

Cosa SBN dovrebbe fare nel breve - medio periodo

Cosa SBN dovrebbe fare nel breve - medio periodo	1
Gruppo di lavoro sulle biblioteche digitali	2
Scopo del documento	3
Apertura di SBN	3
Integrare cataloghi esterni: Indice inteso come “CON-federazione di cataloghi”	3
“Liberare” i record catalografici di SBN	4
Apertura a motori di ricerca “generalisti” (Google, Yahoo e simili)	4
Voci di autorità e interoperabilità	5
Rendere permanentemente accessibili BID e VID SBN: redirect	5
Servizi	6
Fornire indicatori quantitativi sui servizi offerti da SBN nel suo complesso	6
Fornire indicatori quantitativi per decisioni concertate sulle politiche di gestione delle collezioni	7
Allineamento ai principi per lo sviluppo di servizi digitali della Pubblica Amministrazione	7
Realizzazione di API REST	9
Integrazione con servizi come SPID e PagoPA	10
Integrazione con IDEM	11
Digitale: valorizzazione, conservazione e accesso a lungo termine	12
Magazzini digitali: ingest ed elaborazione delle risorse	12
Storage	12
L’Indice SBN e la valorizzazione delle risorse digitali	13
Un esempio: i documenti contenuti nei siti web italiani di carattere pubblico	14

Gruppo di lavoro sulle biblioteche digitali

<https://www.aib.it/struttura/commissioni-e-gruppi/gbdig/>

Componenti in carica (2017-2020)

- Cristian Bacchi (Coordinatore)
- Carla Colombati
- Iolanda Cristaldi
- Bianca Gai
- Cristiana Iommi
- Valdo Pasqui
- Chiara Storti (Referente web)

Consulenti

- Claudio Leombroni
- Maurizio Messina

Consulente esterno

- Pierre Piccotti

Referenti CEN

- Giovanni Bergamin
- Rosa Maiello

Scopo del documento

Il presente documento, elaborato dal Gruppo di lavoro sulle biblioteche digitali, si propone di fornire alcune indicazioni operative di breve / medio periodo per l'evoluzione di SBN, con particolare attenzione alle strutture centrali e alle loro funzioni e servizi, nella convinzione che non sia ulteriormente rinviabile un adeguamento del sistema alle mutate condizioni tecnologiche e alle attuali esigenze degli utenti. Il documento riprende anche temi affrontati dall'AIB a seguito della diffusione a fine del 2017 della versione preliminare del progetto SRI¹. La versione definitiva del Progetto SRI è stata resa pubblica il 4.10.2018 come capitolato tecnico di gara²: questo documento tiene conto delle direzioni già prese e intende formulare proposte nel solco della forte collaborazione che ha sempre caratterizzato le relazioni fra AIB e SBN.

Apertura di SBN

Integrare cataloghi esterni: Indice inteso come “CON-federazione di cataloghi”

L'Indice SBN deve presentarsi all'utente come l'OPAC in cui poter reperire il posseduto di tutte le biblioteche italiane, comprese le biblioteche che non aderiscono alla rete e non catalogano in SBN. A questo fine SBN potrebbe predisporre strumenti e funzioni tali da rendere attraente o conveniente un avvicinamento delle biblioteche estranee alla rete ai servizi SBN. L'uso di standard correnti e di protocolli di larga diffusione e un'accurata valutazione del rapporto costi/benefici sarà condizione del successo di tale offerta.

L'integrazione dei cataloghi esterni ad SBN deve però avvenire a livello di *pubblicazione* dei dati, a valle dei diversi sistemi gestionali: l'OPAC di Indice (e non il back-office che colloquia con i gestionali dei Poli SBN, costituendo la rappresentazione della catalogazione partecipata SBN) dovrebbe periodicamente indicizzare anche cataloghi diversi da quello proprio dell'Indice, e restituire poi i contenuti prevedendo degli appositi filtri di ricerca (solo risultati da Indice SBN, oppure anche altro). Andrebbe dunque sviluppato un servizio di *discovery-tool* unificato, vero fulcro dei servizi, che utilizzi gli standard ed i protocolli recenti più consolidati per la raccolta dei dati, prevedendo anche l'integrazione delle risorse digitali disponibili, che risiedono e restano nei repository di provenienza, tramite l'aggregazione dei *sol*i metadati necessari al loro reperimento.

A titolo di esempio si può considerare il catalogo collettivo swissbib (<https://www.swissbib.ch/>) che integra la Biblioteca nazionale e le Biblioteche universitarie svizzere, oppure il nuovo discovery-tool della British Library (<http://explore.bl.uk>) che integra fonti quali *Research datasets*, *Website*, *Archived website* o *Reports*, oltre a ogni tipo di risorsa digitale e di fonte bibliografica: sia fonti gestite internamente al sistema (*Everything in this catalogue*) sia fonti esterne (*Remote Supply*).

¹ [Osservazioni preliminari al Progetto SRI \(versione 20.11.2017\)](#) documento AIB del 2.1.2018

² [Sistema di Ricerca Integrato e Catalogo generale: Disegno logico del sistema informativo dell'ICCU](#)

I cataloghi integrati in questo modo potrebbero essere quelli di singole biblioteche o, meglio, quelli di reti di biblioteche non colloquianti con SBN, che rispondano a requisiti minimi di qualità. Questo anche al fine di coinvolgere una serie di realtà che si stanno allontanando sempre di più dalla rete nazionale.

Sarebbe opportuno a questo scopo definire un livello di adesione ad SBN più basso degli attuali 4, che prevedono comunque un'interazione con il catalogo: un livello di partecipazione finalizzato alla sola esposizione dei dati nell'OPAC di Indice ma che, comunque e tassativamente, implichi l'adozione dei necessari standard qualitativi relativi al riferimento ai repertori, al livello minimo delle descrizioni, alla coerenza dei dati, all'univocità delle voci di autorità.

Occorre infine ricordare che i poli che si sono aggiunti negli ultimi anni alla catalogazione partecipata sono partiti come "poli vuoti" anche se in realtà avevano un consistente pregresso di record informatizzati e che quindi in realtà l'OPAC dell'Indice SBN non rappresenta nemmeno tutto il posseduto "informatizzato" delle biblioteche aderenti.

“Liberare” i record catalografici di SBN

È necessario ribadire che i dati bibliografici di SBN sono un prodotto della cooperazione e a questa appartengono, e che il catalogo SBN deve essere trattato alla stregua di tutti gli altri dataset pubblici. Attualmente i record di Indice sono esposti con Licenza Attribuzione CC-BY 3.0, ma ciò nonostante non ci sono strumenti per fare export complessivi o scarichi massivi di parti consistenti di dati, che ne consentano il riuso anche in ambienti o domini diversi da quelli in cui sono stati prodotti. Lo sviluppo di queste funzioni appare necessario per adeguarsi alle normative vigenti; un modello potrebbe essere il sistema di export di Wikidata: <https://dumps.wikimedia.org/wikidatawiki/>

Apertura a motori di ricerca “generalisti” (Google, Yahoo e simili)

L'OPAC di Indice, a prescindere dalle possibili iniziative di esposizione dei dati gestiti nel back-office, dovrebbe essere aperto all'indicizzazione da parte dei motori di ricerca di uso più comune (quali Google, Yahoo, Bing o Yandex) integrando, in questo modo, le descrizioni presenti in OPAC tra i risultati delle ricerche effettuate con i più consueti metodi di navigazione nel web.

Per quanto riguarda l'esposizione dei contenuti dei singoli record, nelle pagine HTML del record-display dell'OPAC di Indice potrebbe essere inserito un markup aggiuntivo che esponga una versione semplificata del record bibliografico in base a schemi generici (quali schema.org o Open Graph protocol) rivolti ai crawler automatici: questi metadati aggiunti alle pagine HTML permetterebbero ai motori di ricerca di riconoscere e indicizzare i singoli record dell'OPAC.

Gli stessi dati estratti in base a uno schema generico come schema.org potrebbero essere inseriti nelle pagine HTML del record con una logica Linked Open Data: l'esposizione rivolta ai robot potrebbe essere “negoziata” in base a molteplici formati basati su RDF, quali Turtle, JSON-LD, RDF/XML o N-TRIPLES; e potrebbe eventualmente essere predisposta anche una visualizzazione HTML rivolta all'accesso umano, che presenti i campi dello

schema con i puntatori alle rispettive definizioni (sulla base di standard come l'HTML+RDFa). In questo senso può essere preso ad esempio l'OPAC di Worldcat che, nella visualizzazione dettagliata del record, presenta alla base della pagina una sezione Linked Data.

Si noti che questa ipotesi di esposizione semplificata del record, anche con impostazione Linked Data, non entra nel merito del processo di generazione del record: si tratterebbe di un'estrazione di dati dai record esistenti, in base ad uno schema semplificato quale schema.org.

Mentre, per quanto riguarda l'esposizione di tutti i record ai crawler, potrebbero essere predisposti dei file guida che elenchino le pagine da indicizzare, in base a protocolli quali Sitemaps. Il puntamento ai singoli record dell'OPAC mediante un "Permalink" al display dettagliato è già supportato in Indice e le statistiche di accesso già ne documentano l'utilizzo:³ si tratterebbe di aprire sistematicamente una analoga via d'accesso a tutti i record del catalogo, ad uso dei motori di ricerca e quindi degli utenti finali.

Voci di autorità e interoperabilità

Nella prospettiva dell'interoperabilità tra cataloghi, il primo passo è quello dell'integrazione delle authority; la presenza di link a identificatori esterni consente di affidare alle macchine il matching delle varianti.

Dovrebbe essere possibile inserire nelle voci dell'authority file di Indice, da parte dei catalogatori, le URI degli identificatori delle singole voci nei repertori esterni; al limite rendendoli direttamente ricercabili da SBN.

Per gli autori, sono ormai fonti imprescindibili i riferimenti a VIAF, ISNI e Wikidata (quest'ultimo da preferire a DBpedia in quanto è stato scelto per la manutenzione dei dati strutturati in tutti i progetti Wikimedia e non solo); i riferimenti a tali repertori dovrebbero essere inseriti non solo nelle voci di autorità maggiormente controllate (come i livelli di autorità SUP, LAV e AUF di SBN), la cui forma è già stata accertata, ma anche in quelli con un livello di controllo inferiore. Accade, infatti, che nomi presenti in SBN con un livello di autorità non elevato siano presenti anche in Wikidata ma senza quel reciproco collegamento che aiuterebbe proprio nella definizione della forma corretta⁴.

Su questi temi risultano peraltro attività in corso da parte di ICCU.

Rendere permanentemente accessibili BID e VID SBN: redirect

Sempre nella prospettiva del riuso, oltre che dell'affidabilità dei dati di SBN, deve essere affrontata l'annosa questione dei BID e dei VID che cambiano a seguito delle fusioni di record (per erronee duplicazioni): dovrà essere garantita la permanenza degli identificativi, anche dopo la loro cancellazione e/o sostituzione con altri. Per esempio, in molti altri contesti (come i progetti Wikimedia) esistono meccanismi di redirect dalla versione scartata a quella prescelta.

³ Sezione "Navigazione" > "Accessi" > "Elenco completo" delle statistiche di utilizzo dell'OPAC di Indice: <http://opac.stats.sbn.it/awstats/awstats.pl?config=opac.sbn.it&lang=it>

⁴ A questo proposito è caso esemplare l'iniziativa della Library of Congress: <https://blogs.loc.gov/thesignal/2019/05/integrating-wikidata-at-the-library-of-congress/>

Servizi

Fornire indicatori quantitativi sui servizi offerti da SBN nel suo complesso

Sono necessari indicatori riferiti alla somma dei servizi offerti dalle singole biblioteche o reti di biblioteche aderenti a SBN, affinché il servizio risulti realmente valutabile sia internamente alle biblioteche (verificando quali siano e in quale direzione si possano individuare i margini di miglioramento, e quali servizi possano essere razionalizzati), sia nel confronto con il decisore politico e con le possibili fonti di finanziamento (evidenziando il possibile ritorno di un investimento, in termini di servizi erogati e di visibilità dei risultati), sia per le attività di accountability e rendicontazione sociale.

La selezione dei dati da rilevare dovrebbe far tesoro delle analisi scientifiche prodotte dai gruppi di lavoro concentrati sui servizi, mirando a individuare pochissimi indicatori realmente sostenibili in una prospettiva di aggiornamento costante delle informazioni da parte delle biblioteche.

Un primo esempio di dato da rilevare potrebbe certamente essere la statistica degli accessi al catalogo online, essendo questa senz'altro una componente fondamentale dell'erogazione dei servizi: e in SBN vengono già fornite le statistiche di accesso all'OPAC di Indice, prospettate in base a molteplici criteri di quantificazione;⁵ ma se ai numeri degli accessi all'OPAC di Indice potessero essere aggiunti quelli dell'accesso agli OPAC locali nei Poli, le quantità di sicuro aumenterebbero sensibilmente dando così la misura del reale utilizzo del catalogo condiviso di SBN su tutto il territorio nazionale. Inoltre, se le statistiche dei Poli locali potessero essere prospettate centralmente anche con l'opzione del dato disaggregato, si avrebbe la possibilità di verificare la "geografia delle ricerche" sul territorio. Le statistiche di accesso agli OPAC, grazie agli strumenti di monitoraggio generalmente già presenti nelle installazioni di Polo oltre alle soluzioni esterne di Web analytics disponibili sul mercato, sono forse il dato più facile da prevedere in una prospettiva di costante aggiornamento delle informazioni. Altri potenziali indicatori, il cui aggiornamento richiede un maggiore intervento da parte dei bibliotecari, potrebbero essere i prestiti, i prestiti ILL e le riproduzioni, ormai prevalentemente digitali.

Un importante termine di paragone per la selezione degli identificatori ma soprattutto per la modalità di rappresentazione dei dati si può trovare nelle statistiche IFLA della "Library Map Of The World" (<https://librarymap.ifla.org/>). In queste statistiche gli indicatori presi in esame sono pochi e semplici⁶ (types of libraries, library service points, internet access, full-time equivalent staff, volunteers, registered users, visitors, loans and downloads) e non risultano presenti fra quelli gestiti in Anagrafe SBN: già questi indicatori, raccolti con un adeguato grado di aggiornamento, costituirebbero senz'altro un riferimento prezioso per la pianificazione dei servizi e per l'argomentazione delle richieste di finanziamento.

⁵ <http://opac.stats.sbn.it/awstats/awstats.pl?config=opac.sbn.it&lang=it>

⁶ Si veda la sezione <https://librarymap.ifla.org/data-collection>

I dati sintetizzati sul piano nazionale dovrebbero essere sempre facilmente ottenibili (o linkati) dai principali siti centrali di SBN, quali l'OPAC di Indice e l'Anagrafe; mentre per la loro raccolta centralizzata, che potrebbe avvenire ad opera dell'ICCU o di altro organo da stabilire, dovrebbero essere definite delle linee guida minimali per l'esposizione dei dati da parte delle biblioteche.

Così come avviene per *amministrazione trasparente* in ambito pubblico, le biblioteche potrebbero esporre i dati contenenti i propri indicatori, e tale esposizione potrebbe più probabilmente essere predisposta a livello di Polo.

A questo scopo potrebbe essere definita una procedura di esposizione con Linked Open Data basati su ontologie semplici e condivise: in questo modo l'aggregatore centrale potrebbe raccogliere i dati automaticamente, mentre le stesse fonti di dati potrebbero essere aperte anche al riuso da parte di operatori esterni ad SBN che vogliono dare notizia dei servizi delle biblioteche.

In Indice, poi, è previsto un articolato sistema di monitoraggio e statistiche che permette di estrarre indicatori rispetto ai dati, alle attività dei Poli, alle prestazioni del sistema, alle duplicazioni e agli archivi di autori e titoli; la produzione di questi report, però, è finalizzata solo al supporto delle attività manutentive e dev'essere specificamente richiesta.

Se questi indicatori potessero essere prodotti in modo agile e sintetico, risulterebbero senz'altro utili al miglioramento del servizio finale all'utente - inteso come reale capacità di rendere disponibile il documento - oltre che della gestione del catalogo condiviso (che è alla base di tutti i servizi ma che non è, appunto, il solo servizio).

Fornire indicatori quantitativi per decisioni concertate sulle politiche di gestione delle collezioni

Il tema è stato affrontato e discusso dal Gruppo di lavoro, ma vi si accenna qui sinteticamente, in quanto innesca una serie di considerazioni sulle politiche nazionali per i servizi bibliotecari e bibliografici che esulano dagli scopi di questo documento.

Si ritiene comunque che l'indice potrebbe fornire degli indicatori quantitativi agevolmente reperibili su tipologie documentarie e contenuti del catalogo, in vista di un coordinamento territoriale, a livello dei singoli poli, delle scelte riguardo il posseduto, gli acquisti, gli scarti e le eventuali centralizzazioni del magazzino.

Nel caso dei periodici, in particolare, già in OPAC sono ovviamente presenti per ciascun record i dati di localizzazione nelle diverse istituzioni ma andrebbe valutata l'opportunità di riorganizzare la prospettazione delle consistenze, ove presenti.

Allineamento ai principi per lo sviluppo di servizi digitali della Pubblica Amministrazione

SBN e la maggior parte delle biblioteche che ne fanno parte sono Pubbliche Amministrazioni e come tali coinvolte nel processo di trasformazione digitale della PA italiana. Il Codice per l'Amministrazione Digitale (CAD) affida all'Agenzia per l'Italia Digitale (AgID) il compito di predisporre e verificare l'attuazione del Piano triennale per l'informatica nella Pubblica Amministrazione e l'emanazione delle linee guida e dei documenti per

orientare e governare questo processo di trasformazione. SBN ha intrapreso l'avvio di questo processo con largo anticipo rispetto ad altri settori della PA e comunque l'evoluzione dei suoi servizi deve riferirsi ai principi e alle linee guida definiti dall'AgiD al fine di assicurare la piena adesione agli standard e l'integrazione con gli altri servizi nazionali ⁷.

Assumendo un punto di vista più ampio, a livello europeo occorre ricordare il Regolamento Europeo 910/2014 *Electronic Identification (eID) and Trust Services (AS)*, da cui l'acronimo eIDAS, tradotto in italiano come "Regolamento in materia di identificazione elettronica e servizi fiduciari per le transazioni elettroniche". Questo regolamento definisce un framework con lo scopo di assicurare che la Identificazione Elettronica (eID)⁸ e i servizi fiduciari (AS)⁹ siano interoperabili tra i vari paesi Europei in modo da rendere più sicure ed efficienti le interazioni tra operatori commerciali, cittadini e enti pubblici, indipendentemente dal paese Europeo in cui avvengono.

Il Team per la trasformazione digitale ai primi di settembre 2018 ha pubblicato "La carta dei principi per la condotta tecnologica",¹⁰ contiene una serie di indicazioni che le PA dovranno richiamare e dettagliare nei capitolati di gara per i servizi tecnologici della Pubblica Amministrazione e nelle gare strategiche di Consip. Questi definiscono alcuni requisiti imprescindibili per l'evoluzione dei servizi SBN, tra cui:

- Usare standard aperti per garantire che la tecnologia sviluppata funzioni e comunichi con altre tecnologie e possa essere facilmente aggiornata e ampliata.
- Condividere e riutilizzare dati e tecnologie
- Pubblicare il codice sorgente con licenze Open Source per migliorare la trasparenza, la flessibilità e la responsabilità

Quest'ultimo requisito include due prescrizioni: a) Inserire nel capitolato l'obbligo di rilasciare alla pubblica amministrazione la proprietà intellettuale del software che viene sviluppato ad hoc; b) pubblicare tutto il codice sotto licenza aperta. Si tratta di prescrizioni la cui valenza non è solo di carattere tecnico ma implicano anche scelte strategiche: nel gruppo si è più volte discusso dell'opportunità di un rilascio in modalità open source di tutto

⁷ Tra i documenti emanati da AgiD rivestono particolare rilevanza nel contesto di SBN:

- *Linee Guida Modello di Interoperabilità* [<https://lq-modellointeroperabilita.readthedocs.io/it/latest/>]
- *Il Cloud della PA* [<https://cloud-pa.readthedocs.io/it/latest/>]
- *Linee guida di design per i servizi digitali della PA* [<https://design-italia.readthedocs.io/it/stable/>]
- *Linee guida per l'acquisizione e il riuso del software* [<http://lq-acquisizione-e-riuso-software-per-la-pa.readthedocs.io/it/latest/>]

⁸ L'identificazione elettronica comprende il processo di identificazione (chi sono) e quello di autenticazione (dimostrazione che sono chi dico di essere) per poter ottenere l'accesso ai servizi e alle risorse. Dal mese di settembre 2018 per tutti i paesi EU è obbligatorio riconoscere sistemi di eID notificati alla Commissione europea.

La lista dei sistemi eID notificati è consultabile a

<https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL2018/Country+Overview+-+eID>

SPID è stato notificato alla Commissione Europea nel mese di luglio 2018, l'iter si è concluso con la pubblicazione nella Gazzetta Ufficiale dell'Unione Europea ([GUUE C318](#) e [GUUE C344](#) del 2018) e di conseguenza a decorrere dal 10 settembre 2019 l'identità digitale SPID potrà essere usata per l'accesso ai servizi in rete di tutte le pubbliche amministrazioni dell'Unione, come riportato nella sezione del portale AgiD dedicata a SPI: <https://www.agid.gov.it/it/piattaforme/spid>.

⁹ I servizi fiduciari previsti da eIDAS comprendono la firma digitale (eSignature), la marca temporale (eTimestamp), il sigillo elettronico (eSeal), i servizi di delivery registrati (ERDS) e i certificati qualificati per l'autenticazione dei siti web (QWAC) che assicurano l'autenticità dei siti web.

¹⁰ <https://carta-dei-principi-tecnologici-del-procurement.readthedocs.io/it/latest/>

il software realizzato in ambito SBN, considerando i benefici che potrebbero derivare dalla creazione di una community interessata al suo sviluppo. Le recenti decisioni dell'ICCU che hanno disposto il rilascio in licenza AGPLv3 del software OPAC SBN e SBNWEB¹¹ appaiono quindi pienamente in linea con tali considerazioni. Gli altri due requisiti sono richiamati e sviluppati nei due paragrafi seguenti.

Realizzazione di API REST

Le linee guida di interoperabilità dell'AgID prevedono di utilizzare standard e formati aperti per file e protocolli di comunicazione e l'implementazione delle funzionalità in forma di API documentate. API è l'acronimo di *Application Programming Interface* e indica un insieme di "procedure/funzionalità/operazioni disponibili al programmatore, di solito raggruppate a formare un set di strumenti specifici per l'espletamento di un determinato compito"¹². Una API è una interfaccia che espone delle funzionalità applicative (servizi) che sono necessarie per realizzare altri servizi.¹³ Uno dei principali vantaggi dell'integrazione tra applicazioni basate su API è che tramite questo approccio i servizi vengono definiti in modo astratto, nascondendo le diversità e le complessità dell'implementazione interna. Ad esempio, l'esportazione di una unità di contenuto (come un record bibliografico) in una API può essere definita semplicemente specificando la modalità in cui questa deve essere richiesta (il metodo di interrogazione) e restituita, mentre i due sistemi che colloquiano possono essere sviluppati con linguaggi di programmazione e con finalità completamente differenti.

Questa impostazione del colloquio fra sistemi (o fra moduli del medesimo software) permette anche di avere una maggiore autonomia nella manutenzione, perché sarà sempre possibile modificare e aggiornare le parti componenti dell'architettura senza compromettere la loro interazione.

Inoltre quando la logica applicativa è fornita attraverso API i servizi esposti possono essere utilizzati dal programmatore lato client in modo molto simile all'uso di una libreria locale: cioè come una serie di servizi di cui è necessario conoscere solo le modalità di chiamata e di risposta, senza entrare nel merito del loro meccanismo di funzionamento interno.

¹¹ Pagina nel sito ICCU: <https://www.iccu.sbn.it/it/eventi-novita/novita/OPAC-di-Polo-SBNWeb/>; pagina ICCU su GitHub <https://github.com/IstitutoCentraleCatalogoUnicoBiblio>

¹² Voce di Wikipedia: https://it.wikipedia.org/wiki/Application_programming_interface

¹³ Le interfacce di questo tipo fanno riferimento al modello architetturale Service Oriented Architecture (SOA) nel quale le applicazioni, distribuite in rete, interagiscono tra di loro attraverso servizi che attraverso l'appropriata specifica del loro comportamento esterno nascondono le complessità delle implementazioni interne. Esistono diversi approcci, tecnologie e standard di riferimento per la realizzazione di questi servizi. Uno dei più consolidati, attraverso vari insieme di standard ISO, è il protocollo SOAP (Simple Object Access Protocol) che definisce la struttura dei messaggi avvalendosi della sintassi XML e prevedendo una modalità di interazione di tipo Remote Procedure Call (RPC): i messaggi possono essere trasportati da diversi protocolli sottostanti (es. HTTP, SMTP, TCP).

Recentemente si è affermato per la sua chiarezza e semplicità concettuale il REpresentational State Transfer (REST), uno stile architetturale che non richiede la definizione di uno specifico protocollo come nel caso di SOAP. I Web Services realizzati secondo questo modello usano HTTP come protocollo applicativo (invece SOAP lo usa solo come protocollo di trasporto dei messaggi) basandosi direttamente sulla semantica dei 5 metodi di HTTP (GET, HEAD, PUT, DELETE, POST) attraverso i quali vengono applicate operazioni a risorse identificate mediante URI (Uniform Resource Identifier). Sono ampiamente riconosciuti i vantaggi di manutenibilità, portabilità ed efficienza dei web services di tipo REST per realizzare l'interoperabilità applicativa.

L'Indice SBN garantisce l'interazione con i sistemi di Polo mediante un protocollo applicativo appositamente concepito, denominato SBNMARC e definito tramite uno schema XML.¹⁴ Purtroppo SBNMARC non si è affermato come standard internazionale e il suo utilizzo è rimasto confinato all'interazione dei sw di Polo con l'Indice SBN senza assurgere a standard riconosciuto fuori dai confini italiani.

Le caratteristiche di SBNMARC sono comunque compatibili con un approccio di tipo REST poiché la logica applicativa può essere espressa tramite la semantica HTTP e le operazioni richieste sono di tipo CRUD (Create, Read, Update and Delete). Si potrebbe, dunque, considerare l'opportunità di adeguare l'interfaccia dei servizi SBN, Indice compreso, agli standard tecnologici attuali, attraverso la realizzazione di API REST.

La scelta di esporre un'insieme di Web Services che consentano operazioni più granulari sui dati/oggetti di SBN consentirebbe di avere una maggiore facilità di implementazione dei client applicativi che devono interoperare con l'Indice SBN, una più efficiente gestione delle evoluzioni rendendole trasparenti rispetto alle applicazioni esterne¹⁵ che colloquiano con l'Indice SBN e una maggiore apertura alle integrazioni e agli utilizzi dei dati SBN in contesti non bibliotecari.

Integrazione con servizi come SPID e PagoPA

Nel modello AgID sono comprese alcune piattaforme abilitanti, già pienamente operative, tra cui SPID e pagoPA:

- SPID, il Sistema Pubblico di Identità Digitale, garantisce a tutti i cittadini e le imprese un accesso unico, sicuro e protetto ai servizi digitali della Pubblica Amministrazione.
- PagoPA, il sistema di pagamenti elettronici realizzato per rendere più semplice, sicuro e trasparente qualsiasi pagamento verso la Pubblica Amministrazione

I servizi SBN offerti al pubblico dovrebbero prevedere un adeguamento all'introduzione di SPID, non solo del servizio di prestito interbibliotecario (ILL SBN) ma di qualunque tipologia di servizio online che preveda la registrazione e l'identificazione dell'utente.

Dal momento che verranno trattati dati personali il servizio dovrà essere GDPR compliant (Regolamento UE 678/2016) e trattandosi di sviluppo software dovrà prevedere sin dalla progettazione la protezione *by design* e *by default*. In particolare sarà opportuno prevedere strumenti per la cancellazione dei dati personali e per il diritto di accesso da parte degli utenti.¹⁶

Inoltre l'integrazione dei servizi SBN pubblici con SPID permetterebbe anche di:

¹⁴ Si noti che in SBNMARC, pur avendo una logica di tipo SOAP, i messaggi sono codificati in base ad uno specifico schema XML e trasportati sul protocollo HTTP (stateless) con i comandi GET e POST, pertanto non si seguono gli standard dei Web Services SOAP.

¹⁵ Per esempio con l'aggiunta di nuovi Web Services o avvalendosi della tecnica dell'override dei parametri di chiamata. Inoltre, la realizzazione di API REST permetterebbe di supportare molteplici sintassi e formati di trasferimento dei dati (es. JSON oltre a XML), consentirebbe una gestione dei profili 'autorizzativi più flessibile e configurabile a livello di singole risorse, la distinzione tra servizi aperti/pubblici e servizi autenticati/autorizzati e una maggiore modularizzazione dei servizi riservati ai sw client.

¹⁶ Articolo 25 del GDPR: "Protezione dei dati fin dalla progettazione e protezione per impostazione predefinita". Questi introduce il principio di privacy by design e privacy by default, un approccio concettuale innovativo che impone l'obbligo di avviare un progetto prevedendo, fin da subito, gli strumenti e le corrette impostazioni a tutela dei dati personali.

- prevedere, sin dalla progettazione del software, strumenti di analisi dei comportamenti degli utenti (anonimizzati) al fine di studiare comportamenti ed esigenze
- facilitare l'accesso autenticato alle risorse digitali integrate con l'Indice SBN
- studiare e sviluppare nuove forme di integrazione tra i servizi centralizzati di SBN e quelli locali delle biblioteche dei Poli SBN.

Il servizio PagoPA fornisce la corretta soluzione al problema di come realizzare una piattaforma di pagamenti per le biblioteche aderenti a SBN che eviti i costi e le complicazioni derivanti da accordi e contratti con singoli operatori del settore bancario per l'integrazione con sistemi di pagamento online (es. POS virtuali): una prima applicazione potrebbe realizzarsi per i servizi a pagamento di prestito interbibliotecario e document delivery intermediati dal ILL SBN, e per il pagamento della fornitura di riproduzioni. Nel contesto PagoPA, SBN e ICCU potrebbero svolgere un ruolo di fornitore di servizi nei confronti dei Poli e delle biblioteche aderenti a SBN, sollevandole dagli oneri tecnico-organizzativi che l'architettura PagoPA implica¹⁷.

Integrazione con IDEM

Oltre allo SPID, che è pensato per la comunità dei cittadini, e sempre nell'ambito dell'identificazione dell'utenza, esiste il servizio IDEM [<https://www.idem.garr.it/>], promosso e realizzato dalla Federazione Italiana delle Università e degli Enti di Ricerca per l'Autenticazione e l'Autorizzazione. IDEM fornisce ai ricercatori degli istituti di ricerca, ai docenti e agli studenti delle università e delle scuole la possibilità di usare una sola identità e una sola password per accedere a diversi servizi in rete. IDEM aderisce anche a eduGAIN [<https://www.idem.garr.it/informazioni-tecniche/283-faq-edugain>], il servizio internazionale che connette le "Federazioni di identità" della ricerca e dell'istruzione e permette agli utenti di accedere ai servizi in rete tramite un sistema di *single sign on* (SSO) su scala internazionale usando l'unica password data dalla propria organizzazione di appartenenza.

Considerata la rilevante presenza di biblioteche universitarie e centri di ricerca in SBN l'integrazione dei servizi di identificazione oltre che con SPID deve essere realizzata anche con IDEM.

¹⁷ L'adesione di una PA a PagoPA può essere diretta o indiretta. Nel primo caso la PA assume formalmente a proprio carico tutte le attività tecniche necessarie all'attivazione e al funzionamento. Nel secondo caso la PA demanda le attività tecniche a un intermediario tecnologico che già aderisce al sistema (es. regione, comune, ASL, etc.) oppure ad un partner tecnologico che fornisce esclusivamente servizi ICT alle PA (es. consorzio, banca, etc.) e che non aderisce al sistema ma fornisce alla PA le infrastrutture, i collegamenti, i servizi e il supporto necessari per la connessione a PagoPA e l'attivazione dei servizi di incasso. In particolare l'intermediario o il partner si fanno carico di attivare e gestire il Nodo dei Pagamenti-SPC che è la piattaforma tecnologica per l'interconnessione e l'interoperabilità tra la PA e i Prestatori di Servizi di Pagamento nel rispetto dell'architettura e delle regole tecniche fissate. ICCU potrebbe assumere il compito di gestire il Nodo dei Pagamenti-SPC per conto delle biblioteche aderenti a SBN.

Si consideri che questo non significa che ICCU deve acquisire le risorse (hw e sw) per implementare il nodo dei pagamenti: può fungere da intermediario, individuando un partner che lo gestisca tecnicamente per conto della comunità SBN.

Digitale: valorizzazione, conservazione e accesso a lungo termine

Magazzini digitali: ingest ed elaborazione delle risorse

Magazzini digitali è l'infrastruttura nazionale per la conservazione e l'accesso a lungo termine delle risorse digitali. Dal 2006, anno in cui è iniziata la sperimentazione del deposito legale dei documenti diffusi tramite rete informatica previsto dall'art. 37 DPR 252/2006, il progetto è cresciuto molto fino a comprendere diverse tipologie documentarie che richiedono metodi di raccolta/ingest e di trattamento specifici.

In questo contesto MD si presenta come snodo essenziale dell'ecosistema dei servizi di accesso alle risorse documentarie. Quando si dice che il metaindice (o confederazione dei cataloghi) deve fornire accesso alle risorse digitali indicizzando i soli metadati e senza centralizzare le risorse occorre ricordare che la garanzia di continuità dei servizi nel tempo può essere fornita solo da servizi di accesso nel lungo periodo come MD. Senza un alto livello di integrazione il metaindice si trasformerebbe in brevissimo tempo in un inservibile insieme di link irrisolti.

Storage

Come è noto, mentre gli oggetti digitali possono differire per natura e modalità di produzione e fruizione, la necessità di storage che la loro gestione determina è invece uniforme (e, anzi, dovrebbe essere uniformata in ogni caso). È dunque necessario ipotizzare la costituzione di un'infrastruttura di storage condivisa per la conservazione a lungo termine degli oggetti digitali collegati al catalogo di SBN.

È bene comunque precisare che tale infrastruttura:

- A. non può coincidere direttamente con i servizi di storage IaaS previsti da Cloud SPC Lotto 1 di Consip (<https://www.cloudspc.it/Servizi.html>) che si limitano a fornire lo storage, ma potrebbe avvalersene come strato/componente di livello più basso;
- B. tale servizio inoltre non deve essere confuso con il servizio di Conservazione dei documenti informatici a norma dell'art. 44 del CAD¹⁸, anche questo previsto in ambito Cloud SPC Lotto 1 di Consip, poiché quest'ultimo è orientato esclusivamente alla conservazione digitale di documenti e fascicoli amministrativi prodotti dalle PA, quindi con valore probante e non solo culturale.

Per tutto quanto attiene al digitale collegato ad Indice, sarebbe comunque utile un confronto con i *big providers* sia nella fase di digitalizzazione vera e propria (come avviene per il progetto *Google Books*) sia per la successiva gestione (conservazione e accesso).

18

https://docs.italia.it/italia/piano-triennale-ict/codice-amministrazione-digitale-docs/it/v2017-12-13/_rst/capo3_art4_4.html

Un servizio pilota potrebbe partire proprio dall'esperienza di Magazzini Digitali, che esce rafforzata dal recente inserimento da parte dell'AgID, anche a seguito dell'interlocuzione con BNCf, BNCr, ICCU e AIB, della base dati SBN nel Piano triennale 2019-2021¹⁹. Data la complementarità funzionale fra SBN e MD è lecito attendersi, e vi sono segnali di disponibilità in tal senso, che anche MD venga inserita fra le basi dati di interesse nazionale, essendo questa una condizione essenziale della sostenibilità complessiva del progetto.

Si consideri in questo senso anche la *"Proposta di costituzione di un'infrastruttura di storage condivisa per la conservazione a lungo termine degli oggetti digitali (nativi e non) di interesse culturale"*,²⁰ recentemente avanzata dalla BNCf, che mira a fornire a ciascuna istituzione un servizio di storage condiviso con standard di conservazione e sicurezza omogenei. Sono evidenti, in questa proposta, le economie di scala realizzabili rispetto ai costi necessari perchè ciascun ente si doti di sistemi di storage autonomi, che potrebbero risultare diversi e non interoperabili in dipendenza delle tecnologie scelte e delle procedure operative adottate.

Lo storage dovrà dialogare con gli strati applicativi superiori che potranno essere diversificati a seconda delle esigenze dei singoli materiali, dei singoli istituti, e degli specifici modelli di servizio, anche in base alla *mission* istituzionale, come avviene in Magazzini Digitali.

L'Indice SBN e la valorizzazione delle risorse digitali

A parte alcuni progetti che, se pur in evoluzione, sono concepiti secondo logiche ormai poco funzionali (es. Internet Culturale, che non viene aggiornato dinamicamente), non c'è modo di sapere esattamente cosa è stato digitalizzato nell'ambito delle biblioteche in Italia e da chi.

È necessario quindi, come peraltro già previsto da ICCU,²¹ che per i materiali gestiti in SBN sia nei progetti di digitalizzazione futuri sia in una prospettiva di ottimizzazione di quelli pregressi, venga sempre prevista la creazione in SBN del link alla copia digitale a partire dalla descrizione della copia fisica (affrontando anche le questioni legate alla scelta di gestire il link a livello di edizione o singolo esemplare).

Un'indagine metodica, ad esempio, con grande probabilità mostrerebbe che quasi tutti i periodici italiani storici sono stati digitalizzati. Invece, vista la frammentazione delle interfacce di ricerca, si movimentano - spesso a fatica - i periodici tradizionali (oltre che le persone) anche quando sono in realtà disponibili online. Su questo tema appare di vitale importanza anche stringere accordi con gli editori per il recupero degli archivi di periodici.

¹⁹

https://docs.italia.it/italia/piano-triennale-ict/pianotriennale-ict-doc/it/2019-2021/05_dati-della-pubblica-amministrazione.html#basi-di-dati-di-interesse-nazionale , 5.1.1. - Scenario.

²⁰ https://docs.google.com/document/d/16EwoYptAlss6l-823P_HRG-naKs5cl1kOPEh9CfBDdM/edit

²¹ Si veda la circolare ICCU 2018/01/17, prot. 114/SBN3.3

Un esempio: i documenti contenuti nei siti web italiani di carattere pubblico

La documentazione di fonte pubblica, ormai quasi esclusivamente prodotta e diffusa tramite rete informatica, è spesso di difficile accesso da parte del cittadino o dello studioso, per i seguenti motivi:

- nonostante il lavoro dell'AGID per la diffusione di linee guida sull'organizzazione dei siti della PA e, ancor prima, delle varie norme legate alla trasparenza amministrativa, tali siti spesso non espongono in maniera sufficientemente chiara o semplice la propria documentazione;
- i documenti spesso mancano di metadati e non sono raggiungibili attraverso i motori di ricerca;
- la permanenza del documento su un sito e/o la raggiungibilità del sito stesso non sono quasi mai garantiti;
- i documenti born digital non sono ancora soggetti a deposito legale quindi non vengono gestiti dalle biblioteche.

È dunque necessario rendere ricercabile e disponibile attraverso l'OPAC di Indice la documentazione di fonte pubblica, sia digitalizzata (di cui quindi esiste anche un equivalente cartaceo presumibilmente già descritto in SBN) sia nativa digitale. In entrambi i casi va garantita l'accessibilità a lungo termine ai file. La soluzione ottimale sarebbe quella di legare questo servizio a Magazzini Digitali che proprio in questi mesi sta sperimentando l'harvesting di tali documenti: metadati minimali potrebbero essere automaticamente caricati in SBN, a seguito dei controlli sulle voci di autorità.