



OCTOPUS System. (V.2014DC)

Questo documento descrive l'architettura del sistema "OCTOPUS".

OCTOPUS è un'applicazione progettata come sistema centralizzato per la gestione e conservazione di collezioni di documenti digitali multiformato prodotti e gestiti autonomamente da più istituzioni.

Architettura del sistema.

L'architettura del sistema è basata su tre componenti:

1. L'archivio dei dati (di seguito "*Repository*")
 - a. Provvede allo "storage" dei documenti
 - i. descrizione bibliografica (formato testuale)
 - ii. corpo del documento multiformato (MIME)
2. L'interfaccia applicativa "amministratore – utente" (applicazione Octopus)
 - a. Provvede mediante interfaccia Web-Based a rendere disponibili tutti i servizi previsti dal sistema sia per l'utente finale che per l'amministratore.
 - i. Inserimento documenti
 - ii. Amministrazione e controllo (a livello di singola istituzione) degli inserimenti
 - iii. Ricerca libera (su scelte dell'utente finale)
 - iv. Ricerca basata su liste prestabilite
 - v. Modifica/correzione di documenti
 - vi. Gestione dei diritti di accesso per ogni singolo documento
 - vii. Libero accesso via Internet a chiunque fino alla "*descrizione bibliografica*"
 - viii. Controllo dei diritti di accesso (*visualizzazione e download*) del "*corpo*" del documento per ogni singolo documento
3. Il data base relazionale (MySQL)
 - a. Provvede all'indicizzazione di vari elementi della "*descrizione bibliografica*" del documento. Velocizza la ricerca e l'accesso di

- un documento o insieme di documenti oltre a semplificare la produzione di liste e raggruppamenti di documenti per:
- i. Anno di creazione
 - ii. Tipologia
 - iii. Autore
 - iv. Istituzione
- b. Mantiene informazioni sulle collezioni ed altre relative al controllo dell'accesso da parte dell'amministratore designato per ogni singola istituzione:
- i. Diritti di accesso degli amministratori
 - ii. Informazioni sulle collezioni
 - iii. Informazioni sulle istituzioni

Il “Repository”

Modello logico del documento.

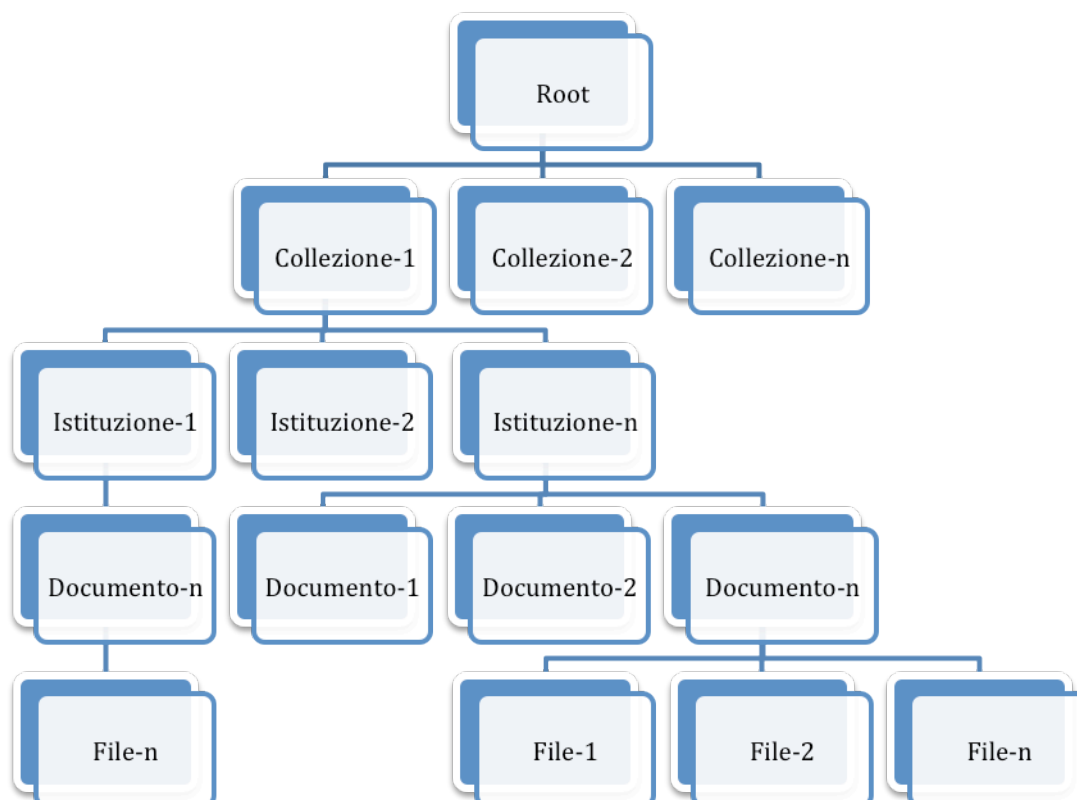
Un elemento essenziale nell'architettura di OCTOPUS è la nozione di documento, un astrazione logica che incorpora i seguenti concetti.

- Ogni documento ha un *unico nome globale* che è definito utilizzando la specifica procedura descritta al punto “*document naming*”.
- Un documento consiste di un numero di componenti. Le due componenti minime essenziali sono la “*descrizione bibliografica*” ed il “*corpo*” del documento.
- La componente bibliografica è disponibile in formato testuale (iso-latin-1), la componente “*corpo*” è disponibile in uno o più *formati* (MIME). Eccezionalmente la componente “*corpo*” può non essere presente.
- La componente “*corpo*” di un dato formato può anche essere strutturata su più *file e sub-directory*. Per esempio il “*corpo*” disponibile in HTML può essere strutturato in pagine e organizzato in modo da contenere sub-directory.

Organizzazione fisica del repository

Fisicamente il “repository” è una porzione di disco con diritti di accesso definiti, non direttamente accessibile dal protocollo “http” e strutturata secondo il modello “File System” dove, a partire da una certa “*Root*”, cartella (mondo Windows) o “*directory*: (mondo UNIX) esiste una *sotto_cartella/directory* per ogni istituzione, collezione e documento ospitato nel sistema.

Tipicamente avremo una struttura del tipo:



Per ogni collezione esiste una sotto_cartella/directory per ogni istituzione definita nell'ambito della collezione.

Ogni singola istituzione ospita una sotto_cartella/directory per ogni documento che gli appartiene.

La sotto-cartella/directory di ogni documento ospita almeno due elementi (file), il primo definito “*descrizione bibliografica*”; il secondo definito “*corpo*” può a sua volta essere un semplice “file” (MIME) o un insieme di file anche strutturati in ulteriori sotto_cartella/directory.

Document naming

Ad ogni documento viene assegnato un codice univoco nel sistema così costruito:

- Codice collezione
- Codice istituzione
- Anno di riferimento (creazione intellettuale del documento)
- Sigla della tipologia del documento
- Numero progressivo

Tipico codice: Codice_Collezione/Codice_Istituzione/anno-sigla-numero

L'interfaccia applicativa “amministratore – utente”

L'interfaccia applicativa è data dal sistema OCTOPUS. L'applicazione realizzata in linguaggio PHP consente mediante la costruzione dinamica di pagine WEB di attivare via “menu” i corrispondenti **servizi** che permettono sia di interagire con il sistema che di produrre pagine dinamiche di risposta ad interrogazioni e/o ad azioni richieste dall'utente utilizzatore del sistema.

L'interfaccia è rivolta principalmente tre tipologie di utenti:

- L'utente amministratore (il bibliotecario dell'istituzione registrata)
- L'utente fornitore/proprietario dei dati (personale di istituzioni registrate)
- L'utente comune/finale (o semplice navigatore internet)

I servizi

I servizi base del sistema sono organizzati in moduli che intergiscono tra loro per lo svolgimento delle funzionalità che il sistema offre.

1. Amministrazione. Servizio riservato agli utenti registrati in qualità di amministratori che consente di svolgere le funzioni amministrative previste per tutte le istituzioni. Ogni singola istituzione può registrare un proprio amministratore; l'accesso al servizio è consentito mediante autenticazione via UserName & Password. Il Servizio consente di:
 - Amministrare i documenti inseriti dagli utenti autorizzati
 - Validazione del documento bibliografico
 - Assegnazione del codice di identificazione
 - Inserimento del documento nel “Repository” e nel data_base
 - Editare un documento esistente in caso di errori rilevati dopo il suo inserimento
 - Rimuovere un documento non più pertinente
2. Inserimento: Servizio riservato agli utenti delle istituzioni registrate. Il servizio e' disponibile dopo autenticazione dell'utente che può essere effettuato in due diverse modalità. Il servizio consente:
 - Di assistere l'utente nella compilazione di un'insieme di form finalizzati alla creazione della “descrizione bibliografica” che caratterizza il documento che l'utente ha intenzione di sottomettere al sistema.
 - Al termine della fase di compilazione e verifica dei dati inseriti, il sistema provvede ad inviare una e-mail di conferma dell'avvenuto invio all'indirizzo di e-mail che l'utente ha inserito nel form stesso e all'amministratore della specifica istituzione perché provveda alla fase di amministrazione.
3. Ricerca libera: Servizio disponibile a qualunque utente internet. Il risultato della ricerca si concretizza con la visualizzazione della “descrizione bibliografica” che consente di visualizzare/scaricare il relativo “corpo” del documento in base ai diritti/modalità di accesso attribuite al documento stesso.

4. Ricerca basata su liste pre-definite: Servizio disponibile a qualunque utente internet. Il risultato della ricerca si concretizza con la visualizzazione della “descrizione bibliografica” che consente di visualizzare/scaricare il relativo “corpo” del documento in base ai diritti/modalità di accesso attribuite al documento stesso.
5. Autenticazione: Servizio per la gestione dei diritti/modalità di accesso ad un servizio (vedi inserimento) o di un singolo documento da parte di un generico utente. L’autenticazione può essere effettuata in due diverse modalità:
 - a. Autenticazione utente: il sistema interroga un “server LDAP” esterno, dedicato alla specifica istituzione che provvede al relativo riconoscimento dell’utente
 - b. Autenticazione indirizzo IP: il sistema rileva l’indirizzo IP della macchina utente che sta interagendo con un determinato servizio e lo verifica con la tabella degli indirizzi IP abilitati all’uso di quel servizio
6. Accesso al “corpo” del documento: Il servizio di accesso, ovvero di visualizzazione/download del “corpo” di un documento verifica che ogni richiesta sia conforme alla politica di accesso stabilita per tutte le istituzioni e al diritto di accesso stabilito per il singolo documento. Allo stato attuale sono previste tre tipologie di diritto di accesso:
 - a. Diritto di accesso:
 - i. Libero – il “corpo” del documento è reso disponibile a tutti
 - ii. Riservato - il “corpo” del documento è reso disponibile solo agli utenti autenticati via LDAP o via indirizzo IP
 - iii. Protetto - il “corpo” del documento è reso disponibile solo a utenti autenticati via LDAP ma che risultano anche essere autore/co-autore del documento.

Il database MySql

Il data base relazionale (MySql) è utilizzato per la creazione di indici su diversi campi che compongono la “descrizione bibliografica” del documento che consentano di velocizzare le operazioni di ricerca e di lista. Inoltre mantiene le informazioni relative alle collezioni, alle istituzioni e ai relativi amministratori “username e password”, oltre che a informazioni per l’assegnazione assistita del “codice documento”.

Uno specifico documento descrive lo schema del data base e delle singole tabelle che lo compongono.

Futuri sviluppi

Il sistema, ha tutte le caratteristiche per poter essere reso operativo in modalità distribuita e federata, in particolare per quanto riguarda i servizi di “inserimento” e di “amministrazione”, permettendo tuttavia di estendere l’interrogazione e la ricerca sull’intera federazione mediante il modulo aggiuntivo (in fase di sviluppo) “OCTOPUS Master Meta Server” che avrà la funzione di rendere trasparente all’utente finale la distribuzione fisica del sistema.