

# La politica di digitalizzazione della Bayerische Staatsbibliothek

**Klaus Kempf**

*Bayerische Staatsbibliothek*

*La Bayerische Staatsbibliothek (BSB), una tra le più importanti biblioteche di ricerca in Europa, con le sue collezioni straordinarie note in tutto il mondo, è una istituzione pioniera nel campo della digitalizzazione, almeno nei paesi di lingua tedesca.*

*Alla fine degli anni Novanta del secolo scorso la BSB ha iniziato una sistematica politica di digitalizzazione, basata su tre diversi tipi di attività: la digitalizzazione project-oriented, la digitalizzazione rivolta all'utenza (digitalizzazione on demand) e la digitalizzazione come moderna politica di conservazione. Con l'affermarsi delle copie digitali il microfilm, inteso come mezzo principale per la conservazione, è stato abbandonato e la copia digitale è diventata la prima forma derivata degli originali. Il tema centrale riguarda i progetti di digitalizzazione finanziati dalla Regione bavarese ma anche da altre fonti, in particolare dalla Deutsche Forschungsgemeinschaft (DFG), l'Associazione tedesca per la ricerca, che da anni supporta speciali programmi per attività di digitalizzazione nelle biblioteche, negli archivi e ora anche nei musei. Ma la maggioranza di quasi un milione di volumi digitalizzati è il risultato della partnership pubblico-privato con Google. La BSB è stata la prima biblioteca continentale europea ad aderire all'iniziativa di digitalizzazione lanciata da Google. Sin dall'inizio, i flussi di lavoro delle attività di digitalizzazione sono stati automatizzati. Il software ZEND, sviluppato dalla stessa Biblioteca e basato sulla filosofia "open source" serve da piattaforma di produzione. Si tratta di un software di tipo modulare e si presta ad un utilizzo molto flessibile; copre tutti i passaggi del workflow, a partire dall'ordine inviato dal singolo utente su internet, fino all'invio automatico della copia originale dal server di produzione al sistema di archiviazione, gestito dal Leibniz Rechenzentrum (LRZ).*

*Anche i flussi di lavoro del progetto Google si basano su ZEND. Nel Centro di scansione della BSB è disponibile ogni sorta di attrezzatura per la scansione, apparecchi fotografici digitali, e anche scanner robotizzati. In tal modo la Biblioteca è in grado di produrre una digitalizzazione massiva (più di mezzo milione di pagine), come anche la digitalizzazione di oggetti singoli (digitalizzazione "di lusso"). L'obiettivo è sempre lo stesso: la creazione di oggetti digitali attrattivi e/o grandi quantità di copie digitali del medesimo argomento che incontreranno l'immediato interesse dell'utente. Nel caso si verifichi una situazione di aumento della produzione, ZEND consente anche la produzione digitale, in special modo la scansione ad opera di terzi. Il principale ostacolo a una più ampia partecipazione delle biblioteche tedesche a progetti*

*di digitalizzazione, e questo vale anche per la BSB, è l'assenza di una politica di digitalizzazione coordinata e finanziata a livello nazionale. La mancanza di un sufficiente supporto finanziario costituisce al momento l'ostacolo maggiore per un coerente progetto di digitalizzazione a livello europeo, e ciò potrebbe costituire la ragione principale del fallimento dell'iniziativa di Europea, che ha avuto un così buon inizio.*

## Introduzione: gli esordi

La Biblioteca nazionale di Baviera (Bayerische Staatsbibliothek - BSB) ha svolto il ruolo di "pioniere" nella digitalizzazione, non solo in Germania, ma anche in Europa. Dal 1997 è responsabile del cosiddetto Centro di digitalizzazione monacense (Münchner Digitalisierungszentrum - MDZ), istituito inizialmente come uno dei due centri nazionali con competenze nel campo della digitalizzazione<sup>1</sup>. Il centro rispondeva a una strategia della Deutsche Forschungsgemeinschaft (DFG)<sup>2</sup> di coordinare fin dall'inizio i progetti di (retro)digitalizzazione a livello nazionale e di orientare verso la scelta di definire e sviluppare standard comuni. Il MDZ ha avuto fino ad oggi un grande e rapido sviluppo ed ha portato a termine con successo più di 230 progetti, molti dei quali in cooperazione con partner, sia ricercatori, sia istituzioni di ricerca impegnate a digitalizzare diversi materiali<sup>3</sup>.

Attualmente sono innumerevoli i progetti di digitalizzazione in Germania. Quasi in ogni "istituto della memoria", quali gli archivi, i musei, le biblioteche, sono entrati in funzione scanner e fotocamere digitali, per trattare diversi contenuti, renderli leggibili dai computer e accessibili in Internet. Tra i centri specializzati il MDZ è di gran lunga il più importante e la BSB ha acquisito grande competenza su tutti gli argomenti che riguardano la digitalizzazione e le tematiche connesse, tanto da essere considerata il punto di riferimento primario in Baviera e in Germania, e in certi casi anche in Europa, per i bibliotecari che hanno bisogno di consigli in materia<sup>4</sup>. In quest'articolo verrà illustrata la politica di digitalizzazione fino ad ora seguita dalla BSB e come poi questa si sia intrecciata con le iniziative di digitalizzazione a livello nazionale, europeo e internazionale; verranno inoltre approfondite dettagliatamente le questioni relative all'organizzazione del lavoro e alle scelte tecnologiche adottate.

<sup>1</sup> Il secondo è stato installato presso la Niedersächsische Staats- und Universitätsbibliothek di Göttingen.

<sup>2</sup> La DFG è un organismo parastatale, cioè del governo tedesco centrale con sede a Bonn, che promuove e sostiene finanziariamente la ricerca, sia universitaria che di altri enti, Max Planck, Helmholtz, ma anche le biblioteche. Vedi in dettaglio: <<http://www.dfg.de/foerderung/programme/infrastruktur/lis/index.html>>.

<sup>3</sup> Cfr. la lista con l'informazione dettagliata sui vari progetti sul sito web del MDZ: <<http://www.muenchener-digitalisierungszentrum.de/index.html?c=kurzauswahl&l=de&adr=regesta-imperii.digitale-sammlungen.de>>.

<sup>4</sup> Numerosi i contatti con il Paul Getty Museum, San Francisco, ma anche con le biblioteche nazionali di Australia, Nuova Zelanda, Singapore, Taipei e con le grandi biblioteche di molti altri paesi asiatici.

## La BSB: la biblioteca tedesca con il più ampio spettro di funzioni

La BSB è la biblioteca di deposito e d'archivio dello Stato federale di Baviera, e può essere considerata come la "Biblioteca nazionale bavarese". Ha le caratteristiche di una biblioteca di ricerca di rango internazionale e conserva raccolte librerie di grandissimo rilievo. Possiede un eccezionale patrimonio di manoscritti medievali (circa 94.000, il quarto al mondo), di incunaboli (circa 20.000, numero pari a quello della British Library) e di edizioni del Cinquecento (la più grande raccolta di lingua tedesca con circa 140.000 titoli). Questo fondo storico, che ammonta complessivamente a circa 1,2 milioni di titoli, non soggetti al diritto d'autore, al cui interno sono numerose collezioni speciali uniche al mondo, è stato il principale motore per l'avvio di progetti di digitalizzazione ad ampio raggio della BSB.

La BSB non è solo una grande biblioteca di ricerca, ma assolve anche ad altre numerose funzioni a livello regionale e nazionale: da una parte è la capofila delle istituzioni presenti nel Consorzio bavarese per il rilascio delle licenze dei media elettronici commerciali, quali gli e-books, le banche dati e le riviste elettroniche, dall'altra presiede il Servizio delle biblioteche bavaresi (Bibliotheksverbund Bayern - BVB), una delle più grandi reti bibliotecarie regionali di lingua tedesca. È responsabile anche di una *library school*, l'Accademia bavarese di biblioteconomia (Bayerische Akademie für Bibliothekswesen - BAB) che costituisce l'organizzazione centrale d'istruzione e formazione per l'intera Baviera, aperta anche a studenti di altri stati tedeschi, a seguito di specifici accordi, e di un istituto molto rinomato per la patologia del libro (Institut für Buchrestaurierung - IBR). Infine, *last but not least*, ha una sezione dedicata al sostegno delle biblioteche pubbliche nelle località piccole e medie della Baviera, la cosiddetta Landesfachstelle für das öffentliche Bibliothekswesen con affiliazioni in altre tre città bavaresi (Würzburg, Nürnberg, Regensburg).

È fondamentale ricordare che in Germania, a differenza della Francia, della Gran Bretagna, ma anche degli Stati Uniti e della Russia, la funzione di biblioteca nazionale tradizionalmente viene svolta da tre grandi biblioteche: la biblioteca nazionale tedesca (Deutsche Nationalbibliothek - DNB) a Francoforte sul Meno e Lipsia, che costituisce il centro bibliografico nazionale e raccoglie e conserva le pubblicazioni in lingua tedesca (pubblicate in Germania e all'estero) e le due grandi biblioteche statali storiche di Monaco di Baviera e di Berlino (Staatsbibliothek zu Berlin).

Tra le funzioni di livello nazionale svolte dalla BSB c'è la gestione degli *authority files* (insieme alla DNB), e quella delle collezioni speciali. La BSB è infatti grande parte del cosiddetto Sondersammelgebietsplan (SSG-Plan), il piano nazionale delle collezioni speciali, al quale partecipano attualmente più di 70 biblioteche con circa 130 collezioni speciali, che riguardano specificamente la letteratura straniera e in lingua straniera nei vari campi di ricerca. Il piano SSG deve poter rendere disponibile in Germania perlomeno un esemplare di ogni opera scientifica rilevante,

indipendentemente dal paese d'origine e dalla lingua di pubblicazione, in modo tale da raccogliere anche le pubblicazioni edite nelle zone più periferiche del mondo in lingue poco conosciute. Nell'ambito del progetto, la BSB assume un ruolo di primo piano e cura otto collezioni del SSG.

La BSB svolge anche altre funzioni biblioteconomiche di livello nazionale: ha un ruolo preminente nella concessione di licenze dei *born-digital media* da parte delle case editrici; offre inoltre, grazie al suo Centro per la pubblicazione elettronica (ZEP), esistente dal 2008, anche la piattaforma tecnico-organizzativa per coadiuvare gli sforzi intrapresi, in particolare dai ricercatori e studiosi di scienze umane, e /o da organizzazioni di ricerca indirizzati a garantire l'accesso all'e-Publishing ai vari Istituti della storia tedesca presenti all'estero – Deutsche Historische Auslandsinstitute (DHI).

### **Le basi della politica di digitalizzazione della BSB**

Fin dall'inizio la politica di digitalizzazione della BSB ha potuto contare su solide basi per quanto riguarda i finanziamenti: la titolarità amministrativa della BSB è della Regione bavarese, non del governo centrale di Berlino. La Baviera provvede al finanziamento principale, ma la BSB riceve, rispetto ad altre organizzazioni di cultura bavaresi ed anche ad altre biblioteche tedesche, notevoli finanziamenti da diverse fonti. Si tratta soprattutto di finanziamenti di progetti che vengono in maggior parte erogati – su domanda – dalla Deutsche Forschungsgemeinschaft (DFG).

Per la digitalizzazione inoltre sono previsti ulteriori finanziamenti speciali della Regione di Baviera con i quali la BSB può realizzare progetti specifici di digitalizzazione di determinate parti delle raccolte librerie, o può in cooperazione con altri enti scientifici fornire il suo *know how* in materia e/o procedere alla digitalizzazione dei fondi di un altro ente considerato come partner tecnico, sulla base della presentazione, per entrambi i partner, di un piano finanziario completo. Nel caso di un progetto con un partner che intende digitalizzare un proprio fondo la BSB conserva poi una copia digitale per sé, arricchendo così le sue raccolte. Esempi significativi di collaborazione si sono avuti con il progetto *Regesta Imperii*, realizzato con l'Accademia delle scienze di Magonza (Mainz) o anche la digitalizzazione di tutte le pubblicazioni fino ad ora apparse dei Monumenta Germaniae Historica (MGH) la cosiddetta dMGH<sup>5</sup>.

Analogamente, quando un utente richiede la digitalizzazione a proprie spese, parziale o completa, di un singolo testo presente nel patrimonio bibliografico della BSB, la Biblioteca, nel soddisfare la richiesta, realizza comunque una copia intera del testo che va ad arricchire il suo patrimonio di copie digitali<sup>6</sup>.

<sup>5</sup> Cfr. i rispettivi link, per esempio della dMGH: <<http://www.mgh.de/dmgh/>>.

<sup>6</sup> Il servizio di digitalizzazione *on demand* è ormai completamente automatizzato: grazie al software tool *Erato*, l'utente può inserire autonomamente su Internet le sue richieste di digitalizzazione e

Un ulteriore motivo per digitalizzare il patrimonio bibliografico è quello della conservazione: nel 2002 la BSB ha modificato la sua strategia di conservazione, abbandonando la scelta del microfilm come supporto prioritario di conservazione, a favore dei testi digitalizzati a questo precipuo scopo. Di questi ultimi può essere eventualmente prodotta per motivi di sicurezza un'ulteriore copia su microfilm. L'abbandono del microfilm come forma privilegiata di conservazione secondaria fu all'epoca percepito tra le biblioteche tedesche come un fatto rivoluzionario, e in parte criticato. Oggi questo procedimento ha trovato numerosi imitatori, e il livello di sicurezza, attribuito finora al microfilm per la sua resistenza all'invecchiamento, è ormai sotto vari aspetti relativizzato. Anche l'utenza mostra di apprezzare i vantaggi di lavorare con una copia digitale invece che con un microfilm. Infine, non appare possibile altra soluzione per le pubblicazioni *born digital* che quella di tentare un'archiviazione digitale.

### Digitalizzazione con Google

Un discorso a parte merita la partecipazione della BSB al progetto *Google books*<sup>7</sup>. Dopo la Bodleian Library di Oxford, la BSB è stata la seconda grande biblioteca in Europa a partecipare a questo progetto mondiale di digitalizzazione delle pubblicazioni non soggette a copyright. La partecipazione ha avuto inizio con un accordo di cooperazione pubblico-privato conclusosi nel 2007. È importante sottolineare che è stata la società Google a rivolgersi alla Biblioteca, e non il contrario. Nell'accordo *non-disclosure* la BSB ha potuto far valere il proprio punto di vista sul progetto, con soddisfazione di entrambe le parti<sup>8</sup>.

La BSB ha deciso autonomamente quali testi digitalizzare con Google. Alla scadenza del progetto si prevede che siano a disposizione, liberamente accessibili su Internet, circa un milione di volumi non soggetti a copyright provenienti dai fondi della BSB. Tutti i fondi speciali, in linea di massima quelli con anno di pubblicazione precedente il 1600, sono stati esclusi dal progetto<sup>9</sup>. In particolare, l'esclusione riguarda i manoscritti medievali, gli incunaboli e la maggior parte delle cinquecentine. Tra i fondi speciali sono comprese anche altre tipologie di materiali, come

sapere immediatamente quanto costerà il servizio desiderato. Questo servizio, che viene offerto anche in lingua inglese, è molto apprezzato e fa risparmiare alla Biblioteca spazio e tempo. Vedi in dettaglio: Thomas Busch et al., *ERaTo: vom Papierformular zum multifunktionalen Bestelltool*, «ABI-Technik», 29 (2009), n. 4, p. 227-233.

<sup>7</sup> Per una sintesi del progetto vedi: <[http://de.wikipedia.org/wiki/Google\\_Books](http://de.wikipedia.org/wiki/Google_Books)>; cfr., anche, più ampiamente e in particolare sul significato del progetto per il mondo bibliotecario, Edgar Jones, *Google books as a General Research Collection*, «Library Resources & Technical Services», 52 (2010), n. 2, p. 77-89.

<sup>8</sup> Nel frattempo hanno aderito al progetto altre biblioteche, come quelle catalane, le biblioteche del Belgio, della Francia (non la BNF, ma la grande biblioteca civica di Lione) e della Svizzera, ed anche le biblioteche nazionali dell'Austria (ÖNB) e della Repubblica Ceca, oltre che le due biblioteche nazionali centrali italiane di Firenze e di Roma.

<sup>9</sup> Sono state date a Google solo alcune cinquecentine preselezionate.

carte geografiche, partiture ovvero spartiti (prima del 1600) o anche foto (fino al limite temporale imposto dalla normativa sul copyright per l'accesso in internet). Sono stati inoltre esclusi i fondi con problemi di conservazione.

Dall'altra parte, entrambi i partner devono poter disporre di volta in volta, e nel tempo, di una propria copia digitale che possa essere integrata negli strumenti di ricerca, cioè nell'OPAC della BSB, ma anche nei diversi servizi web, come i portali d'informazioni e per materia o soggetto, le cosiddette *Virtuelle Fachbibliotheken* (ViFa), e deve poter essere archiviata autonomamente. La cessione delle copie digitali a terzi, per fini commerciali, è vietata per ambo le parti; è invece permesso a tutti i ricercatori nel mondo di utilizzarle gratuitamente<sup>10</sup>.

La concreta cooperazione con Google è iniziata nell'anno 2008/2009, dopo una lunga fase preparatoria. Da allora, ogni settimana vanno in scansione fino a 5.000 opere, e più o meno la stessa quantità torna indietro; ogni singolo volume che viene dato a Google e poi rientra in Biblioteca viene registrato in una banca dati di *workflow* (WDB) appositamente creata per questo scopo. In questo modo, da una parte è possibile in qualunque momento un controllo complessivo e dettagliato dei fondi non disponibili al momento per gli utenti, dall'altra vengono registrate nella banca dati WDB tutte le informazioni sulle singole opere che sono state digitalizzate o che per vari motivi non sono state ancora digitalizzate da Google. In tutto vengono spostati ogni settimana circa 20.000 volumi prelevati dagli scaffali e poi di nuovo collocati e/o classificati. Questo impegnativo lavoro logistico si svolge in maniera quasi invisibile per l'utente. Poiché il principio ultimo era, ed è, che il livello di qualità del servizio della Biblioteca, cioè - in particolare - la fornitura rapida e senza intoppi delle opere richieste dall'utente, non deve risentire in nessun caso del progetto di digitalizzazione.

Un gradito effetto collaterale del progetto di digitalizzazione con Google non può non essere ricordato: attraverso la digitalizzazione più o meno estesa di tutte le pubblicazioni non soggette a copyright, edite tra il 1600<sup>11</sup> e il 1875, il progetto ha anche consentito una grande revisione generale dei fondi, sempre auspicata, fin dalla fine della guerra, ma mai realizzata.

<sup>10</sup> La BSB riceve da Google la cosiddetta *library digital copy*, della stessa qualità della copia digitale che Google stesso si assicura a conclusione del processo di digitalizzazione. Il grosso vantaggio di essere partner di Google consiste nel fatto che fin quando durerà il progetto la BSB sarà partecipe dei miglioramenti che Google apporta via via alle copie digitali, grazie ad un lavoro permanente sulla qualità. Così non solo Google, ma anche la BSB, potranno disporre di una (frequentemente migliorata) *library digital copy* dei sui libri, ed anche di una versione full text trattata da Google con OCR di quasi ogni volume della Biblioteca digitalizzato tramite l'accordo. Nel frattempo, con il suo software OCR, Google è persino riuscito a trovare un'accettabile soluzione per la decifrazione dei caratteri *fraktur*, cioè dei caratteri utilizzati normalmente in Germania fino alla fine del XIX secolo, che finora hanno rappresentato un ostacolo quasi insuperabile per ogni tipo di software OCR.

<sup>11</sup> Come si è detto, nel frattempo la BSB ha dato a Google anche qualche cinquecentina, allo scopo di accelerare la digitalizzazione di questo fondo così importante e tanto richiesto dalla comunità dei ricercatori.

Alcune opere credute perdute sono state ritrovate, così come in qualche caso sono state individuate, catalogate e digitalizzate anche opere non ancora registrate nei cataloghi.

Ormai sono stati scannerizzati quasi un milione di volumi non soggetti a *copyright*, le cui copie digitali sono offerte sia da Google, sia dalla BSB, 24 ore su 24, per la ricerca e il *download* in modalità *open access*. Fino ad ora la collaborazione con Google si è svolta, a parere della BSB, senza alcun attrito. I timori riguardo al danneggiamento degli originali si sono rivelati assolutamente infondati. Nel corso della cooperazione è migliorata considerevolmente la qualità – all’inizio problematica – di alcune opere digitalizzate (per esempio per pagine mancanti o annerite). Poiché Google lavora continuamente al miglioramento dei prodotti, per le opere fino al 1700 è ora in grado di fornire la digitalizzazione a colori, e per tutte anche una versione completa trattata con OCR. La BSB e i suoi utenti traggono grandi vantaggi dalla collaborazione con Google.

I risultati parlano da soli: gli accessi e il *download* delle opere digitalizzate ormai sono ogni giorno in media più di 2000. Allo stesso tempo, gli originali vengono richiesti sensibilmente meno, sia in prestito che in consultazione, a tutto vantaggio di una migliore protezione delle raccolte<sup>12</sup>. L’utente trova su Internet l’opera di suo interesse e ne è contento. Nei pochi casi in cui la qualità dell’opera digitalizzata non è risultata soddisfacente e/o manchino parti (quali ad esempio le carte allegare, i cosiddetti *fold-outs*, che non vengono scansionati da Google), le opere in questione vengono digitalizzate ancora una volta (per intero) dalla BSB stessa nel suo Centro di scansione e poi offerte su Internet in questa forma completa<sup>13</sup>.

Anche un altro importante risultato merita di essere ricordato: grazie alle opere digitalizzate si è ormai costituita, sotto diversi aspetti e per svariati temi, una “massa critica” di documenti tematici pertinenti. Questa enorme quantità di dati digitali tra di loro comparabili, fa parte del fenomeno dei cosiddetti *big data*, cioè della produzione, ma anche dell’uso e riuso, di grosse quantità di dati utilizzabili per la ricerca nei campi più svariati, come fonti primarie che consentono l’elaborazione di

<sup>12</sup> In un’indagine statistica sullo sviluppo dei prestiti locali e interbibliotecari dei fondi storici della BSB, il Dipartimento servizio all’utenza nel 2012 ha rilevato che il numero dei prestiti nell’arco di tempo compreso tra il 2008 e il 2011 per quanto riguarda – in un primo caso – le opere pubblicate tra 1501 e l’anno 1700 e – in un secondo caso – tra il 1701 fino al 1840, si è ridotto considerevolmente. Nel primo caso la riduzione ammonta al 32,4 % e nel secondo caso arriva persino all’incredibile percentuale del 64,3 %!

<sup>13</sup> In realtà, siccome la digitalizzazione delle opere contenenti *fold-outs* è tanto faticosa (e costosa), la redigitalizzazione sistematica da parte della BSB è stata fatta solo in caso di finanziamento certo, come è avvenuto nell’ambito del progetto VD18, di digitalizzazione delle opere dell’Ottocento a livello nazionale, grazie al quale sono state digitalizzate le mappe su tavole ripiegate contenute in 900 opere (per un totale di 4.586 immagini). Come già detto, il *workflow* in questi casi è particolarmente faticoso: si calcola in media due ore, per cui è possibile svolgere questo intervento solo nei casi eccezionali e/o quando la biblioteca dispone dei mezzi finanziari mirati a questo scopo.

nuove linee di ricerca, o meglio di metodologie di lavoro, fino a poco tempo fa inimmaginabili<sup>14</sup>.

### L'organizzazione della digitalizzazione: strutture e processi di lavoro

Prima dell'avvio della digitalizzazione di massa<sup>15</sup> i responsabili della BSB erano già convinti che la nuova attività si sarebbe dovuta fondare su un sistema automatizzato. L'obiettivo era quello di costruire un workflow, dalla richiesta dell'opera da digitalizzare fino all'archiviazione della copia digitale. Lo ZEND (Zentrale Erfassungs - und Nachweisdatenbank), cioè la banca dati centrale per la registrazione e il controllo, uno strumento sviluppato in più moduli e basato su software open-source, si è assunto l'incarico di fungere da "piattaforma di produzione". Lo ZEND è sviluppato in modo modulare. La distribuzione di workflow eseguita con esso comprende in ogni modulo anche l'aspetto del controllo di qualità. Allo ZEND può essere connesso principalmente anche un illimitato numero di gestori esterni, che per esempio prendono in consegna (dopo apposita autorizzazione ed autenticazione) il completamento di sottoprocessi, come la scansione e/o – in certi casi anche contemporaneamente – l'elaborazione OCR. In questo modo al Centro di scansione è possibile realizzare una gestione delle capacità di produzione relativamente flessibile. Nel caso in cui le capacità produttive dell'Istituto si rivelino momentaneamente insufficienti, è possibile attingere rapidamente alle capacità di aziende specializzate sul mercato<sup>16</sup>.

Tre ulteriori principi, che sin dall'inizio sono stati presi in considerazione, non possono non essere ricordati:

1. *Nessuna risorsa digitale senza record bibliografico e senza metadati bibliografici.* Questo principio è valido dai tempi del microfilm, quando si sono svilup-

<sup>14</sup> Sul fenomeno del "big data" nel campo delle scienze umanistiche e il possibile ruolo delle biblioteche nel futuro in questo campo vedi, con ulteriori citazioni bibliografiche, Klaus Kempf, *Der Sammlungsgedanke im digitalen Zeitalter. L'idea della collezione nell'età digitale*, Lectio magistralis tenuta all'Università degli studi di Firenze il 5 marzo 2013, Fiesole: Casalini, 2013, p. 125-131.

<sup>15</sup> Si parla di digitalizzazione di massa nella BSB quando in un progetto vengono scannerizzate più di 500.000 pagine. Tutte le altre iniziative, meno voluminose sono considerate "digitalizzazioni-boutique".

<sup>16</sup> Le altre biblioteche tedesche che si occupano in modo sistematico e con una certa ampiezza della digitalizzazione dei loro patrimoni bibliografici usano come "piattaforme di produzione", nel senso sopra indicato, o un prodotto *open-source* come lo ZEND, chiamato Goobi, che è il risultato (e *work in progress*) della cooperazione di un gruppo di biblioteche tedesche – ora riunito in una associazione che porta lo stesso nome – con le biblioteche-leader, Niedersächsische Staats-und Universitätsbibliothek di Göttingen e Sächsische Landesbibliothek Staats-und Universitätsbibliothek (SLUB) di Dresden (vedi più in dettaglio la relativa rivista online: <<http://www.goobi.org/en/>>), oppure ricorrono, come la Landes-und Universitätsbibliothek Sachsen-Anhalt di Halle, ad un prodotto commerciale, quello della ditta Semantics GmbH (vedi: <<http://digitale.bibliothek.uni-halle.de/>>) che offre funzionalità simili allo ZEND o al Goobi.



pati “milioni di fosse di testo”, poiché la catalogazione dei testi microfilmati non teneva il passo con la produzione dei microfilm. Sono nate così enormi quantità di dati che, essendo privi dei metadati contenutistici e di appartenenza, non risultavano né accertati – cosa che riguarda la singola opera – né rintracciabili. Nella BSB, quando nessun record bibliografico (completo) è disponibile *online*<sup>17</sup> viene registrato nel catalogo almeno un breve record bibliografico “intermedio” (con l’indicazione della segnatura), come succede per esempio nel caso dei manoscritti medievali. Sarà poi lo ZEND ad accedere a questo record e a riunirlo alla risorsa digitalizzata, garantendo l’inserimento del link che provvede all’accesso alla risorsa digitalizzata direttamente a partire dal record presente sull’OPAC locale della BSB. Si è deciso infatti di offrire all’utente un record unico, riutilizzando il record esistente per l’opera stampata anche per la registrazione della sua versione digitalizzata. L’obiettivo è evidente: l’utente deve vedere a colpo d’occhio se di un’opera esiste anche una versione digitale e deve essere indotto a ricorrere prioritariamente alla versione digitale per ottenere un servizio immediato, in modalità “self-service”, ma ancora di più per far risparmiare costi e risorse (umane) alla biblioteca, per esempio riducendo i prestiti.

2. *Le risorse digitali devono essere prodotte con un alto standard qualitativo.* Vale il principio: si digitalizza una volta sola (possibilmente con una risoluzione di 300 – 600 dpi) e con quella copia digitale si conseguono gli scopi e gli obiettivi più disparati<sup>18</sup>. Naturalmente la velocità, e indirettamente anche i costi, della scansione vengono determinati in modo decisivo dalla qualità e dalla risoluzione scelta. Quanto più alta è la risoluzione, tanto più lenta (e costosa) è la scansione.

<sup>17</sup> La conversione della maggior parte dei vari cataloghi (alfabetici) cartacei nella BSB è stata realizzata dagli anni Ottanta fino al 2006 in tre grandi fasi. Successivamente, dal 2009 fino al 2012, sono stati “automatizzati” anche gli ultimi cataloghi alfabetici cartacei rimasti, quelli che contengono i record delle opere in lingua cinese, giapponese e coreano. Cfr. dettagliatamente: Werner Holbach, *Retrokonversion und Retrokatalogisierung in den Bayerischen Staatlichen Bibliotheken*, «BFB», 34 (2006), 3, p. 206-228; Tabery/Werner Holbach, *Die Retrokonversion der Ostasienkataloge der Bayerischen Staatsbibliothek*, «ZfBB», 59 (2012), 1, p. 12-19.

<sup>18</sup> Nel progetto della BSB *Digi20* che prevede la digitalizzazione di letteratura ancora protetta da copyright pubblicata da case editrici selezionate (piccole e medie, soprattutto nel campo della storia), anche le case editrici ricevono una copia digitale del “loro libro” per costruirsi man mano un archivio digitale. Questo progetto merita particolare attenzione anche per altre novità nel campo della digitalizzazione: il contratto di cooperazione stipulato tra la BSB e le singole case editrici prevede infatti la digitalizzazione e la messa a disposizione (da parte della Biblioteca) in modalità *open access* delle opere che hanno passato un *moving wall* di 3-5 anni (a seconda della decisione della casa editrice). Le case editrici ricevono, per la loro rinuncia all’usufrutto del copyright, un compenso *una tantum* da parte della DFG. In più, mantengono il privilegio di poter vendere le edizioni stampate parallelamente all’offerta gratuita degli stessi titoli su Internet in modalità *open access*. Si è potuto constatare che le copie stampate vengono comunque vendute, e le vendite risultano persino aumentate, grazie alla “pubblicità” dell’offerta digitale (*open access*).

D'altro lato, è anche chiaro che sottoporre a scansione un'opera, diverse volte e con risoluzioni diverse, non solo causa uno stress enorme al supporto, ma comporta anche un aggravio di costi, poiché il prelievo del libro dallo scaffale e la sua ricollocazione incidono in modo considerevole sulla manodopera. Per questo motivo è importante trovare un accettabile compromesso tra la qualità della scansione e la necessità di rispettare il piano dei costi predisposto per l'utilizzo dei fondi messi a disposizione da terzi.

3. *Ogni digitalizzazione deve servire per tenere sotto controllo lo stato di conservazione dell'opera.* L'IBR, infatti, non solo è presente con i propri collaboratori nel Centro di scansione per la digitalizzazione delle opere più preziose, al fine di effettuare all'occorrenza interventi preventivi di urgenza, ma rilascia anche una *expertise* (su richiesta) per tutti i lavori di scansione. Sotto questo punto di vista, la digitalizzazione può essere considerata parte integrante della politica di conservazione dei fondi della Biblioteca.

Com'è già detto, tutti i lavori di digitalizzazione nella BSB sono sottoposti ad un workflow uniforme. Si tratta di una combinazione di elementi centralizzati e decentrati. Centralizzato in una propria unità organizzativa, il cosiddetto Scanzentrum (Centro di scansione) che appartiene al Dipartimento acquisti, catalogazione e digitalizzazione della BSB, fa parte della grande spina dorsale di produzione (tradizionale e digitale) della Biblioteca; la scannerizzazione e la fotografia digitale vengono eseguite con tutti i necessari lavori di preparazione e rifinitura, dal controllo dello stato di conservazione (eventualmente insieme all'IBR), fino alla preparazione della fattura per il singolo utente che richiede una copia digitale. La messa a disposizione dei materiali da una parte, e la ricezione, ossia la risposta alle richieste dell'utenza, dall'altra - nel caso di fornitura di singole copie digitali - viene svolta di volta in volta attraverso l'opera di reparti addetti alla distribuzione del materiale o di dipartimenti specialistici. Questa forma organizzativa è stata mantenuta e si è rivelata più vantaggiosa rispetto alla teoria pura, che tiene separati gli approcci centralizzati da quelli decentrati. Verso l'interno è garantita un'organizzazione efficiente e uniformemente strutturata per lo svolgimento del lavoro; verso l'esterno, cioè di fronte all'utente, la biblioteca entra in contatto attraverso "un viso", cioè quello della persona addetta alla sezione (speciale), responsabile della soddisfazione dell'utente, che rende giustizia alla concezione di un servizio integrato o - come si dice in tedesco - di "servizio ad una mano" (oppure, in inglese, "one stop shopping")<sup>19</sup>.

<sup>19</sup> Su questa idea del servizio in biblioteca vedi più ampiamente Klaus Kempf, *Zugang zum Wissen. Die Bibliothek als Ort der Verfügbarkeit und Vermittlung von Information*, in: *Die Teßmann. Friedrich-Teßmann-Sammlung der Österreichischen Akademie der Wissenschaften (1957 – 2012). Landesbibliothek Dr. Friedrich Teßmann (1982-2012)*, Wien-Bozen, 2012, p. 129-141.

### Excursus: allestimento tecnico del Centro di scansione

Il Centro di scansione della BSB appartiene, come già detto, ai “pionieri” iniziatori della digitalizzazione non solo nelle biblioteche tedesche, ma d’Europa, e dispone degli strumenti più efficienti e moderni che si possono trovare nei centri di digitalizzazione di tutta Europa. Dal 1997 a oggi ha contribuito allo sviluppo e all’efficace utilizzazione di parecchie generazioni di scanner e altrettante di fotocamere (digitali). Fino ad oggi il Centro di scansione ha mantenuto la sua duplice caratteristica: da un lato funziona come laboratorio per l’innovazione e lo sviluppo tecnologico, dall’altro è il luogo della produzione digitale di routine. Al momento ha in dotazione 23 scanner e fotocamere digitali di diverso tipo e con diverse capacità di prestazione. Particolare attenzione meritano i quattro *ScanRobot* della ditta austriaca Treventus di Vienna, che lavorano in modo particolarmente efficiente e sono in grado di elaborare, per ogni progetto, la scansione di opere per un totale di 2.000 pagine l’ora<sup>20</sup>. Per i progetti particolarmente delicati vengono utilizzati i cosiddetti *Grazer Kameratische*, una particolare attrezzatura di scansione per opere ipersensibili e/o in uno stato di conservazione precario. Quest’attrezzatura speciale è stata sviluppata dai colleghi della biblioteca universitaria di Graz, in Austria, in collaborazione con una ditta privata ed è usata in Biblioteca per la scansione di manoscritti e stampe preziose. La sua capacità di risoluzione arriva a 600 dpi. Per le “taglie forti” e insolite, quindi soprattutto per carte geografiche, opere grafiche o anche riviste, viene utilizzato lo scanner del formato AO con una risoluzione fino a 400 pixel.

Tutti gli scanner menzionati possono riprodurre in modalità “a colori”. Per la scansione delle filigrane viene utilizzato uno scanner, frutto di una realizzazione speciale compiuta con l’Istituto Fraunhofer, che si trova unicamente allo Centro scansione della BSB. Un salto di qualità nell’equipaggiamento di scanner e fotocamere si è ottenuto con l’acquisizione dell’attrezzatura in 3D, grazie alla quale lo Scanzentrum ha portato la BSB in una nuova dimensione del lavoro digitale. Le sfide particolari degli oggetti in 3D richiedono, oltre a uno specifico equipaggiamento tecnico, anche copie digitali di grande qualità con tempi e costi sostenibili.

### Il workflow della digitalizzazione in dettaglio

Un ottimale *workflow* deve essere di norma:

- Complessivo
- Efficiente
- Privo di ridondanza
- Trasparente

<sup>20</sup> Cfr. < <http://www.treventus.com/index.html>>. Il numero delle pagine scannerizzate all’ora effettivamente raggiungibile dipende fortemente dall’oggetto/libro “in produzione” e dalle sue caratteristiche. La ditta, sul suo sito web, riferisce persino di 2.500 pagine l’ora.

Lo strumento tecnologico, lo ZEND, garantisce che tutte le fasi di lavoro - dall'ordinazione alla realizzazione della copia d'archivio - si svolgano secondo uno schema uniforme quanto più possibile efficiente e privo di ridondanza. Questo non è di certo un automatismo, ma è il risultato di molti anni di lavoro di tutte le unità organizzative partecipi del *workflow*. In regolari incontri settimanali si lavora continuamente al miglioramento e all'ulteriore ampliamento del *workflow*. Particolare attenzione è stata dedicata dall'inizio ad una quanto più ampia possibile trasparenza del *workflow*. Anche questo scaturisce dalle esperienze di un passato recente, quello dei progetti di scansione dei microfilm. La disponibilità immediata degli oggetti della digitalizzazione, cioè la loro reperibilità nel *workflow* (a favore di qualche utente interessato), era ed è il primo obiettivo.

Di seguito la sequenza dei singoli passi del processo di lavoro pilotato dallo ZEND:

- **Primo passo:** dopo aver deciso la scansione di un documento, viene compilato dall'unità organizzativa responsabile un modulo d'ordine, nel quale sono indicati alcuni parametri da rispettare per la digitalizzazione (per esempio la risoluzione, il numero di pagine per immagine, se il volume può essere sottoposto ad un'apertura di 180° o meno, ecc.) e la sua segnatura.
- **Secondo passo:** l'opera con il modulo d'ordine (stampato) allegato, arriva al Centro di scansione. L'addetto allo scanner, tenuto conto di cosa si richiede nell'ordine, procede alla scansione del documento, facendo estrema attenzione alla sua segnatura, che dovrà alla fine inserire nel nome del file, come riconoscimento dell'opera stessa.
- **Terzo passo:** i dati dell'opera digitalizzata creati dallo scanner sono trasmessi alla memoria di produzione primaria. Lo ZEND ha accesso a questa memoria e amministra tutti i metadati bibliografici, tecnici, strutturali e amministrativi del contenuto digitalizzato, al quale viene assegnato un nome di file definitivo.
- **Quarto passo:** vengono elaborate le immagini, preparate le versioni web ecc. (si tratta di processi automatizzati all'interno dello ZEND). Per gli utenti ci sarà così la possibilità di accesso ai contenuti digitali. Una via di accesso - dall'esterno - sono i normali motori di ricerca e/o i cosiddetti portali d'informazione con standard OAI (per esempio il portale "Biblioteca bavarese online", "Chronicon"). Anche il network delle biblioteche della Baviera, ossia il portale "Gateway Bayern", fornisce un accesso via web a queste risorse. Per il materiale protetto da copyright viene offerto invece soltanto un accesso *in house*.
- **Quinto passo:** lo ZEND provvede anche alla generazione dell'Uniform Resource Number (URN) cioè, alla consegna di un identificatore persistente sul livello nazionale. Gli URN vengono amministrati normalmente dalla relativa biblioteca nazionale, cioè in Germania dalla Deutsche Nationalbibliothek (DNB). La consegna nel caso concreto, però, avviene a livello locale, attraverso l'istituzione richiedente che sta producendo un documento online che deve persistere: in questo caso la BSB.

A questo punto l'URN ed il relativo Uniform Resolving Number (URL) deve essere inserito nel record bibliografico del catalogo unico del network delle biblioteche bavaresi (BVB-OPAC). Lo ZEND prende i metadati bibliografici dal catalogo locale online (BSB-OPAC).

- **Sesto passo:** viene integrato il link che collega il contenuto digitalizzato nel master record bibliografico che si trova nel catalogo unico del (BVB-OPAC).
- **Settimo passo:** si procede alla strutturazione del contenuto digitalizzato tramite il TOC-Editor e/o, in alcuni casi, mediante l'uso di un software OCR che produce in modo automatico versioni fulltext ricercabili<sup>21</sup>.
- **Ottavo passo:** lo ZEND trasferisce tutti i dati per il loro salvataggio a lungo termine nell'archivio (memoria dei dati secondaria) presso il centro calcolo (Supercomputer center) Leibniz (LRZ).

Anche i contenuti digitalizzati che provengono dalla produzione di Google vengono immagazzinati nello ZEND e li trasformati, ossia vengono messi a disposizione sullo ZEND nei singoli servizi del Web (OPAC; Portali di informazione per materie) ed archiviati.

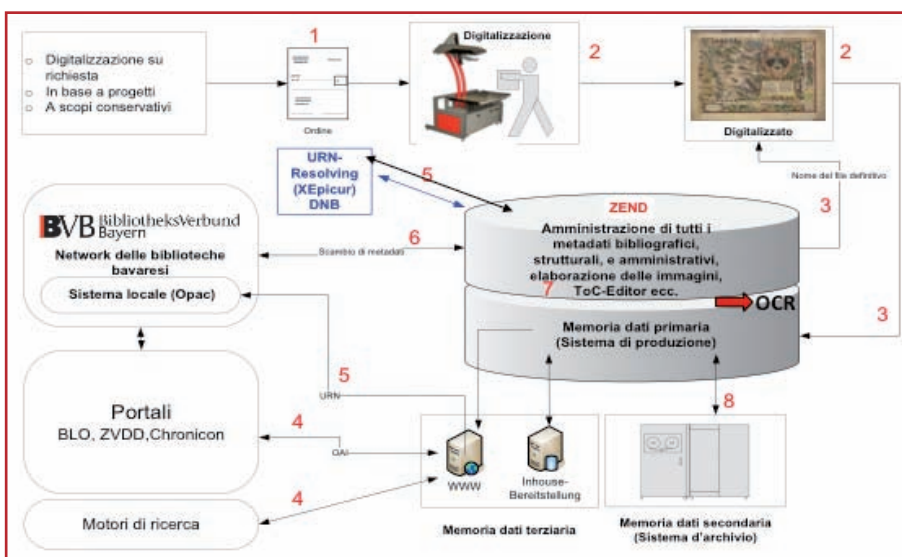


Grafico 1. Schema riassuntivo del workflow

<sup>21</sup> L'uso di routine del software OCR viene ostacolato da vari problemi, non ultimo quello dei caratteri gotici della scrittura *Fraktur* sui testi tedeschi storici. Come si è detto, solo Google ha sviluppato un software OCR che consente una lettura del testo con una tolleranza d'errori accettabile. La BSB effettua una copia completa di OCR nell'ambito del progetto *Digi20*, anche se la procedura non appare priva di problemi. Sui libri usciti recentemente, forti tracce d'uso ostacolano decisamente la digitalizzazione, ossia la realizzazione di una versione del testo completa di OCR. Si torna quindi alla questione "gestione costi": se non è utilizzabile la copia della Biblioteca, bisogna ricorrere – con dispendio di tempo e denaro – ad esemplari in possesso della casa editrice (le sue copie d'archivio) oppure rivolgersi ad altre biblioteche.

## Risultati e conclusioni

Le prestazioni e i risultati della politica di digitalizzazione della BSB si deducono meglio dai contributi ai nuovi portali digitali intersettoriali. Il portale di Europeana, attivato nel 2011, ha ricevuto finora dalla Germania circa un milione di record di opere a stampa (libri). Nel portale culturale nazionale tedesco della Biblioteca tedesca digitale (DDB) circa il 64% di tutte le opere testuali sono della BSB. E tuttavia questi numeri non dicono ancora niente, o molto poco, sulla soddisfazione dell'utente. È fondamentale, come già prima evidenziato, che il patrimonio della BSB dal 1600 fino al 1875 circa, sia ormai in gran parte digitalizzato grazie alla collaborazione con Google. Attraverso entrambi i progetti di digitalizzazione di massa, che sono sotto la responsabilità della BSB e sovvenzionati dal DFG, anche la digitalizzazione delle cinquecentine di lingua tedesca (oltre 40.000 opere) e degli esemplari unici degli incunaboli (circa 9.000 titoli) il limite di tempo è ormai spostato indietro fino a comprendere opere stampate fino al 1452. A breve inizierà la digitalizzazione sistematica dei manoscritti medievali. A ciò si aggiungono inoltre i progetti di digitalizzazione (di massa) condotti in autonomia dalla BSB, come il già citato progetto *Digi20*, grazie al quale, per la prima volta, viene spostato lo spartiacque entro il XX secolo<sup>22</sup>.

Per attivare veramente su tutto il territorio la digitalizzazione nelle biblioteche, negli archivi e nei musei, in tutti gli istituti della memoria, c'è bisogno di finanziamenti di gran lunga più consistenti. Per creare una massa davvero critica di contenuti digitalizzati, liberamente utilizzabili, servono alla Germania, secondo una prima stima approssimativa, almeno 500 milioni di euro, ma gli istituti di cultura tedeschi sono ben lontani dall'ottenere finanziamenti di tale entità. In considerazione del disastroso stato finanziario di quasi ogni stato federale tedesco, solo lo stato centrale può effettivamente agire più liberamente e disporre delle risorse necessarie.

Ciò che attualmente manca è una chiara assunzione di responsabilità per avviare una politica di digitalizzazione sostenuta e sovvenzionata che comprenda a livello nazionale tutti i settori e tutte le regioni. Poche o nulle sono le speranze che un progetto interstatale di digitalizzazione a livello europeo venga promosso anche dall'Europa a fianco di Europeana.

Non sono state soddisfatte finora neanche le speranze di trovare in Germania uno sponsor o un partner privato, come è Google per la BSB. A fronte di questo panorama è possibile che la digitalizzazione del nostro patrimonio culturale si protragga molto più a lungo di quanto si sia pensato. D'altro lato, l'"informatizzazione" del

<sup>22</sup> Fra gli altri progetti di digitalizzazione di massa promossi dalla BSB bisogna ricordare i progetti *Archicon*, (<<http://www.bayerische-landesbibliothek-online.de/bsb-archicon>>), che riguarda la digitalizzazione di disegni/manoscritti d'architettura in senso ampio, e *Altsinica* (<<http://ostasien.digital-sammlungen.de/>>), cioè la digitalizzazione di uno dei fondi storici di materiale cinese più importanti nel mondo, grazie ai quali molti materiali speciali, presenti in gran numero nelle grandi biblioteche storiche, come la BSB, purché privi di copyright, sono stati digitalizzati e appaiono liberamente disponibili in rete.

nostro mondo progredirà comunque in modo molto rapido. Per questo motivo è indubbiamente in crescita il numero di documenti digitali autentici, i cosiddetti *born digital documents*, così dinamici che già forniscono in svariati campi scientifici tutte le informazioni rilevanti e attuali sia *online*, sia a stampa.

È vero ormai che ciò che non è disponibile in un tempo prevedibile e leggibile dalle macchine non esiste, perché rimane vittima della “legge di Internet”: ciò che non è in rete, non esiste.

*The Bayerische Staatsbibliothek (BSB) one of the most important research libraries in Europe with outstanding collections of world wide reputation is at least in the German speaking world one of the pioneers of digitisation. Just in the late nineties of the last century the BSB started a systematical digitisation policy. The last is widely based on three columns of activities: Project-oriented digitisation, user driven digitisation (= digitisation on demand) and last but not least, digitisation motivated by a modern preservation policy; in the BSB with the coming up of digital copies the microfilm as the major preservation remedy was abandoned and the digital copy became the first secondary form of the original items. The most important argument is the digitisation by doing projects which are extra funded by its stakeholder, the Free State of Bavaria or by other third parts, in particular the German research Association, the DFG which runs since years special programs for digitisation in libraries, archives and now also for museums. But the major part of almost one milion digitised volumes is a result of the public-private partnership with Google. The BSB was the first continental european library which took part at the Google digitisation initiative. From the beginning on the digitisation workflows in the BSB are computerbased. The open source based and own developed software tool ZEND serves as production platform. It's modular and can be used in a very flexible way. It covers all the workflow steps, from the order (on internet by an interested single user) until the automatically sending of the mastercopy from the production server to the archiv system which is technically run by the Leibniz super computer center (LRZ). Also the workflows of the Google-project are ZEND-based and -handled. In the BSB own scanning center there are all types of scanning facilities and digital cameras available, also some robot scanner. So the library is able to do mass digitisation (= more than a half a milion of pages) as well as the digitisation of outstanding single objects (= boutique digitisation). The objective is always the same, the creating of attractive digital objects and/or critical masses of digital copies of the same argument which will find immediate user interest. In the case of stressed capacity situation ZEND allows also the digital production, especially the scanning by third parties. The biggest bottle neck for a stronger input in digitisation for all libraries in Germany, on principal also for the BSB, is the absence of a nationally coordinated and centrally funded digitisation policy. The missing of an sufficient financing support forms at the moment also the big obstacle for a consequent digitisation policy on a European level and could be subsequently the main reason for the failure of the well started Europeana initiative.*

L'ultima consultazione dei siti web è avvenuta nel mese di giugno 2013.