



FONDAZIONE
LELIO E LISLI BASSO ISSOCO

Alcune questioni metodologiche e quesiti su ordinamento, conservazione e schedatura di archivi privati in formato digitale.

Note preliminari.

Innanzitutto va segnalato che è stato fatto un backup con la documentazione così come è stata consegnata dai soggetti produttori; backup che sarà conservato anche alla fine dell'ordinamento, in quanto è evidente che, a differenza degli archivi cartacei, una volta che si interviene sui file (spostandoli o rinominandoli) non è più possibile risalire o individuare cosa sia originale e cosa "critico".

Si tratta di un fondo non eccessivamente grande (circa 3.500 file), con documentazione sostanzialmente del 2011 (probabilmente fondi molto più grandi presenterebbero diverse questioni su gestione e ordinamento).

Denominazioni dei file.

Tranne rare eccezioni, dichiarate in fase di schedatura, al momento non sono stati modificati i nomi dei file; questo però pone il problema della gestione di nomi file discorsivi e lunghi in vista della conservazione. In Windows, un singolo percorso (ad esempio C:\ITA_FLLB_REF_SUB01_02\nomefile.doc) non può superare i 260 caratteri. Per questo motivo può talvolta verificarsi un errore durante la copia di un file con un nome molto lungo. In questo senso si pone allora la questione su come conciliare l'esigenza tecnica con il mantenimento di informazioni nella forma originaria.

Unità archivistiche.

Un quesito preliminare riguarda le unità archivistiche, che in un fondo digitale coincidono con *directory* (definibili come locazione del file system in cui sono presenti file o altre directory); ci si è posti la domanda se, in fase di inventariazione, possano essere definite "fascicoli" o se ci sia un termine più appropriato per designarle.

Ordinamento.

Ci si è trovati di fronte a due tipi di documentazione:

- documentazione organizzata dal soggetto produttore in vista della donazione all'istituto conservatore;
- copia di backup di cartelle di sistema di un Pc (cartelle *Desktop* e *Documenti*).

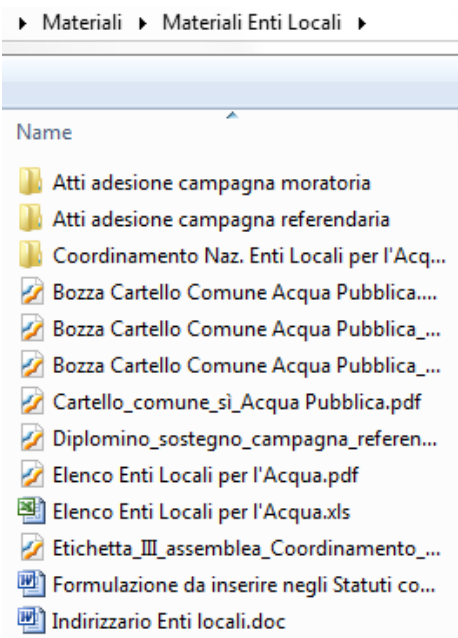
Nel primo caso si è sostanzialmente mantenuta l'organizzazione originaria, cercando di adattarla a una struttura archivistica, ovvero considerando le directory più "alte" dei livelli (serie o sottoserie) e le directory in esse contenute come unità archivistiche.

In linea di massima sono stati utilizzati i nomi delle directory originarie come denominazioni di livelli o unità archivistiche; quando è stato ritenuto opportuno cambiare i nomi delle directory di livello alto ("Serie" o "sottoserie"), perché troppo generici o poco rispondenti a criteri archivistici, il nome originario è stato comunque segnalato nella scheda descrittiva (nel campo *Ordinamento e struttura*).

Nel secondo caso – si segnala che al momento non è stato svolto un vero e proprio ordinamento di tale documentazione, quanto piuttosto una mappatura dettagliata – sarebbe necessario procedere invece a un lavoro di smistamento file per file; non si ritiene infatti opportuno pensare di offrire percorsi tipo "desktop" o "backup". Inoltre si è proceduto allo scarto (eliminazione) di file "di servizio", come eseguibili di programmi, directory di sistema (vuote), ecc.

File sciolti.

Molto spesso all'interno delle directory (che in fase di ordinamento sono state considerati dei livelli) erano presenti sottodirectory (considerate delle unità archivistiche, o se così si possono chiamare, dei fascicoli) e documenti sciolti.



[nell'es. Subfondo 01, Serie 3, sserie 1, situazione originaria]

Tali documenti sciolti sono stati inseriti in nuove directory denominate Miscellanea, sia in fase di schedatura, che fisicamente sui file.

Naturalmente in fase di inventariazione è stato segnalato (nel campo *Note*) che la directory non è originale (ma in realtà si può dedurre anche dal fatto che il titolo dell'unità archivistica/fascicolo è critico piuttosto che originale).

Date.

Pochi sono i documenti datati (a parte la posta elettronica), ovvero con la data riportata nel testo o nella denominazione del documento stesso.

Nella definizione degli estremi cronologici di un'unità archivistica (directory) - naturalmente nel caso in cui la maggior parte dei documenti fosse priva di date esplicitata nel testo - è stata considerata la data riportata nelle "Proprietà" di ogni file, che permette di recuperare le date di creazione e Ultima modifica.

Si tratta in realtà di una datazione labile (che in realtà può anche essere modificata), però risulta utile per dare una datazione di massima a file altrimenti senza data. In fase di schedatura è stato naturalmente verificato che tale data (è stata scelta la data Ultima modifica) fosse compatibile con la creazione dell'archivio; può infatti capitare che un file assuma come data quella del backup o dello spostamento da un supporto all'altro, quindi non pertinente con il fondo.

In fase di inventariazione si è scelto di riportare (nel campo *Note alla data*) la provenienza della datazione (utilizzando per convenzione la formula "data Ultima modifica").

Condizionamento (se così si può chiamare per documenti digitali).

Problema di come ordinare fisicamente i file all'interno delle directory (di norma i sistemi operativi li ordinano alfabeticamente); si può selezionare l'ordine per data (Data Ultima modifica, sembra il più appropriato).

Naturalmente questo sembra un ordinamento molto labile (rispetto all'ordine cronologico che può essere dato a documenti cartacei). Per renderlo stabile sarebbe da rinominare, numerandolo, ogni singolo documento (ad es. prima la minuta, poi la versione definitiva, poi la versione in pdf, ecc.), ma risulta molto oneroso in termini di tempo/lavoro (non paragonabile al condizionamento di documenti cartacei).

Inoltre l'inserimento di una numerazione che preceda il nome del file, ripropone la questione dei nomi dei file troppo lunghi.

Inventariazione.

Si è ritenuto necessario attribuire a ogni livello e ogni unità archivistica un Reference code (previsto dalle Isad-G), da riportare sull'inventario e conseguentemente sulle directory fisiche.

Nello specifico, il Reference code attribuito risulta così composto:

ITA FLLB REF

(Italia - Fondazione Basso - Fondo Campagna referendaria...)

SUB01 (subfondo 01: Segreteria nazionale del Forum italiano dei movimenti per l'acqua)

SUB02 (subfondo 02: Legambiente)

01 (indicazione di serie o sottoserie, con due cifre)

002 (indicazione delle unità archivistiche, a tre cifre)

Quindi

ITA FLLB REF SUB01 04 003

sarà

Subfondo 1, Serie 4, fasc. 3.

Posta elettronica.

Il fondo comprendeva scambi di e-mail in vari formati:

1. e-mail singole, come file in formato eml

Il nome del file è risultato composto da una stringa contenente i seguenti dati:

[nome account] Oggetto – mittente – data – numero.eml

Si è presentato il problema di avere un ordinamento della corrispondenza; i file erano in ordine alfabetico e la data dell'Ultima modifica dei file risultava essere quella del salvataggio degli stessi.

Si è perciò deciso di spostare la data contenuta nella denominazione del file (fortunatamente nel formato anno/mese/giorno) all'inizio del nome file, in modo da presentare le e-mail in ordine cronologico. Naturalmente ciò è stato possibile dato il numero non eccessivo dei file da doversi rinominare (in quanto non è stata trovata una soluzione in tal senso per utilizzare un'applicazione in grado di rinominare in automatico i file secondo le necessità).

Si pone ora la questione sulla denominazione del file che risulta problematica in quanto composta da un numero eccessivo di caratteri. (vedi sopra). La proposta è lasciare come nome file solo la data (tutti gli altri dati sono all'interno della mail, quindi non si perde alcuna informazione). Una seconda ipotesi, più orientata verso l'utenza, è lasciare data e oggetto originale (per permettere all'utente di aprire solo le e-mail di interesse); ma in questo caso il nome file risulta nuovamente troppo lungo; sarebbe quindi necessario sintetizzarlo, facendo un'operazione soggettiva e onerosa in termini di tempo.

2. backup da gestore di posta Thunderbird e backup da gestore di posta Outlook.

Si è scelto di importare tutto (anche il carteggio proveniente da Outlook) su gestore di posta Thunderbird.

Questa soluzione non dà problemi nel senso dell'ordinamento fisico della posta.

Dal punto di vista della schedatura archivistica, sono stati creati due unità archivistiche, uno per la posta in arrivo, una per la posta inviata; oppure tante unità archivistiche quante sono le cartelle tematiche create dal soggetto produttore.

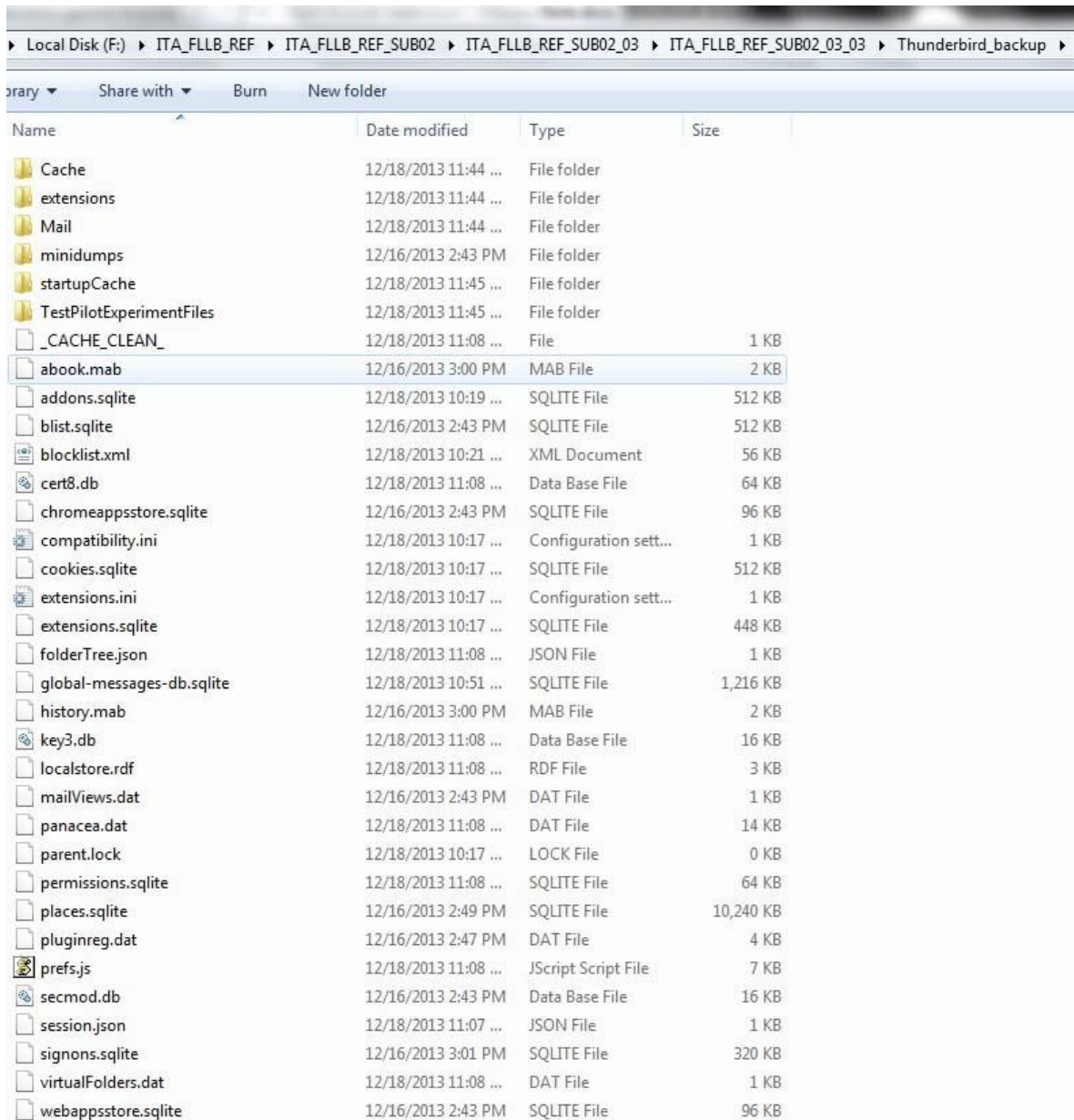
Posta elettronica / Client di posta elettronica.

Analisi su alcune questioni metodologiche riscontrate nel caso in cui il soggetto produttore consegni la propria posta elettronica in formato di backup di un client di posta elettronica (ovvero un programma "che consente di gestire la composizione e l'organizzazione di e-mail da parte dell'utente del servizio, nonché la ricezione e la trasmissione da e verso un server di posta" – definizione da Wikipedia).

Si pone la questione su come creare un legame intuitivo tra schedatura/inventariazione e documentazione fisica (file):

- fisicamente si possiede un backup del gestore di posta (nel caso specifico di Thunderbird), che sarà copiato sul supporto dove si conserva la documentazione

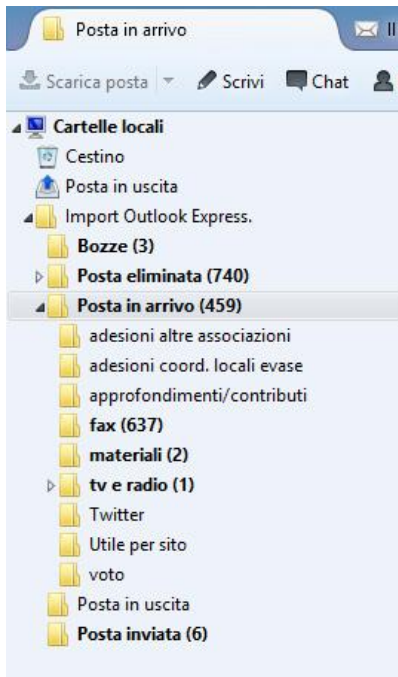
[es. che riporta la barra degli indirizzi recante il Reference code e i file in oggetto]



Name	Date modified	Type	Size
Cache	12/18/2013 11:44 ...	File folder	
extensions	12/18/2013 11:44 ...	File folder	
Mail	12/18/2013 11:44 ...	File folder	
minidumps	12/16/2013 2:43 PM	File folder	
startupCache	12/18/2013 11:45 ...	File folder	
TestPilotExperimentFiles	12/18/2013 11:45 ...	File folder	
_CACHE_CLEAN_	12/18/2013 11:08 ...	File	1 KB
abook.mab	12/16/2013 3:00 PM	MAB File	2 KB
addons.sqlite	12/18/2013 10:19 ...	SQLITE File	512 KB
blist.sqlite	12/16/2013 2:43 PM	SQLITE File	512 KB
blocklist.xml	12/18/2013 10:21 ...	XML Document	56 KB
cert8.db	12/18/2013 11:08 ...	Data Base File	64 KB
chromeappsstore.sqlite	12/16/2013 2:43 PM	SQLITE File	96 KB
compatibility.ini	12/18/2013 10:17 ...	Configuration sett...	1 KB
cookies.sqlite	12/18/2013 10:17 ...	SQLITE File	512 KB
extensions.ini	12/18/2013 10:17 ...	Configuration sett...	1 KB
extensions.sqlite	12/18/2013 10:17 ...	SQLITE File	448 KB
folderTree.json	12/18/2013 11:08 ...	JSON File	1 KB
global-messages-db.sqlite	12/18/2013 10:51 ...	SQLITE File	1,216 KB
history.mab	12/16/2013 3:00 PM	MAB File	2 KB
key3.db	12/18/2013 11:08 ...	Data Base File	16 KB
localstore.rdf	12/18/2013 11:08 ...	RDF File	3 KB
mailViews.dat	12/16/2013 2:43 PM	DAT File	1 KB
panacea.dat	12/18/2013 11:08 ...	DAT File	14 KB
parent.lock	12/18/2013 10:17 ...	LOCK File	0 KB
permissions.sqlite	12/18/2013 11:08 ...	SQLITE File	64 KB
places.sqlite	12/16/2013 2:49 PM	SQLITE File	10,240 KB
pluginreg.dat	12/16/2013 2:47 PM	DAT File	4 KB
prefs.js	12/18/2013 11:08 ...	JScript Script File	7 KB
secmod.db	12/16/2013 2:43 PM	Data Base File	16 KB
session.json	12/18/2013 11:07 ...	JSON File	1 KB
signons.sqlite	12/16/2013 3:01 PM	SQLITE File	320 KB
virtualFolders.dat	12/18/2013 11:08 ...	DAT File	1 KB
webappsstore.sqlite	12/16/2013 2:43 PM	SQLITE File	96 KB

- la posta elettronica è visibile aprendo il client di posta elettronica (nel caso specifico Thunderbird)

- le cartelle in cui la posta è organizzata (es. posta in arrivo, posta inviata, cartelle create dal soggetto produttore) sono visibili sul programma di posta elettronica



- le unità archivistiche (corrispondenti alle cartelle di cui sopra) saranno ordinate, schedate e numerate (nel nostro caso con il software Gea);



La questione – che nella pratica sta dando dei problemi di metodologia – verte sul fatto che vari fascicoli (la corrispondenza via mail) con una loro segnatura archivistica siano fisicamente conservati insieme in una directory di backup del client di posta elettronica, a cui è impossibile attribuire fisicamente una segnatura archivistica. Tale segnatura (numero di fascicolo) risulterebbe visibile solo aprendo il programma di posta elettronica.



[in questo es. alla cartella "Materiali" è stato attribuito un Reference code (ipotetico)]

Inoltre ci si chiede anche se sia lecito intervenire sulle cartelle del programma [come nell'es. di lato] rinominandole, per assegnare loro un percorso archivistico (es il Reference code).

Posta elettronica / Client di posta elettronica / Ordinamento

Un interrogativo riguarda il caso in cui sia evidente che il soggetto produttore abbia creato – all'interno del client di posta elettronica - delle cartelle per organizzare la corrispondenza. Il lavoro da parte del soggetto produttore non è però stato portato a compimento: ci sono cartelle contenenti poche e-mail e soprattutto e-mail con lo stesso tema nelle cartelle di sistema *Posta in arrivo* o *Posta inviata*. Ci si pone il quesito se sia lecito, in fase di ordinamento, completare l'organizzazione della posta elettronica così come impostato dal soggetto produttore e quindi smistare le e-mail relative a temi per cui esiste una cartella specifica [come si può vedere nell'es. sopra], conservate in *Posta in arrivo* o *Posta inviata*.

Posta elettronica / Link su e-mail.

Nelle mail ci sono spesso link a pagine Web; al momento dell'ordinamento si è verificato che tali link sono funzionanti (bisogna però considerare che la posta risale al 2011). Per non perdere le informazioni (sarebbe come perdere il contenuto di una lettera) sarebbe necessario salvare le pagine Web presenti nelle e-mail (ma anche nei documenti), ossia sotto forma di documento HTML consultabile anche offline con qualsiasi browser Web.

Tale lavoro, naturalmente possibile, risulta piuttosto oneroso in termini di tempo.

Scarto.

Come già segnalato sono stati eliminati eventuali file di sistema o programmi; file vuoti (molto rari).

Relativamente alla posta elettronica, dalle e-mail consegnate originariamente sono state eliminati i file di errore di consegna messaggio e di conferma lettura.

Consultabilità.

Trattandosi di documentazione molto recente, si presentato dei problemi sull'accesso alla consultabilità.

Relativamente ai dati sensibili, si pone la questione su come tutelare i soggetti interessati. Ad esempio, riguardo la posta elettronica, è stato esplicitamente chiesto dai soggetti produttori di tutelare la riservatezza delle persone iscritte alle mailing list organizzate dai Comitati, ovvero di non rendere visibili i nominativi degli iscritti.

Il fondo in questione è inoltre ricco di documenti (tabelle, liste, ecc.) contenenti dati personali (indirizzi, recapiti, ecc.) di attivisti, volontari, sottoscrittori di iniziative.

È evidente che il contenuto di tali tabelle può essere interessante al fine dello studio del movimento (numero di aderenti, provenienza, ecc), ma rimane l'interrogativo su come salvaguardare i dati e la riservatezza personale.

Altra questione è come dare in consultazione tale fondo agli utenti (a prescindere dei problemi esposti al punto precedente). Si potrebbe far consultare su hard disk esterno, però facendo in modo che nessun file possa essere modificato (quindi solo se possibile impostare una consultazione in sola lettura). Sarebbe necessaria però anche una postazione di consultazione che abbia inibito l'accesso al web o a eventuali Usb pendrive, per evitare la copia non controllate del fondo o parti del fondo.

(Simona Luciani – Archivio storico Fondazione Basso)