

# Papyrus

*Un vero software bibliografico per Macintosh*

di Francesco Dell'Orso

**P**apyrus della Research software design (RSD, Oregon) è un programma annoverabile nella famiglia dei *bibliography formatting software* o *citation managers* di cui "Biblioteche oggi" si è occupata molte volte negli ultimi cinque anni e anche di recente.<sup>1</sup> Per anni presente con la sola versione DOS, Papyrus si è guadagnato fama e mercato grazie alle capacità intrinseche, a un costo unitario mai superiore ai 100 dollari USA e ad una licenza d'uso originale. Questi fattori hanno contribuito, per esempio, a fargli mettere radici in università britanniche dove le biblioteche sono interessate a servirsi del programma in maniera istituzionale e dunque acquisiscono una speciale licenza "di sito" per offrire agli studenti, ai docenti e al personale il prodotto, ed insieme ad esso istruzione, assistenza, aggiornamento per l'uso. Il contratto per la licenza è eccentricamente variegato. In brutale sintesi: fino a un massimo di quattro database lo si può usare su quante macchine si vuole (che non siano in rete), quindi, diciamo, quattro individui con un database a testa diverso possono usare

lecitamente quattro versioni di Papyrus.

La versione Macintosh uscita nel 1999 era molto attesa. Per il suo creatore<sup>2</sup> era una sorta di voto, dal momento che la RSD aveva sempre lavorato con dei Macintosh e produceva un prodotto DOS: un esercizio spirituale? Questa edizione onora l'impegno e regala al parterre – è noto – sensibilissimo dei maccofili una vera versione Mac e non il trasferimento di una versione Windows (fra l'altro in preparazione, ma senza data prevista). Papyrus fa difatti uso di specificità Mac come: AppleEvents per comunicare con altri programmi esterni (come correttori ortografici e videoscrittura); funzione di *drag and drop* usata in profondità; help contestualissimo – anche a palloncino da sé itinerante nelle varie parti della zona di lavoro; compatibilità col WorldScript della Apple per trattare alfabeti non latini come ebraico e cirillico. Per i dati anagrafici e i requisiti di sistema si veda la tabella 1.

In sintesi, caratteristiche originali, notevoli e positive di Papyrus sono: le immagini (per ora solo se in formato PICT) possono venire

mostrate, stampate ed esportate contestualmente al record; legami orizzontali e verticali sono imoianstabili fra voci e fra record: si tratta di legami navigabili, utilizzabili in ricerca, stampa ed esportazione. Per contro, in negativo: manca di azioni trasversali su più database; l'operatività in rete è limitata a una lettura multipla; non ha client di ricerca Z39.50; è lento. Sarà la sofisticatezza di varie opzioni ad essere responsabile di una lentezza delle operazioni cruciali (salvare, ordinare, ricercare...) che non può sempre venire accolta da chi lavora come un contributo alla meditazione distesa?

L'installazione si fa pianamente, ma non viene distribuito alcun database con dati di prova. Si possono stabilire parole d'ordine di sicurezza a tre livelli: per regolare le aggiunte alla base di dati, le correzioni e le modifiche delle impostazioni generali. Buona parte della manutenzione di un database si fa da sistema operativo, dall'interno solo creazione, risanamento, vista di statistiche del database.

Il database si presenta semplificato in dati e indici, ma in realtà è articolato all'interno prevedendo i citati legami fra record e fra voci.

Liste ed indici (per non dire *authority files* e *thesauri*) sono un argomento importante. Servono in ricerca e in immissione, possono velocizzare le operazioni e contribuire al controllo di qualità riducendo gli errori di scrittura e di scelta dei termini. Quando vengono concepiti con intelligenza, rivelano la loro autonomia e non si riducono a dati derivati dai record. Questo ed altro è assicurato dai programmi specialistici per la gestione di cataloghi di biblioteca, mentre normalmente non rientra fra le caratteristiche dei BFS. Papyrus spicca in tale panorama per una concezione più compiuta e per i risultati conseguiti.

Di sistema sono forniti indici/liste

Tab. 1 - Dati anagrafici, requisiti di sistema, limiti fisici

**Prodotto:** Papyrus Macintosh v. 8.07 (dati compatibili anche regressivamente con v. 7.0.16 DOS)  
**Versione in sola lettura:** sì  
**Versione demo:** sì, completamente funzionante fino a 200 record  
**Versione in rete:** no, ma consente lettura multipla  
**Assistenza:** support@rsd.com  
**Prezzo:** ca 90 \$ US se download di programma e manuali via Internet, 100+5 se inviati su cd, 140 e spedizione se si ordinano cd e anche i 4 manuali in 2 volumi  
**Licenza (alquanto originale):** senza limiti per una macchina; su macchine – non in rete – diverse il limite è di 4 database (vedere a: <http://www.rsd.com/Site8.html>)  
**Produttore:** Research Software Design, 2718 Kelly Street, Suite 181, Portland, OR 97201 (USA) (tel. 01-503-796-1368, fax: 01-503-452-8920), e-mail: info@rsd.com  
**Sito www:** www.rsd.com con varie rubriche per problemi, FAQs, contributi degli utenti, filtri e formati supplementari, aggiunte dell'ultim'ora.  
**Requisiti:** qualsiasi Macintosh (68K or PPC); OS: MacOS >= 7.0 (uso pronunciato di Drag & Drop; Apple Events usato per comunicare con word processors e correttori ortografici, tratta alfabeti come ebraico e cirillico)  
**RAM:** 4 Mb  
**Hard-disk:** ca 10 Mb  
**Numero di database:** illimitato su una macchina, per licenza fino a 4 su macchine diverse  
**Database utilizzabili contemporeamente:** 1  
**Numero di record per database:** illimitato (16 milioni...)  
**Numero di caratteri per record:** illimitato (32.000 per campo)  
**Numero di caratteri per campo:** 32.000 a lunghezza variabile (fino a 1.000 autori e 4.000 keyword)  
**Numero di tipi di record:** 16 (se ne possono creare in più)  
**Numero di campi:** 59 di sistema a lunghezza variabile e non modificabili; se ne possono creare altri e si possono decidere gli attributi  
**Numero di stili di citazione:** ca 100 di sistema e se ne possono creare di nuovi  
**Numero di filtri di importazione:** ca 150 (e se ne possono creare di nuovi)  
**Lunghezza delle chiavi indicizzabili:** ca 240 caratteri  
**Indici liste per immissione dati e ricerca:** 3 di sistema (nomi, parole chiave e titoli di riviste) + 1 glossario generico + una lista per ogni campo definito – dal sistema o dall'utente – come indicizzato per intero, ossia non parola-per-parola, ad es.: editore, città.

per nomi, parole chiave, titoli di riviste, più un glossario globale che accoglie i termini di tutti i campi non indicizzati. Tutti i campi indicizzati hanno la loro lista, che, questa sì, è un mero riflesso del contenuto dei campi.<sup>3</sup> Ma le citate tre liste di sistema<sup>4</sup> possono avere voci che non sono legate ai record, importate anche dall'esterno. In queste è possibile aggiungere, fondere e correggere voci senza dovere intraprendere un'operazione globale sui record, ma agendo direttamente in un punto solo e con effetto immediato sui record collegati. In certi casi – ISSN, *call number*,

URL e titoli di riviste con le loro abbreviazioni – è previsto che alle voci siano aggregabili delle note. Tutte le liste sono utilizzabili durante l'immissione dei dati nelle schede; parole chiave e titoli di riviste controllano (convalidano) l'aggiunta di voci nuove dandone avviso. Tutte le liste sono utilizzabili in ricerca, e quelle di sistema mostrano anche il numero totale dei record collegati.

Come già accennato, la sofisticatezza risiede nella possibilità di avvantaggiarsi dei legami fra parole chiave, e fra record distinti. I legami sono di tipo e di profondità di-

versi con possibile discendenza gerarchica annidata (fig. 1).

Certi tipi di legami fra le voci sono già previsti (generico/specifico, parenti, antonimi), altri se ne possono aggiungere. I legami fra i record consentono ad esempio di collegare periodici che si succedono, diverse edizioni, documento contenente e contenuto anche a più di un livello di parentela. Le funzioni di ricerca, stampa ed esportazione possono utilizzare questi legami. È così possibile impostare una ricerca per cosiddetta esplosione dei termini – in gergo “cerca un BT e tutti i suoi NT” ovvero “un termine e quelli ad esso subordinati” (*downposting*) – oppure si possono irretire un'opera e le sue altre manifestazioni. La ricerca può infatti venire indirizzata per tipo di legame, ad esempio per cercare “quel titolo e solo le sue traduzioni” oppure “quel periodico e i suoi precedenti”, “un termine e i sinonimi”; si possono indicare anche quanti livelli di legami considerare, in profondità verticale o estensione orizzontale che sia.

Complessivamente è una gestione piuttosto ricca, efficiente, con attenzione all'autonomia e alla specificità delle voci, che introduce elementi di relazionalità e di controllo delle intestazioni non comuni in questi programmi. Si rimpiange che la stampa specifica a queste liste sia povera: o si stampa tutta la lista o l'arborescenza di un termine, e anche che poco sia stato fatto per il controllo della struttura tesauroale: i legami sono immediatamente bidirezionali e nessun termine può fare riferimento a se stesso, ma a parte ciò uno è libero di complicarsi la vita ponendo come termine maggiore di un altro un suo minore o utilizzandone uno escluso.

Si possono trattare documenti di sedici tipi diversi e se ne possono aggiungere fino a cento, in ogni documento c'è posto per cin- ➤

Fig. 1 - La lista delle parole chiave con legami di vario tipo e vario livello, riutilizzabili in ricerca

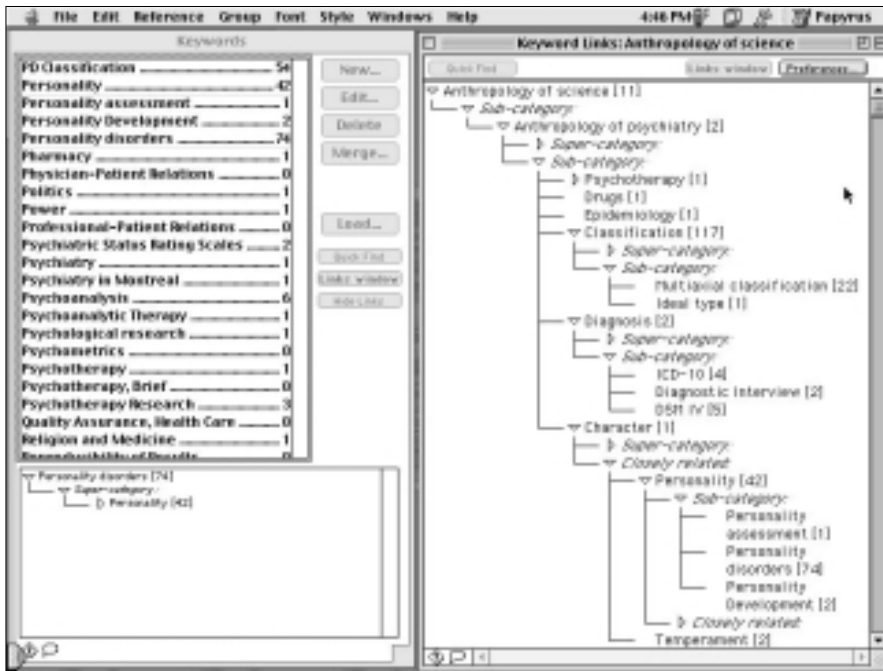
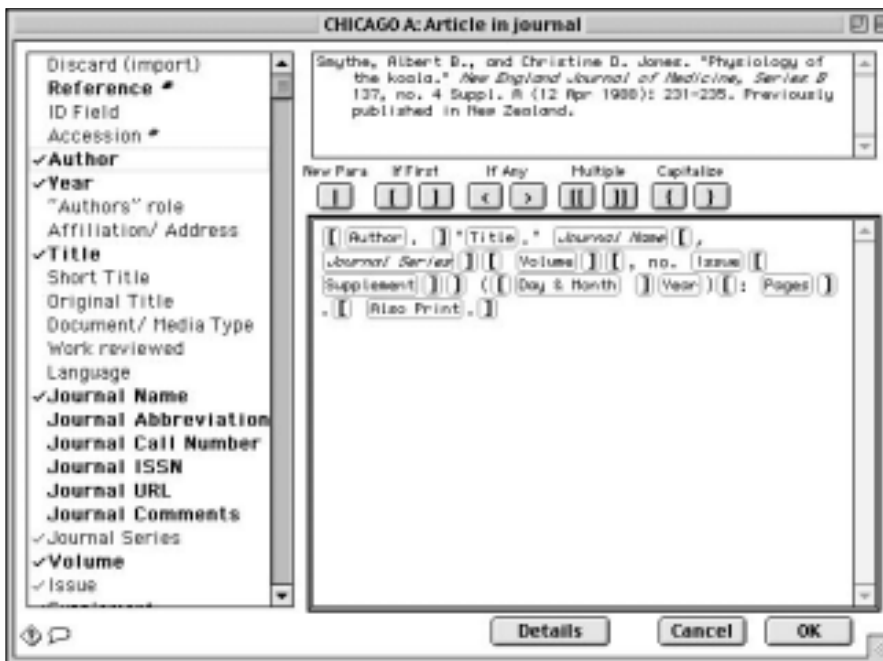


Fig. 2 - Scheda per immettere dati in un record: un articolo di rivista



quantanove campi già definiti e fino a cento per le aggiunte. I campi definiti hanno i loro attributi non modificabili, ma per quelli che si aggiungono sta invece a noi ca-

ratterizzarli liberamente per tipo di contenuto (numero, data, testo, nome), tipo e modo di indicizzazione (tutto il campo come una stringa, oppure parola per parola),

ripetibilità, immissione obbligatoria, correzione ortografica (grazie a software esterno) ecc.

Internet è tenuta in conto nella misura in cui: si possono attivare – con il browser preferito – indirizzi URL registrati nelle schede; si stampano record in formato HTML con URL attivi (in realtà tutto un record diventa un legame “cliccabile”), pubblicabili dunque subito nel web; si possono trattare risorse elettroniche – dunque anche di rete – come tipi di documenti. Altre entità esterne al database, file sonori, grafici non PICT, testi scritti, tabelle ecc. sono sempre attivabili col mouse, cliccando sul loro nome, e così ricorrendo automaticamente al programma (videoscrittura...) che li gestisce.

L'immissione di dati nelle schede (fig. 2) si può avvantaggiare – oltre che di normali funzioni di copia – della descritta consultazione di tutte le liste di campi indicizzati con convalida delle parole chiave e due titoli di rivista aggiunti per la prima volta. Per intercettare i doppietti ci sono due criteri, più o meno sofisticato, ma non si possono alterare. Le correzioni globali permettono di aggiungere, rimpiazzare, distruggere, spostare il contenuto di un campo all'interno di un gruppo di record selezionati, non di modificare il tipo di record, e, pur senza fare usare i caratteri jolly, permettono di tenere sotto controllo, accettare o rifiutare, ogni correzione.

L'altra forma di immissione dei dati, l'importazione, riceve un'attenzione particolare in Papyrus: l'autore vi dedica pagine e pagine appassionate dei suoi manuali, sciornandovi la coratella degli sforzi fatti e sempre da fare per acchiappare i dati (de)strutturati nelle forme più varie. Per farlo si usa lo stesso approccio metodologico di EndNote e cioè: per leggere un record in entrata si fa come se lo si dovesse produrre in uscita, il linguaggio

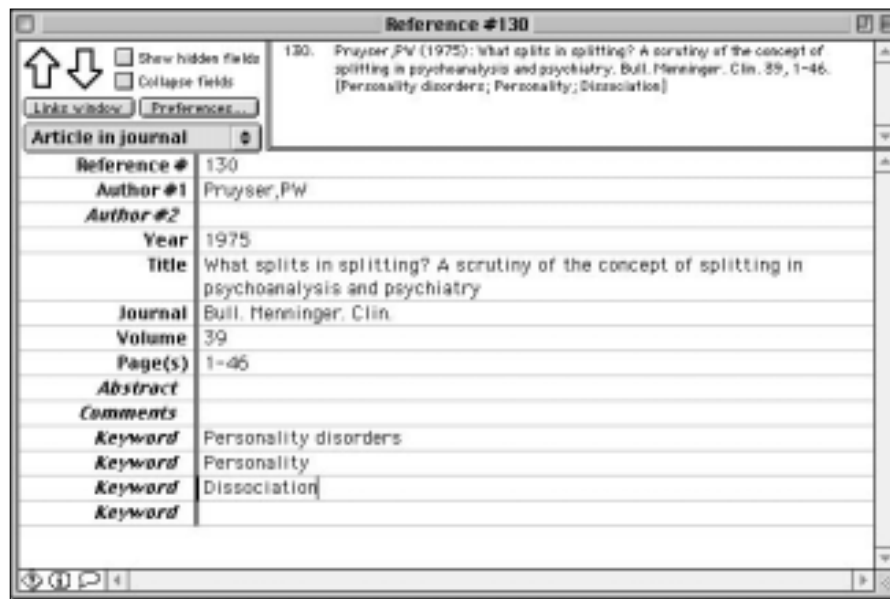
per decodificare ciò che si deve importare è il medesimo che si usa per formattare l'output. Semplice e ingegnoso assunto che ci lascia con poche istruzioni e con tanta fantasia, genialità e pazienza da applicare. Si arriva a potere importare formati ad etichette, o discorsivi, fondendo, sopprimendo, aggiungendo e frammentando il contenuto dei campi. L'approccio di Papyrus e di EndNote dà il meglio proprio quando il formato d'ingresso è discorsivo e quando si devono frammentare i campi: in realtà tutta la lettura è un'interpretazione e frammentazione. Si dispone comunque di funzioni già pronte per un primo trattamento dei dati<sup>5</sup> e molti filtri già pronti per decine di database commerciali, per i dati provenienti dai prodotti concorrenti e per l'importazione automatica da BioMedNet Evaluated Medline. Poi magari, invece, per importare un classico formato delimitato (CDF), occorre farlo interamente da sé, il che comunque non è immediato (anche se RSD si dichiara disponibile a farlo gratis...). Ancora una volta MARC e ISO 2709 non sono considerati. Anche l'operatività dell'importazione è curata: si può leggere un record alla volta confermandone l'ingresso o no, intercettare e decidere la sorte dei duplicati, vedere in anteprima lo svolgimento e l'esito della fase di digestione dei record, ma attenzione: godersi tutte le opzioni porta ad una velocità operativa pari all'osservazione in vivo del dischiudersi di una margherita con rugiada colante.

Il catalogo si presenta così come lo si disegna e dunque anche la lista tabulare con un record per riga, poniamo, è uno dei formati da impostare e che consente poi di scorrere il catalogo. Le liste dei campi indicizzati sono uno dei modi grazie a cui si può consultare e, quando si tratta di soggetti, titoli e nomi, ricercare il contenuto.

**Fig. 3 - Finestra di ricerca con pulsanti per gli operatori e opzione per i legami fra voci da considerare**



**Fig. 4 - Costruzione di un formato di visualizzazione**



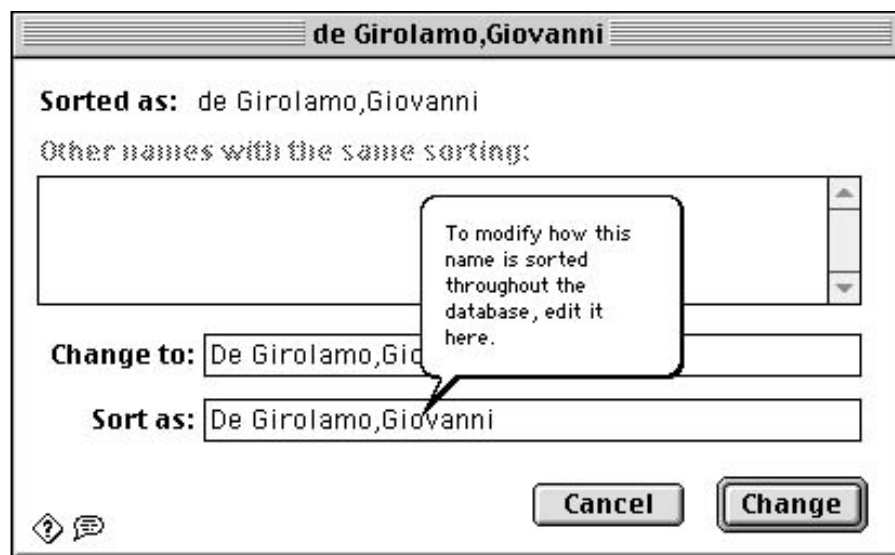
Naturalmente c'è anche la modalità di ricerca più tecnica con la composizione di espressioni nutrite di operatori logici e parentesi, troncamenti a dritta e a manca, carattere jolly interno (m\*ssa) ed è assistita da bottoni, liste, promemoria (fig. 3). "Tutti i campi sono ricercabili, ma alcuni lo sono meglio di altri" e dunque sarà più rapido cercare

per i campi "indicizzati" come numero di record, nome, data, editore, titolo. Il soccorso di una ricerca su tutto il testo non viene negato, ma esclude sia le parole chiave che i titoli di riviste. Mancano modalità di ricerca in linguaggio naturale o per approssimazione (semantica, ortografica, fonetica...) o con classificazione dei risultati. ➤

E per ora nessun client per una ricerca secondo il protocollo Z39.50 viene offerto (verrà, di sicuro). Espressioni e risultati di ricerche possono venire accantonati. Da rammentare la già descritta ricerca che si avvale dei legami fra parole chiave e fra record.

Papyrus appartiene ai BFS, dunque i suoi cento stili predefiniti non mancano e altri se ne possono creare, il linguaggio è abbastanza efficace all'interno di poche istruzioni formali (fig. 4) che tuttavia meriterebbero una spiegazione più formale nelle centinaia di pagine dei manuali. Manca di "SE ... ALLORA", ossia di un comando per porre clausole condizionali, ma molte ne incorpora, come i prodotti parenti, all'interno

Fig. 5 - Assegnazione ad un nome di persona di una forma per l'ordinamento diversa da quella di visualizzazione



di decine di opzioni preparate da selezionare col mouse (se abbreviare o no un prenome, se apporre "p." alle pagine ecc.). Presenta subito l'effetto delle istruzioni che si danno grazie alla visualizzazione contestuale e dinamica di un record formattato. Si possono ottenere un formato a colonne, titolazione corrente daccapo o a piè, stampe con intestazioni in esponente - non a più di un livello -

e con i record su bordati, indici con rimando al numero di record.<sup>6</sup> Per l'output la

scelta di formati è di larghe vedute, prevede: RTF, Word, WordPerfect, Nisus Writer, TeX/LaTeX, MacWrite. Per contro manca una visualizzazione delle stampe in anteprima a video: occorre produrre un file e aprirlo.

Il *sort* dell'intero database è limitato a due opzioni: autore o numero di scheda, ma nelle stampe ogni stile può avere il suo criterio di ordinamento e lì il numero di livelli annidati è "senza limite", sotto nostro controllo. Accuratezza apprezzabile e non comune viene riservata al trattamento dei nomi che possono avere una forma *sort* diversa da quella in chiaro, per esempio "Lafontaine" può valere per ordinare un nome scritto "de La Fontaine, Jean" (fig. 5).

Quanto all'altra forma di output che consiste nell'esportazione dei dati, sono previste delle opzioni standard in direzione di altri BFS, mentre le forme delimitate o etichettate restano a cura di chi se ne vuole servire.

La formattazione del manoscritto offre un'integrazione fra programma e videoscrittura alquanto unilaterale: è Papyrus, e i ponticelli software da lì edificati, che si preoccupano

pa di fare arrivare dentro al documento i segnalini, copiandoceli, cioè non si agisce restando dentro al word processor. Buone sono le capacità di alterazione delle citazioni testuali caso per caso: per cui si possono distinguere stesso nome e stesso anno, cambiare differenzialmente ogni singolo richiamo, impostare la prima citazione diversa dalle successive. Il documento formattato rimane una copia dell'originale che non viene toccato.

La documentazione è molto buona, scritta con spirito sia umoristico che didattico: ci sono vere e proprie lezioni che fanno avanzare ed apprendere. Il manuale introduttivo è completato da uno di riferimento. Si lamenta la mancanza di sezioni, ancorché brevi, che illustrino una volta per tutte, in modo formale ed esauriente, le funzioni tecniche che si avvalgono di istru-

zioni codificate come ricerca, formattazione e importazione. I manuali sono talmente curati che hanno un costo a parte, la versione proposta scaricabile da Internet o su cd li contiene in formato HTML da usare, completi, come help all'attività in linea.

Complessivamente: un ottimo prodotto, da raccomandare a chi ha un Macintosh e sa apprezzare la finezza e la lentezza – meglio verrà. ■

#### Note

<sup>1</sup> I nomi dei prodotti citati sono marchi registrati delle rispettive case produttrici. Non intrattengo nessun rapporto commerciale o promozionale con la Research Software Design. Manoscritto terminato il 27 gennaio 2000. Per una breve, molto elogiativa, recensione di Papyrus, si veda: ELLEN M. QUARDOKUS, *Papyrus 8.0.6 for Macin-*

*tosh*, "HMS Beagle: The BioMedNet Magazine", 1999, issue 67 (in linea dal 26 Novembre, 1999), disponibile presso: <<http://www.biomednet.com/hmsbeagle/67/reviews/sreview>>.

<sup>2</sup> Dave Goldman: autore e factotum, boss della ditta, estensore dei manuali, moderatore della lista di discussione.

<sup>3</sup> Che un campo sia indicizzato, è un attributo di sistema per alcuni campi, e va assegnato esplicitamente per i campi che vengono creati.

<sup>4</sup> La gestione della lista dei nomi è più ristretta, ad esempio voci a sé possono esistere solo nella misura in cui dei nomi sono stati importati con una forma sort diversa, cfr. oltre e in figura 5.

<sup>5</sup> Ad es. plug-in per l'ordinamento alfabetico delle etichette.

<sup>6</sup> Lo spazio di manovra qui può rivelarsi angusto: se le voci in esponente sono estratte da un campo indicizzato parola per parola, esse non potranno che essere parole singole, il *sort* non è stato adeguatamente separato dall'indicizzazione per fini di ricerca allora che è una questione di output.