cui il documento viene più volte trasformato fino a ottenere tutti i formati necessari per la pubblicazione del libro, sembra tagliare fuori la figura dell'impaginatore, che non trova più un ruolo nel nuovo flusso di lavorazione.

In realtà non è affatto così: come abbiamo visto, la gestione delle immagini è tutt'altro che risolta in maniera automatica e ci sarà bisogno di grafici capaci di intervenire sulla loro lavorazione. Inoltre rimane in sospeso la definizione del foglio stile per la resa

grafica dei formati digitali: il CSS delle collane dovrà comunque essere creato dai professionisti della presentazione dei documenti.

Considerato che la fase di impaginazione è in molti casi esterna alla casa editrice, i service editoriali potrebbero valutare di specializzarsi nella gestione del nuovo flusso di lavoro, affiancando ai software tradizionali del mestiere anche delle nuove competenze nella gestione dei file tramite gli elaboratori di testo e proponendosi come figure versatili in grado di gestire la pagina secondo le esigenze e le modalità proposte dall'editore in base alla strada scelta per la produzione dei suoi libri.

Anche in questo caso, l'editore dovrà ben valutare le risorse professionali con cui

abitualmente collabora, e tenere in considerazione le sue opportunità concrete quando opterà per una soluzione o per un'altra riguardo al modo in cui intende gestire il processo di lavorazione.

Tutto ciò premesso, se si decide di produrre il foglio stile attraverso un programma come InDesign ci si trova davanti a non pochi interrogativi, che rischiano di aggiungere complessità a un quadro già intricato. Infatti, generare l'XML dal software Adobe, oltre a comportare una trasformazione verso il DocBook meno diretta e più complessa rispetto quella possibile a partire da un elaboratore di testo, è complicato anche per via del professionista che se ne dovrebbe occupare.

Come abbiamo osservato nel paragrafo precedente, il redattore non ha difficoltà nell'individuare tutti gli elementi necessari da inserire in un foglio stile che definisca la struttura del testo: per mestiere si occupa proprio della gestione di tali aspetti. L'impaginatore riveste invece un altro ruolo, per definizione votato alla presentazione del testo, e si interessa meno della sua semantica. Senza considerare poi che il foglio stile per come lo abbiamo descritto è uno strumento che dimostra la sua efficacia nel momento in cui viene messo a disposizione dell'autore, che può utilizzarlo per organizzare e strutturare il contenuto secondo le esigenze gestionali stabilite dall'editore: l'impaginatore non è quasi mai l'interlocutore diretto dell'autore, e anche questo è un compito che spetta al redattore.

Come fare quindi a delegare la produzione di un foglio stile riguardante strettamente la struttura e la semantica del testo a un professionista competente riguardo tutt'altri aspetti, che dovrebbe realizzarlo con un software non ideale proprio perché pensato per rispondere a esigenze differenti e che, infine, non è neanche il professionista a cui spetta il compito di lavorare con gli autori? Non va dimenticato, inoltre, che in ogni caso, l'autore non lavorerebbe mai direttamente in InDesign ma comunque in un elaboratore di testo: il processo dunque non si alleggerirebbe ma comporterebbe un carico di lavoro molto maggiore sull'impaginatore, che si troverebbe investito di competenze e responsabilità non propriamente sue e che – a differenza di quelle che abbiamo individuato

per il redattore – non è detto che sia utile acquisisca.

La strada più interessante sembra dunque essere quella segnata dalla lavorazione del libro a partire da un miglior sfruttamento dei documenti prodotti con gli elaboratori di testo. Gli impaginatori dovranno quindi assumere sì nuove competenze, ma relative ad ambiti e ruoli da loro già rivestiti, riguardanti la cura delle immagini e la produzione e manutenzione dei fogli stile per la resa grafica dei formati digitali.

Un'ulteriore possibilità per gli impaginatori è quella di offrire all'editore un servizio efficace sul catalogo. Il percorso che abbiamo descritto, infatti, presuppone che si tratti di libri ancora da pubblicare, ma naturalmente un editore ha tutto l'interesse di

riproporre in versione digitale anche libri già pubblicati.

A questo proposito tutti gli aspetti legati ai cambiamenti del processo di lavorazione vengono in qualche modo a mancare, perché di fatto si tratta di recuperare il testo lavorato secondo le modalità usuali e adattarlo alle necessità delle trasformazioni per i formati digitali. L'impaginatore potrebbe quindi assumere un ruolo cruciale riguardo al riversamento in digitale del catalogo: è il professionista più indicato a riprendere in mano i file dei libri già pubblicati, ai quali ha lavorato lui stesso, e ad adattarli secondo le indicazioni proposte dall'editore riguardo alla versione digitale degli stessi titoli.

Suggerimento *Il riversamento del catalogo nei formati digitali assume un aspetto rilevante se si considerano le*

opportunità di far rivivere tutti quei titoli che non sono abbastanza forti da sostenere i costi di una ristampa ma che potrebbero ancora trovare un pubblico in digitale.

Altre figure professionali

A seconda del libro su cui si lavora potrebbe essere necessario considerare l'intervento di altri professionisti nella produzione dei file per la pubblicazione. È il caso, per esempio, di traduttori e curatori: come gli impaginatori sono figure generalmente esterne alla casa editrice, ma il loro lavoro impatta sulla struttura e sul contenuto del testo in maniera molto più significativa.

Visto il tipo di rapporto di collaborazione che lega tali figure all'editore, potremmo considerarli quasi in una posizione

intermedia tra autore e redattore: nel primo caso perché – di fatto – contribuiscono in qualche modo alla creazione del contenuto, nel secondo perché da loro si può pretendere un certo grado di competenza tecnica che non può sempre essere richiesta a un autore.

Il loro ruolo non riguarda in alcun modo la creazione dei fogli stile o degli altri strumenti necessari per il nuovo flusso produttivo, ma si concentra su un loro utilizzo puntuale ed efficace, che può essere preteso dall'editore proprio in virtù del tipo di rapporto professionale che si instaura.

Traduttori e curatori potrebbero e dovrebbero essere in grado di lavorare con precisione sulla base del foglio stile che l'editore fornisce loro. L'acquisizione delle competenze tecniche necessarie per l'utilizzo

corretto dell’elaboratore di testo secondo le esigenze del nuovo flusso è auspicabile e necessaria: come non si penserebbe mai di rivolgersi a un impaginatore che non sa lavorare con efficienza con gli strumenti professionali propri del suo ruolo, analogamente a questi collaboratori andrebbero richieste non solo competenze specifiche sulla materia, ma anche capacità di inserirsi correttamente nel flusso di lavoro dell’editore, senza comportare un aumento del carico per il redattore, che sarebbe costretto a investire del tempo a correggere gli errori di struttura nei documenti prodotti da traduttori e curatori.

Questo discorso ha un peso diverso sulla figura del traduttore rispetto a quella del curatore. Quest’ultimo, infatti, spesso si limita a supervisionare l’opera dal punto di

vista del contenuto, senza grandi interventi concreti sul testo, ma più spesso discutendo con la redazione degli aspetti più rilevanti.

Il traduttore, invece, deve di fatto produrre un nuovo testo, e l'editore riceve da lui un documento redatto con un elaboratore di testo: di fatto è come se rivestisse il ruolo dell'autore ma con un maggior carico di vincoli e responsabilità tecniche più simili a quelle del redattore.

È legittimo quindi chiedere a queste figure professionali un certo grado di competenza, predisponendo anche per loro guide e supporti per l'utilizzo dei nuovi strumenti di lavorazione.

E ora, dove li metto?

Nel decidere di avviare la produzione sistematica di libri digitali non è sufficiente pianificare gli aspetti tecnici sul come fare: occorre anche una precisa riflessione sul come collocare questa nuova parte della produzione all'interno della linea editoriale.

Pubblicare libri digitali non significa limitarsi a riproporre in formato elettronico i contenuti che di solito si stampano: le possibilità offerte dalla tecnologia permettono di ripensare le forme e i metodi per confezionarli e diffonderli, e il ruolo di un editore dovrebbe essere anche trovare le vie migliori per valorizzare al meglio le forme di fruizione che i diversi strumenti di lettura

propongono, sia su carta sia in formati digitali.

Tutti gli investimenti necessari per acquisire le competenze che permettono di avviare e gestire la nuova produzione devono essere intesi da questo punto di vista, e non come un mero problema tecnico da risolvere senza le adeguate premesse culturali.

In questo senso è ancora più difficile generalizzare e suggerire linee guida comuni: ogni editore fa caso a sé, conosce il proprio mercato e valuta di conseguenza gli obiettivi da raggiungere attraverso gli ebook. È possibile però fare una serie di considerazioni utili, su cui ognuno possa ragionare in base alle proprie necessità.

Tutti i miei libri vanno bene in digitale?

È una delle prime questioni da affrontare. Non tutti i contenuti si prestano allo stesso modo per la pubblicazione in digitale: pensiamo per esempio a opere che trovano nella stampa una valorizzazione che l'ebook non permette (un esempio, i libri fotografici). Ciò non significa che per questo tipo di libri non ci sia alcun vantaggio nell'essere digitalizzati: vale più che in altri casi, però, la necessità di ripensare attentamente il contenuto e di valutare quale sia la forma migliore da fargli assumere.

Anche testi apparentemente semplici propongono questioni sottili e per questo richiedono una riflessione ancora più

accurata. Questo è chiaro se ripensiamo al processo descritto per assegnare in maniera efficace il corretto valore semantico al testo in lavorazione: al di là delle complessità tecniche, l'approccio anche nei confronti del testo semplice si evolve verso una creazione e lavorazione dei contenuti molto più consapevoli, lasciando spazio a eventualità di perfezionamento ulteriori che consentiranno in futuro uno sfruttamento delle possibilità dei libri digitali sempre migliore.

Un altro aspetto di questo tema è l'opportunità di proporre certi contenuti in digitale così come sono strutturati su carta. È lecito pensare che la strutturazione di determinati argomenti su carta non possa essere semplicemente replicata in digitale e viceversa: per certe opere varrebbe la pena di pensare a come proporre il contenuto in modo diverso

a seconda del supporto di lettura (cartaceo o digitale) a cui è destinato.

Ognuna di queste decisioni impatta su tutte le scelte tecniche, che possono essere fatte solo a seguito delle riflessioni sul proprio catalogo e sui modi più adatti per valorizzarlo attraverso nuove forme di pubblicazione. Decidere di ottenere dalla stessa lavorazione tutti i formati necessari comporta la necessità di realizzare *ad hoc* ogni progetto che abbia bisogno di allontanarsi da questa uniformazione; d'altra parte è piuttosto difficile pubblicare sistematicamente libri digitali senza standardizzare in qualche modo il flusso lavorativo.

Si tratta di decisioni strettamente legate al tipo, alla quantità e ai destinatari dei testi che si producono, oltre che agli obiettivi che

si intendono raggiungere con la loro digitalizzazione. Non per tutti la scelta di pubblicare ogni titolo sia su carta sia in digitale è la più giusta: potrebbe essere meglio, per esempio, intervenire solo su alcuni, prendendosi il tempo di immaginare un modo diverso di proporre il resto del catalogo.

Una collana per gli ebook o gli ebook nelle collane?

Conseguenza di questo discorso è la questione del posizionamento dei libri digitali tra (o nelle) le collane dell'editore. Proviamo a prospettare quattro scenari diversi.

- Intendo produrre di ogni mio libro stampato anche la copia digitale.

- Non intendo produrre sistematicamente ebook, ma solo alcuni titoli stabiliti di volta in volta a seconda dell’opportunità.
- Intendo avviare il recupero dei titoli più datati del mio catalogo che potrebbero avere una nuova vita grazie all’ebook.
- Intendo produrre alcuni titoli solo ed esclusivamente in digitale.

Il primo caso è quello relativamente più semplice. Una volta stabilito che ogni libro verrà pubblicato in forma cartacea e digitale, la definizione delle collane dell’editore non cambia, ma si arricchisce di uno o più formati per ogni libro prodotto. Il contenuto rimane invariato e la collocazione all’interno della collana a cui appartiene quel contenuto, di conseguenza, continua a essere valida anche per l’ebook.

Il secondo caso è più delicato. In assenza della produzione sistematica, ogni progetto di libro digitale si carica di maggiori investimenti dal punto di vista sia della produzione sia della promozione: diffondere la notizia che si producono tanti libri digitali è più semplice che dover promuovere un singolo titolo, o almeno permette di distribuire l'onere della promozione su tutti i titoli piuttosto che farla gravare su uno solo; allo stesso modo, dover lavorare *ad hoc* ogni titolo comporta il dover riconsiderare ogni volta gli strumenti per la progettazione e per la trasformazione (fogli stile, fogli di trasformazione e così via).

Suggerimento *È anche vero che gli strumenti più adatti alla promozione di un libro digitale – pensiamo*

soprattutto al Web – non sono onerosi come quelli per il libro stampato.

Trattandosi di pubblicazioni così sporadiche, la collocazione all'interno di una collana potrebbe ancora derivare dalla relazione con il libro cartaceo. Rimane da valutare quanto lo stesso contenuto formalizzato in maniera radicalmente diversa sulla carta e sul digitale (perché magari arricchito con materiali multimediali o con elementi di interattività, per pensare agli aspetti più macroscopici) possa rientrare nella stessa collana del libro stampato senza forzarne i confini.

In altre parole: quando la forma influenza in modo importante la fruizione del contenuto, non sta di fatto influenzando il contenuto stesso, facendolo diventare altro rispetto alla versione cartacea? I due oggetti possono convivere comunque nello stesso

contenitore/collana? Questo aspetto è così fortemente legato a ciò che ogni editore intende per collana e all'uso che ne fa da non poter essere oggetto di generalizzazioni ulteriori: occorre tuttavia tenerlo presente per non rischiare di indebolire con scelte poco ragionate la propria identità editoriale.

Il terzo caso potrebbe riguardare editori con un ampio catalogo e un gran numero di titoli fuori produzione: alcuni di essi potrebbero ancora avere un mercato ma non essere in grado di sostenere i costi di una ristampa. Questa scelta richiede un diverso approccio alla produzione, che non influisce sulle lavorazioni in corso ma richiede investimenti per la rilavorazione delle pubblicazioni preesistenti. Ciò premesso, e a meno di optare per il recupero integrale di una o più collane, si potrebbe ragionare sulla

possibilità di creare una collana dedicata ai riversa menti in digitale, per dare un'identità precisa e una certa uniformità al progetto, che consentirebbe anche di studiare la maniera più efficace di riproporre in modo innovativo i contenuti che si vanno a recuperare.

La quarta ipotesi è ancora diversa. Sganciarsi completamente dalle necessità e dai costi della stampa potrebbe permettere sperimentazioni a vario livello: dalla pubblicazione di titoli molto specifici o molto innovativi, al tentativo di proporre diverse modalità per la diffusione dei contenuti, oltre che allo sfruttamento più efficace possibile delle opportunità dei formati digitali. In questo caso potrebbe essere interessante realizzare delle collane pensate apposta per gli ebook.

C’è però un ulteriore scenario da tenere in considerazione: l’eventualità che l’idea di collana stessa si indebolisca, se non che scompaia del tutto, di fronte alla diversificazione delle forme che il contenuto può assumere. L’identità di una collana è data, oltre che dai contenuti che propone, anche dal suo confezionamento, dalle caratteristiche fisiche e grafiche del libro: carta, rilegatura, stile delle copertine e degli interni e così via.

Chiunque sia mai entrato in libreria riesce a riconoscere a colpo d’occhio le collane dall’identità più forte: l’aspetto, ricominciando a pensare il libro come oggetto, è determinante per la caratterizzazione dell’identità dell’editore e delle sue emanazioni rappresentate dalle collane.

Ma proprio perché con l’ebook abbiamo imparato a vedere oltre all’oggetto-libro, dobbiamo pensare alla collana in maniera più astratta, come un collante di idee e non come una confezione fisica uniforme per dei contenuti. Da questo punto di vista si tratta di un cambiamento nel modo di intendere la propria identità editoriale non indifferente.

Chi fa da sé fa per tre

A seconda delle decisioni prese riguardo alla produzione dei libri digitali si dovrà adattare il flusso lavorativo. In particolare si dovrà riflettere sull’eventualità di mantenere il processo di creazione degli ebook all’interno della casa editrice o di delegarlo all’esterno.

Oltre a delle ragioni di natura economica che esulano dal nostro discorso (ma non sono certamente secondarie), bisogna valutare quale sia la via che più si addice al tipo di scelta editoriale che abbiamo fatto nell'avviare la produzione di ebook.

Rimane il fatto che impadronirsi internamente delle competenze necessarie almeno a supervisionare il processo, se non a gestirlo nel dettaglio, non solo è auspicabile ma anche consigliabile per non rischiare di perdere il controllo del libro.

In questo senso la figura interna alla casa editrice che più deve arricchirsi di nuove competenze è – come abbiamo visto – quella del redattore.

Distribuzione

Ormai è deciso: si intendono produrre dei libri digitali, si è scelto il metodo più adatto e si sono individuate le figure professionali necessarie a portare a termine il progetto. Infine, gli ebook sono pronti.

In questo paragrafo vedremo alcune delle strategie di distribuzione per i libri digitali, basandoci su delle esperienze concrete in grado di mostrarci le evoluzioni e gli adattamenti del mercato editoriale riguardo al nuovo prodotto.

Dove li vendo

A questo punto del discorso occorre accennare a uno strumento che diventa

fondamentale per la distribuzione degli ebook: il sito web dell'editore. Per poter vendere i libri digitali dal proprio sito occorrerà – se già non la si possiede – predisporre una struttura di e-commerce appropriata, e naturalmente offrire l'opportunità di accedere al proprio catalogo attraverso delle schede promozionali per i libri.

Il sito web riveste oggi per qualunque azienda un'importanza fondamentale per la promozione della propria attività. Per l'editore costituisce un'ulteriore possibilità di proporsi come creatore e diffusore di contenuti, appropriati ai settori editoriali ricoperti e alle aree di specializzazione che lo riguardano. Di certo non ci si può permettere di improvvisare un sito senza un'attenta pianificazione: anche questo strumento deve essere progettato con cura, in

modo da rispecchiare l’identità dell’editore nella maniera più efficace.

Il prezzo

Anche le decisioni per determinare il prezzo degli ebook sono dettate dalle peculiarità dell’editore e del contesto in cui si trova a operare. In questa fase instabile e in continua evoluzione, in cui il mercato è ancora lontano dall’avere dei contorni ben delineati, esistono alcuni grandi attori capaci di sostenere una politica di prezzo molto aggressiva, che non per tutti è proponibile. Tuttavia, finché l’ebook non assumerà tra i lettori lo stesso valore del libro cartaceo, grazie a una maggiore consuetudine col mezzo e con le sue possibilità e al riconoscimento del suo valore culturale, le decisioni sul prezzo dei libri digitali saranno ancora dettate da

ragionamenti differenti da quello del valore dell’ebook in sé.

Per comprendere meglio questo aspetto procederemo adesso con l’osservazione di alcuni esempi concreti riguardo la distribuzione e la vendita dei libri digitali, che ci permetteranno di comprendere meglio quanto finora abbiamo descritto in linea di principio.

Editori

Iniziamo l’osservazione a partire da alcuni grandi editori americani. Ci riferiamo tali esempi perché negli Stati Uniti il mercato è più maturo rispetto a quello europeo, e perché i diversi numeri a cui si rivolge permettono riflessioni di respiro ben più ampio rispetto alla realtà europea, troppo frammentata per

poter essere utilizzata come esempio (pensiamo solo alle differenze linguistiche e legali).

O’Reilly Media

O’Reilly Media (http://oreilly.com), importante casa editrice americana di informatica (oltre ai libri gestisce un gran numero di siti web sull’argomento), propone ai lettori varie formule per l’acquisto dei libri digitali e un ampio ventaglio di servizi sui contenuti. Osserviamo per esempio la scheda di un libro sul sito dell’editore (Figura 5.1).

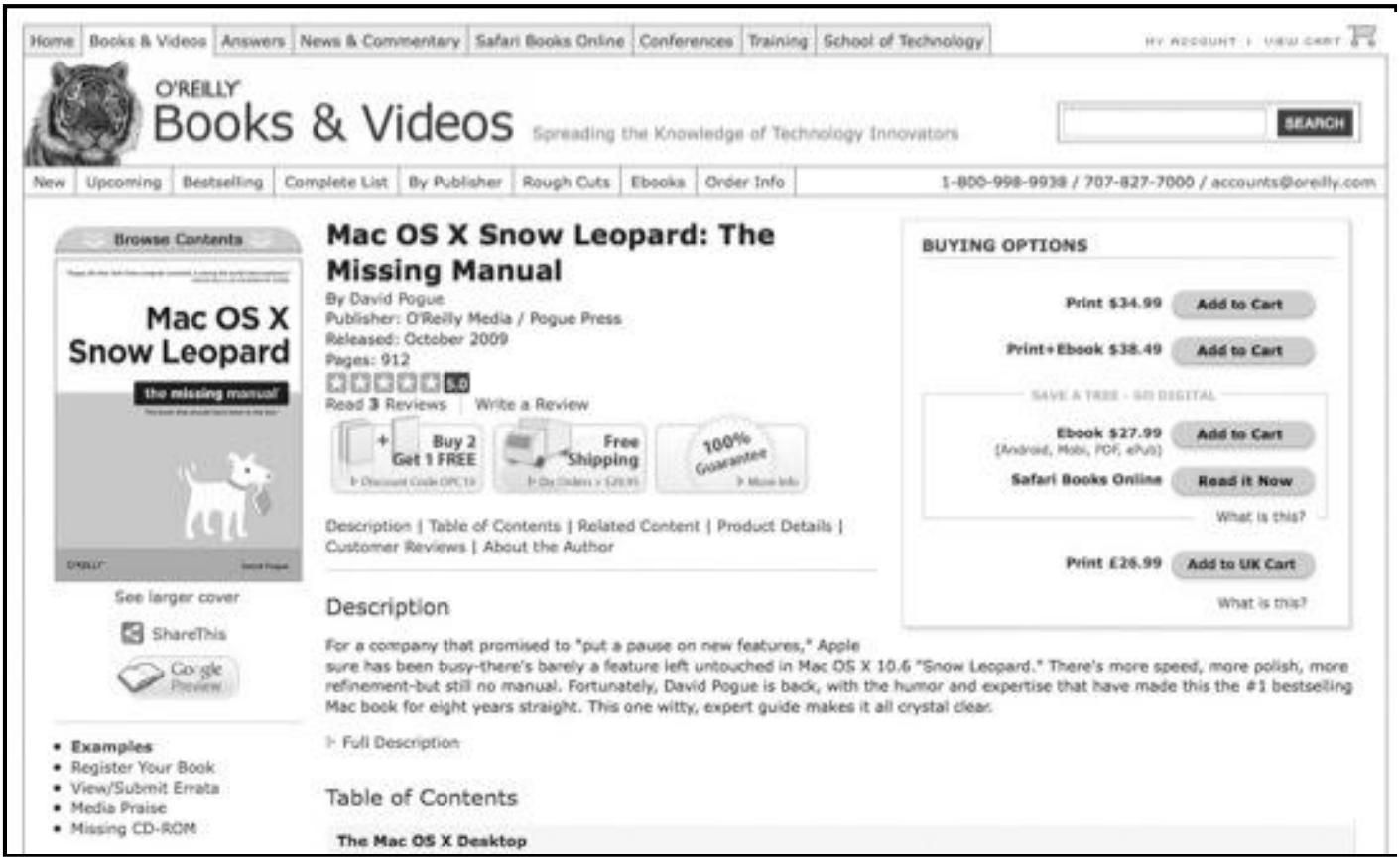

Figura 5.1 Una scheda libro sul sito dell'editore O'Reilly Media.

Nel box sulla destra sono possibili differenti modalità di acquisto:

- libro cartaceo;
- libro cartaceo ed ebook;
- solo ebook con possibilità di consultazione attraverso Safari Books Online.

Il libro digitale può essere acquistato sia insieme al libro cartaceo (costa il 10% in più rispetto al solo libro cartaceo) sia singolarmente, in una formula che propone un pacchetto comprendente tutti i formati digitali in cui è disponibile il libro più l'accesso al contenuto tramite il servizio Safari Books Online – di cui parleremo tra poco – al 20% in meno rispetto al prezzo di copertina del solo libro stampato.

L'offerta del solo libro digitale, oltre all'essere graficamente in evidenza insieme ai vantaggi ecologici dell'ebook, propone una finestra di supporto che illustra all'utente tutto ciò a cui può accedere con l'acquisto del pacchetto di formati digitali.

All'indirizzo http://oreilly.com/ebooks/ l'editore offre un gran numero di

informazioni relative ai formati che propone, ai dispositivi di lettura su cui possono essere utilizzati, alla propria politica sul DRM e a un insieme di servizi: per esempio, nel caso di aggiornamenti relativi all'ebook acquistato, l'editore provvederà a sostituire gratuitamente il libro digitale con l'ultima versione (il servizio è fornito solo per il formato PDF).

L'iniziativa Rough Cuts (http://oreilly.com/roughcuts/) è un perfetto esempio di possibilità derivata dalle nuove modalità di lavorazione del libro. Permette infatti al lettore di accedere ai contenuti di libri ancora in lavorazione e spesso in fase di scrittura: li si può leggere, scaricare in PDF e stampare, pagando una cifra differente a seconda della formula desiderata. Questo consente di avere un feedback

immediato sull'interesse riguardo al libro ancora prima che questo venga pubblicato. A seconda delle esigenze del titolo, inoltre, può offrire la possibilità di creare delle vere e proprie comunità online riguardo all'argomento trattato, per raccogliere suggerimenti e suggestioni che possono essere integrati nella versione definitiva del libro.

Riferimenti *O'Reilly propone inoltre una vastissima scelta dei suoi titoli in forma di applicazioni per iPhone, sia gratuite sia a pagamento. Queste applicazioni racchiudono all'interno anche il libro in formato ePub: lo si può estrarre dall'applicazione e leggerlo su qualunque altro reader in grado di interpretarlo.*

Safari Books Online (http://my.safaribooksonline.com/) è un servizio in abbonamento offerto in collaborazione con Pearson Publishing, un colosso dell'editoria americana attivo soprattutto nel settore universitario, e permette l'accesso a una vastissima gamma di contenuti tramite la sottoscrizione di un abbonamento (Figura 5.2). Non si tratta solo di libri – per i quali sono disponibili tutti i più diffusi strumenti di supporto alla lettura digitale – ma anche di video, i già citati *rough cuts*, modalità di accesso dai dispositivi mobili e un certo numero di download al mese da consultare anche offline. È possibile gestire il proprio account organizzandolo in cartelle e salvando annotazioni e segnalibri utilizzati durante la lettura.

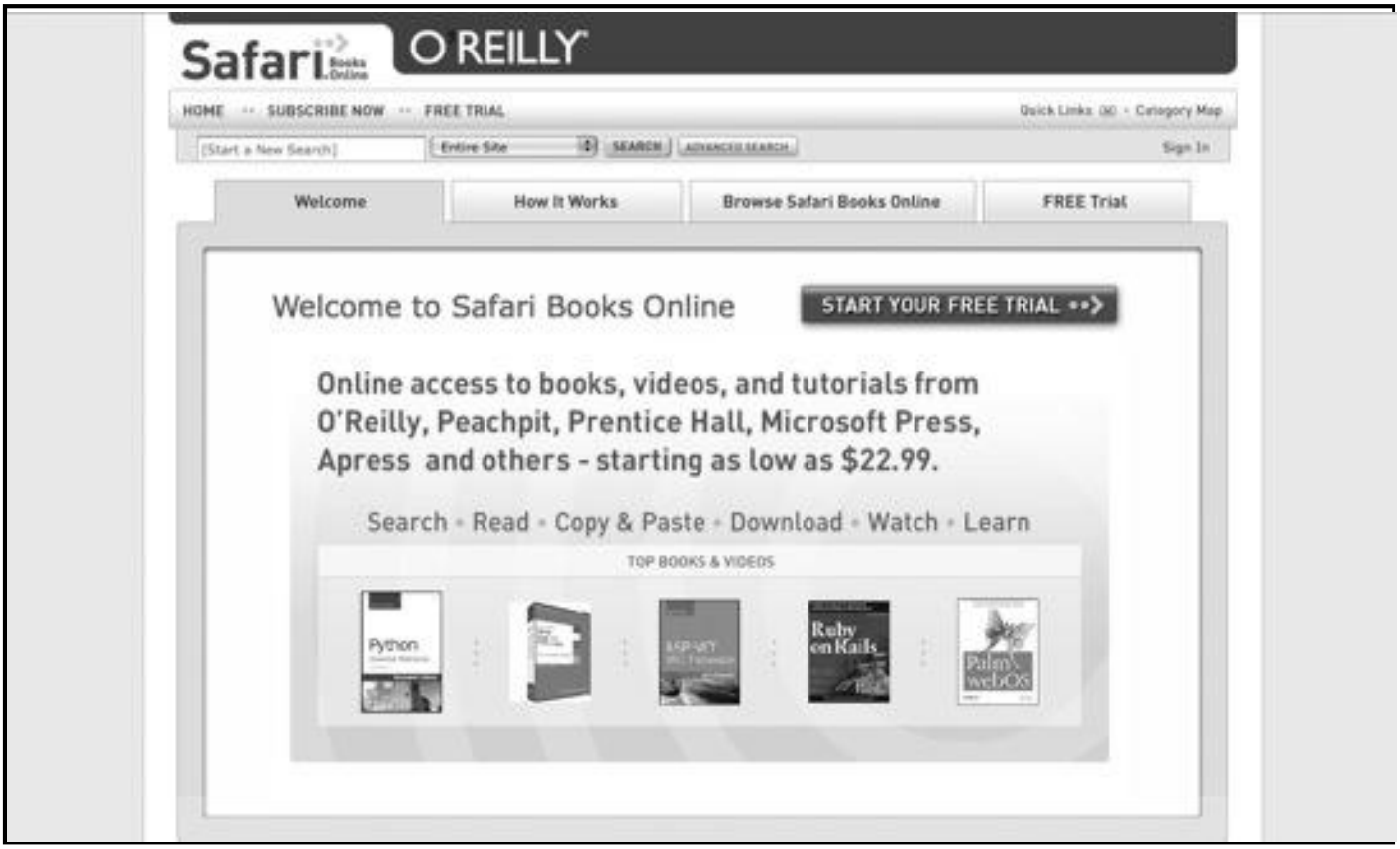

Figura 5.2 La pagina di benvenuto di Safari Books Online.

Simon & Schuster

Simon & Schuster (http://www.simonandschuster.com/) è uno dei più grandi editori americani di narrativa e varia. Per trovare gli ebook sul suo sito dobbiamo accedere all'apposita sezione nel menu di navigazione a sinistra nella pagina dedicata ai libri. A

differenza di O'Reilly, Simon & Schuster non propone l'acquisto di tutte le versioni dello stesso libro da un'unica scheda, ma distingue l'ebook dagli altri formati (hardcover e audiolibro), proponendo per ciascuno una pagina diversa (Figura 5.3). Anche riguardo al prezzo non ci sono differenze: il libro digitale costa quanto quello stampato, e non esistono formule di acquisto del titolo in più versioni.

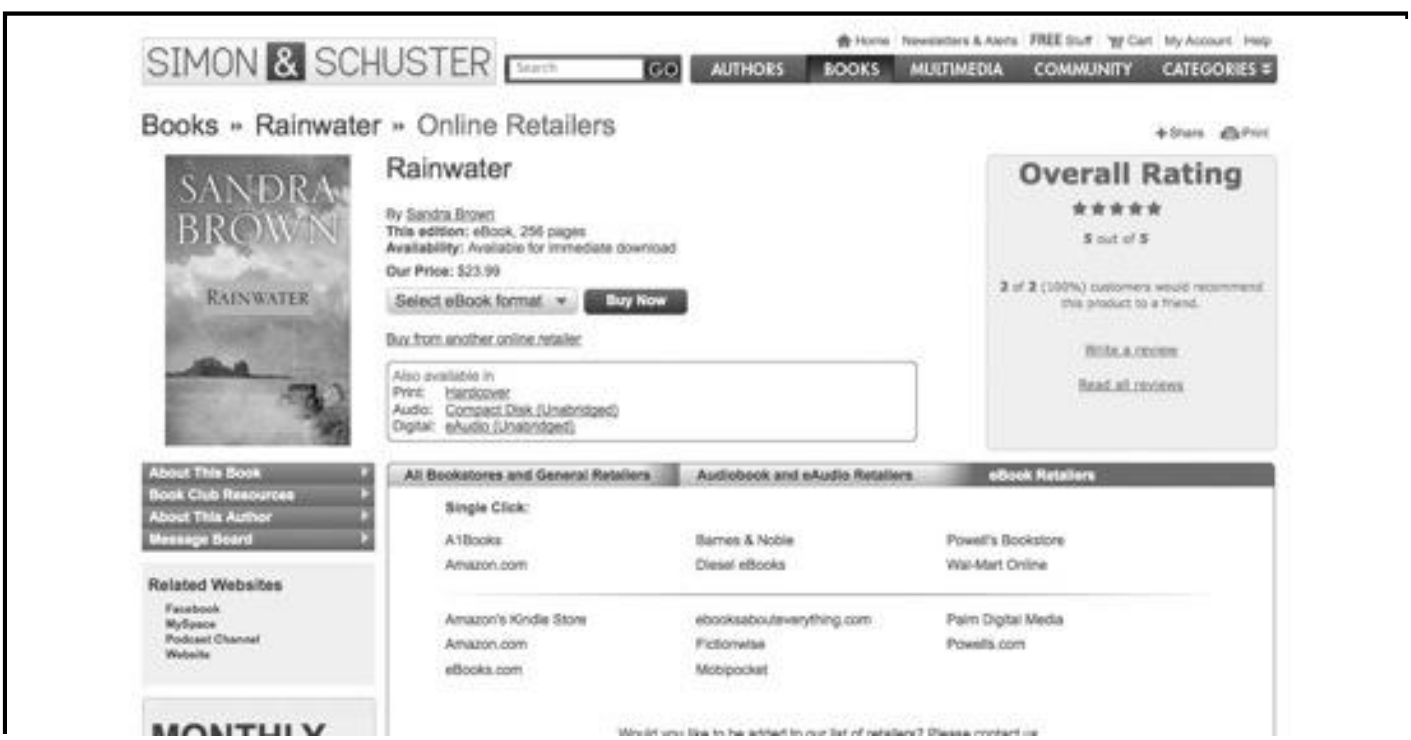

Figura 5.3 Una scheda ebook sul sito di Simon & Schuster.

Sul sito viene proposto un unico formato per il download (generalmente l'ePub), ma si può acquistare lo stesso titolo in digitale da un gran numero di altre librerie online in molti altri formati (come per esempio mobi e AZW).

Suggerimento *Simon & Schuster propone anche diverse applicazioni per iPhone, dai corsi di lingua ai cruci-verba, tutte a pagamento.*

Penguin

Il gruppo Penguin (http://us.penguin-group.com) è un colosso internazionale dell'editoria in lingua inglese. Per trovare la sezione dedicata agli ebook sul suo sito occorre selezionare *Multimedia Books* dal

menu a sinistra della home page. Come nel caso di Simon & Schuster, per ogni edizione del libro esiste una scheda differente, e anche il prezzo non varia tra la carta e il digitale. Penguin propone un'applicazione gratuita per iPhone dalla quale gestire il carrello, accedere ai contenuti multimediali proposti dall'editore e scaricare e leggere alcuni estratti dei libri.

Random House

Random House (http://www.random-house.com), gigante dell'editoria internazionale attivo in un'ampissima gamma di settori di produzione, riunisce gli ebook in una delle categorie del suo sito, insieme ai generi narrativi e ad altri servizi, come la possibilità di personalizzare i libri che si intende acquistare. Così come Simon &

Schuster e Penguin, propone una scheda diversa per ogni versione in cui il titolo viene declinato, senza differenze di prezzo tra il libro di carta e l'ebook. Ciò che non è immediatamente chiaro dalla scheda libro è quali formati siano disponibili per il download: nella pagina di aiuto sugli ebook viene specificato che si tratta di ben sei formati (PDF, LIT, PRC, PDB, Sony ed ETI).

Riferimenti *ETI è un formato supportato solo da dispositivi eBookwise (www.eBookwise.com).*

Tiriamo le somme

Da questi pochi esempi che non pretendono certamente di esaurire il panorama di modalità distributive da parte degli editori, possiamo ricavare alcuni elementi di riflessione.

- L’architettura del sito deve essere in grado di valorizzare il catalogo di libri digitali, specialmente quando si opta per pubblicare tante schede libro quante sono le versioni del titolo proposte.
- Gli strumenti di supporto (FAQ, help) sono fondamentali in questa fase di sviluppo del mercato, e necessari per orientare il lettore nella scelta dell’ebook adatto alle sue esigenze.
- Proporre nuove formule di acquisto e/o servizi dedicati alla lettura digitale può essere una mossa vincente sia per dare rilievo alla propria produzione di ebook, sia per aprire nuove prospettive relative al ruolo di editore come fornitore di contenuti e servizi su di essi.

Librai che diventano editori

Come abbiamo già osservato, i libri digitali non sono di per sé una novità: se ne parla da più di un decennio, ma solo adesso sembra che la tecnologia sia sufficientemente matura da portare a una reale diffusione degli ebook.

Il Big Bang dell'editoria digitale è avvenuto pochi anni fa, quando un device più economico rispetto a qualunque altro allora sul mercato e con tecnologia wireless integrata ha incontrato una massa di contenuti tale da renderne finalmente davvero interessante l'acquisto. Parliamo di Kindle e dell'immensa libreria digitale messa in piedi dal suo stesso promotore: Amazon.

Rivoluzione nella rivoluzione, è un *libraio* che ha dato avvio a tutto il processo, proponendo un dispositivo veramente efficace e alla portata di molti più lettori, nonostante i suoi numerosi limiti, e occupandosi di creare e distribuire anche i contenuti adatti per il suo utilizzo.

Amazon

Nata come libreria online, Amazon (http://www.amazon.com) è ora la compagnia di commercio elettronico più grande degli Stati Uniti. Grazie ad accordi con numerosi editori, Amazon ha potuto accedere al loro catalogo, digitalizzarlo e metterlo in vendita secondo le sue politiche di prezzo, rendendolo così immediatamente disponibile per il download diretto su Kindle (Figure 5.4 e 5.5).

amazon.com

Hello. Sign in to get personalized recommendations. New customer? Start here.

Your Amazon.com | Today's Deals | Gifts & Wish Lists | Gift Cards

Your Account | Help

Shop All Departments | Search | Categories | Cart | Wish List

Kindle Store | Buy A Kindle | Kindle Books | Newspapers | Blogs | Magazines | Accessories | Discussions | Manage Your Kindle | Kindle Support

Amazon Kindle's Publishing Program

Thank you for your interest in Amazon Kindle's Publishing Program. All the tools to create and upload your content are free of charge. Once you make your titles available for Amazon Kindle, you'll receive payments for every title sold. Depending on the number of publications you'd like to convert, your technical expertise and your existing relationship with Amazon, there are several ways to make your content available:

Kindle Books

- Small publishers and authors can take advantage of Amazon's self-service tools through the Amazon.com Digital Text Platform (DTP).
- Publishers who have a relationship with an Amazon Kindle vendor manager should read more about Kindle Title Manager.
- Publishers and individuals who are comfortable working with HTML can use **KindleGen** (see the **KindleGen** section below) to create content, and eBookbase to upload files.
- Publishers who have many titles to convert, but may not have the time or technical expertise, might want to consider outsourcing to a conversion house.

For more information on any of these options, or to see a complete explanation of how to make a book available on Amazon Kindle, please see the Amazon Kindle Publishing Guidelines.

If you have further questions regarding how to sell your books on Amazon Kindle, please contact us at digitalrights@amazon.com.

Kindle Newspapers, Magazines, Blogs & News Feeds

- Publishers can self-publish content as a blog or news feed using Kindle Publishing for Blogs beta.
- Publishers interested in making other types of periodicals available for Amazon Kindle should contact digital-publications@amazon.com.

Figura 5.4 La pagina su Amazon sui progetti per pubblicare e distribuire sul suo sito libri digitali.

digitaltextplatform beta | My Shelf | My Reports | My Account | Help | Sign out

amazon.com

Hello, lethia.

Filter | Search

Add new item

Title | Author | Publication Date | Status

New Title 1 | Draft

1. Enter Product Details — Incomplete
2. Confirm Content Rights — Incomplete
3. Upload & Preview — Media not uploaded
4. Enter Suggested Retail Price — Incomplete

Enter Product Details

Title * New Title 1

Description

Authors * click Add/Edit to add contributors | Add/Edit

Publisher

ISBN

Language * <Select language> | Pub Date

Search Keywords enter keywords separated by commas, e.g. dogs, training

Categories * click Add/Edit to add categories | Add/Edit

Edition Number

Series Title

Series Volume

DRM * Enable digital rights management (DRM) / Do not enable digital rights management

Product Image

Terms and Conditions | Privacy Notice © 1996-2007, Amazon.com, Inc. or its affiliates

Figura 5.5 Il servizio Digital text platform permette a chiunque di pubblicare e distribuire libri digitali su Amazon.

La digitalizzazione di massa di titoli da parte di Amazon è stata possibile anche grazie ai vantaggi economici per gli editori derivanti dagli accordi stretti con la libreria online riguardo i ricavi su ogni copia venduta, molto più elevati rispetto a quelli abituali sulle copie stampate. Amazon vende l'assoluta maggioranza dei suoi ebook entro il prezzo di $ 9.99, che rende difficile la concorrenza.

Sono tuttavia in atto alcuni cambiamenti rispetto alla posizione se non di monopolio certamente predominante da parte di Amazon nel mercato degli ebook. Alla luce delle preoccupazioni da parte degli editori nel lasciare troppa autonomia nella gestione

del loro catalogo digitale a un unico attore, altri distributori di contenuti si sono fatti avanti e cercano di intaccare la posizione di mercato di Amazon. Sony con un nuovo store, altre grandi catene di librerie fisiche, come Barnes & Noble, di cui parleremo nel paragrafo successivo: si cerca in generale una via per sottrarre mercato e contenuti al padrone di Kindle.

Per questo Amazon stessa sta giocando d'anticipo: in previsione dell'ingresso di Apple come produttore di dispositivi (iPad) e librerie online (l'annunciato iBook Store), intende rivedere gli accordi con gli editori nel corso del 2010, per rendere ancora più vantaggiosa la distribuzione tramite la sua libreria e assicurarsi la possibilità di utilizzare tecnologie come l'innovativo

sintetizzatore vocale *Text-to-Speech* integrato nei suoi ereader.

Barnes & Noble

Barnes & Noble, la maggiore catena di librerie degli Stati Uniti, è una diretta concorrente di Amazon sia come libreria online sia come distributore di libri digitali (Figura 5.6). Anche Barnes & Noble produce e commercializza un ereader, Nook (ne parleremo nel Capitolo 6), ha un'ampia proposta di ebook e crea applicazioni per diversi dispositivi mobili (come iPhone e BlackBerry).

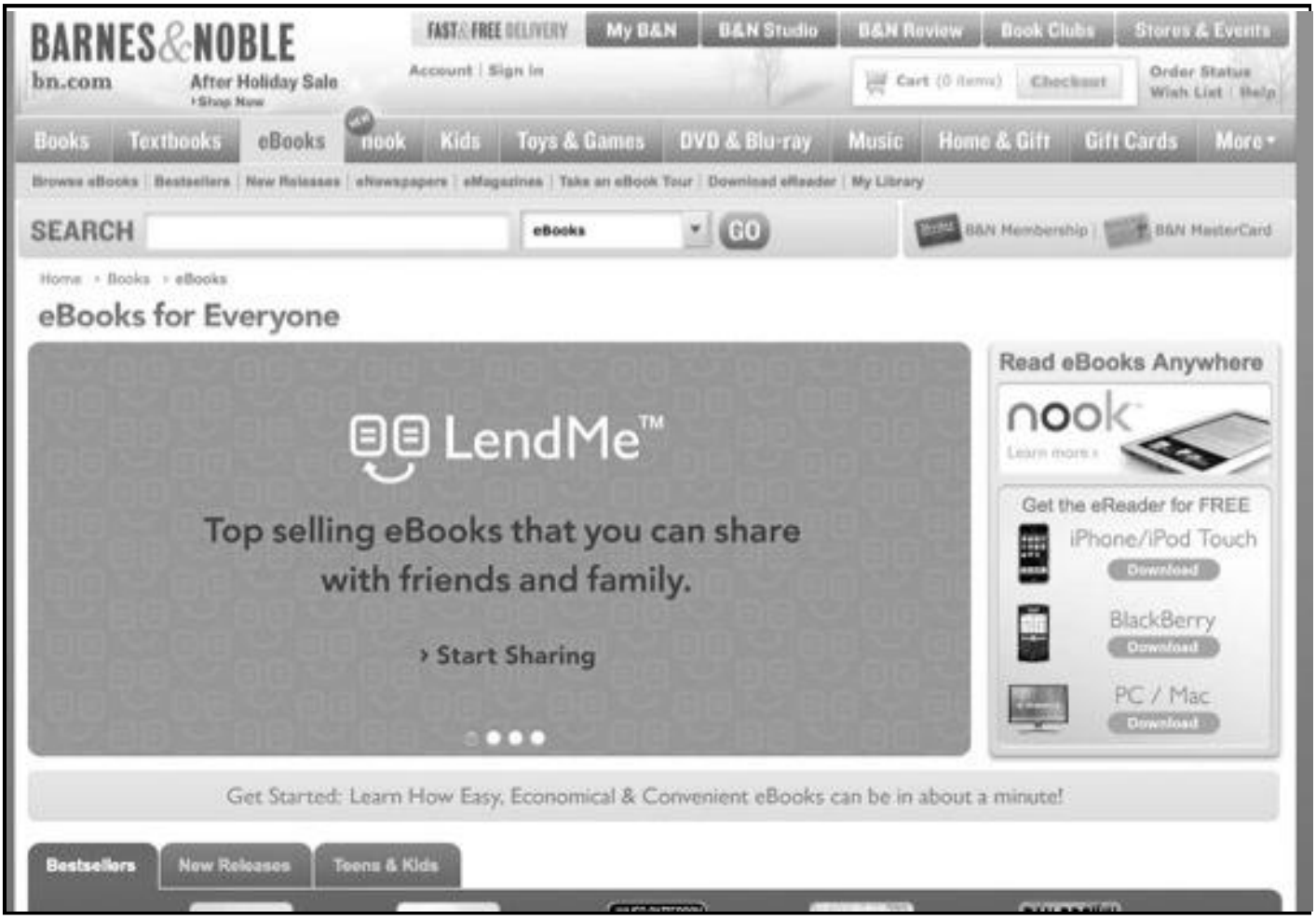

Figura 5.6 La pagina di Barnes & Noble dedicata agli ebook.

Nella sezione dedicata alla vendita dei libri digitali propone anche un gran numero di strumenti di supporto (FAQ, pagine di informazioni, aiuto), la possibilità di gestire il proprio account e i download e di scaricare le applicazioni per la lettura degli ebook per

PC, Mac e diversi dispositivi mobili. Anche Barnes & Noble pratica una politica dei prezzi che tende a mantenere il costo dei libri digitali per i lettori inferiore ai dieci dollari; come Amazon, inoltre, dedica una sezione del sito al suo ereader, Nook.

Autori editori di se stessi

I libri digitali permettono a tutti, in maniera sempre più immediata, di autopubblicare le proprie opere. In qualunque forma si intenda distribuirle – come applicazioni per iPhone o ebook in ogni tipo di formato – il digitale consente di annullare la mediazione dell'editore e di pubblicare direttamente libri senza passare per il suo filtro. L'autopubblicazione demanda esplicitamente il compito di selezionare i contenuti ai lettori.

Se questo può essere un vantaggio, perché offre infinite possibilità per la circolazione delle idee e un'ideale "democrazia editoriale" in cui ognuno può raccontare la propria storia, in realtà non fa che aumentare il sovraccarico di dati a cui siamo costantemente sottoposti, in cui facciamo spesso fatica a trovare le informazioni che ci interessano, perdendole nel "rumore" di quelle inutili, scorrette o poco accurate.

La mediazione dell'editore è ancora necessaria sotto diversi punti di vista. In primo luogo come garante della qualità dei contenuti: l'editore si assicura dell'accuratezza delle informazioni e dell'organicità del testo, oltre che della sua cura redazionale e formale.

I servizi che permettono agli autori di diventare editori di se stessi sono sempre più numerosi: come abbiamo osservato (Figura 5.5), Amazon stessa permette a chiunque di pubblicare e distribuire sulla sua libreria online libri digitali scaricabili direttamente su Kindle. Nei due paragrafi seguenti vedremo invece degli esempi di siti nati appositamente per l'autopubblicazione di opere inedite.

Smashwords

Smashwords (https://www.smashwords.com/) è una piattaforma di pubblicazione e distribuzione di ebook pensata appositamente per gli autori, dove questi ultimi hanno – gratuitamente – pieno controllo delle loro opere, del prezzo e delle possibilità di promozione. L'autore

ricava da ogni copia venduta intorno all'85% del prezzo di copertina: molto più di quanto guadagnerebbe con qualunque editore. Smashwords è presente nella sezione *eBooks* di Barnes & Noble e nei cataloghi delle maggiori applicazioni di lettura su iPhone (Stanza ed eReader, per esempio). Il sito distribuisce anche le opere digitali dei cataloghi degli editori che ne fanno richiesta.

Lulu

Lulu (http://www.lulu.com) è un'azienda che offre l'opportunità di autopubblicare le proprie opere sia in formato cartaceo sia digitale. Permette la pubblicazione multiformato, la possibilità di pubblicare attraverso un pacchetto che comprende sia il libro stampato sia l'ebook e fornisce all'autore un completo sistema di strumenti di supporto per

scegliere al meglio la soluzione adatta alle proprie esigenze (Figura 5.7).

Figura 5.7 La sezione eBook su lulu.com.

Quando arriva anche il motore di ricerca

Nel primo capitolo abbiamo fatto alcune considerazioni riguardo al ruolo dei motori di ricerca che possono essere intesi sotto molti aspetti come nuovi editori. Nel prossimo futuro potrebbe non trattarsi di una semplice definizione ma di un dato di fatto. Vediamo come.

Google Editions

In occasione della fiera del libro di Francoforte del 2009, Google ha rivelato i suoi piani commerciali riguardo l'attività di digitalizzazione dei libri del progetto Google Books (http://books.google.com). Tali piani riguardano l'apertura – nel primo semestre del 2010 – della li breria digitale di Google,

con il nome di Google Editions. Con questa soluzione gli editori partner di Google potranno mettere in vendita i loro ebook attraverso modalità di pagamento di cui il motore di ricerca stesso si farà garante, potenziando gli account dei propri utenti.

Google annuncia di non avere intenzione di permettere il download dei contenuti, ma di voler rendere disponibile l'accesso agli stessi da qualsiasi dispositivo – dallo smartphone al notebook, passando per l'ereader – in grado di connettersi alla Rete.

Si tratta di una prospettiva su cui tenere l'attenzione puntata: considerata la diffusione degli strumenti di ricerca di Google e la loro semplicità di utilizzo, l'aggiunta di una libreria online ai servizi già esistenti sarà piuttosto interessante, specialmente

pensando a una eventuale sinergia con altri progetti come Google Docs.

Ma in effetti cosa vendo?

Se, come abbiamo appena visto, il libro digitale abbatte ogni intermediario tra autore e pubblico e chiunque è potenzialmente in grado di pubblicare il suo libro senza bisogno di dover passare da un editore, viene da interrogarsi su quale sia – o su quale debba diventare – il ruolo di quest'ultimo nel nuovo mercato dei libri digitali.

Oltre all'autorevolezza, che continuerà a rimanere appannaggio dell'editore, in quanto indiscusso selezionatore e garante dei contenuti e della cura formale e strutturale del testo, per l'editore si aprono nuove

ipotesi per ridefinire e valorizzare le sue competenze e il suo peso culturale.

Con l'inserimento dei libri digitali all'interno della propria linea editoriale, si potranno trovare nuove strade per la diffusione dei contenuti che permettano di sfruttare al massimo le opportunità offerte dal media digitale, affiancandolo nel modo più efficace alle pubblicazioni su carta.

Inoltre, come abbiamo visto nel caso di O'Reilly Media, l'editore potrà proporsi come un fornitore di servizi sul contenuto stesso: ciò significa studiare le nuove esigenze di lettura e riuscire a intercettare le necessità dei propri destinatari, proponendo strumenti e supporti che rendano l'esperienza di lettura quanto più semplice, divertente e

proficua possibile, a seconda del target di riferimento.

DISPOSITIVI PER LA LETTURA DEGLI EBOOK: EREADER

I primi dispositivi per la lettura degli ebook fanno la loro comparsa nel 1999. Da allora la tecnologia ha compiuto enormi passi avanti: se l'editoria digitale sta vivendo un momento di particolare vitalità è dovuto anche alla diffusione di dispositivi sempre più efficienti e con caratteristiche di volta in volta più in sintonia con le esigenze di lettura.

In questo capitolo prenderemo in esame alcuni tra i dispositivi più interessanti in

commercio al momento in cui si scrivono queste pagine. Dal discorso sui dispositivi possono partire spunti di riflessione che vanno oltre alle caratteristiche tecniche, che rischiano di diventare obsolete a velocità sorprendenti. Entriamo nel merito.

Il libro sullo schermo

Il punto di partenza per affrontare il tema dei dispositivi tecnologici per la lettura digitale è stato in parte affrontato nel Capitolo 1, quando abbiamo chiarito come anche il libro stesso sia in realtà una tecnologia e non un mezzo naturale per l'espressione dei contenuti. Il vantaggio del libro è la consuetudine che ha conquistato nelle nostre abitudini culturali: leggere un libro è

un’azione talmente comune da essere considerata quasi scontata. Il libro è un oggetto così familiare da permettere una completa immedesimazione e concentrazione nella lettura, dimenticando il mezzo per favorire la totale attenzione sul contenuto. Questo è il primo ostacolo che un ereader deve affrontare: si tratta di una tecnologia che non può contare sul senso di familiarità che il libro ha costruito in più di cinquecento anni di diffusione.

Oltre agli aspetti in qualche modo soggettivi che determinano la diffidenza nei confronti della lettura da un dispositivo elettronico (sensazioni tattili, olfattive, un rapporto “fisico” col libro, che può essere manipolato in vari modi, scarabocchiato e così via), gli ereader incontrano anche delle difficoltà oggettive dettate dal tipo di

tecnologia che utilizzano: una tra queste è l'affaticamento della vista dato dalla lettura prolungata su schermi retroilluminati. A questo scopo è stata sviluppata una tecnologia in grado di ridurre tale difficoltà: l'inchiostro elettronico.

Epaper: quasi come la carta

L'epaper è una tecnologia che simula la resa dell'inchiostro su carta su uno schermo. A differenza dei monitor comuni che utilizzano la retroilluminazione per rendere visibili i pixel, l'epaper riflette la luce come la carta normale e visualizza testo e immagini per un periodo di tempo indeterminato senza consumare elettricità.

L'epaper, rispetto ai display convenzionali, permette una lettura più confortevole:

questo è dovuto alla stabilità dell'immagine sullo schermo, che non ha bisogno di essere aggiornata costantemente, e soprattutto al fatto che lo schermo riflette la luce ambientale anziché emetterne lui stesso. Un display con tecnologia epaper può essere letto tranquillamente anche in piena luce solare senza che l'immagine perda di luminosità. Al momento in cui si scrive, sull'epaper non è possibile visualizzare colori.

Riferimenti *L'inventore di questa tecnologia è stato Joe Jacobson nel 1996, fondatore della compagnia E-Ink.*

Una delle tecniche più diffuse per l'inchiostro elettronico prevede l'utilizzo di piccole sfere all'interno dello schermo. Ciascuna sfera è colorata per metà di bianco e

per metà di nero e le due parti possiedono rispettivamente una carica elettrica negativa e una positiva. Grazie a dei campi elettrici è possibile orientare le sfere e ottenere così che le parti interessate dello schermo cambino colore. In questo modo si possono realizzare supporti sottili che richiedono alimentazione solamente nel momento in cui la configurazione delle sfere dev'essere variata: i dispositivi riescono così a essere molto leggeri e ad avere un'autonomia piuttosto elevata, valutabile in voltate di pagina (l'unico momento in cui il dispositivo consuma energia muovendo le sfere) anziché in tempo di accensione.

Specializzati o multifunzione?

Per leggere libri in formato digitale esiste un’infinita varietà di soluzioni. I vincoli principali sono la compatibilità del formato del file con l’hardware e il software che dovrà interpretarlo. In questo capitolo ci occuperemo dei dispositivi per la lettura degli ebook distinguendo tra dispositivi dedicati, ossia con funzioni mirate in modo quasi esclusivo alla lettura digitale, e dispositivi multifunzione, in grado di svolgere molte altre operazioni e allo stesso tempo di proporre sufficienti strumenti per un’esperienza di lettura completa e soddisfacente.

Dispositivi dedicati

Come abbiamo osservato più volte, i dispositivi dedicati sono quelli progettati specificamente per la lettura di libri elettronici. Possono svolgere altre funzioni, ma generalmente si tratta di attività utili o complementari alla lettura. Nei paragrafi seguenti vedremo i modelli più diffusi e interessanti soffermandoci più sulle funzionalità di lettura che sui dettagli tecnici di ciascuno.

Kindle

Kindle è l'ereader sviluppato da Amazon e lanciato sul mercato statunitense nel 2007 (Figura 6.1). Il nome indica sia l'hardware sia il software utilizzato tanto sul dispositivo

quanto sul computer con il quale si intende sincronizzarlo. Esistono attualmente tre modelli di Kindle: Kindle, Kindle 2 e Kindle DX, mentre il software è disponibile per varie piattaforme, inclusa un'applicazione per iPhone. Nell'ottobre 2009 Kindle 2 è stato lanciato anche sul mercato internazionale, con alcune differenze specialmente riguardo alla connessione senza fili.

Kindle utilizza la tecnologia E-Ink per lo schermo, e permette di scaricare contenuti attraverso Amazon Whispernet, utilizzando, negli Stati Uniti, la rete a banda larga fornita dall'operatore Sprint; la versione internazionale utilizza quella di AT&T.

Riferimenti *Amazon Whispernet è una connessione wireless pensata appositamente per Kindle, e consente di accedere al Web per scaricare libri e*

altri contenuti senza dover passare da un provider Internet esterno. Negli Stati Uniti Whispernet è accessibile gratuitamente; la versione internazionale sfrutta la tecnologia 3G integrata nel dispositivo.

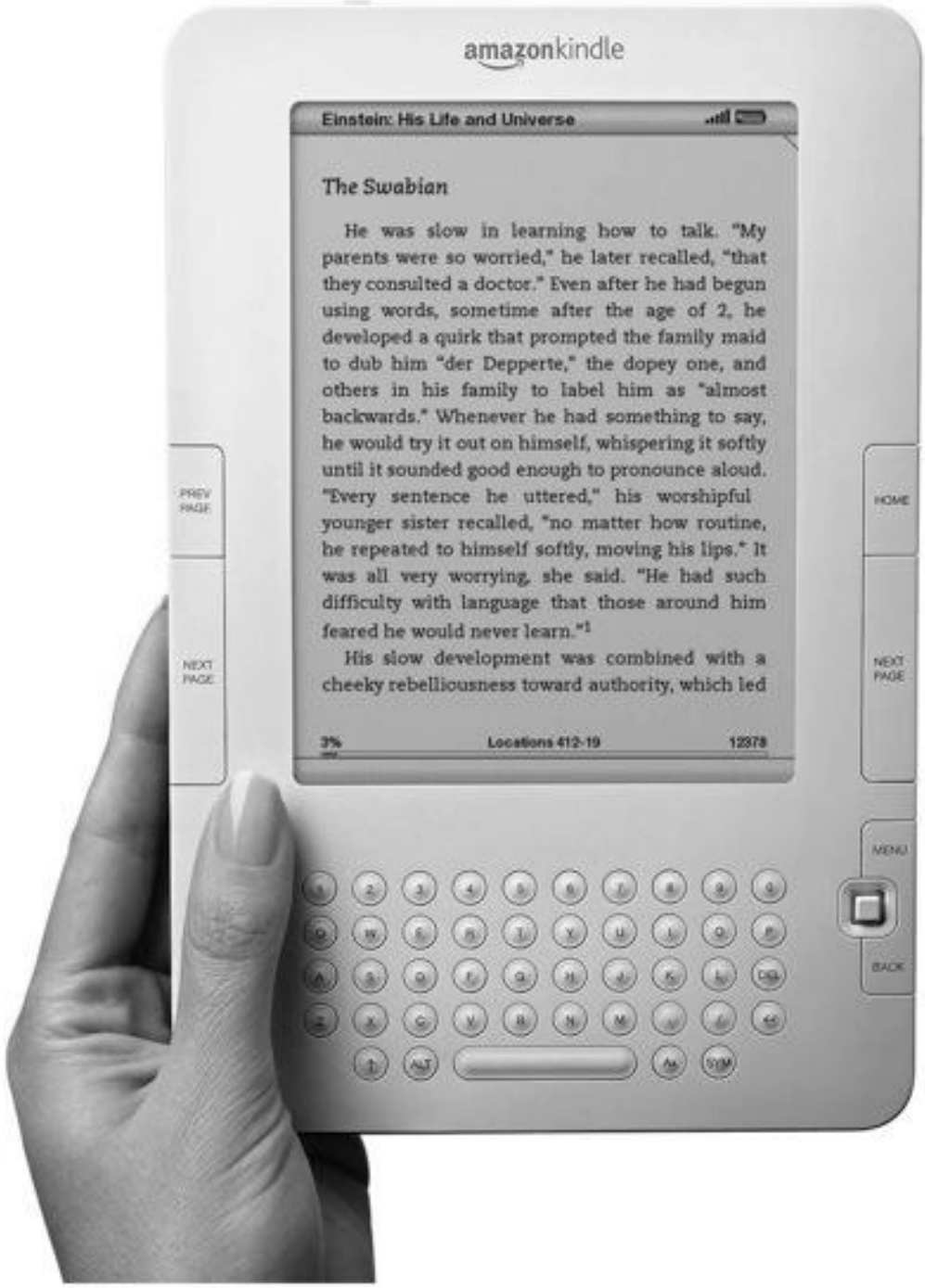

Figura 6.1 Kindle 2.

Il software

Kindle, come abbiamo detto, è anche il nome del software installato sull'ereader Amazon; ma allo stesso tempo si riferisce anche a una serie di applicazioni che permettono di sincronizzare il contenuto di Kindle con altre tecnologie. In particolare:

- l'applicazione Kindle per iPhone e iPod touch consente di sincronizzare i contenuti di Kindle e del proprio account su Amazon con i due dispositivi Apple attraverso la tecnologia Whispersyn. Non è tuttavia possibile acquistare i libri direttamente dall'applicazione;
- Kindle per PC è un'applicazione gratuita per leggere i libri scaricati con l'ereader anche sul computer, sincronizzando gli acquisti effettuati su Amazon con il software.

Il dispositivo

Kindle 2, lanciato sul mercato nel febbraio del 2009, possiede uno schermo da 6" con tecnologia E-Ink (16 livelli di scala di grigi) e una memoria interna da 2 GB non espandibili (di cui 1,4 GB utilizzabili dall'utente) in grado di contenere circa 1500 libri non illustrati. Nell'ottobre del 2009 Amazon ha lanciato un modello di Kindle 2 pensato appositamente per il mercato internazionale: la difficoltà maggiore in questo caso è costituita dalla disponibilità della rete telefonica per l'utilizzo della connessione a Internet.

Riferimenti *Il modello DX di Kindle, disponibile anche per il mercato internazionale, possiede uno schermo di 9,7 " e una memoria da 4 GB non espandibili (di cui circa 3,3 GB utilizzabili). Integra anche un accelerometro*

che può modificare l'orientamento dello schermo a seconda dell'orientamento del dispositivo.

La differenza tra il Kindle 2 per il mercato statunitense e quello per il mercato internazionale, dal punto di vista hardware, sta proprio nell'utilizzo di un differente standard di connessione mobile. Oltre a questo, mentre la versione americana permette di navigare sull'intero Web attraverso il browser sperimentale integrato nel dispositivo, la stessa cosa non è possibile con i Kindle acquistati in Europa.

Riferimenti *A differenza delle reti Wi-Fi, Whispernet non necessita della connessione a un punto di accesso per navigare sul Web. La connessione (negli Stati Uniti) è gratuita grazie agli*

accordi tra Amazon e il gestore della rete (Sprint).

I contenuti per Kindle possono essere scaricati direttamente dal Kindle Store di Amazon nel formato AZW; possono anche essere caricati sull'ereader contenuti in altri formati, come per esempio il mobi. I termini di utilizzo degli AZW impediscono il trasferimento dei file ad altri utenti. È possibile sottoscrivere abbonamenti a numerosi quotidiani e riviste e blog disponibili in formati adatti per la lettura su Kindle.

Kindle supporta fin dalla sua prima versione gli ebook nel formato Mobipocket non protetto (MOBI o PRC), semplici file di testo (TXT) e naturalmente il formato proprietario di Amazon (AZW) coperto da DRM. Gli ePub non sono supportati dal Kindle 2, mentre un recente aggiornamento del

firmware (novembre 2009), permette di leggere i PDF nativamente.

Ogni Kindle dispone di un indirizzo di posta elettronica, personalizzabile: questo viene sfruttato per il servizio via mail che consente di convertire diversi formati immagine (JPG, GIF, PNG e BMP) in AZW, così come file HTML e DOC. Il servizio si occupa di inviare all'account di posta elettronica o direttamente sull'ereader il file formattato correttamente per Kindle (nel secondo caso dietro una formula di pagamento stabilita in base al peso del file).

Riferimenti *Kindle supporta anche file audio in formato MP3 e audiolibri Audible, che devono essere trasferiti sulla memoria di Kindle tramite cavo USB.*

A partire dal Kindle 2 è possibile scaricare libri nei formati diversi dall'AZW utilizzando il browser sperimentale di cui il dispositivo è fornito (solo negli Stati Uniti, per ora). I collegamenti ipertestuali all'interno dei file in questi formati possono essere utilizzati per "linkare" risorse online ma non libri conservati nella memoria di Kindle.

Si possono inserire segnalibri e appunti sui contenuti salvati su Kindle. Mentre si sta leggendo una pagina, un menu di opzioni permette di accedere alla ricerca sul dizionario integrato. Il software del dispositivo, inoltre, è in grado di ricordare l'ultima pagina letta di ogni libro aperto. Le pagine possono essere salvate in forma di appunti ciascuna in un singolo file di testo poi trasferibile tramite cavo USB.

I libri acquistati sul Kindle Store sono sempre disponibili e riscaricabili in qualsiasi momento. Il backup del dispositivo comprende gli appunti presi sugli ebook acquistati e la memorizzazione dell'ultima pagina letta.

In via sperimentale, Kindle supporta una funzione di lettura automatica dei libri, chiamata *Text-to-Speech*. I contenuti digitali vengono interpretati automaticamente da una voce maschile o femminile di cui si può regolare la velocità di lettura; le pagine scorrono automaticamente con il procedere del testo letto. Ci si può spostare nel testo avanti e indietro, salvando anche dei segnalibri.

Riferimenti *Nella lettura di riviste e quotidiani, Kindle consente di salvare integralmente gli articoli che ci interessa leggere in un secondo*

momento. Sottoscrivendo un abbonamento, gli aggiornamenti della pubblicazione verranno scaricati automaticamente con cadenza regolare corrispondente all'uscita cartacea.

Sull'ereader è presente un dizionario integrato, il *New Oxford American Dictionary*. Le definizioni delle parole compaiono automaticamente nella parte bassa dello schermo posizionando il cursore sul termine interessato. È inoltre possibile consultare Wikipedia attraverso un accesso predefinito.

Amazon ha annunciato nel gennaio 2010 di voler aprire Kindle ad applicazioni sviluppate da terze parti in vendita attraverso un negozio online dedicato, in maniera molto simile a quanto Apple fa per iPhone con l'App Store, come vedremo nelle pagine seguenti. Il tentativo – a parte il chiaro

intento di concorrenza riguardo ai dispositivi Apple – mostra interesse verso un ampliamento delle possibilità di Kindle attraverso una maggiore interattività.

Pro e contro

Il lancio di Kindle ha rappresentato un vero e proprio scossone per il mercato dei libri digitali. È stato il primo lettore a essere veramente portatile e accessibile a una fetta più vasta di pubblico rispetto ai concorrenti allora molto più costosi in circolazione. Ma senza dubbio il vantaggio più grande è stata l'associazione con la vastissima libreria di titoli digitali di Amazon, unita all'accesso Wi-Fi allo store. I limiti riguardano principalmente il formato proprietario dei file AZW che ne vincola la fruizione alle soluzioni tecnologiche Amazon. Inoltre l'acquisto

diretto dal dispositivo è possibile solo sul Kindle Store.

Sony

Sony propone tre modelli di ereader dalle varie funzionalità e caratteristiche tecniche: i reader Touch, Pocket e Daily Edition.

Touch Edition – PRS-600SC

La versione Touch (Figura 6.2), come si può dedurre dal nome, sfrutta uno schermo sensibile al tocco (6") e la tecnologia E-Ink. Tra le funzioni, la possibilità di inserire segnalibri, modificare i font, utilizzare la tastiera virtuale o la penna inclusa per prendere appunti o sottolineare le parti del testo di cui si vuole tenere traccia; dispone inoltre di un dizionario integrato. Supporta PDF, DOC, TXT, vari formati di immagine ed ePub, e

può essere sincronizzato con il software Digital Editions di Adobe (se ne è parlato nel Capitolo 3). Compatibile sia con PC sia con Mac, è anche in grado di riprodurre file audio compressi nei formati MP3 e AAC. La memoria integrata di 512 MB è espandibile grazie allo slot per le schede SD o per le Memory Stick PRO Duo. Non è dotato di connettività Wi-Fi, ma grazie al Reader Library Software – di cui parleremo tra poco – può accedere allo store di Sony e alla collezione di alcune biblioteche.

Pocket Edition – PRS-300SC

Questo modello (Figura 6.2) ha uno schermo leggermente più piccolo (5") e senza tecnologia touch. Supporta gli stessi formati della versione Touch e può essere anch’esso sincronizzato con l’applicazione

Reader Library Software, oltre che con Adobe Digital Edition. Rispetto alla versione Touch le funzionalità di lettura sono limitate alla sola aggiunta di segnalibri: non è possibile prendere appunti o evidenziare parti del testo. Non dispone del dizionario integrato né della possibilità di riprodurre file audio o immagini. La memoria integrata non è espandibile.

Daily Edition – PRS-900SC

La versione Daily (Figura 6.2) del reader Sony è l'unica a disporre di un accesso Wi-Fi alla Rete (solo negli Stati Uniti, per ora), che consente di accedere allo store di Sony, ad alcune biblioteche digitali e a Google Books per scaricare nuovi libri senza bisogno di utilizzare il computer. Permette anche di abbonarsi a quotidiani e riviste. Lo schermo,

con tecnologia touch, è leggermente più ampio rispetto agli altri modelli: 7".

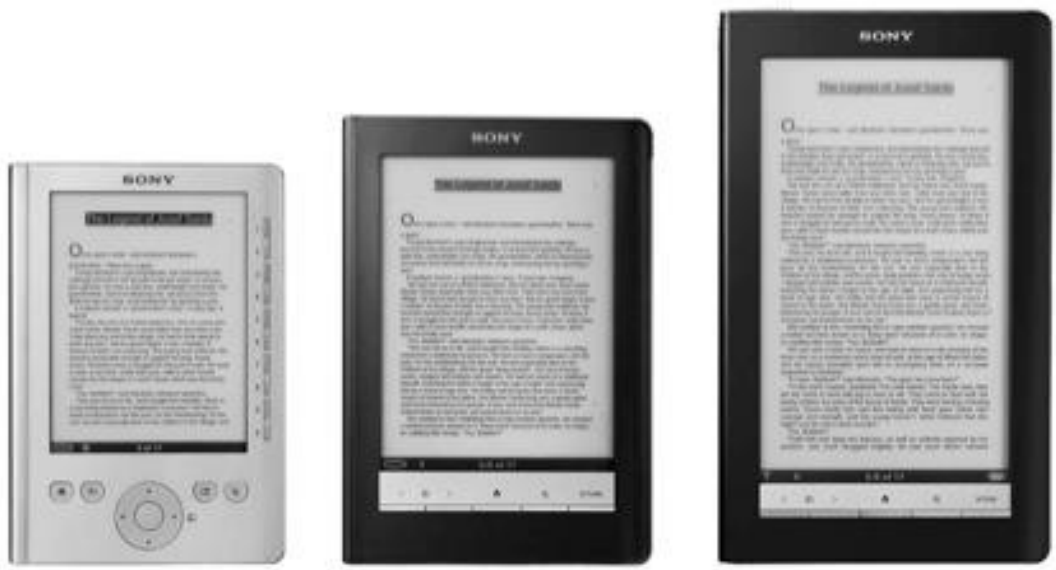

Figura 6.2 Da sinistra verso destra, i lettori Sony Pocket, Touch e Daily Edition.

Reader Library Software

L'applicazione, utilizzabile gratuitamente e a prescindere dal possesso di un ereader Sony, è scaricabile all'indirizzo

http://ebookstore .sony.com/download/. Grazie al software è possibile accedere sia allo store, da cui acquistare le ultime novità pubblicate, sia – se si è in possesso della tessera di iscrizione – al prestito di alcune biblioteche. Inoltre, grazie a un accordo con Google, si può accedere al servizio Google Books e scaricare sul computer titoli liberi dal diritto d'autore, trasferirli direttamente sull'ereader o leggerli dal computer, dopo aver attivato un account. Grazie a Reader Library si possono trasferire gli appunti presi con il reader sul computer.

Riferimenti *Attraverso il software è possibile anche abbonarsi a dei feed RSS trasferibili poi sul lettore.*

Pro e contro

I lettori Sony hanno dalla loro parte la possibilità di accedere a più librerie online: oltre allo store dedicato, a Google Books e alla sua sterminata collezione di titoli gratuiti perché non più coperti dal diritto d'autore, Sony ha stretto accordi con Borders, la seconda libreria degli Stati Uniti dopo Barnes & Noble (di cui parleremo più avanti). Sui dispositivi di fascia bassa il principale limite è l'assenza di connettività Wi-Fi.

IREX

IREX è un produttore europeo di ereader che si differenzia dai suoi concorrenti per la creazione di una tecnologia proprietaria per gli schermi, proponendo fin dai primissimi

modelli monitor con 16 livelli di scala di grigio e tecnologia touch.

Gli ereader di IREX sono iLiad Book Edition e IREX Digital Reader.

Riferimenti *Il primo iLiad è stato prodotto nel 2006: IREX è stato un pioniere riguardo alla tecnologia per gli ereader. Nella seconda versione, l'anno successivo, i miglioramenti puntarono soprattutto a una maggiore durata della batteria.*

iLiad Book Edition

iLiad Book Edition permette di annotare documenti e di salvare e trasferire appunti su PC o su un altro iLiad; oltre ai libri digitali, è possibile utilizzarlo per l'enigmistica, la lettura di fumetti, per prendere appunti

come su un blocco note o fare schizzi e disegni grazie allo schermo touch. La memoria è espandibile attraverso schede Compact Flash o MMC; si può inoltre collegare al lettore una memoria USB e consultarne direttamente il contenuto.

iLiad supporta i formati PDF, HTML, TXT, JPG, BMP, PNG e PRC (Mobipocket) e viene venduto con cinquanta titoli precaricati liberi dal diritto d’autore.

IREX Digital Reader

La serie Digital di IREX si distingue soprattutto per le dimensioni dello schermo (10,2"). Integra la connettività wireless e uno schermo con tecnologia touch e inchiostro elettronico. Comprende una memoria SD da 1 GB, sostituibile con schede con capacità maggiori.

Nel settembre 2009 è stato lanciato negli Stati Uniti un nuovo modello della serie Digital (il DR800, Figura 6.3), con uno schermo da 8,1" (come iLiad), molto compatto e senza alcun pulsante a eccezione della flipbar, che permette di voltare le pagine. Tra i formati supportati figurano PDF ed ePub. Il reader è in grado di connettersi alla Rete grazie a un modulo 3G e a un accordo con Verizon. Il DR800, grazie al sistema operativo Linux, consente di installare sul dispositivo anche reader sviluppati da terzi, espandendo così le funzionalità del lettore: è il caso del FBReader, di cui si è parlato nel Capitolo 3.

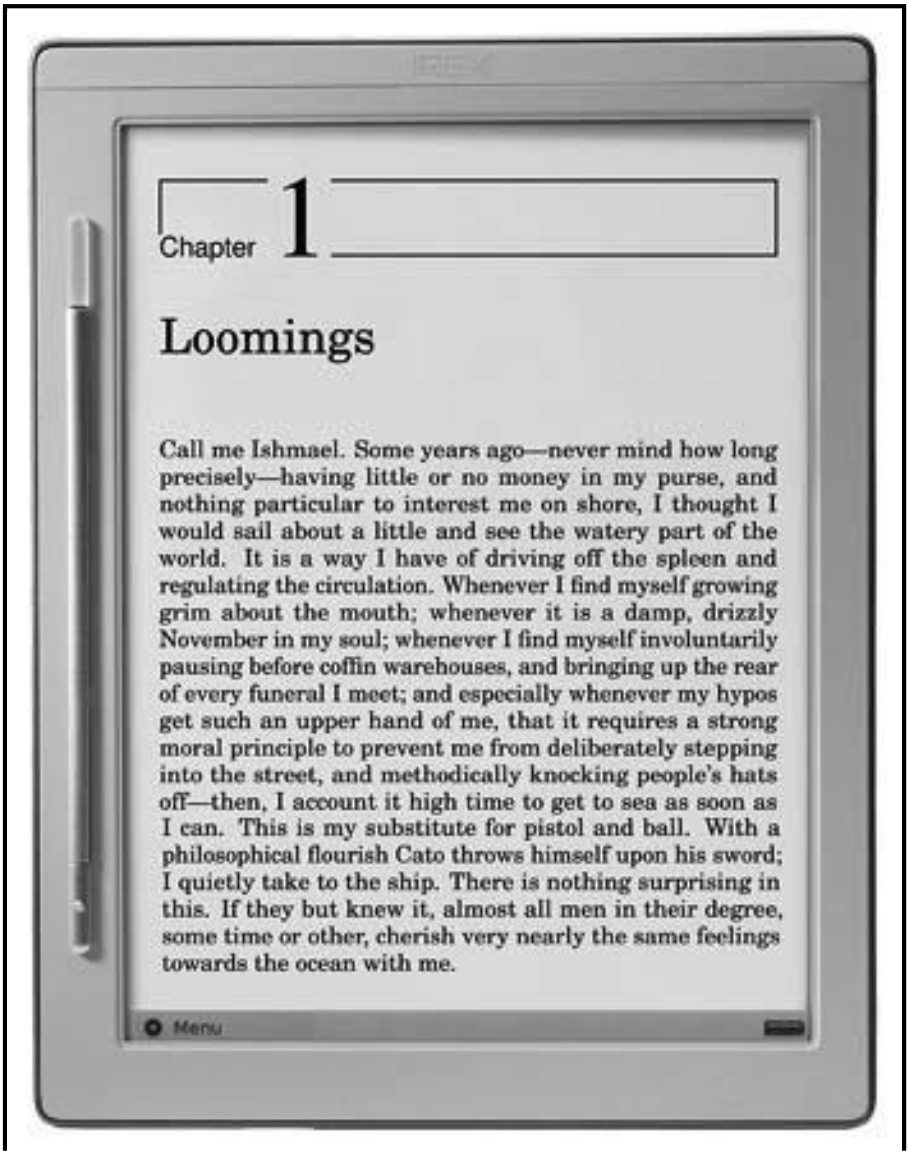

Figura 6.3 Il modello DR800.

Riferimenti *Il modello di connessione gratuita e diretta dal dispositivo ricalca l'accordo tra Sprint e Amazon per Kindle: anche Verizon è un operatore di telefonia mobile statunitense.*

Il reader può connettersi direttamente allo store di Barnes & Noble per l'acquisto di ebook, e a quello di NewsPaperDirect per i quotidiani e le riviste. Gli accordi con i due store non sono esclusivi, al contrario di quanto avviene tra Amazon e Kindle: ciò rende possibile futuri ampliamenti nell'offerta accessibile direttamente dal dispositivo, nel momento in cui altri store volessero farne parte.

Pro e contro

Ricchi di funzionalità sofisticate e di forti potenzialità specie per quanto riguarda la creazione di contenuti da parte del lettore, gli ereader di IREX hanno come principale svantaggio il prezzo, il più elevato nel panorama dei dispositivi dedicati. Le dimensioni, anche se in molti casi possono essere un

vantaggio (specie nella visualizzazione di formati statici come il PDF), diminuiscono la portabilità dei dispositivi, rendendoli più adatti a un uso professionale o come strumenti di studio.

Bookeen

Bookeen è una compagnia francese produttrice di ereader creata nel 2003 da due veterani del settore degli ebook reader (Laurent Picard e Michaël Dahan), impegnati sulla tecnologia fin dal 1998. Attualmente la compagnia ha due modelli di punta: il Cybook Gen3 e il Cybook Opus.

Cybook Gen3

Rilasciato nel 2007, il Cybook Gen3 (Figura 6.4) sfrutta la tecnologia E-Ink e supporta un gran numero di formati: PDF,

MOBI, PRC, PalmDOC (PDB), HTML, TXT, JPG, GIF, PNG e anche file audio in MP3. Riguardo ai file protetti, è in grado di visualizzare i Mobipocket coperti da DRM e, con un aggiornamento del firmware, anche ePub o PDF con protezione Adobe, perdendo però così il supporto ai Mobipocket coperti da DRM.

Riferimenti *L'incompatibilità tra i DRM Adobe e Mobipocket è data dalla limitazione stabilita da quest'ultima a rappresentare file MOBI protetti su dispositivi dedicati per la lettura di ebook che prevedono il supporto di altri formati DRM, come appunto quello di Adobe.*

Il Cybook Gen3 viene venduto con numerosi titoli in PDF prevaricati da FeedBooks

(http://www.feedbooks.com), un sito da cui si possono scaricare molti titoli di pubblico dominio e opere originali di autori esordienti.

Una delle caratteristiche più interessanti del dispositivo è che consente di caricare font TrueType, ampliando la scelta di caratteri da utilizzare per la visualizzazione dei contenuti (massimo 12).

Riferimenti *Alcuni font OpenType funzionano, ma non sono supportati ufficialmente.*

Lo stile del lettore è decisamente minimalista, con uno schermo da 6" senza retroilluminazione. Il sistema operativo è basato su Linux: è open source e scaricabile dal sito Bookeen. Il software per la gestione del dispositivo del computer è il Bookeen Multi-

format eBook reader, che può essere utilizzato anche come reader desktop. La memoria interna è espandibile attraverso schede di memoria SD.

Cybook Opus

Il Cybook Opus (Figura 6.5), dal punto di vista del software, è molto simile al Gen3. Supporta un numero minore di formati (PDF ed ePub anche coperti da DRM), ha una memoria integrata da 1 GB espandibile con schede di memoria micro SD fino a 4 GB e fornisce la stessa possibilità di personalizzazione di layout e caratteri per la presentazione dei contenuti.

Tuttavia, le caratteristiche più interessanti del dispositivo stanno nelle sue dimensioni: appena 5" di schermo, un ereader decisamente tascabile. Inoltre dispone di un accelerometro: percependo l'orientamento del lettore modifica di conseguenza la visualizzazione della pagina.

Figura 6.5 Il Cybook Opus.

Pro e contro

Pensati per un lettore interessato alla lettura di libri digitali senza troppe esigenze ulteriori, gli ereader di Bookeen puntano tutte le loro risorse verso l'estrema semplicità sia nell'hardware sia nel software: design minimale e funzioni ridotte all'essenziale costituiscono allo stesso tempo la forza e il punto debole di questi lettori. Perfetti per chi vuole solo leggere e non è interessato ad acquistare e scaricare libri in meno di 60 secondi sul suo dispositivo, sfruttano la marcia in più data dall'open source.

Nook

Nook è il lettore lanciato dalla più grande libreria americana, Barnes & Noble. A differenza della maggior parte degli altri lettori ha

un doppio schermo: uno con tecnologia E-Ink per la lettura e un altro, nella parte inferiore del dispositivo, a colori e con tecnologia touch per navigare nel lettore e accedere allo store (Figura 6.6).

Figura 6.6 Nook: il secondo schermo nella parte bassa del dispositivo è touch e a colori.

Nook supporta nativamente i formati ePub e PDF. Inoltre permette di accedere alla Rete non solo sfruttando una connessione 3G ma anche tramite Wi-Fi, gratuita se si accede dagli hot spot delle librerie B&N.

La memoria interna è di 2 GB, espandibile tramite schede di memoria microSD. Consente di prendere appunti sui libri che si stanno leggendo e trasferirli insieme ai file su altri dispositivi. Nook utilizza Android come sistema operativo: è quindi open source. Inoltre può essere provato nelle librerie prima di essere acquistato.

Riferimenti *Nook è al centro di una controversia legale: Spring Design, società produttrice di un altro ereader – Alex – ha fatto causa a Barnes & Noble per appropriazione illecita di segreti industriali per quanto riguarda il*

design di Nook. Spring rivendica la paternità della tecnologia basata sul doppio schermo.

Pro e contro

Nook è un dispositivo particolarmente innovativo nel mercato degli ereader: il doppio schermo associato all'utilizzo di un sistema operativo open source come Android potrebbe rappresentare l'inizio di un'interessante serie di sperimentazioni. Attualmente non si ha notizia di un eventuale lancio sul mercato internazionale.

Altri ereader

Descrivere tutti gli ereader attualmente in circolazione non è possibile sia per ragioni di spazio sia per la continua evoluzione del mercato in questa fase. Se siete interessati a

scoprire una gamma più ampia di modelli, l'esplorazione può partire dalla pagina dei dispositivi compatibili con Adobe Digital Editions (http://www.adobe.com/products/digitaleditions/devices/), dove si trovano, per esempio, anche reader pensati per i mercati dell'Estremo Oriente e altri disponibili solo per il mercato statunitense. Ancora, il lemma di Wikipedia "List of e-book readers" (http://en.wikipedia.org/wiki/List_of_e-book_readers) propone un'interessante tabella riassuntiva di un gran numero di dispositivi con un confronto puntuale delle funzionalità. Tra i più interessanti vogliamo citare il COOL-ER e eSlick, quest'ultimo già citato nel Capitolo 3 a proposito del software di lettura per PDF Foxit.

Riferimenti *Nel gennaio 2010 si è svolto a Las Vegas il CES, un evento*

internazionale sul design, lo sviluppo e la distribuzione di prodotti elettronici, durante il quale sono stati presentati alcuni nuovi modelli di ereader. I più affermati sono stati illustrati in queste pagine, ma suggeriamo di tenere sotto controllo nuovi modelli interessanti in uscita nel corso del 2010.

Dispositivi multifunzione, ovvero la “Good Enough Revolution”

Robert Capp, Senior Editor di *Wired*, ha scritto nell’agosto del 2009 un articolo riguardo le nuove tecnologie dal titolo “The Good Enough Revolution: When Cheap and Simple Is Just Fine”. La tesi di fondo è che la

tecnologia ha dato una spinta tale alle nostre abitudini da cambiare in maniera radicale ciò che ci aspettiamo dai prodotti e dai servizi. Si preferiscono la flessibilità, la convenienza e la velocità, anche a fronte di una minor pulizia, ampiezza e dettaglio del prodotto. È un cambiamento trasversale, che non interessa solo ambiti direttamente coinvolti dalla tecnologia.

L'esempio più evidente riguarda l'industria musicale. Inizialmente, le case discografiche sottovalutarono il formato MP3 perché convinte che l'oggettiva migliore qualità dei CD non potesse essere neanche lontanamente avvicinata da quella – scarsa – dei formati compressi, trascurandone in questo modo il gran numero di aspetti positivi. La riduzione delle dimensioni dei file musicali ha facilitato il loro utilizzo sui computer ma

soprattutto su Internet, senza considerare i dispositivi portatili che permettono di portare con sé un'enorme quantità di musica digitale e di condividerla con i propri amici. Evidentemente, questa serie di caratteristiche può offrire benefici maggiori rispetto al parametro della qualità utilizzato per valutare fino ad allora la musica.

Riferimenti *Potete leggere l'articolo su Wired per intero all'indirizzo http://www.wired.com/gadgets/miscellaneous/magazine/17-09/ff_goodenough*.

In realtà, i libri stessi fanno parte della rivoluzione del Good Enough. Se pensiamo ai codici miniati, i libri a stampa erano di qualità inferiore, meno accurati e di scarso pregio. Ma proprio questo li rendeva alla

portata di molte più persone, e la velocità con cui potevano essere prodotti e diffusi non aveva paragoni.

Gli ereader, i dispositivi progettati quasi esclusivamente per la lettura di libri digitali, sono una tecnologia che permette di svolgere il compito per cui è stata pensata a un livello qualitativamente elevato. Tuttavia, nell'ottica del Good Enough, hanno lo svantaggio di essere ancora molto costosi e poco flessibili, proprio perché così specializzati.

C'è di più. Se dovessimo dare ascolto ai continui allarmi sulla sempre più scarsa diffusione della lettura (in Italia e nel mondo), dovremmo chiederci se un dispositivo costoso e concepito per svolgere un'unica funzione (o meglio, diverse funzioni legate a

un’unica attività) sia lo strumento giusto per la diffusione dei libri digitali. È qui che entrano in campo i dispositivi multifunzione.

Console portatili, telefoni cellulari e altri dispositivi in grado di compiere le operazioni più diverse contemporaneamente hanno trovato il modo di conquistarsi letteralmente uno spazio nelle tasche di un numero sempre maggiore di persone. Non sempre si tratta di dispositivi di alta qualità, ma compensano con la flessibilità e una relativa economicità rispetto alla gamma di funzioni proposte.

Tutto questo determina una paradossale sovrabbondanza di dispositivi dai quali è possibile eseguire applicazioni per la lettura di libri digitali: non sono i migliori supporti sui quali leggere, non sempre vantano tutte

le funzioni di un dispositivo dedicato, ma gli esempi che proporremo qui di seguito competono in maniera più che eccellente con gli ereader visti nelle pagine precedenti.

Senza dubbio un monitor con tecnologia E-Ink che può effettuare tutte le operazioni più utili per la lettura è più funzionale a questa attività; non sottovalutiamo però la possibilità di riunire in un unico dispositivo un gran numero di funzioni che non coinvolgono unicamente i libri digitali. Proprio questa versatilità potrebbe fare da traino per tutti quei consumatori che non acquisterebbero un costoso ereader ma, già in possesso di un altro dispositivo, potrebbero considerare di caricarci anche qualche ebook.

iPhone

iPhone è il telefono cellulare Apple, anche se definirlo in questo modo appare piuttosto riduttivo. Di fatto, con il suo schermo multitouch, l'accelerometro, la possibilità di essere continuamente connesso alla Rete (via UMTS o Wi-Fi), il GPS e le migliaia di applicazioni disponibili sull'App Store, iPhone ha rappresentato una vera e propria rivoluzione nel mondo della telefonia mobile e dei dispositivi tecnologici portatili in generale.

Riferimenti *Lanciato nel 2007 e arrivato alla terza generazione, iPhone è uno strumento dai numerosissimi possibili utilizzi, amplificati dall'enorme numero di applicazioni presenti gratuitamente o a pagamento sull'App Store.*

È possibile sviluppare applicazioni e metterle in vendita nello stesso negozio online di Apple. Funzioni simili sono disponibili sull'iPod touch che però, non essendo un telefono, può contare sulla connessione Wi-Fi e non UMTS.

Se gli ereader hanno conquistato quasi due milioni di americani con ottime previsioni di crescita, nella stessa nazione le persone in possesso di uno smartphone capace di leggere libri digitali sono circa 84 milioni, di cui circa una cinquantina usano iPhone.

Riferimenti *Un articolo del New York Times, consultabile all'indirizzo http://www.nytimes.com/2009/11/18/technology/18-reader.html, propone interessanti considerazioni a riguardo.*

Proprio questo ha spinto compagnie come Amazon, Barnes & Noble e altre a realizzare la propria applicazione per iPhone in grado di leggere gli ebook. Un'altra strategia piuttosto diffusa è la vendita dei libri stessi sottoforma di applicazioni: per fare un esempio, potreste trovare sull'App Store l'applicazione "Editoria digitale" e scaricare il libro in questa forma anziché in uno dei formati ebook precedentemente descritti.

La differenza non è di poco conto: gli ebook "classici" sono perfetti per qualunque dispositivo di lettura ne supporti il formato, ma la creazione di un'applicazione specifica per uno smartphone permette di pensare per ogni titolo tutta una serie di funzioni *ad hoc* impossibili da gestire su un ereader dedicato, e di sfruttare, per esempio, le particolari potenzialità tecniche del dispositivo per

rendere ancora più interessante l’esperienza di lettura.

Riferimenti *Una ricerca condotta sui trend riguardo la telefonia mobile da Flurry, una compagnia specializzata in questo tipo di indagini negli Stati Uniti, rileva che nel mese di ottobre 2009 un’applicazione su cinque tra le nuove comparse sull’App Store aveva a che fare con i libri digitali.*

Tecnicamente, rispetto a un ereader, iPhone ha uno schermo più piccolo (3,5"), a colori, retroilluminato e con tecnologia touch; le funzioni per la lettura variano a seconda dell’applicazione che si decide di utilizzare. Sull’App Store le applicazioni relative ai libri digitali sono raccolte sotto la categoria *e-Books*: vediamo le principali.

Stanza

Stanza è un’applicazione gratuita per la lettura di ebook, quotidiani e altre pubblicazioni digitali. Sviluppata da Lexcycle, supporta i formati ePub, eReader, LIT, Amazon Kindle, Mobipocket e PalmDoc, oltre a HTML, PDF, DOC e RTF.

Il software è fornito in due versioni: per iPhone e iPod touch e per Desktop (di questa versione abbiamo già parlato nel Capitolo 3). Stanza per iPhone consente di organizzare e gestire la propria libreria e di personalizzare al massimo l’interfaccia di lettura intervenendo sulla presentazione del testo. Può collegarsi a un gran numero di cataloghi online (anche personalizzati) dai quali scaricare migliaia di libri sia gratuiti sia a pagamento, a seconda del catalogo a cui si fa

riferimento, e permette di aggiungere i propri cataloghi preferiti.

eReader

La versione 2.1.1 – l'ultima disponibile al momento in cui si scrivono queste pagine – è legata all'omonimo software di cui si è già parlato nel Capitolo 3. Consente di accedere alle librerie dei siti eReader.come Fictionwise.com ma non vincola al download da questi siti: è possibile caricare e leggere anche contenuti personali o scaricati da altri siti. L'unico formato supportato al momento è l'eReader PDB. Il software è in lingua inglese e gratuito.

BookShelf

Si tratta di un reader a pagamento in grado di leggere libri digitali in formato TXT, HTML, Palm, Mobi privi di DRM. Tra le

altre possibilità, permette di aprire link e indirizzi di posta elettronica contenuti negli ebook lanciando le relative applicazioni (Safari e Mail); per scaricare i libri ha bisogno di appoggiarsi all'applicazione Desktop collegata (ShelfServer), da cui può accedere a diversi cataloghi online.

B&N eReader e Kindle

B&N eReader è l'applicazione per la lettura degli ebook della libreria americana Barnes & Noble. Consente di accedere (via Safari) a un vastissimo store di libri digitali, di personalizzare le opzioni di lettura ed è fornita anche in versione Desktop per Mac e PC.

Riferimenti *Barnes & Noble rende disponibile anche l'applicazione Barnes & Noble Store, dalla quale si può navigare nel negozio online riempiendo un*

carrello che sarà trasferito su Safari al momento di concretizzare l'acquisto.

Kindle è l'applicazione di Amazon per iPhone. Permette di sincronizzare i libri salvati sull'account nella libreria online con il dispositivo e – se lo si possiede – con il proprio Kindle, conservando memoria dell'ultima pagina letta a prescindere dal dispositivo sul quale è stata visualizzata. Le opzioni di lettura richiamano molto da vicino quelle di Kindle stesso, anche dal punto di vista dell'interfaccia. Gli aquisti su Amazon possono essere effettuati tramite Safari.

Altre applicazioni

Le applicazioni per la lettura su iPhone sono molto più numerose di quelle elencate in queste pagine, e non sono

necessariamente solo strumenti per la lettura, ma possono rappresentare veri e propri strumenti promozionali. Per averne un'idea approfondiamo la lista con un elenco sintetico.

- Galera eReader: gratuita, è un'applicazione creata da una delle case editrici facenti parte del maggiore gruppo editoriale latinoamericano (Grupo Editorial Record). Mette a disposizione anteprime delle novità in pubblicazione, aggiornate periodicamente e con possibilità di collegarsi direttamente alle schede dei libri sui siti degli editori.
- Classic: a pagamento, raccoglie una libreria di titoli fuori dal diritto d'autore e li ripropone con una veste grafica molto curata e funzionalità di lettura pratiche e intuitive.

- BookReader: a pagamento, supporta diversi formati come PDF, TXT, ePub, HTML, RTF e FB2. Permette di scaricare i libri direttamente dal Progetto Gutenberg, oltre che di caricare i propri documenti.
- i2Reader Lite: gratuito, oltre a supportare un gran numero di formati è disponibile in diverse lingue, tra cui anche il russo. Esiste anche una versione a pagamento, con un maggior numero di funzioni.

Non ultimo, come già ricordato, esiste un vastissimo numero di libri venduti direttamente in forma di applicazione, acquistabili e scaricabili dall'App Store come se fosse una vera e propria libreria digitale. Questa prospettiva è interessante perché consente potenzialmente a chiunque non solo di pubblicare il proprio libro in formato

digitale, ma anche di ideare nuove forme per il contenuto, arricchendo l'esperienza di lettura attraverso l'interattività di iPhone.

Riferimenti *Alcuni esempi interessanti di ciò che si può fare con un libro-applicazione sono dei titoli per bambini, che includono sia una voce narrante pre-registrata sia la possibilità di registrare la propria voce per raccontare la storia: http://itunes.apple.com/it/app/pocket-story-the-smelly-sprout/id338986079*.

Progettare applicazioni per iPhone richiede capacità di programmazione che possono sembrare distanti dalle competenze di un editore. Questo approccio può essere visto sotto una luce diversa se si pensa al modo in cui venivano considerati i siti web non

troppi anni fa, mentre adesso, come abbiamo visto, sono diventati uno strumento fondamentale per l'editore, imprescindibile per la diffusione e la distribuzione dei contenuti.

iPad

Il 27 gennaio 2010 Apple ha presentato un nuovo device che potrebbe cambiare le carte in tavola nel settore dei dispositivi di lettura multifunzione (e non): iPad.

In poco più di un centimetro di spessore, uno schermo con tecnologia Multi-Touch da 9.7" nasconde un processore altamente performante associato al sistema operativo di iPhone, che permette totale compatibilità con le centinaia di migliaia di applicazioni già esistenti; Wi-Fi e 3G assicurano il

massimo delle prestazioni per la connessione Internet.

Riferimenti *iPhone SDK, l'ambiente di sviluppo per le applicazioni per iPhone, permette di programmare anche nuove applicazioni* ad hoc *per iPad.*

Al di là delle specifiche tecniche del dispositivo, è la presenza di una nuova applicazione tra quelle integrate, iBooks, a destare maggiore interesse in campo editoriale (Figura 6.7). iBooks è pensata appositamente per leggere ebook e gestire la propria biblioteca digitale. La novità fondamentale è l'integrazione nella stessa applicazione della libreria online, iBookstore, che sul modello dell'iTunes Store consentirà la distribuzione e la vendita di libri digitali in formato ePub.

Rispetto a ciò che permette Amazon sullo store di Kindle (esplicitamente indicato da Apple come modello ispiratore), Apple ha stipulato accordi molto convenienti con gli editori ai quali consentirà di vendere in prima persona i propri libri digitali, lasciando stabilire loro i prezzi.

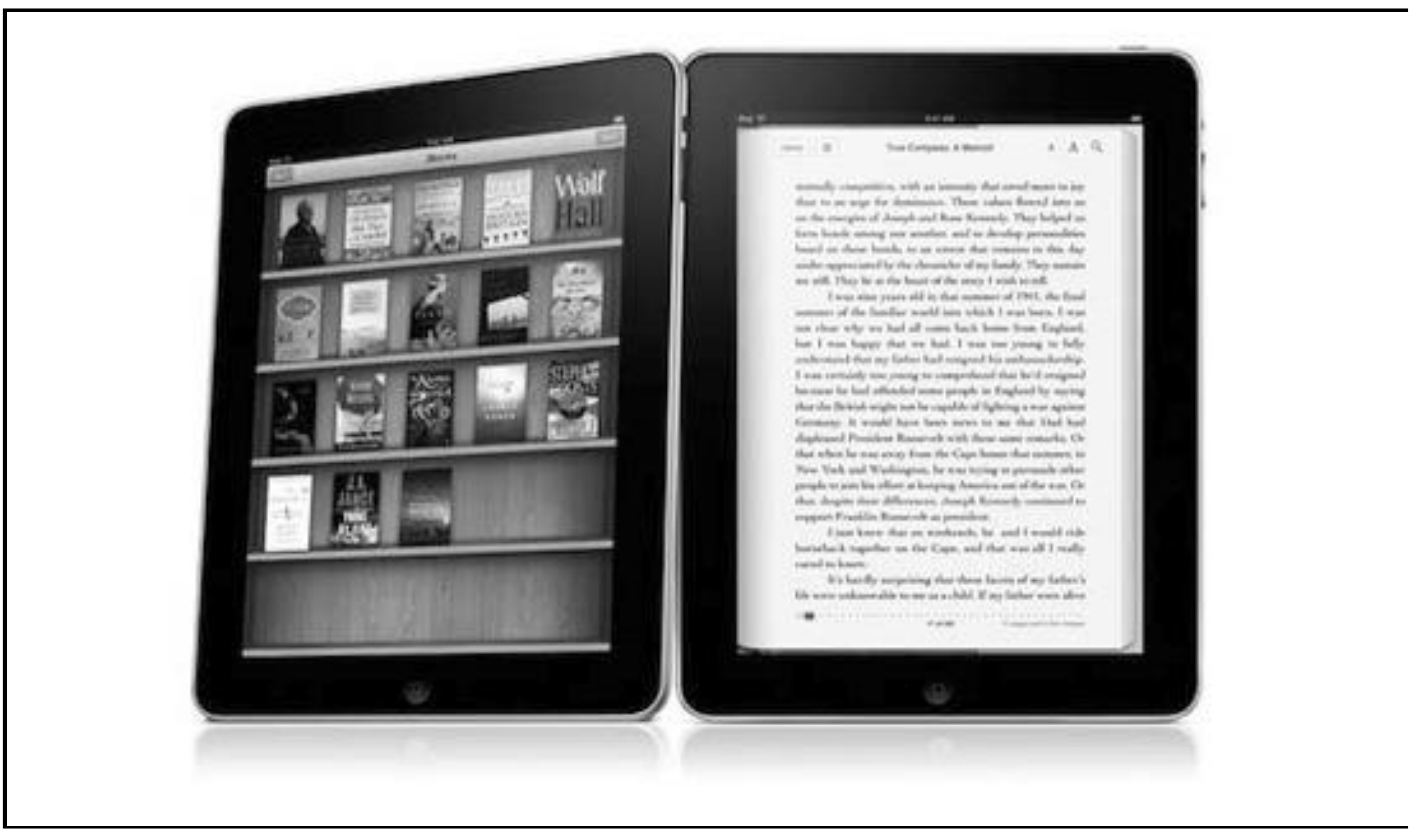

Figura 6.7 iPad, il nuovo dispositivo Apple con l'applicazione iBooks integrata.

Riferimenti *Gli editori che già fanno parte del progetto sono Penguin, HarperCollins, Simon & Schuster, Macmillan e Hachette Book Group.*

L'integrazione di iBooks e iBookstore su iPad combina la formula iTunes per i contenuti audio e video con il successo indiscutibile dell'App Store e di iPhone: iBookstore potrebbe riuscire a replicare per i libri quella struttura di servizi già collaudata da iTunes e in grado di tenere testa alla pirateria, portando nelle tasche del lettore un dispositivo con enormi potenzialità per la lettura digitale capace di svolgere tante operazioni diverse quante sono le applicazioni nell'App Store. Inoltre, con l'apertura di uno store dedicato ai libri digitali, potrebbe venire meno una caratteristica svantaggiosa della distribuzione di libri-applicazione sull'App

Store: infatti, la corrispondenza libro/applicazione impediva, per esempio, di navigare tra le annotazioni della biblioteca in maniera trasversale, e quindi di avere uno sguardo d'insieme sulla propria raccolta di libri digitali.

Alcune criticità

L'annuncio di iPad da parte di Apple ha creato grandi aspettative sotto molti punti di vista: per i cambiamenti possibili nel mercato dei netbook, per l'attesa di innovazioni rispetto a ciò che già permette iPhone e per le nuove opportunità di un dispositivo dalle dimensioni maggiori in grado di sfruttare diversamente tutte le funzioni dello smartphone di Cupertino. La presentazione del device ha in parte deluso tali aspettative, lasciando però spazio a futuri miglioramenti

che potranno rendere iPad uno strumento potenzialmente rivoluzionario.

Dal punto di vista della lettura digitale segnaliamo alcune criticità.

- La durata della batteria e l'affaticamento della vista: l'ampio schermo retroilluminato di iPad al momento non può competere con gli ereader basati sulla tecnologia E-Ink quanto a durata (si parla di circa dieci ore); nonostante il colore sia un'opportunità importante per gli ebook, l'impossibilità di leggere a lungo senza dover ricaricare il dispositivo e senza un eccessivo affaticamento della vista potrebbero costituire un fattore determinante.
- Il costo: decisamente elevato al momento del lancio (da $ 499 a $ 899, a seconda della

capacità di memoria e delle possibilità di connessione).

- La chiusura del dispositivo: non potendo archiviare documenti in memoria, non è chiaro se attraverso iBooks si potranno caricare e leggere anche ebook di cui l'utente è già in possesso (senza passare per altre applicazioni come Stanza, che come abbiamo visto lo permette).

Android

Android è un sistema operativo open source per dispositivi mobili basato su Linux e sviluppato inizialmente da Google. Ciò che lo rende particolarmente interessante è proprio la scelta di Google di entrare nel mercato della telefonia mobile, grazie ad accordi presi con alcuni tra i principali

produttori di dispositivi. La sua struttura aperta permette sia di sviluppare applicazioni sia un altissimo livello di personalizzazione del proprio device.

I software per la lettura di libri digitali su Android non sono ancora numerosi come quelli per iPhone, ma propongono un sufficiente ventaglio di opzioni per soddisfare le principali necessità di lettura.

Aldiko

Aldiko è un reader per Android da cui è possibile scaricare libri da cataloghi online, gestire la propria collezione di download anche attraverso tag personali e personalizzare al massimo l’esperienza di lettura. Consente anche di utilizzare strumenti di ricerca nel testo e scegliere il dizionario online in cui ricercare il significato di un termine.

FBReader

Abbiamo già parlato di FBReader nel Capitolo 3: ricordiamo qui la grande versatilità di questo lettore, disponibile – tra le tante piattaforme – anche per Android.

Sony PSP e Nintendo DS

Non solo telefoni cellulari e smartphone possono diventare dispositivi adatti alla lettura dei libri digitali, ma anche alcune tra le console portatili più note, come il Nintendo DS (in tutte le sue versioni) e la PSP della Sony.

Per il Nintendo DS, oltre ad alcune applicazioni ancora piuttosto acerbe in grado di interpretare prevalentemente file TXT o HTML, la più interessante opportunità di lettura è offerta da HarperCollins, un editore

che ha deciso di creare un pacchetto di cento libri liberi dal diritto d'autore e di venderli sotto forma di cartuccia di gioco, con la possibilità di creare i propri segnalibri e alcune opzioni di personalizzazione delle modalità di lettura. L'esperimento è interessante perché mostra l'attenzione di un grande editore verso l'eventualità di utilizzare strumenti hardware certamente non pensati apposta per leggere come ulteriori stimoli a diffondere la lettura, studiando nuovi modi per proporre contenuti e cultura.

Riferimenti *Potete avere maggiori informazioni riguardo la cartuccia di gioco HarperCollins all'indirizzo http://www.nintendo.co.uk/NOE/en_GB/games/nds/100_classic_book_collection_10283.html*. *Le*

altre applicazioni a cui abbiamo accennato possono essere utilizzate a patto di possedere particolari cartucce in grado di contenere delle schede di memoria microSD, capaci di far funzionare software amatoriali e non ufficiali.

Sulla Sony PSP è possibile leggere solo installando un homebrew (così vengono definite le applicazioni amatoriali di cui parlavamo anche per il Nintendo DS), Bookr, che permette di leggere file PDF, TXT e RTF e di salvare dei segnalibri. Indirettamente, grazie al browser integrato sulla console, si può accedere ai servizi di archiviazione e lettura online per i libri digitali: data l'ampiezza dello schermo della PSP, potrebbe essere un'occasione di lettura in più da tenere presente.

Riferimenti *A partire dal 2010 è disponibile per PSP l'applicazione Digital Comics, che permette di leggere e scaricare i fumetti di un gran numero di case editrici direttamente sul dispositivo, collegandosi al PlayStation Store. Tutte le informazioni sul programma sono disponibili all'indirizzo http://it.playstation.com/psp/hardware/accessories/detail/item229746/Digital-Comics/.*

IL COPYRIGHT E I DRM

In questo capitolo prenderemo in esame alcuni aspetti relativi al diritto d'autore e alle modalità di protezione dei libri digitali. Naturalmente non abbiamo la pretesa di esaurire un argomento così complesso in poche pagine.

Ci limiteremo a considerare quegli elementi che permettono di sviluppare riflessioni generali sul tema, gli stessi che dovrebbero essere tenuti in considerazione quando si decide di pubblicare degli ebook.

La crisi del copyright

Il termine *copyright* indica, nel mondo anglosassone, l'insieme di normative relative al diritto d'autore. In Italia si utilizza per indicare le norme vigenti a riguardo, nonostante non ci sia una coincidenza esatta tra i due concetti.

Terminologia *Il* diritto d'autore *stabilisce la posizione giuridica soggettiva dell'autore di un'opera d'ingegno, a cui viene riconosciuta la facoltà esclusiva di diffusione e sfruttamento dell'opera stessa, e in ogni caso il diritto a essere indicato come autore anche nel caso in cui abbia concesso a terzi la facoltà di sfruttamento economico (diritto*

morale d'autore). Il diritto d'autore è tipico dei paesi in cui l'ordinamento giuridico si basa sulla civil law*; il copyright, invece, è diffuso nei paesi della* common law.

Il copyright, letteralmente "diritto di copia", si occupa di regolamentare i diritti sulla riproduzione delle opere d'ingegno. Alla base della concezione classica del copyright stanno il costo e la difficoltà di riprodurre e diffondere sul territorio le opere: questi aspetti (nel mondo editoriale) potevano essere gestiti facilmente solamente dagli editori, dietro compenso o attraverso la cessione dei diritti da parte degli autori.

Terminologia *Negli Stati Uniti è previsto il concetto di* fair use*, che lascia spazio alla riproduzione delle opere per scopi didattici o scientifici,*

inquadrando come promozione del progresso della scienza e delle arti la citazione non autorizzata o l'incorporazione del materiale protetto da copyright in un'opera di un altro autore. In Europa e anche in alcuni Paesi della common law il fair use o non è regolamentato o non ha la stessa applicazione sistematica.

Tale concezione entra in crisi con la diffusione dei computer e di Internet: si tratta di tecnologie pensate per copiare e diffondere i dati in modo economico e veloce e questa opportunità impatta direttamente sull'idea che realizzare una copia di un libro sia un'attività costosa, tecnicamente complessa e dalla difficile distribuzione. Prima della loro diffusione di massa, la violazione del copyright era pressoché impossibile per chiunque

non fosse un editore: in sostanza, tutelava gli editori da violazioni che potevano essere commesse solamente da altri editori.

Il copyright entra in crisi nel momento in cui viene sistematicamente applicato ai fruitori delle opere: se per un editore difendere i diritti su un'opera dalle speculazioni di altri editori è un'operazione più che legittima, lo è meno quando l'estensione di tale difesa va a limitare i diritti del lettore sul libro di cui è diventato proprietario acquistandolo legittimamente.

Se il problema del copyright colpisce l'editoria su carta soprattutto attraverso il fenomeno delle copie illegali (fotocopie, in particolare nel settore scolastico e universitario) e in misura minore in forma di copie

digitali, parlando di editoria digitale la questione assume altre proporzioni.

Riferimenti *Il fenomeno delle copie digitali illegali di libri cartacei non è così allarmante come nel caso della pirateria audiovisiva: realizzare una copia digitale di un libro stampato è una procedura laboriosa, che richiede un grande dispendio di tempo ed energia, considerato che il metodo più diffuso è la scansione manuale delle pagine. Il risultato, per di più, è spesso di scarsa qualità. In Italia l'associazione che tutela i diritti di riproduzione delle opere librarie e periodiche è l'AIDRO, che si occupa di contrastare la pirateria e di gestire i diritti di riproduzione in fotocopia delle opere per uso professionale o commerciale, offrendo un*

sistema di licenze in grado di soddisfare le diverse esigenze di mercato.

Copiare un ebook, come un qualunque altro tipo di file, non richiede più di qualche clic anche dovendone forzare la protezione, e diffonderlo via mail o tramite una rete peer to peer è un'operazione da pochi secondi. Il paradosso è che il legittimo acquirente di un file protetto è di fatto il soggetto più limitato nell'utilizzare il file di cui è per diritto il *proprietario*.

Il cortocircuito nell'applicazione rigida del copyright a un libro digitale si verifica proprio nel momento in cui va a limitare la proprietà dell'acquirente rispetto all'oggetto. Se si acquista un libro stampato si diventa proprietari dell'oggetto fisico, che può essere regalato o prestato, nelle forme che si ritengono più opportune: quello che il lettore ne

fa una volta acquistato non è oggetto del copyright, che si occupa di salvaguardare la paternità delle idee che il libro veicola ma non limita l'utilizzo dell'oggetto fisico.

Tuttavia, nel caso del libro digitale, questo confine è più sottile: sia perché, come abbiamo visto, un ebook perde la fisicità per trasformarsi in contenuto "puro", sia perché la riproducibilità di un file è di ben altra semplicità rispetto a quella della carta.

DRM: la gestione dei diritti sui file digitali

I *Digital Rights Management* (DRM) sono letteralmente dei sistemi per la gestione dei diritti digitali. Si tratta di una tecnologia

mediante la quale il titolare del diritto d'autore può esercitare e amministrare tale diritto in ambiente digitale, attraverso la possibilità di proteggere, identificare e tracciare le proprie opere.

Riferimenti *I DRM possono essere assimilati alla filigrana delle banconote, che ne impedisce la falsificazione: aggiungono infatti delle informazioni nascoste ai file con lo scopo di regolamentarne l'utilizzo.*

I file digitali vengono codificati e criptati tramite i DRM, così da controllarne la diffusione. L'utilizzo del file può essere limitato nel tempo o nelle destinazioni d'uso, oppure stabilito a priori attraverso una licenza d'accesso che viene fornita a parte all'utente finale.

Il funzionamento dei DRM è basato sulla crittografia: il contenuto è leggibile solo se si è a conoscenza della chiave di cifratura, e questo è possibile a patto di essere legittimi proprietari del contenuto.

Attraverso tali sistemi si può identificare la copia originale del file e di conseguenza le eventuali copie illegali (certificazione di proprietà), fare in modo che il file protetto sia riproducibile solo attraverso lettori capaci di riconoscerne le informazioni di codifica ed eventualmente un numero di volte stabilito in fase di acquisto (controllo d'accesso) e risalire al possessore dei file originali, per individuare e prevenire le violazioni del diritto d'autore in relazione all'utilizzo delle nuove tecnologie (controllo delle copie illegali).

In Italia la legge permette ai titolari di diritti d'autore e di diritti connessi di applicare sulle opere d'ingegno misure di protezione tecnologiche efficaci: esse comprendono tecnologie, dispositivi o componenti con la funzione di impedire o limitare atti non autorizzati dal titolare dei diritti.

Riferimenti *La legge sul diritto d'autore è la n. 633 del 22 aprile 1941, in materia di "Protezione del diritto d'autore e di altri diritti connessi al suo esercizio"*.

In caso di elusione del DRM la legge prevede sanzioni penali fino alla reclusione (le pene più severe sono applicate a illeciti riguardo al software); inoltre vengono sanzionate pesantemente anche un gran numero di attività riguardo attrezzature,

prodotti o componenti con finalità prevalente di eludere le misure tecnologiche di protezione. Anche la totale rimozione abusiva o l'alterazione delle informazioni elettroniche poste sulle opere d'ingegno è perseguibile penalmente. Sanzioni amministrative, infine, sono previste per chi acquista o noleggia questo tipo di tecnologie.

Riferimenti *Secondo la legge, la presenza di sistemi di DRM non può impedire al possessore dei diritti di effettuarne una copia privata per uso personale: a questo scopo prevede un "equo compenso" che consiste in una somma imposta sul prezzo di apparecchiature idonee a registrare contenuti audio o video e sui relativi supporti vergini. Questo compenso viene*

incamerato dalla SIAE, che poi lo ripartisce tra i titolari del diritto d'autore.

Criticità

I sistemi di DRM normalmente vengono forzati nel giro di pochi minuti, a volte giorni. Raramente mesi. Non è perché le persone che li inventano sono stupide. Non è perché le persone che li forzano sono intelligenti. Non è perché c'è un errore nell'algoritmo. Alla fine, tutti i sistemi di DRM condividono la stessa vulnerabilità: forniscono al
loro aggressore il testo cifrato, il cifrario e la chiave. A questo punto il segreto non è più tale.

Si esprime così Cory Doctorow in un suo discorso al gruppo di ricerca sul DRM di Microsoft, cercando di dimostrare l'inefficacia dei siste mi di protezione dei file digitali. Le sue opinioni riguardo copyright e DRM sono

radicali e spesso provocatorie: tuttavia ci consentono alcuni spunti di riflessione piuttosto interessanti.

Riferimenti *Potete trovare il discorso da cui è tratta la citazione e altri saggi riguardo copyright e DRM in* Content, *disponibile per il download gratuito all'indirizzo http://www.apogeonline.com/libri/9788850310852/scheda*.

- I sistemi di protezione sono studiati per difendere i file non dagli utenti più smaliziati o peggio dai criminali informatici, ma dagli utenti medi, quelli che nella maggior parte dei casi hanno acquistato legalmente il file e vedono limitare le opportunità di utilizzo di una loro proprietà in nome della protezione del diritto d'autore. A volte

queste limitazioni sono così invadenti da spingere l'utente a preferire un download illegale ma privo di protezione, perché si sente raggirato da un sistema di protezione che gli impedisce di fare l'uso che più ritiene opportuno di un file che è una sua proprietà a tutti gli effetti.

- Il copyright deve adattarsi ai nuovi scenari dettati dalla tecnologia. Se il progresso tecnologico ci mette a disposizione strumenti sempre più efficaci per la riproduzione e la diffusione dei dati, entrando in collisione con gli aspetti fondanti del copyright classico, è necessario che questo si evolva in maniera da ridefinire scenari più adeguati, che tengano conto delle esigenze e possibilità attuali della riproducibilità delle opere d'ingegno.

- I sistemi DRM proprietari obbligano a riprodurre i file a cui vengono applicati attraverso i dispositivi che sono in grado di interpretarli. Questo non avvantaggia né il distributore dei contenuti – che in questo modo si preclude spazi di diffusione per il suo prodotto – né l'utente finale, che si sente costretto a utilizzare quel determinato contenuto su un dispositivo sostanzialmente imposto.

Nel mercato ancora in evoluzione dei libri digitali, ci sono pareri molto discordanti riguardo all'opportunità di applicare i DRM agli ebook. Da un lato si teme il fenomeno della pirateria e la perdita di lettori e di vendite sia della versione digitale sia di quella cartacea del libro. Dall'altro si preferiscono diverse forme di protezione dei file, basate più che sulle restrizioni tecniche

(peraltro non inviolabili) sui meccanismi sociali della Rete.

Possiamo fare a questo proposito due esempi illustri: Amazon, con Kindle e il suo formato proprietario e protetto AZW ha optato per una tecnologia che limita fortemente l'utilizzo del file legalmente acquistato dall'utente; O'Reilly propone i suoi libri in più formati e liberi da DRM, per precisa scelta editoriale.

La decisione di applicare o no restrizioni nell'utilizzo dei libri digitali, come altre scelte che abbiamo descritto, va ponderata con molta attenzione. Si tratta di una scelta per certi versi "politica": se l'applicazione del DRM, di fatto, non impedisce il proliferare delle copie illegali, le ragioni che spingono ad adottarlo riguardano la convinzione che –

se anche il file viene piratato – sono state comunque tentate tutte le possibilità per proteggerlo.

Ma se è pressoché impossibile evitare che un file finisca in una rete peer to peer o avere come minimo il controllo sulla sua diffusione, allora forse bisogna provare a valutare il fenomeno pirateria da un nuovo punto di vista. La lettura è un'attività prevalente: non si possono fare altre cose mentre si legge un libro, a differenza di quanto accade per esempio con la musica. Se chi scarica un libro illegalmente per il solo gusto di farlo poi non lo legge, di fatto non si è perso un lettore: si tratta di una persona che con ogni probabilità non avrebbe mai acquistato quel libro. Se quel libro viene letto si è comunque conquistato un lettore, che magari la prossima volta sarà tentato di acquistare un

libro: si tratta di convincerlo che l'acquisto gli offre qualcosa di più del download illegale. Come?

Facciamo un esempio. Pochi secondi dopo la sua pubblicazione su iTunes, il negozio Apple di musica online, un brano musicale è già disponibile per il download illegale anche nelle reti peer to peer, gratis. Eppure su iTunes le vendite sono comunque più che buone. Questo perché attraverso iTunes, per una manciata di centesimi, l'utente non ha solo il brano musicale, ma un insieme di servizi: scarica il brano senza code di download, il file è completo di metadati precisi e accurati, gli viene proposta una serie di scelte musicali compatibili con i suoi gusti secondo le sue abitudini di acquisto, può leggere le opinioni di altri che prima di lui hanno acquistato quel contenuto e così via.

Il lancio di iBooks, l'applicazione Apple che consente sia la lettura degli ebook sia l'accesso all'iBookstore, potrebbe replicare questo modello anche per i libri digitali, offrendo agli editori una piattaforma efficace per la distribuzione degli ebook.

È in questa chiave che occorre ripensare il proprio ruolo editoriale: come fornitori di servizi sui contenuti, oltre che come fornitori dei contenuti stessi. Se l'utente può raggiungere il contenuto in vari modi, compresi quelli illegali, occorre trovare nuove formule che lo incoraggino a rivolgersi agli editori per acquistare i libri digitali.

L'esperienza di Doctorow è interessante: l'autore rilascia sul Web tutti i suoi libri in formato digitale gratuitamente, lasciando i suoi lettori liberi di farne ciò che più

ritengono opportuno. Allo stesso tempo li pubblica su carta presso il suo editore, e non ha mai avuto motivo di lamentarsi delle vendite. Sarebbe interessante poter verificare se e quanto questa strategia sia efficace come vero e proprio strumento di marketing.

Se guardiamo al mercato musicale, dove la diffusione del digitale ha già raggiunto un certo grado di maturità, possiamo notare che si è passati dalla presenza di sistemi DRM fortemente limitanti a forme di protezione basate su tutt'altre strategie.

iTunes inizialmente prevedeva l'applicazione di sistemi di protezione che limitavano – tra le altre cose – la riproduzione del file su un numero di computer prestabilito. Adesso questo tipo di restrizioni

non vengono più applicate e si è passati a un tipo differente di deterrente per la diffusione illecita della musica digitale: i social DRM. Vediamo di cosa si tratta nel paragrafo seguente.

Un’alternativa: il social DRM

Il social DRM consiste nell’includere all’interno del file che si acquista alcune informazioni sul proprietario (nome e indirizzo e-mail, per esempio), in maniera da poter risalire a lui nel caso in cui il documento venga immesso nelle reti pirata o semplicemente ne venga fatto un uso illegale. A differenza dei sistemi DRM classici, il social DRM non limita l’utilizzo del file in alcun modo, ma agisce da deterrente contro il file sharing. Anche in questo caso non si tratta affatto di una difesa inviolabile: come i DRM classici

può essere rimosso. Tuttavia, visto che la sua presenza non limita nell'utilizzo del file, viene percepito come meno invasivo e non spinge tanto quanto il DRM classico al tentativo di rimuoverlo o forzarlo.

Possiamo considerare l'aspetto social del DRM anche in un'altra chiave: come dicevamo, la maniera più efficace per evitare che gli utenti scarichino illegalmente ebook o altri tipi di file è quella di diventare fornitori di un servizio migliore sui contenuti, che diventi di gran lunga più accattivante di un download gratuito ma illegale. In questo senso si può ragionare sul rafforzare le relazioni sociali con i propri lettori, anche attraverso le reti sociali che il Web ci mette a disposizione: il rapporto di fiducia che ne deriva, unito al valore dei servizi proposti sui contenuti potrebbe diventare la base di un

sistema di protezione veramente “social”, fondato sull’autorevolezza dell’editore e sulla valorizzazione del suo ruolo culturale.

Creative Commons

In materia di copyright non si possono trascurare le licenze di tipo Creative Commons. Queste nascono per regolamentare la distribuzione gratuita delle opere di ingegno e permettono ai detentori dei diritti di trasmetterne alcuni al pubblico e di conservarne degli altri, attraverso diverse tipologie di contratto. L’intenzione è quella di evitare alcuni dei problemi che insorgono riguardo la diffusione e la condivisione delle informazioni con le attuali leggi sul copyright.

Riferimenti *In Italia le licenze Creative Commons sono state tradotte e adattate a partire dal 2003.*

Le licenze Creative Commons sono di sei tipi, e stabiliscono esplicitamente quali sono i diritti riservati, modificando la regola alla base del copyright che per default prevede che lo siano tutti. Si basano su quattro specifiche.

- (BY): indica che l'attribuzione dell'opera (l'indicazione dell'autore) è obbligatoria, così che sia possibile individuarne la paternità.
- *Non uso commerciale* (NC): non consente usi commerciali dell'opera, che rimangono esclusiva dell'autore.

- *Non opere derivate* (ND): non consente elaborazioni dell'opera che dev'essere diffusa senza alcuna alterazione.
- *Condividi allo stesso modo* (SA): qualsiasi utilizzo e diffusione dell'opera deve essere fatto secondo le stesse condizioni scelte dall'autore (in pratica non è possibile alterare la licenza originale).

Riferimenti *Le licenze possono essere consultate online all'indirizzo http://creativecommons.it/licenze*.

Queste specifiche vengono combinate in vari modi, dando origine alle seguenti sei licenze:

- Attribuzione (BY)
- Attribuzione – Non opere derivate (BY-ND)

- Attribuzione – Non commerciale – Non opere derivate (BY-NC-ND)
- Attribuzione – Non commerciale (BY-NC)
- Attribuzione – Non commerciale – Condividi allo stesso modo (BYNC-SA)
- Attribuzione – Condividi allo stesso modo (BY-SA)

È importante sottolineare che le licenze Creative Commons non sono sostitutive del copyright, ma servono per declinare in più varianti la formula standard “Tutti i diritti riservati”. Ognuna di esse si estende per tutta la durata del diritto d’autore.

Il caso Google Library Project

Concludiamo questo capitolo accennando a un caso che ha fatto e fa tutt'ora discutere riguardo questioni di copyright e libri digitali. Si tratta della causa legale che coinvolge il Google Library Project (ora Google Libri).

Sul finire del 2004 Google lanciò il servizio Book Search, o Ricerca Libri nella versione italiana: il programma consisteva in una serie di accordi con diverse biblioteche per avviare la digitalizzazione di libri, inclusi titoli protetti dalle leggi statunitensi sul copyright, presenti nelle collezioni di tali biblioteche. L'unica cautela, in questo caso, era che le opere non erano rese visibili ma solo

ricercabili e – quando trovate – ne venivano mostrati solo brevi stralci.

Riferimenti *Le collezioni in realtà comprendevano un gran numero di libri anche di editori europei, coperti dalle leggi sul diritto d'autore; questo ha contribuito a rendere molto più controversa la questione, date anche le differenze a livello normativo.*

L'Authors Guild of America e l'Associazione degli editori americani hanno separatamente citato in giudizio Google, sostenendo che tale digitalizzazione fatta senza autorizzazione violasse il copyright: l'idea era che di fatto si costruisse un'iniziativa commerciale, creando una banca dati di testi senza chiedere l'autorizzazione e senza remunerare i detentori dei diritti.

Riferimenti *I cinque editori membri della AAP (*Association of American Publishers*) sono McGraw-Hill Companies, Pearson Education, Penguin Group, John Wiley & Sons e Simon & Schuster.*

In risposta a tale accusa, Google ha ribattuto che la digitalizzazione dei libri e la visione di frammenti o poche righe degli stessi sono autorizzate dal principio del fair use previsto dalla legge americana sul copyright. Piuttosto che risolvere la controversia legale riguardo all'applicazione del fair use, le parti hanno optato per una transazione.

Riferimenti *La transazione permette di concludere una controversia legale senza l'intervento di un tribunale o di una giuria che decida in favore di una*

delle parti, evitando così i costi e i rischi di un processo.

Il primo accordo raggiunto attraverso la transazione regolava sostanzialmente tre situazioni:

- il destino dei volumi digitalizzati fino a quel momento;
- in quali casi Google è autorizzato a digitalizzare senza un'autorizzazione preventiva nuovi volumi e quali utilizzi potrà farne;
- i termini economici delle future digitalizzazioni in accordo con gli aventi diritto.

I libri già digitalizzati che non saranno rimossi, quelli fuori commercio digitalizzati in futuro senza autorizzazione, i libri in commercio saranno digitalizzati previa autorizzazione di autori ed editori. L'accordo

regolava le condizioni generali per una serie di utilizzi, commerciali e non. In questa seconda categoria rientrano la consultazione delle copie digitali all'interno delle biblioteche proprietarie delle collezioni, l'accesso ai disabili visivi e la possibilità per alcuni ricercatori di utilizzare la banca dati per analisi relative alle loro ricerche.

Gli utilizzi commerciali ricoprono quasi tutte le ipotesi di sfruttamento di un'opera su Internet. Si tratta di quattro tipologie.

- Il *display use* consente all'utente di accedere a parti limitate del libro riprodotto all'interno di un frame da cui è possibile accedere ad altri servizi sia interni sia esterni al sistema Google.
- La possibilità per gli utenti finali di acquistare le versioni digitali dei libri, anche se

non è chiaro in quali formati, permettendo agli editori di fissare un prezzo dal quale Google detrarrà una sua percentuale.

- L'accesso all'intera banca dati sarà disponibile per sottoscrittori istituzionali, come per esempio le biblioteche, con formule di abbonamento già sperimentate nel caso delle riviste.
- Le università potranno acquistare licenze di riproduzione di brani di libri per inserirli in dispense per gli studenti, messe a disposizione online o su carta.

Questo insieme di utilizzi commerciali definisce di fatto un ingresso di Google nella distribuzione dei contenuti librari, che non può passare inosservato per via delle dimensioni e potenzialità di un attore del calibro di Google.

Riferimenti *Si pensi anche a Google Editions, presentato nel Capitolo 5.*

L'annuncio dell'accordo ha dato origine a un dibattito anche in Europa, non solo tra autori ed editori, ma anche tra librerie e biblioteche. Tra i problemi di natura giuridica, tecnica o economica, si possono individuare tre aspetti principali: la modalità scelta di gestione dei diritti; il rovesciamento del principio di uso sempre autorizzato; il rischio di monopolio dell'informazione da parte di Google.

Un accordo simile, anche se stipulato negli USA, finisce per influenzare anche autori ed editori europei. Il motivo può essere così riassunto: Google, attore ricco e potente, ottiene una licenza generale che gli consente ampie libertà a meno che non siano gli aventi diritto a reclamare il dovuto; il costo

della ricerca del contenuto da rimuovere è scaricato sugli aventi diritto che devono occuparsene per evitare potenziali abusi, sgravando Google da questo onere.

Inoltre, il fatto che il discrimine chiave sia che il libro sia considerato in commercio o fuori commercio è piuttosto ambiguo e l'accordo rimane vago su questo aspetto. La definizione di "fuori commercio" negli USA è differente da quella europea: negli Stati Uniti si definisce un libro *non commercially available* quando non è reperibile in un "canale commerciale abituale" per gli USA. L'ambiguità sta proprio nello stabilire cosa sia un canale commerciale "abituale".

Riferimenti *Una clausola nell'accordo, infine, desta non poche perplessità: la possibilità per Google di escludere – a sua discrezione – alcuni*

libri da uno o più utilizzi. Uno strumento che rischia di intaccare i basilari principi della libertà di espressione.

È proprio a seguito di tali perplessità che, sul finire del 2009, è stato stipulato un nuovo accordo, che non si applica più a tutti i libri stranieri ma solamente a quelli pubblicati in USA, Regno Unito, Canada e Australia, comprendendo comunque tutti i libri registrati presso il Copyright Office di Washington, ovunque siano stati pubblicati e in qualunque data.

Nel nuovo accordo sono presenti miglioramenti nella qualificazione delle opere in commercio e nel trattamento delle opere “orfane”, ossia di quelle non rivendicate dagli aventi diritto. In particolare, un libro è definito in commercio se venduto in una qualsiasi parte del mondo e acquistabile

nelle nazioni citate sopra; inoltre, prevale la qualificazione dell'opera come "in commercio" da parte dell'editore a prescindere da quella di Google, che dovrà anche impiegare i redditi prodotti da quei libri stessi per ricercare i proprietari dei diritti ancora ignoti.

Tuttavia, molti editori europei non trovano ancora soddisfacente questo nuovo accordo: a far discutere è proprio l'inclusione dei libri registrati presso il Copyright Office di Washington. Viene fatto notare, infatti, che fino agli anni Ottanta questa era una pratica indispensabile per tutelare le opere straniere per gli Stati Uniti, pertanto non si tratta – come erroneamente valutato – di un numero ristretto di titoli, ma di una quantità decisamente rilevante. Le associazioni degli editori italiani, tedeschi, austriaci e svizzeri

hanno presentato le proprie obiezioni alla Corte di New York; in modo analogo si sono mosse le associazioni francese e spagnola.

CALIBRE: AL SERVIZIO DEL LETTORE

Abbiamo visto molti metodi professionali per produrre e gestire ebook, ma non abbiamo accennato alle necessità meno impegnative che può avere chi desidera leggere i libri digitali.

Vediamo dunque una breve panoramica di un software – Calibre – in grado di soddisfare in maniera più che efficiente le esigenze di chi è interessato agli ebook come il lettore.

Meglio di un coltellino svizzero!

Calibre (http://calibre-ebook.com) è un software gratuito, open source e multipiattaforma per la gestione di una biblioteca digitale. Tra le sue numerose funzioni troviamo la possibilità di convertire i più vari formati di ebook, sincronizzare la libreria con i dispositivi di lettura, scaricare news dal Web e convertirle nel formato ebook desiderato, visualizzare il libro digitale e accedere online alla propria collezione tramite server.

Riferimenti *Calibre è soggetto a frequentissimi aggiornamenti: in queste*

pagine si fa riferimento alla versione 0.6.37 del software.

L'interfaccia del software è piuttosto intuitiva: nella barra dei pulsanti nella parte superiore della finestra troviamo le principali funzioni per iniziare a gestire la nostra collezione di libri digitali (Figura A.1).

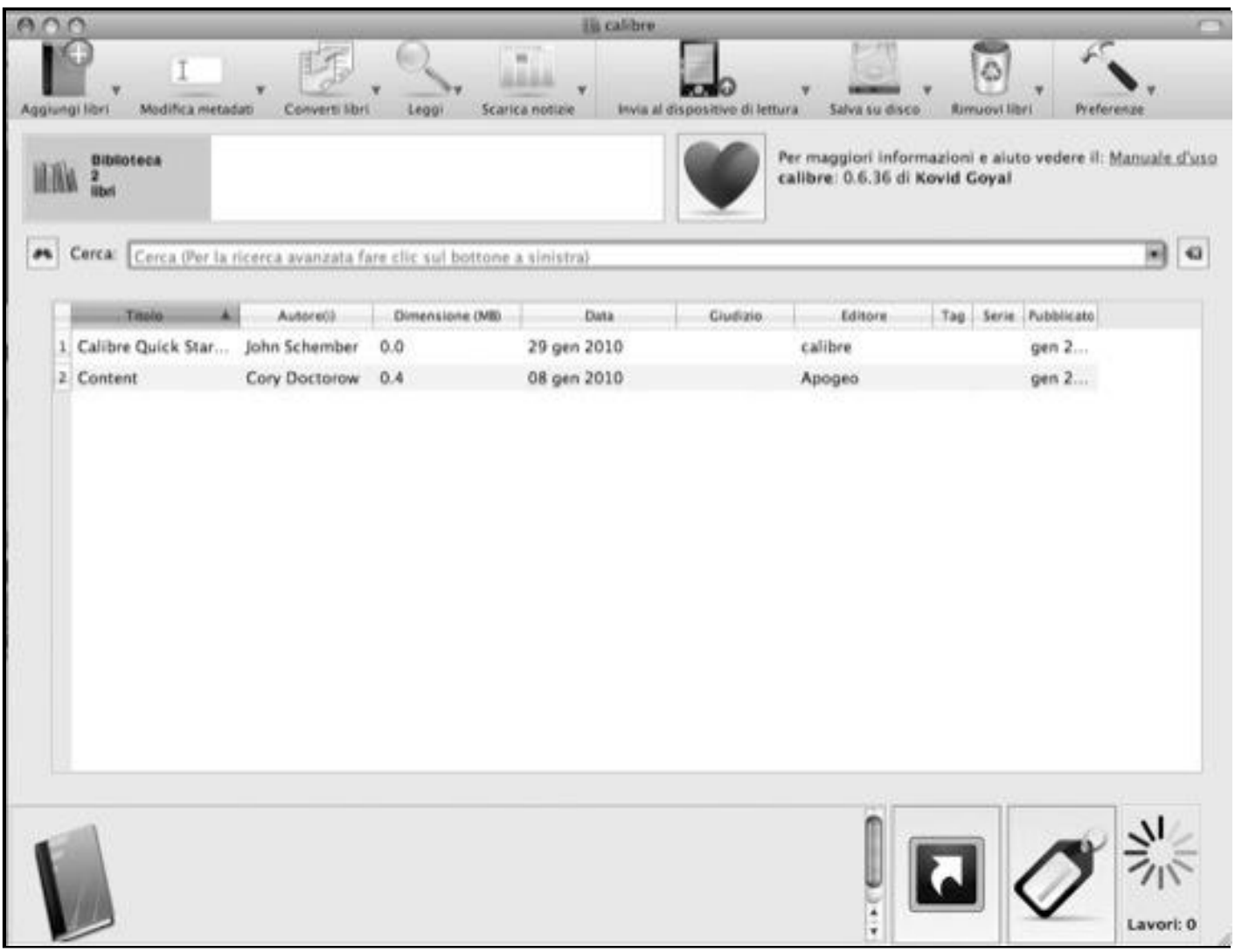

Figura A 1 La schermata iniziale di Calibre.

Vediamo alcune delle funzionalità più interessanti del programma.

- *Scarica notizie*: attraverso questo pulsante è possibile accedere a un archivio di fonti di news sul Web preimpostate e anche aggiungerne di personalizzate. Una volta individuate le fonti che più ci interessano si può avviare il download: i contenuti scaricati compariranno nella libreria, pronti per essere convertiti nel formato di cui abbiamo bisogno e inviati al nostro ereader.
- *Invia al dispositivo di lettura*: consente di sincronizzare la libreria digitale con i dispositivi di lettura. Calibre è compatibile con un numero vastissimo di ereader, ed è in grado di riconoscerli automaticamente senza grosse difficoltà.

Suggerimento *Un clic con il tasto destro su uno dei titoli nella libreria permette, tra le altre scelte, di leggere il titolo selezionato, avviando il reader integrato di Calibre.*

- *Modifica metadati*: consente di gestire tutte le informazioni sia del singolo libro sia dell'intera libreria con un elevato grado di dettaglio.

È possibile visualizzare il contenuto della libreria facendo clic sul pulsante con la freccia nella barra della parte bassa della finestra (Figura A.2).

Figura A.2 La visualizzazione della libreria attraverso le copertine. Nella parte bassa della finestra, a sinistra, potete vedere alcune informazioni sul libro selezionato.

Le funzioni di conversione dei file (pulsante *Converti libri*) sono veramente interessanti: permettono un buon grado di personalizzazione a seconda del formato che si intende ottenere e riescono a gestire un numero davvero ampio di formati, soddisfacendo le esigenze di un gran numero di dispositivi (Figura A.3).

Il pulsante *Converti libri* permette anche, attraverso il menu accessibile dalla piccola freccia rivolta verso il basso, di generare un catalogo della propria libreria in diversi formati (ePub, mobi, XML o CSV), che può essere caricato su un server e consultato online attraverso qualunque lettore sia in grado di interpretarlo, per esempio Stanza.