

**LE DOTAZIONI INFORMATICHE DELL'ARCHIVIO STORICO DELLA BANCA
D'ITALIALE DOTAZIONI INFORMATICHE DELL'ARCHIVIO STORICO DELLA BANCA
D'ITALIA**

di Angelo Battilocchi

Intervento al convegno

*“La memoria d'impresa - Storia, archivi,
organizzazione, marketing”*

Villalago di Piediluco - Terni, 21-23 aprile 1999

La gestione dell'archivio storico della Banca d'Italia è stata sempre caratterizzata dall'utilizzo di tecnologie avanzate. Ciò è avvenuto anche grazie alla posizione funzionale della struttura che, essendo pienamente inserita all'interno dell'organizzazione aziendale, ha potuto beneficiare dell'evolversi delle dotazioni di cui si avvaleva l'amministrazione. In alcuni casi anzi, come nell'adozione della tecnologia digitale per la riproduzione documentale, l'archivio storico è all'avanguardia anche all'interno della banca.

La riproduzione su microfilm

Un esempio di ricaduta positiva delle scelte aziendali sul settore archivio è dato dalla microfilmatura dei documenti, che iniziò subito dopo la costituzione della Sezione storica, alla fine degli anni sessanta, parallelamente al lavoro di riordino dei fondi e di stesura dei primi inventari ed elenchi di consistenza. Si applicava così immediatamente all'archivio storico l'esperienza da poco avviata per l'archivio corrente che prevedeva la ripresa su microfilm della corrispondenza in arrivo e in partenza. La microfilmatura degli archivi correnti era legata, sia detto per inciso, ad un sistema di classificazione informatizzata degli atti mediante parole chiave e codici di argomento, che sia pure archivisticamente non ortodosso (e alla prova dei fatti inadatto) rappresentava per l'epoca un primo interessante esempio di applicazione di tecniche informatiche alla gestione documentale.

La microfilmatura della corrispondenza corrente fu sospesa quando venne modificato il sistema di classificazione. Proseguì invece la ripresa dei documenti della Sezione storica, fino al recente subentro della tecnologia digitale. Il risultato di questi quasi trenta anni di lavoro è racchiuso in circa 9.000 bobine di microfilm in due formati (16 e 32 mm) che coprono il 55% dell'attuale consistenza dell'archivio.

Durante questo periodo si sono avvicendate diverse tecnologie, dal planetario con ripresa dall'alto dei documenti e indicizzazione meccanica ai Reliant, dotati di alimentazione semiautomatica e predisposti per la ricerca diretta dei fotogrammi. Si sono inoltre acquisiti una professionalità e un *know how* nel trattamento delle carte d'archivio che costituiscono un patrimonio prezioso al di là dell'evolversi dei sistemi di acquisizione.

L'inventariazione elettronica

Nel 1988, in vista del centenario della Banca, celebrato nel 1993, ha avuto inizio il progetto di inventariazione elettronica integrale dei fondi. Fu costituito un gruppo di studio che si avvale della consulenza di archivisti esterni e fu elaborata una scheda di rilevazione piuttosto analitica, tuttora in uso. Per l'inventariazione fu scelto il prodotto di information retrieval BASIS, in grado di indicizzare tutte le parole contenute nei campi; fu quindi scartata l'idea di affidarsi all'individuazione di parole chiave. Il programma e i dati risiedevano su un elaboratore remoto di grandi prestazioni: di nuovo un esempio di come l'archivio abbia potuto beneficiare significativamente dei vantaggi dati dal suo pieno inserimento nella rete informatica dell'istituto.

Il tracciato record comprende 92 campi¹, suddivisi in tre aree: dati originali, dati aggiuntivi e dati di collocazione degli originali e di riferimento al microfilm.

Nel primo gruppo sono compresi intestazione, estremi cronologici, dati di segnatura; nel secondo il campo note, che contiene una descrizione in forma libera del contenuto, ed altri campi significativi per la ricerca, quali mittenti, destinatari e citazioni. Del terzo gruppo fanno parte i dati di collocazione fisica degli originali, i riferimenti al microfilm e un campo riservato alle annotazioni di carattere archivistico.

La finalità del progetto era quella di conciliare i vantaggi dati dall'informatica per l'accesso alle informazioni con la rigorosa applicazione del metodo storico per quanto riguarda le chiavi di ordinamento, basate rigorosamente sulle segnature e sulle provenienze

¹ Il numero indubbiamente elevato deriva dalla presenza di molti campi tecnici richiesti dal sistema informatico. Va inoltre precisato che naturalmente non tutti i campi vengono necessariamente compilati in tutte le schede.

per il materiale classificato, e su aggregazioni omogenee per quanto riguarda la documentazione priva di segnatura.

Il sistema, che oggi copre 23 tra i fondi più importanti permette quindi di offrire risposte sia a chi si rivolge all'archivio con la classica domanda "cosa c'è sul tale argomento?" sia a chi segua un approccio di ricerca più correttamente funzionale, di tipo deduttivo, che parta dal quadro istituzionale per arrivare alle materie. La ricerca informatizzata in archivio dovrebbe infatti assomigliare il più possibile alla consultazione di un ideale inventario tradizionale dotato di note e indici infiniti. In altre parole occorre salvaguardare la sistematicità delle informazioni, pur sfruttando appieno l'opportunità di un accesso di tipo "random". E' di vitale importanza che si mantengano ben presenti queste due esigenze perché lo strumento risponda pienamente al fine: se infatti nel caso di ricerche puntuali (un nome, un toponimo, una ragione sociale) la ricerca casuale è estremamente efficace, quando si tratti di ricercare concetti o materie di definizione più sfumata una conoscenza dettagliata delle funzioni e delle competenze degli uffici produttori è requisito irrinunciabile.

Per migliorare la fruibilità di quest'ultima modalità di accesso si è recepita di recente nel sistema la struttura ad albero prevista dalle ISAD (G), con una scheda di fondo, una per sottoserie² ed una per fascicolo o sottofascicolo. Per ora la scheda di fondo non è integrata nel sistema, ma coincide con le descrizioni contenute nella "Guida all'archivio storico" pubblicata nel 1993. Con l'ultimo aggiornamento del programma di archiviazione, realizzato in occasione di recenti integrazioni al sistema di riproduzione ottica dei documenti, sarà possibile inserire nel sistema file di tipo WORD, che permetteranno l'immissione di descrizioni di lunghezza adeguata e con formattazione più leggibile. Nella scheda del fondo quindi, oltre agli elementi informativi di carattere generale sul complesso documentale saranno elencate in forma tabellare le serie e sottoserie che lo compongono, contraddistinte da un codice che permetterà l'accesso immediato alle schede di sottoserie e ai record di dettaglio.

Le schede di sottoserie, già in buona parte redatte, contengono informazioni specifiche sulle caratteristiche comuni degli aggregati, e in determinati casi (documentazione ripetitiva ed omogenea) possono essere sostitutive delle schede di dettaglio sulle singole unità archivistiche.

² L'utilizzo del termine sottoserie è dovuto al fatto tradizionalmente con il termine serie in Banca d'Italia si fa riferimento ad aggregati tipologici (pratiche, copialettere e registri).

Il passaggio al disco ottico

Il quadro dell'evoluzione informatica dell'archivio storico è stato completato nel 1996 con l'adozione della tecnologia digitale per la riproduzione dei documenti, in sostituzione della microfilmatura.

Il progetto è stato suddiviso in tre fasi, i cui punti principali sono i seguenti:

- acquisizione del sistema di base
- integrazione di uno scanner per l'acquisizione delle veline dei copialettere, che presentano rilevanti problemi di trattamento a causa del supporto estremamente fragile e delle qualità della scrittura che è spesso sbiadita
- acquisizione di uno o più juke box di dischi ottici per tenere in linea le immagini e di uno scanner di tipo planetario per i grandi formati.

La prima fase si è conclusa nel 1996; la seconda è in questi giorni in fase di collaudo; la terza sarà presumibilmente avviata entro il 2000. La suddivisione del progetto in step successivi, oltre a facilitare l'assimilazione delle nuove tecnologie nel flusso di lavoro si è rivelata molto utile perché ha permesso di attendere il consolidarsi di nuove risorse HW e SW in una fase in cui il campo della gestione documentale elettronica è oggetto di continue innovazioni.

Con il progetto disco ottico si è costituito un sistema pienamente integrato di ricerca e consultazione. Dal lato della produzione le due fasi di inventariazione elettronica e di riproduzione delle carte, pur restando ovviamente distinte, sono ricomprese in un processo gestionale governato dalla stessa struttura informatica. Specularmente, dal lato della ricerca, sono state unificate le fasi di consultazione degli strumenti di corredo e di accesso ai documenti, assoggettate anch'esse alla medesima procedura.

Con l'occasione si è abbandonato il vecchio sistema BASIS/plus e i dati sono stati migrati in un database gestito dal pacchetto ARCHEA, che si avvale dell'information retrieval FULCRUM. Si è così fatto un significativo passo in avanti verso la semplificazione dell'accesso alle informazioni. Il prodotto BASIS, nella versione adottata, era infatti poco fruibile da parte di terzi, perché necessitava di interrogazioni in linguaggio nativo e si presentava con l'interfaccia a caratteri tipico dei software pre-Windows.

Uno dei capisaldi del progetto disco ottico è stato invece proprio l'obiettivo di offrire la massima indipendenza nello svolgimento delle ricerche.

L'utente riceve una password per l'accesso al database, ovviamente con privilegi limitati alla consultazione, ed ha a disposizione un personal computer collegato alla rete locale e dotato di uno schermo da 20 pollici ad alta definizione. Tramite apposite maschere di ricerca, più o meno complesse, lo studioso è libero di navigare all'interno del sistema seguendo i propri percorsi di ricerca. Individuati dai dati originali o dalle descrizioni aggiuntive i record che gli interessano può accedere direttamente alle immagini dei relativi documenti, che sono memorizzate su CD, cliccando sulla lista delle immagini presente nel record stesso. In una prossima fase i CD prodotti, che ora sono consultabili in locale e vanno quindi inseriti manualmente nel lettore, saranno tenuti in linea avvalendosi di uno o più juke box. Si sarà allora realizzato integralmente il progetto: sarà come aver consegnato virtualmente le chiavi dei depositi agli utenti lasciandoli liberi di prelevare dagli scaffali ciò che desiderano. Naturalmente la consultazione da parte dei terzi avviene in piena sicurezza per quanto riguarda l'eventuale accesso a documenti riservati, dal momento che per ogni record è determinato un livello di consultabilità che va da un grado "zero", in cui l'intera unità archivistica è consultabile a un grado "uno" in cui le immagini sono totalmente inibite, passando per livelli intermedi che prevedono la "scrematura elettronica", ovvero la sottrazione alla consultazione di singoli documenti contenuti in un fascicolo.

La richiesta di copie cartacee dei documenti è stata anch'essa automatizzata, realizzando una notevole economia di risorse rispetto alla precedente procedura di stampa da microfilm. L'utente, durante la consultazione, accantona su un proprio database le immagini dei documenti che desidera siano riprodotti, e il personale dell'archivio provvede in un'unica soluzione alla stampa, che avviene con una procedura batch. Su ogni pagina riprodotta appaiono stampati i riferimenti codificati del fondo, della serie, del numero di corda e di carta del documento stesso. E' anche possibile la distribuzione delle riproduzioni di documenti direttamente in forma digitale su un CD, con il quale è anche fornito un applicativo di visualizzazione.

Inoltre il sistema provvede all'aggiornamento automatico di statistiche di sala studio, tenendo memoria, per ogni utente, delle unità archivistiche consultate e delle copie richieste. Le statistiche automatiche vengono integrate manualmente nei casi, ancora frequenti di consultazione su altri supporti (carta, microfilm).

Alcuni dati quantitativi e metodologici

L'obiettivo finale del progetto è quello di trasferire in formato digitale tutta la documentazione dell'archivio storico, riprendendo dagli originali i documenti non ancora microfilmati e convertendo le bobine di microfilm, per un totale stimato di circa trenta milioni di immagini, che saranno racchiuse tre serie di circa tremila CD ciascuna³. Delle tre serie una viene usata in sala di studio, una è a disposizione dell'ufficio e una costituisce una copia di riserva.

Attualmente sono state digitalizzate circa 2.500.000 immagini cominciando dai fondi maggiormente consultati. Nel '98 il 50% delle consultazioni si è avvalso del nuovo sistema.

L'architettura, del tipo Client server, è basata su personal computer collegati in rete locale. Come è noto il sistema client/server consente di sfruttare appieno le potenzialità di memoria di ogni elemento del sistema, aumentando la velocità complessiva e riducendo il fabbisogno dell'unità centrale, che nel nostro caso è un semplice personal computer con una quantità di RAM e di spazio disco maggiorate. Il server, che ha sinora usato il sistema operativo Unix, sta proprio in questi giorni per essere aggiornato con la migrazione a Windows NT. Sono attive, oltre al server, diverse postazioni dedicate :

- alla consultazione (3 riservate gli esterni)
- alla conversione microfilm
- alla scansione del cartaceo
- al controllo di qualità e la riproduzione delle immagini
- alla masterizzazione dei cd.

Nel server risiedono il database definitivo (che contiene circa 120.000 schede di inventario), il database di lavoro, dove sono archiviati i record e le immagini prodotti durante la scansione, i database provvisori degli studiosi, dove possono essere accantonati schede e immagini per la riproduzione, ed i database di controllo degli accessi e di elaborazione delle statistiche.

³ La cifra è stimata sull'utilizzo attuale del CD ; se si dovesse passare ad altri supporti, come il DVD la quantità di dischi sarebbe notevolmente minore. Il supporto fisico di memorizzazione non rappresenta comunque un problema rilevante in quanto la parte più impegnativa è quella della digitalizzazione delle carte che impone una elevata manualità, mentre i riversamenti da supporto a supporto sono del tutto automatici. Il riversamento su nuovi supporti è un'operazione da mettere comunque in conto, stante la durata non ancora ben definita delle unità di memorizzazione.

Gli scanner impiegati uniscono ad un'alta velocità di acquisizione una particolare flessibilità nel trattare i diversi formati delle carte storiche e la garanzia di evitare danni agli originali. Come accennato sopra, in un prossimo futuro sarà anche acquisito uno scanner per formati superiori all'A3, da utilizzare in prevalenza per la ripresa dei numerosi registri di grande formato.

I microfilm vengono convertiti da due macchine in grado di gestire i formati a 16 e a 32 mm. I convertitori, agendo su una serie di parametri, sono in grado di raggiungere un buon grado di qualità, per quanto le operazioni di taratura risultino piuttosto complesse.

Tutte le immagini, sia quelle convertite da microfilm, sia quelle scandite dal cartaceo, vengono riprese in bianco e nero a 200 Dpi e danno luogo a file Tiff, con compressione del gruppo 4. Questi parametri sono adeguati al tipo di documentazione trattata (archivi in massima parte novecenteschi e dattiloscritti) dal momento che i file prodotti comportano un'occupazione media di circa 30 Kb per la corrispondenza ripresa dal cartaceo e di un massimo di 200 Kb per le immagini derivanti da registri (convertiti da microfilm e ripresi a doppia pagina).

Le immagini provenienti dagli scanner e dai convertitori di microfilm sono memorizzate provvisoriamente sui dischi del server, collegate a record del cosiddetto database di lavoro che contengono i dati identificativi dell'unità logica (Fondo, serie, numeri di corda, fascicolo e sottofascicolo).

Segue una fase di controllo di qualità, effettuata a campione, e facilitata da un applicativo che per ogni record segnala le immagini la cui occupazione in Kb eccede o non raggiunge alcuni limiti prefissati. In questa fase l'operatore ha la possibilità di rimandare le singole immagini scartate alla riscansione.

L'ultimo passo è rappresentato dalla produzione del CD, in triplice esemplare, e al collegamento delle immagini alle relative schede contenute nel database definitivo di ricerca. Sul CD viene anche memorizzato un file di testo con dati identificativi delle immagini contenute.

Sono possibili ulteriori interventi nel caso in cui durante la consultazione si riscontrino immagini poco leggibili che siano sfuggite ai controlli: un apposito applicativo provvede alla rilettura del CD e al loro riversamento nell'ambiente di lavoro per la sostituzione delle immagini deteriorate. Si procede quindi ad una nuova masterizzazione.

Conclusioni

La Banca d'Italia ha ormai acquisito un decennio di esperienza nel trattamento informatico del proprio archivio storico. Dieci anni sono nel campo dell'informatica un periodo sufficientemente lungo per poter trarre qualche indicazione da questa esperienza.

- Tra le principali fonti di incertezza nell'utilizzo delle tecnologie informatiche sono senz'altro da annoverare le migrazioni imposte dall'evoluzione dei sistemi software: su questo punto l'esperienza di tre riversamenti di dati (BASIS K, BASIS plus, ARCHEA) non ha comportato operazioni particolarmente complesse, né tantomeno danni alla base dati; anzi, come si è cercato di dimostrare sopra, ha portato notevoli vantaggi.
- Il servizio offerto all'utenza ha sicuramente beneficiato in misura significativa dall'aiuto dell'informatica, che come si è detto è ormai lo strumento di corredo principale dell'ASBI, anche se il supporto dell'esperienza dell'archivista che conosce l'archivio e la struttura aziendale è senz'altro insostituibile. Il ruolo dell'archivista rimane poi fondamentale nella fase di progettazione dei nuovi sistemi, in quanto solo con un dialogo stretto tra informatici e archivisti si può dar vita ad architetture valide.
- L'informatizzazione degli archivi storici è premessa indispensabile per essere attrezzati ad accogliere il flusso di documentazione proveniente dagli archivi di deposito, che sarà tra breve composto in buona parte da file.