



Ministero dei beni e delle attività culturali e del turismo

BIBLIOTECA NAZIONALE CENTRALE

50122 Firenze – Piazza Cavalleggeri, 1 – Tel. 055/249191 – <http://www.bncf.firenze.sbn.it>

AVVISO PUBBLICO PER MANIFESTAZIONE D'INTERESSE AL RILASCIO DELL'AUTORIZZAZIONE ALL' HARVESTING DEI DOCUMENTI CONTENUTI NEI SITI WEB DI INTERESSE CULTURALE

Nell'ambito del programma triennale del MIBACT finanziato ai sensi dell'art. 1, commi 9 e 10, della L. 190/2014 (Legge di stabilità 2015), la [Biblioteca Nazionale Centrale di Firenze](#) è capofila di **un progetto di raccolta periodica (*harvesting*) dei documenti contenuti nei siti web di istituzioni di interesse culturale**, e di conservazione in [Magazzini Digitali](#).

Il progetto ha lo scopo di **assicurare la conservazione a lungo termine e l'accesso (nei limiti della vigente normativa sul Diritto d'Autore e diritti connessi) ai documenti digitali** pubblicati nei siti delle istituzioni di interesse culturale, a partire da quelli della Pubblica Amministrazione, e che si configurano come la naturale evoluzione delle raccolte fisiche delle biblioteche.

In attesa del completamento del **“Regolamento sul deposito dei documenti diffusi tramite rete informatica”** - previsto dall'art. 37 del D.P.R. 3 maggio 2006 n. 252 - che renderà obbligatorio il deposito legale di tali documenti, **è necessario richiedere ai titolari dei siti web interessati un'esplicita autorizzazione** al fine di garantire alla BNCF la facoltà di effettuare la raccolta automatica (*harvesting*) dei documenti contenuti nei siti web di interesse culturale; tale autorizzazione potrà prevedere anche la possibilità che la BNCF offra l'accesso in rete alla raccolta.

La raccolta riguarderà, in una prima fase, documentazione di natura “pubblica” appartenente ad una delle seguenti tipologie:

1. Pubblicazioni ufficiali dello Stato
2. Altre pubblicazioni dello Stato
3. Pubblicazioni dei Ministeri
4. Pubblicazioni degli organi centrali e periferici del MIBACT
5. Pubblicazioni di altri Enti e Istituzioni statali o a carattere nazionale
6. Pubblicazioni delle Regioni e degli altri enti locali territoriali

B N C F

Posta elettronica: Ufficio di Direzione: bnc-fi@beniculturali.it - Posta certificata: mbac-bnc-fi@mailcert.beniculturali.it

7. Pubblicazioni delle Università (ad esclusione di quelle già rientranti in Magazzini Digitali)
8. Pubblicazioni presenti su siti web di altre istituzioni di interesse culturale

Non si esclude la possibilità di estendere il campo di applicazione del progetto ad altri tipi di documentazione su richiesta degli interessati, purché sia rispettato il principio della continuità con le raccolte esistenti delle biblioteche depositarie.

Enti e istituzioni produttori di tale documentazione saranno contattati dalla BNCF tramite e-mail, dall'indirizzo harvesting@depositolegale.it, con la richiesta di compilare un modulo con i dati relativi al sito web, al suo responsabile istituzionale e al suo referente tecnico.

Chi volesse aderire, anche spontaneamente, è pregato di **compilare l'apposito [modulo online](#)**.

Resta ferma la possibilità da parte di BNCF di non procedere all'*harvesting* di taluni domini non in linea con i principi enunciati.

Specifiche tecniche

La raccolta automatica (*harvesting*) e la catalogazione dei documenti digitali pubblicati nei siti italiani di interesse culturale avverrà tramite la piattaforma [Archive-it](#) di Internet Archive.

L'*harvesting* partirà da un *seed*, da intendersi come l'URL corrispondente al livello di dominio più utile per la raccolta dei documenti giudicati di interesse in base all'appartenenza alle categorie di cui sopra, ad es.:

<http://www.bv.ipzs.it/>

oppure

<http://www.istat.it/it/prodotti/pubblicazioni>

Per consentire l'*harvesting*, i siti devono prevedere nel proprio [protocollo di esclusione robot](#) un'apposita eccezione per il bot di Archive-it: **archive.org_bot**

L'*harvesting* di siti o sezioni di sito ad accesso limitato è possibile se vengono fornite alla BNCF le credenziali, ma in ogni caso l'*harvesting* non può operare se il sito fa uso di [CAPTCHA](#).

BNCF

Posta elettronica: Ufficio di Direzione: bnc-fi@beniculturali.it - Posta certificata: mbac-bnc-fi@mailcert.beniculturali.it

Per i siti che contengono teche digitali bisognerà verificare le modalità più adeguate di trattamento.

I file [WARC](#), risultanti dall' *harvesting* e contenenti eventuali metadati aggiunti in fase di catalogazione, saranno archiviati in Magazzini Digitali.

La documentazione archiviata sarà visibile agli utenti della BNCF, dalla rete interna della biblioteca, in apposite collezioni predisposte su Archive-it.

In fase di compilazione del modulo di adesione è possibile rilasciare anche il consenso per l'accesso pubblico (da parte di qualsiasi utente del Web) alla raccolta.

Si prevede, inoltre, in una fase successiva, di rendere visibili le descrizioni dei documenti raccolti nel [catalogo di BNCF](#).

Contatti

Per qualsiasi informazione è possibile contattare:

Chiara Storti

BNCF - Servizi informatici

tel. 055 24919219

chiara.storti@beniculturali.it

IL DIRETTORE
Dott. Luca Bellingeri

BNCF

Posta elettronica: Ufficio di Direzione: bnc-fi@beniculturali.it - Posta certificata: mbac-bnc-fi@mailcert.beniculturali.it