

Gli open data delle biblioteche italiane

<http://anagrafe.iccu.sbn.it>

<http://opendata.anagrafe.iccu.sbn.it>

I numeri dell'Anagrafe

- 13.484 biblioteche censite
- 17.441 indirizzi
- circa 230.000 ricerche nel 2014
- circa 6.500.000 di schede sintetiche presentate nel 2014

(dati aggiornati al 27 febbraio)

Modalità di aggiornamento

- applicativo web gestionale
 - interventi di collaboratori autenticati
 - interventi del personale ICCU (circa 2.000 nel 2014)
- invio di file secondo un tracciato XML (1.244 nel 2014)

Formato di scambio

- *biblioteche.xsd*, nato nel 2002 e arrivato alla versione 1.6
- usato regolarmente da Regione Veneto, UNBCE della CEI, Regione Campania
- presto tra noi la Regione Liguria, grazie ad una trasformazione XSLT ormai quasi pronta
- in questi giorni l'UNBCE ha aggiornato i dati di 463 biblioteche di ente ecclesiastico

Come *aprire* i dati?

- lo schema XML non sembra adatto
 - è stato sviluppato per lo scambio di dati con i partner istituzionali
 - richiede sviluppo software piuttosto pesante
- è sembrato più opportuno sviluppare un'estrazione *ad hoc*, più semplice e leggera
 - classificazione W3C: 3 stelle
 - licenza CC0

Quali dati *aprire*?

- cinque data set
 - anagrafica e territorio (CSV)
 - contatti (XML)
 - patrimonio (XML)
 - fondi speciali (XML)
 - tipologie (CSV)
- il codice ISIL identifica univocamente una biblioteca in ognuno dei data set

Anagrafica e territorio

- probabilmente i dati più appetibili per il grande pubblico
- TomTom li ha utilizzati nelle sue applicazioni *Routes e Places*
- anagrafica, indirizzo, codici ISTAT (comune e provincia), CAP, coordinate, contatti principali

Dettagli tecnici, 1

- strumenti di sviluppo
 - linguaggio Java (SDK 1.6)
 - ambiente di sviluppo Eclipse 3.8
 - sistema operativo Debian 7 “Wheezy”
 - gestione delle versioni con *git*
 - *git* repository su Google Code

Dettagli tecnici, 2

- librerie Java
 - JDOM 1.1.3 per la costruzione dei file XML
 - MySQL Connector 5.0.7 per accesso JDBC al database
 - Apache Log4j 1.2.17 per il logging
- altri strumenti
 - LibreOffice Base per la messa a punto delle query SQL

Tempi e costi, approssimativi

- 32 ore/uomo (8 giorni lavorativi, 4 ore al giorno) per il prototipo
- 20 ore/uomo (4 giorni lavorativi, 5 ore al giorno) per la versione attualmente operativa
- costi:
 - software = 0
 - lavoro = 624 euro \approx 52 ore a 12 euro l'ora

Buone pratiche

- ICCU, ICAR e DGA hanno messo a fattor comune le (poche) risorse umane informatiche formalizzando una collaborazione
- questa collaborazione ha reso possibile questo risultato (e anche altri)
- la collaborazione e la condivisione di conoscenze e competenze è importante almeno quanto la tecnica