

1. Introduzione

Per quanto ricca e completa possa essere una collezione d'erbario, il suo potenziale contributo scientifico e la sua efficienza saranno notevolmente ridotti se non si dispone di un sistema di catalogazione e se l'accesso alle informazioni è limitato. Per chi è disposto ad investire il proprio tempo, lo sviluppo di una banca dati ("*database*" o base di dati) dei *taxa* e/o delle collezioni può aiutare a risolvere molti problemi, come la compilazione di liste o l'allestimento di pubblicazioni e monografie. Inoltre, se appropriatamente sviluppato, un sistema informativo può aiutare a focalizzare e gestire l'attività di ricerca di un erbario (FILER, 2001)

Negli ultimi 20 anni i curatori di erbari, orti botanici, musei di storia naturale hanno iniziato ad allestire banche dati sulle loro osservazioni e collezioni. Se all'inizio le basi di dati furono impiegate solo come strumento di gestione delle collezioni, negli ultimi anni è sorta la consapevolezza che questi strumenti possono essere visti come una fonte di informazioni di enorme valore per la ricerca sulla biodiversità.

La gestione di una generica collezione di campioni biologici è una attività di carattere museale. Una volta che sono raggiunte le condizioni di conservazione appropriate, si può dire che il materiale si documenta da solo: l'informazione sull'accessione è conservata con il materiale, è l'accesso all'informazione è subordinato all'accesso al materiale. Lo scopo di un sistema di gestione è, da un lato facilitare la localizzazione del materiale, dall'altro svincolare almeno una parte dell'informazione dal materiale stesso (BERENDSOHN & NIMIS, 2000).

L'informazione può essere codificata e gestita in molti modi e molte forme; nei sistemi informativi essa prende la forma di dati. I dati sono generalmente molto più stabili del sistema che li gestisce: il sistema può cambiare nel tempo, ad esempio si può passare da uno schedario cartaceo ad un *database* informatizzato, ma i dati devono necessariamente essere conservati.

2. Vantaggi nell'utilizzo di una banca dati

Prima della diffusione dei computer, ed in parte ancora oggi, le informazioni relative alle collezioni erano conservate in schedari cartacei: l'accesso a queste era legato all'accesso fisico all'archivio. Ciò ha comportato un notevole svantaggio nella consultazione e nell'utilizzo di una gran mole di dati. Con l'aumentare in dimensioni delle collezioni, il processo di ricerca si faceva lungo e difficoltoso e in qualche caso poteva divenire complesso anche ricordarsi dove cercare.

I primi mezzi di schedatura computerizzata si affidavano a singoli archivi elettronici (*file*) con una precisa struttura. Questi venivano usati semplicemente per contenere i dati relativi allo schedario cartaceo senza alcuna integrazione fra le informazioni di varia natura relative ai campioni biologici. Nella sostanza le operazioni da svolgere rimanevano le medesime: anche in questo caso bisognava sapere dove erano memorizzati i *file* e come rintracciare i dati in essi contenuti. Per fare ciò venivano impiegati programmi dedicati sviluppati in maniera "artigianale", generalmente compilati "ad hoc" e adattati di volta in volta alle richieste degli utenti.

L'impossibilità di integrare le informazioni e la mancanza di modelli di riferimento per la costituzione di una banca dati fece sorgere l'esigenza di avere un sistema di conservazione dei dati nuovo ed indipendente dalla natura dei dati stessi e nuovi metodi per lo sviluppo di basi di dati.

Si giunge così al modello attuale di sistema informativo in cui viene distinta una parte passiva dove sono contenuti i dati (il *database* vero e proprio in senso stretto) ed una parte attiva che li gestisce ed integra le informazioni di varia natura (il "*Database Management System*", DBMS).

L'utilizzo dei dati attraverso un tale sistema permette di controllare la ridondanza dell'informazione, ottenere un'integrazione fra i dati, condividerli e tenere sotto controllo gli aspetti legati alla sicurezza.

La gestione di una collezione biologica attraverso una base di dati comporta inoltre numerosi vantaggi pratici, legati ad esempio alla rapidità con cui è possibile effettuare le ricerche e compilare liste ordinate per uno o più criteri, oppure alla possibilità di integrare l'informazione con dati provenienti da altre fonti. Vedremo in seguito come l'informazione contenuta in una collezione biologica sia un importantissimo strumento nello studio della biodiversità. Inoltre, la condivisione delle informazioni attraverso Internet rappresenta una ulteriore possibilità di sviluppo, permettendo il confronto fra le diverse realtà scientifiche, fatto che di per sé è un bene irrinunciabile.

In realtà, accanto agli innegabili vantaggi, esistono anche delle difficoltà di cui bisogna necessariamente tenere conto: innanzitutto la quantità di dati da inserire nei *database* è immensa, e comporta un notevole dispendio di tempo ed energie. La precisione necessaria per l'inserimento dei dati è significativamente maggiore rispetto ai comuni schedari e la tentazione di aggiungere informazioni, giustificata dalla rapidità di inserimento e di adattamento della banca dati alle proprie esigenze, si trasforma in un'arma a doppio taglio quando i dati devono essere effettivamente immessi nel sistema. Proprio per questo alcuni ricercatori affermano che il computer aumenta la mole di lavoro, anziché ridurla (LUNOVA ET AL., 2001). A nostro avviso, però, queste ultime affermazioni sono eccessive perché l'utilizzo razionale e consapevole di una banca dati informatizzata rappresenta comunque, per una collezione biologica, un grande, spesso irrinunciabile, vantaggio.

3. Struttura ed organizzazione di una banca dati

3.1 Concetti generali

Come precedentemente riportato, un sistema informativo è composto da due parti: *database* e *Data Base Management System* (DBMS). Il *database* è definito come un insieme di dati integrati e immagazzinati su un supporto di memoria di massa, che possono essere gestiti da vari programmi applicativi: è una collezione strutturata di informazione quale rappresentazione di un oggetto reale. I dati che compongono un *database* devono possedere una struttura tale da poter essere gestiti da un sistema informatizzato. Il *Data Base Management System* (DBMS) è invece il sistema *software* per la gestione delle basi di dati, attraverso cui si realizza l'aggiornamento, manutenzione, e consultazione dell'informazione: si tratta in sostanza della parte attiva di un sistema in cui il *database* è la parte passiva (SIGNORE, 1995; PAVOLETTI, 2001).

Il modo con cui i dati sono fisicamente archiviati nella memoria del computer ha un'importanza relativa per l'utente finale. Il DBMS, infatti, utilizza un modello astratto dei dati per rappresentare le informazioni archiviate in un modo comprensibile all'utente che assicura una rapida manipolazione e recupero dei dati senza dover ricordare dove essi siano fisicamente immagazzinati.

Negli ultimi anni sono stati proposti diversi modelli di banche dati, ma il *modello relazionale* è quello attualmente più diffuso. Questo modello presenta all'utente le informazioni in una forma familiare di tabelle con righe e colonne. Ogni riga rappresenta un oggetto, ogni colonna rappresenta un attributo.

BOX-SCHEDA – DATABASE RELAZIONALI

Una **base di dati relazionale** è definita come un insieme di dati strutturati e permanenti raggruppati per insiemi omogenei in relazione fra loro, organizzati con la minima ridondanza per essere usati da applicazioni in modo controllato (SIGNORE, 1995). Questo modello è stato ideato da F. CODD nel 1970, ed è stato oggetto negli anni 70 di un intenso lavoro di sviluppo culminato (con lo standard ANSI X3.135-1989) nell'implementazione del linguaggio **SQL** (*Structured Query Language*), linguaggio deputato alla interrogazione ed alla gestione della base di dati, nel 1989.

L'American National Standard Institute (ANSI) ha messo a punto una proposta denominata standard **ANSI-SPARC** (Figura 1) secondo la quale un *database* deve essere articolato in tre livelli: interno, logico, esterno. Il **livello interno** (o fisico) è legato al modo con cui i dati sono fisicamente memorizzati su supporto magnetico; il **livello esterno** (o vista logica) descrive il modo con cui i dati vengono visti dall'utente, mentre il **livello logico** è la rappresentazione dell'intera struttura del *database* secondo un modello ad alto livello, indipendente dal DBMS impiegato. Il passaggio da un livello all'altro, detto *mapping*, si sviluppa in modo da garantire l'indipendenza logica fra i livelli grazie alla quale i programmi applicativi non vengono influenzati dalle modifiche apportate allo schema logico e tantomeno a quello fisico: gli utenti ed i programmi applicativi possono ignorare – grazie a tale indipendenza - i dettagli della costruzione del *database*. Ancora, l'indipendenza logica permette di interagire con il livello esterno senza modificare quello logico (aggiungere o modificare le *viste* esterne senza toccare lo schema logico), mentre l'indipendenza fisica consente di agire sul sistema di memorizzazione senza intaccare le relazioni dello schema logico ed esterno e viceversa.

Lo sviluppo di una banca dati passa attraverso 4 fasi:

Analisi dei requisiti, in cui vengono esaminate le esigenze dell'utente acquisendo una conoscenza concreta (le entità che vanno memorizzate, ciascuna con le diverse proprietà e le eventuali relazioni presenti tra le entità), una conoscenza astratta (la struttura della precedente conoscenza con le eventuali restrizioni sui suoi valori possibili e sui modi in cui questi cambiano nel tempo, cioè vincoli d'integrità), ed una procedurale (tutte le operazioni basilari possibili che possono riguardare le varie entità o più entità contemporaneamente. Tali operazioni potranno poi essere combinate per ottenere risultati più complessi).

Nel caso di un erbario, tra le conoscenze concrete avremo i dati fisici sul campione d'erbario e sulla sua conservazione, i dati sull'evento di raccolta, e naturalmente i dati nomenclaturali. Tra le conoscenze astratte vorremo che l'altitudine non sarà al di sotto dei -10 m s.l.m. ed oltre gli 8000 m slm. Tra le conoscenze procedurali avremo invece le richieste specifiche degli utenti, come ad esempio: data una località, elencare tutti i campioni che da essa provengono, e che sono stati raccolti prima di una certa data.

Progettazione concettuale, si occupa della rappresentazione ad alto livello della realtà di interesse: è una rappresentazione indipendente dal sistema di catalogazione, che potrebbe anche non essere uno strumento elettronico, ed è legata alla struttura stessa della realtà che si vuole affrontare. Il modello concettuale è da considerarsi relativamente stabile nel tempo, a meno che non sia la realtà stessa a cambiare. Il risultato della progettazione concettuale è detto *Information Model*: attraverso un processo di astrazione, le entità simili (o istanze) presenti nella realtà in esame vengono identificate come oggetti e raggruppate in classi; le loro caratteristiche vengono identificate come proprietà, mentre tra le varie classi vengono stabilite le diverse associazioni.

Progettazione logica, che porta alla creazione di uno schema logico, ossia la descrizione della banca dati in un linguaggio comprensibile al DBMS. È a questo punto che entra in gioco la struttura del software. Il modello logico è il linguaggio che è impiegato: i più comuni sono: quello relazionale, quello gerarchico e quello a rete.

Progettazione fisica, che porta a definire le strategie di memorizzazione dei dati e le modalità di accesso: questo aspetto è intimamente legato al tipo di DBMS impiegato.

3.2 Organizzazione di un sistema informativo per un erbario

Un sistema informativo sviluppato per un erbario rientra nella grande categoria delle banche dati realizzate per le collezioni biologiche. Anzi, dal punto di vista logico, per esempio, non vi è differenza alcuna fra un sistema per un orto botanico e per un erbario: ambedue gestiscono i dati relativi ad organismi biologici secondo vari livelli di integrazione.

L'informazione relativa ad un campione biologico è sostanzialmente di due tipi: una legata al campione stesso, l'altra attinente alla categoria tassonomica a cui il campione è ascritto. Nel primo

caso i dati variano da campione a campione, nel secondo i dati variano solamente quando si cambia la categoria tassonomica. Nella maggior parte dei casi, il livello tassonomico di riferimento è quello della specie. Quindi, riassumendo, il campione biologico possiede un insieme di attributi propri ed un altro insieme condiviso con tutti i campioni che appartengono alla medesimo *taxon*.

Appare conveniente e logico suddividere il *database* di una collezione biologica in due sottoinsiemi: quello relativo ai singoli campioni e quello relativo alle informazioni sulle specie. L'integrazione fra questi due sottoinsiemi avviene attraverso il *Database Management System*.

Per quanto riguarda la parte relativa ai singoli campioni, nel un *database* destinato alla gestione di una collezione biologica l'informazione sarà suddivisa in sei grandi aree tematiche (OLSVIG-WHITTAKER & BERENDSOHN, 2000) che vengono di seguito riportate.

- 1) Evento di raccolta
- 2) Sito di raccolta
- 3) Determinazione tassonomica
- 4) Informazioni sul campione
- 5) Posizione dei campioni nella collezione
- 6) Gestione dei campioni

Ognuna di queste aree può essere sviluppata in vario modo, e con vari livelli di approfondimento, purché si conservi la centralità dell'unità o campione (Figura 2). Le collezioni biologiche hanno esigenze di archiviazione molto differenti in base al tipo di oggetti di cui si occupano: tenendo conto che possiamo spaziare da virus a fossili di dinosauri, le categorie di informazioni sopra descritte rappresentano l'informazione minima necessaria per la gestione di una collezione, espressa nel modo più generale.

Per quanto riguarda la gestione di un erbario, la banca dati conterrà innanzitutto quelle informazioni che vengono di solito riportate sui cartellini identificativi dei campioni: ricordiamo quanto detto ad inizio capitolo, cioè che in una collezione biologica il campione si documenta da solo. In realtà le potenzialità di uno strumento informatizzato consentono di immagazzinare molti più dati dello stretto necessario, e soprattutto di integrarle con molte altre fonti di informazione, a cominciare dai riferimenti bibliografici, per arrivare ad interfacciarsi con i sistemi informativi geografici, o con i dati sulla coltivazione delle stesse specie negli Orti Botanici.

Per quanto riguarda la parte relativa al livello tassonomico di riferimento, nel *database* verranno distinte sei sezioni, che vengono di seguito riportate.

- 1) Posizione tassonomica
- 2) Ecologia
- 3) Corologia
- 4) Caratteristiche morfologiche e funzionali
- 5) Caratteristiche cenologiche
- 6) Conservazione

La posizione tassonomica individua il *taxon* sulla base della sua appartenenza alle differenti categorie tassonomiche di ordine gerarchico superiore. Le restanti sezioni descrivono il comportamento ecologico, fitogeografico e cenologico del *taxon*, le sue peculiarità morfologico-funzionali e quanto esso sia soggetto al rischio di scomparsa. Le informazioni relative a queste parti possono essere codificate e descritte in vari modi come proposto da differenti autori (ad esempio ELLENBERG, 1974; LANDOLT, 1977; PIGNATTI, 1982; CHEVILOTTE, 1998).

4. Sistemi disponibili

Sono disponibili in commercio numerosi prodotti *software* per la gestione di banche dati: alcuni sistemi sono stati sviluppati appositamente per le collezioni biologiche, ma la maggior parte è costituita da sistemi generici per lo sviluppo di basi di dati.

4.1 Sistemi non dedicati

Tra questi, le soluzioni in commercio vanno naturalmente dai prodotti più semplici di utilizzo personale o per pochi utenti, o applicazioni *desktop*, ai più complessi *database* utilizzabili in rete, o applicazioni *server*.

I *database* da *desktop* offrono una soluzione economica e semplice ai problemi meno complessi di gestione e manipolazione dei dati. Il loro nome deriva dal fatto di essere stati progettati per funzionare su personal computer, *desktop* appunto. Si tratta di prodotti molto comuni, come Microsoft Access e FileMaker Pro, che presentano numerosi vantaggi pratici: innanzitutto il loro costo è in genere contenuto, e l'utilizzo è piuttosto semplice, basato su un'interfaccia grafica comprensibile, che non rende necessaria una conoscenza profonda del linguaggio SQL, che svolge la funzione di gestione e di interrogazione del *database*.

La maggior parte di questi sistemi offre inoltre la possibilità di pubblicare i propri risultati direttamente su Internet.

I prodotti principali sono di seguito elencati.

MICROSOFT ACCESS è forse il software più comune per banche dati da *desktop*: è basato sulla familiare interfaccia Microsoft ed è corredato da una guida in linea ben strutturata. Uno dei suoi punti di forza è l'integrazione pressoché completa con gli altri prodotti del pacchetto Office e la rapidità di sviluppo delle soluzioni più semplici. Infine, può essere messo in relazioni con applicazioni server di cui fornisce l'interfaccia utente (Figura 3).

FILEMAKER PRO è estremamente popolare fra gli utenti Macintosh, ma sta rapidamente diffondendosi anche fra gli utenti PC. Offre un'interfaccia grafica intuitiva, nascondendo molti degli aspetti più complessi di un DBMS, pur essendo compatibile con applicazioni server complesse. L'ultima versione in commercio è la 8.

4DIMENSION è un software multiplatforma per banche dati e sviluppo di applicazioni indipendenti. flessibile ed orientato alla condivisione in rete dei progetti. Anche esso offre una facilità d'uso.

LOTUS APPROACH - Lotus offre il DBMS *Approach* come parte del pacchetto software SmartSuite, contenente varie applicazioni, ed è progettato in competizione con il pacchetto Office.

VISUAL FOXPRO 9.0 della Microsoft non è propriamente un *database*, ma piuttosto un ambiente di sviluppo per DBMS. Richiede pertanto una certa abilità di programmazione, e non si presenta certo come un prodotto per principianti: tuttavia offre una soluzione molto flessibile ed economica che permette di creare prodotti altamente personalizzati.

Le applicazioni *desktop*, accanto agli innegabili vantaggi di praticità d'impiego e basso costo, risultano tuttavia inadeguate quando le esigenze della banca dati si fanno più complesse. Nonostante quanto venga in genere pubblicizzato, è molto raro infatti che una di queste applicazioni sia davvero in grado di sopportare modifiche contemporanee da parte di più di un utente. In linea generale, se una banca dati deve essere usata da più di una persona, è il caso di rivolgersi ad applicazioni di tipo *server*, cioè un software basato sulla presenza della struttura informativa in unità centrale di gestione a cui accedono i vari utenti attraverso collegamenti di rete da terminali periferici. Allo stesso modo, conviene scegliere questa soluzione se le esigenze di sicurezza aumentano o se la banca dati deve essere impiegata attraverso Internet.

Le applicazioni *server* offrono l'opportunità di gestire grandi quantità di dati in maniera rapida, efficiente e soprattutto in un ambiente multiutente, anche se presentano un costo piuttosto elevato.

I vantaggi principali vengono dalla loro potenza, flessibilità, e dalla capacità di gestire virtualmente ogni tipo e quantità di dati. Per i programmatori sono disponibili interfacce utente che permettono di creare applicazioni in maniera relativamente semplice, pur essendo necessarie competenze altamente specializzate. Il prodotto finale, se non basato su un'applicazione *desktop*, è spesso di utilizzo complesso anche per gli utenti.

I prodotti che dominano il mercato sono Oracle e Microsoft SQL Server; attualmente, con il crescente sviluppo di applicazioni web, si stanno diffondendo i sistemi "open source", come MySQL e PostgreSQL, versatili e complessi almeno quanto i prodotti commerciali, presentano l'indiscutibile vantaggio di essere gratuiti. Sono particolarmente indicati per lo sviluppo di basi di dati accessibili via internet, e molti siti web sono oggi sviluppati **facendo uso di una soluzione comunemente denominata LAMP: computer con sistema operativo Linux, server http di Apache con il linguaggio PHP e base di dati gestita da MySQL.**

4.2 Sistemi dedicati

Molte istituzioni scientifiche hanno sviluppato le proprie applicazioni per la gestione delle collezioni biologiche. Si tratta di *database* più o meno complessi e basati su vari sistemi, ma che presentano dei vantaggi derivanti dal fatto di essere prodotti dedicati.

Essendo stato progettati *ad hoc*, infatti, questi prodotti cercano di coprire tutte le esigenze che i gestori di collezioni biologiche possono incontrare.

HORTI Sviluppato per la gestione degli Orti Botanici italiani, questo sistema si presta anche alla catalogazione di collezioni di erbario; la struttura relazionale è articolata in numerose tabelle, come sempre per ridurre la ridondanza. La coerenza dell'informazione tassonomica è garantita dal riferimento costante ad un server centrale che fornisce le liste di taxa e convalida le proposte di aggiunta.

BG-BASE – Applicazione per PC sviluppata per gestire informazioni relative ad ogni tipo di collezione biologica. L'obiettivo è sviluppare uno standard che assicuri compatibilità fra le diverse istituzioni, ma che al tempo stesso sia modificabile ed adattabile alle esigenze degli utenti. Il sistema attualmente in commercio è uno strumento potente che facilita le operazioni di inventario e permette di documentare, schedare e curare le collezioni, che in questo modo acquistano anche un grande valore per scopi di ricerca, conservazione ed educazione. BG-BASE è compatibile con i principali standard internazionali, incluso ITF. I principali centri di sviluppo sono il Giardino Botanico di Edimburgo e l'Holden Arboretum. Per la gestione dell'informazione, essa viene suddivisa in 6 grandi categorie: gestione delle collezioni, tassonomia e nomenclatura, distribuzioni, bibliografia, conservazione, risorse umane. Vi è stato affiancato di recente un modulo per la realizzazione di mappe degli orti botanici, BG-MAP.

<http://www.rbge.org.uk/BG-BASE//welcome.htm>

BRAHMS – è un sistema di *database* per botanici, curatori di erbari, e chiunque lavori con dati botanici. E' utilizzato per mettere insieme e gestire dati botanici e creare vari prodotti, da semplici elenchi, moduli di prestito, grafici e cartine a checklist complesse, inventari di biodiversità, pubblicazioni tassonomiche. Il progetto Brahms è iniziato nel 1985 presso il Dipartimento di Scienze Botaniche dell'Università di Oxford, e oggi è impiegato in numerosi erbari d'Europa, Africa, Asia e Americhe. L'ultima versione del software include moduli per la gestione di immagini, mappe, campioni e sequenze di DNA, indirizzi web e dati bibliografici.

<http://www.brahms.co.uk>

FLORIN – è un sistema informativo progettato per gestire un ampio spettro di dati sulle piante: tassonomia e nomenclatura; distribuzione geografica, erbari, collezioni viventi, informazioni dettagliate sulle piante raccolte in natura, dati bibliografici. E' strutturato in forma di *database* relazionale e varie applicazioni satellite come interfaccia GIS, generatore di etichette, browser internet, etc.

<http://www.florin.ru/florin/>

HERBAR – *database* in formato MS-Access sviluppato per la gestione dell'erbario del Giardino Botanico di Madrid. Permette l'immagazzinamento di tutti i dati identificativi degli esemplari, informazioni geografiche, controllo delle revisioni. Gestione dei prestiti, stampa di etichette con e senza codici a barre, integrazione con MS-Word per generare etichette, liste, moduli.

<http://www.rjb.csic.es/herbario/herbar.htm>

BIBMASTER – Il Giardino Botanico di Madrid ha messo a punto un'applicazione per la gestione della nomenclatura e della bibliografia. Permette di realizzare vari tipi di liste bibliografiche, tassonomiche, di campioni d'erbario. Basato sul modello relazionale, permette l'immissione rapida di dati, riconosce gli errori nomenclaturali, permette la gestione degli ibridi, e il controllo delle abbreviazioni dei nomi di piante e d'autori.

<http://www.rjb.csic.es/bibmaste/bibcaract.htm>

TRACY – in uso presso l'erbario del Texas, è stato creato perché l'utilizzo di Internet ha reso possibile un grande scambio di dati tra gli erbari, ma evidentemente ciò richiede che i dati siano conservati in una qualche forma digitale. TRACY è l'interfaccia di un *database* relazionale sviluppato in forma di tabelle Paradox, e accessibili attraverso il motore Borland: ciò permette sia l'accesso diretto alle tabelle, sia l'accesso via SQL e la creazione di query e di report. In questo modo il programma può essere messo in comunicazione con qualunque server SQL in commercio con il minimo sforzo.

<http://www.csdl.tamu.edu/FLORA/input/inputsys.html>

University of California Davis Herbarium Management System - Sviluppato in MS-Access, questo sistema assolve 4 funzioni:

Gestisce il traffico in entrata e in uscita dall'erbario (accessioni e prestiti)

Crea le etichette per l'erbario conservando l'informazione

Gestisce la biblioteca dell'erbario

Permette la gestione di liste di personale, liste di distribuzione, etc.

<http://herbarium.ucdavis.edu/herbaccess/databaseinfo.htm>

BIOLINK – è progettato per assistere chi lavora con informazioni basate su campioni e su taxa, principalmente tassonomi, ecologi, curatori di collezioni biologiche e biogeografi. Gestisce le informazioni basate sul taxon, come la nomenclatura, distribuzione, classificazione, ecologia, morfologia, iconografia, letteratura. Le informazioni centrate sul campione includono siti di raccolta, date, localizzazione nella collezione, prestiti, accessioni.

<http://www.ento.csiro.au/biolink/index.html>

4.3 Standard internazionali

Un requisito fondamentale dei dati riguardanti collezioni biologiche è la possibilità di scambio con altre istituzioni. A tale scopo sono stati sviluppati alcuni standard internazionali, sui quali si basa la struttura di gran parte dei sistemi dedicati che abbiamo preso in esame.

ITF (International Transfer Format)

E' uno standard derivante da accordi internazionali attraverso il quale è possibile scambiare informazioni che riguardano piante vive tra istituzioni botaniche, principalmente orti botanici. Rispetto alla prima versione, pubblicata nel 1987, la versione 2.0 permette lo scambio di campi a lunghezza variabile e l'omissione dei campi che non contengono dati. Ogni file inizia con l'identificatore "startfile" e include tutte le informazioni associate ad un'Accessione; ogni accessione può contenere uno o più campioni individuali.

<http://www.bgbm.fu-berlin.de/TDWG/acc/itf2tdwg.htm>

HISPID (Herbarium Information Standards and Protocols for Interchange of Data)

HISPID è un formato standard per lo scambio di informazioni elettroniche sui campioni d'erbario. E' stato sviluppato da un comitato di rappresentanti dei principali erbari d'Australia.. Si tratta di uno standard basato sull'accessione: sebbene molti campi si riferiscono ad attributi dei taxa, essi vanno intesi come applicati al campione rappresentato dal record, non al taxon per se. Lo scambio di informazioni tassonomiche, bibliografiche, nomenclaturali, conservazionistiche, non è contemplato da questo standard a meno che non si riferisca al campione. I campi trattati in questo standard coprono la maggior parte delle attività di erbari e giardini botanici, e sono organizzati in gruppi di informazioni simili.

<http://www.rbgsyd.gov.au/HISCOM> <= il link porta ad una pagina bianca

4.4 Sistemi integrati

Esiste al mondo una quantità immensa di dati sulla biodiversità: questi dati sono stati e sono continuamente raccolti, elaborati e conservati nei formati più diversi. Si calcola che le collezioni biologiche contengano nel loro insieme circa 3×10^9 esemplari di esseri viventi; di questi, circa il 5-10% è schedato in formato elettronico. Si tratta quindi di una gran mole di dati sulla biodiversità, che sono tuttavia dispersi, eterogenei, privi di integrazione (SOBERON, 1999).

In sostanza, il mondo scientifico possiede dei dati sulla biodiversità ma manca di un vero sistema informativo globale. Così, accanto ai sistemi per la gestione di erbari singoli, sono state avanzate varie proposte al fine di unificare i sistemi informativi in un'unica rete, attraverso cui sia garantita la comunicazione tra le istituzioni scientifiche, e venga assicurata la circolazione dell'informazione.

Disciplinare l'informazione in un unico sistema di accesso ha infatti molteplici vantaggi: la combinazione di informazioni provenienti da aree tematiche diverse favorisce la comprensione dei fenomeni e le scoperte scientifiche; un'interfaccia comune può presentare agli utenti un ampio spettro di argomenti conosciuti e non, ed inoltre la spinta verso un sistema unico rappresenta un incentivo per gli sforzi per il miglioramento delle procedure e la loro standardizzazione sarà limitata la duplicazione degli sforzi.

Un progetto diretto a tal fine è stato messo a punto in Australia Occidentale, dove i *database* di gestione degli erbari sono collegati fra loro in una rete (FloraBase) sotto la supervisione dell'Erbario di Perth. Tale rete rappresenta oggi l'informazione più recente sulla flora dello Stato (WESTERN AUSTRALIAN HERBARIUM, 1998).

Stati Uniti e Messico condividono un sistema informativo transnazionale la cui funzione è quella di riunire e integrare l'informazione proveniente dai numerosi *database* sparsi nelle diverse istituzioni scientifiche dei due Paesi. Il prototipo di questo network è detto NABIN (North America Biodiversity Information Network) e si basa sullo standard ANSI/NISO Z39.50, utilizzato per l'interscambio di dati bibliografici e geografici; attraverso questa rete informatica i ricercatori possono consultare contemporaneamente un gran numero di banche dati connesse a Internet.

In Europa il progetto BIOCISE ha come obiettivo finale la creazione di un portale unico per le collezioni biologiche di ogni tipo: istituti di ricerca, musei, etc (BERENDSOHN ET AL. 2000); è stato seguito dal progetto BIOCASE (BERENDSOHN ET AL. 2001), sempre diretto alla creazione di un sistema informativo europeo. Non è uno standard internazionale, quanto piuttosto un portale web realizzato con lo scopo di mettere in comunicazione le collezioni biologiche d'Europa: ogni banca

dati locale dovrebbe restare nella sua forma originaria mentre il sistema Biocase permette di accedere a tutti i dati come se fossero un'unica risorsa.

4.4.1 I sistemi integrati in Italia

Ultimamente si sta sviluppando in Italia un'altra iniziativa che propone a livello nazionale un sistema integrato chiamato anArchive (PANFILI, PESCIAIOLI E VENANZONI, 2004). Tale iniziativa, coordinato dall'Università di Perugia, raccoglie in parte le esperienze maturate nel progetto ARCVeG (BURBA et al. 1992) della Società Italiana di Fitosociologia per un sistema informativo della vegetazione italiana. Il progetto ARCVeG, sviluppatosi all'inizio degli anni '90, proponeva alla comunità scientifica dei botanici italiani un software di pubblico dominio con cui poter immagazzinare, integrare ed elaborare le informazioni relative ai rilievi fitosociologici. L'utilizzo da parte di molti ricercatori di questo strumento informatico avrebbe permesso di costituire un cospicuo nucleo di informazioni sulla vegetazione italiana, uniformate dal punto di vista della loro struttura e accessibile a quanti volessero farne uso. Purtroppo, l'idea si rivelò prematura e, dopo alcuni tentativi iniziali con dei workshop presso l'università di Ancona, non venne successivamente sviluppata, restando solamente un brillante e solitario tentativo di proposta di sistema integrato per la vegetazione per l'Italia.

Il progetto anArchive [VEDI BOX ANARCHIVE] si sviluppa sul solco di ARCVeG, alla luce delle acquisizioni tecnologiche degli ultimi dieci anni: la diffusione capillare della connettività di rete che permette di mantenere le informazioni su un server centrale lavorando in remoto e la disponibilità in rete di prodotti e linguaggi accessibili a qualsiasi utente senza alcun costo (soluzioni open source o comunque a costo zero). Questo sistema permette di integrare le informazioni sulla flora e vegetazione italiana in una unica struttura informativa in grado di gestire le collezioni di erbario, i dati bibliografici e i rilievi fitosociologici associati anche alle relative informazioni geografiche. Il database, infatti, è accessibile via internet e consente alle istituzioni che ne facciano richiesta di inserire e consultare i dati relativi a collezioni di erbario (Figura 12), studi fitosociologici e cartografia floristica. Il sistema provvisto, infine, di un motore GIS interno che permette anche di visualizzare la distribuzione dei campioni consultati a livello regionale e nazionale.

Il nucleo centrale è articolato in una parte server (dove sono depositati i dati), ed una parte client (che consente la gestione dei dati attraverso una connessione di rete). Il programma server si basa su sistemi operativi, programmi; l'inserimento dei dati avviene attraverso una serie di applicazioni specifiche: *herbArchiver* (Figura 13) permette ad esempio l'inserimento dei dati di erbario mentre la gestione delle check-list tassonomica è affidata a *Linnaeus* (Figura 14). Tutte le applicazioni essendo basate sul linguaggio Java(TM) sono indipendenti dal sistema operativo utilizzato; la consultazione può anche avvenire via web utilizzando esclusivamente tecnologie standard pertanto anch'essa è indipendente dal sistema operativo.

L'interscambio dei dati tra il server e i client avviene tramite un dialetto XML (le cui specifiche sono pubbliche) per avere una maggiore interscambiabilità dei dati indipendentemente dal formato originale.

BOX-ANARCHIVE VENANZONI R. PANFILI E. L'obiettivo del progetto anArchive è quello di fornire uno strumento flessibile per la gestione di dati botanici di diverso tipo tramite la rete internet. Attualmente le sezioni più avanzate riguardano la gestione delle check-list (necessarie per l'inserimento dei dati), dei campioni di erbario e della flora, dei rilievi fitosociologici e la generazione di carte di distribuzione dei dati inseriti. Sin dalla fase di progettazione del sistema si è scelto di utilizzare tecnologie il più possibile aperte, in particolare ogni qual volta l'alternativa open source lo permetteva si è optato per quest'ultima. Tecnicamente il sistema è composto da una parte server che archivia ed elabora i dati e da una serie di soluzioni client che consentono di interagire con il server.

Tutta l'architettura del server poggia su una soluzione di tipo Java Enterprise, disponibile per tutti i sistemi operativi più comuni; le altre due componenti del sistema server sono il container per servlet

Tomcat del progetto Apache e il motore per basi di dati PostgreSQL (vedi testo), entrambe i server sono open source e punti di riferimento nelle rispettive aree, oltre a questo sono facilmente intercambiabili con prodotti analoghi.

L'accesso ai dati è permesso tramite due distinte modalità una per la consultazione e la richiesta dei prestiti tramite web e l'altra per l'immissione e la modifica dei dati tramite applicazioni client appositamente realizzate.

L'accesso via web è possibile utilizzando qualsiasi browser compatibile con gli standard attuali. I percorsi di accesso ai dati sono stati disegnati in modo tale per cui è possibile rintracciare le informazioni necessarie partendo da punti differenti. È ad esempio possibile trovare un cartellino di erbario, che può essere eventualmente chiesto in prestito, partendo sia da un modulo di ricerca diretto in cui si possono specificare diversi parametri (fig X1) che da una vista gerarchica delle specie presenti nelle check-list del sistema (fig X2). La disponibilità dei dati è comunque bilanciata dalla riservatezza degli stessi: particolari campioni possono essere privati e quindi visibili soltanto da parte di persone autorizzate, in altri casi il dato può essere pubblico a meno della localizzazione esatta del punto di raccolta. Un'altra parte del sistema accessibile via web riguarda la generazione di cartine di distribuzione dei campioni, queste vengono costruite a partire da dati presenti in diverse raccolte all'interno del sistema, anche di tipo eterogeneo come erbari e collezioni di rilievi della vegetazione.

I client per l'immissione dei dati sono specializzati per i diversi compiti: gestione delle check-list, erbari e rilievi della vegetazione, non creano una copia locale delle informazioni ma modificano direttamente quelle presenti nel database centrale: utilizzando questo approccio si rende facile la collaborazione all'interno di gruppi di lavoro formati da più persone che possono così condividere i dati in tempo reale. Le applicazioni client sono dei programmi java e non hanno bisogno di installazione, possono funzionare su qualsiasi sistema operativo che abbia a disposizione una macchina virtuale Java, attualmente sono inclusi: Windows, Linux e Mac OS X.

Linnaeus (fig X3) è il programma che permette di gestire le liste di specie presenti nel server; un utente con diritti di accesso particolari può inserire e modificare specie presenti nella check-list a cui ha accesso, il sistema può avere contemporaneamente più check-list diverse che contengono dati di specie diverse e ogni amministratore di check-list può modificare solamente dati presenti nella sua lista. Ogni specie può avere un numero qualsiasi di sinonimi, eventualmente anche pro-parte e la gestione della sinonimia è automatizzata: è possibile far diventare un sinonimo nome valido portando automaticamente al livello di sinonimo il vecchio nome valido.

HerbArchiver (fig X4) è il gestore dei dati di erbario: ogni utente ha un proprio codice per l'accesso e tramite questo può elaborare i dati di uno o più erbari. L'inserzione del singolo dato è assistita da parte del programma in diverse fasi, l'aiuto più evidente si ha durante l'immissione del nome della specie: una funzione di completamento automatica fornisce dei suggerimenti e se il nome scelto è un sinonimo rende possibile la scelta del nome valido corrispondente. Dalla stessa applicazione è possibile gestire la stampa dei cartellini secondo diversi modelli personalizzati e l'immissione nel sistema di foto collegate ai campioni di erbario.

L'ultima applicazione si occupa dei rilievi fitosociologici: vegArchiver ha delle funzionalità analoghe alla sua controparte per gli erbari come ad esempio l'immissione assistita dei nomi delle specie, che in questo caso vengono inserite all'interno di un numero di strati personalizzabile dall'utente.

Una banca dati nazionale di questo tipo, una volta a regime, sarà uno strumento molto potente per la diffusione delle conoscenze scientifiche: permette ad esempio di mettere a disposizione una gran mole di informazioni sulle collezioni di erbario, con la possibilità di richiedere campioni in prestito. Le collezioni, infatti, rimangono depositate presso le singole istituzioni, che si occuperanno della gestione dei prestiti. Inoltre la condivisione dei dati fitosociologici potrebbe essere un'occasione per riunire le informazioni sulla vegetazione italiana, da sempre piuttosto frammentarie, in un sistema integrato e di libero accesso.

5. Utilizzazione dei dati

Accanto agli impieghi puramente gestionali, la banca dati di una collezione biologica restituisce un gran numero di informazioni i cui utilizzi possono essere molteplici a seconda del tipo di utenti (FELINKS ET AL., 1999).

Possiamo identificare quattro principali gruppi di utilizzatori:

1. Sistematici e tassonomi
2. Utenti del settore ambientale
3. Ricercatori nel campo delle biotecnologie
4. Pubblico

Per quanto riguarda la ricerca scientifica in sistematica e tassonomia, la consultazione di banche dati relative alle collezioni biologiche è di fondamentale importanza. La nomenclatura scientifica, in botanica e zoologia, si basa sulla conservazione dei campioni cui il nome è stato attribuito, e sulla circolazione dei campioni fra le istituzioni scientifiche per confrontarli e per studiarne la variabilità. I dati sulla raccolta degli esemplari servono a restringere il campo di indagine ad una particolare area geografica, oppure a valutare l'attendibilità di una determinazione.

Il sistema informativo di una collezione biologica può aiutare quindi ad entrare in possesso del materiale desiderato con il minore dispendio di tempo ed energie. Allo stesso modo per una prima selezione del materiale da esaminare è utile includere nella banca dati del materiale iconografico.

Nel campo delle biotecnologie, le collezioni biologiche sono utili principalmente nell'offrire il materiale certificato per l'impiego diretto o come strumento di ricerca per analisi chimiche, ma anche per confronto.

Le diverse categorie di utenti necessitano di dati organizzati in varie forme: l'approccio più generale descrive per grandi linee cosa è presente in una regione e in un certo periodo, ed è richiesto principalmente dai settori della didattica e dell'amministrazione. Nel settore ambientale invece verranno richiesti dati con una maggiore integrazione e magari con informazioni statistiche basate su dati precisi di presenza/assenza. L'assenza di informazioni su un taxon o un'area è a sua volta un'informazione, potendo essere d'aiuto nell'indirizzare le ricerche. I dati delle collezioni d'erbario possono essere utilizzati per identificare habitat con alte potenzialità di ospitare flora rara o poco comune. Se integrati con rigorose ricerche sistematiche e di campo, i dati degli erbari sono un utile primo passo nell'identificazione di taxa di interesse conservazionistico, nel produrre dati geografici sulle località ove i campioni sono stati raccolti e sugli habitat ai quali si associano, e nell'identificare le lacune sulla distribuzione e sugli habitat che necessitano di ulteriore investigazione (MACDOUGALL ET AL., 1998).

A questi livelli, l'informazione tassonomica può essere di natura generale, raggruppata anche per categorie sovraspecifiche: è necessario un livello superiore di dettaglio per valutare i cambiamenti nel tempo nella flora di una regione oppure dinamiche di popolazione.

Un livello di dettaglio ancora maggiore è richiesto nelle ricerche autoecologiche, in cui la banca dati può servire per catalogare ogni genere di informazione qualitativa su singoli taxa.

I dati provenienti da una collezione naturalistica possono essere molto importanti per fini conservazionistici, e per una valutazione della biodiversità di una regione: un erbario ad esempio contiene una gran mole di informazioni riguardanti la distribuzione di organismi vegetali nello spazio e nel tempo. Ogni tipo di collezione museale fornisce, in questo senso, una dimensione storica a complemento degli studi sul campo. L'analisi temporale di una collezione di erbario può, con le dovute cautele, servire a documentare l'andamento della biodiversità vegetale di una regione fornendo informazioni sulla presenza delle specie, e di conseguenza sulle dinamiche di espansione o contrazione degli areali. (SCHAFFER ET AL. 1998)

Gli utenti del settore ambientale focalizzano l'attenzione sui dati di tipo geoeologico, temporale, tassonomico: in particolare, nella consultazione di una banca dati, le loro ricerche sono basate sui primi due criteri, mentre il dato tassonomico viene inteso come la connessione per altri livelli informativi. Gli utenti di questo settore necessitano inoltre di dati in forma di inventari floristici al fine di evidenziare, ad esempio, le aree con maggior ricchezza di specie, oppure la presenza di entità di particolare pregio.

Tali inventari rivestono un'importanza cruciale anche per quel che riguarda la divulgazione scientifica e la pubblicizzazione di istanze ambientali.

I sistemi informativi geografici (GIS) possono essere impiegati per esplorare la biodiversità nelle varietà selvatiche di piante coltivate: le collezioni di germoplasma di queste varietà contengono una gran quantità di informazioni sulla distribuzione geografica delle specie, e gli erbari rappresentano una ulteriore fonte di dati. L'analisi dei dati geografici può essere utile per indicare nuove aree di interesse per la raccolta ed approfondire le conoscenze sulla distribuzione della biodiversità intraspecifica ed interspecifica (JONES ET AL., 1997).

Il *database* deve essere articolato in modo da presentare all'utente il dettaglio richiesto, preferibilmente a livello di unità o campione; l'accesso primario è generalmente basato su criteri temporali e geo-ecologici, mentre la tassonomia è utilizzata come anello di congiunzione tra le informazioni. L'impiego dei dati provenienti da collezioni biologiche nello studio della biodiversità è un campo relativamente inesplorato, ed attualmente gli studi sono diretti per lo più a valutare l'attendibilità dei dati stessi, e la coerenza con la realtà che rappresentano (PONDER ET AL., 2001; HAWKES, 1997; STORK ET AL., 1996).

La banca dati di un erbario può rappresentare pertanto il punto di partenza di uno studio sulla biodiversità, fornendo risposte a domande come:

Qual è la distribuzione delle piante in un regione?

Quali sono le aree a maggior ricchezza e varietà di specie?

Quanto accuratamente questi dati spiegano la biodiversità della regione?

In quali aree è necessario intensificare gli sforzi di raccolta?

Questi dati possono essere di supporto per una politica conservazionistica?

Le banche dati integrate hanno lo scopo di riunire le informazioni provenienti da più fonti in un unico sistema informativo, ad esempio a livello europeo come il progetto BIOCASE (<http://www.biocase.org/default.shtml>), tentando di superare uno dei principali problemi della comunità scientifica, cioè la carenza di circolazione delle informazioni. Il valore scientifico dei dati conservati nelle collezioni europee è immenso, ma largamente sottoutilizzato a causa della sua dispersione ed eterogeneità: un progetto integrato di *database* a livello continentale permetterebbe l'accesso alle informazioni da parte di ricercatori di ogni disciplina. Da non sottovalutare è anche la necessità di uniformare i dati dal punto di vista semantico, in modo che tutta l'informazione biologica d'Europa sia espressa nello stesso modo.

RIVISTE

DATA & KNOWLEDGE ENGINEERING, Elsevier Science

BIODIVERSITY AND CONSERVATION, Kluwer Academic Publishers

CONSERVATION BIOLOGY, Blackwell Science Inc

TRENDS IN ECOLOGY AND EVOLUTION Elsevier Science

TAXON International Association for Plant Taxonomy

BIBLIOGRAFIA CITATA

BERENDSOHN W.G. & NIMIS P.L., 2000 The complexity of collection information. Pp. 13-18 in:
BERENDSOHN, W. G. (ed.), Resource Identification for a Biological Collection Information

- Service in Europe (BioCISE). - Botanic Garden and Botanical Museum Berlin-Dahlem, Dept. of Biodiversity Informatics.
- BERENDSOHN, W. G., ANAGNOSTOPOULOS, A., HAGEDORN, G., JAKUPOVIC, J., NIMIS, P. L., VALDÉS, B., GÜNTSCH, A., PANKHURST, R. J. & WHITE, R. J. (1999): A comprehensive reference model for biological collections and surveys. *Taxon* 48: 511-562.
- BERENDSOHN W.G. 2001: ENHSIN & BioCASE - Access to European Collections. 17th annual meeting of the Taxonomic Databases Working Group (TDWG 2001), Sydney November 2001, Abstract Volume
http://plantnet.rbg Syd.gov.au/bioforum/TDWG_program/tdwg_abstracts.html
- CHE VILLOTTE H., 1998 Herbarium collection *database* - Data conceptual model. SIMES – 961620 Système d'Information Multimédia Pour l'Environnement Subsaharien.
- CODD E.F., 1970. A Relational Model of Data for Large Shared Data Banks. *Communications of the ACM*, **13** (6): 377-387
- CONN B.J., 2001. HISPID3 – Herbarium Information Standards and Protocols for Interchange of Data . draft 1.4
- ELLENBERG, H. (1974): Zeigerwerte der Gefässpflanzen Mitteleuropa. *Scripta Geobotanica* 9: 1-97.
- FELINKS BIRGIT, ANDREA HAHN, LINDA OLSVIG-WHITTAKER AND WOUTER LOS Users and uses of biological collections, Pp. 20-32 in: Berendsohn, W. G. (ed.), *Resource Identification for a Biological Collection Information Service in Europe (BioCISE)*. - Botanic Garden and Botanical Museum Berlin-Dahlem, Dept. of Biodiversity Informatics
- GOLDSTEIN R.C., STOREY V.C., 1999. Data abstractions: why and how? *Data & Knowledge Engineering*, 29: 293-311
- HAWKES J.G., 1997. A *database* for wild and cultivated potatoes. *Euphytica* 93, 155-161
- JONES PETER G., STEPHEN E. BEEBE JOE TOHME, NICHOLAS W. GALWEY, 1997. The use of geographical information systems in biodiversity exploration and conservation. *Biodiversity and Conservation* 6, 947-958
- LANDOLT E., 1977 – Okologische Zeigerwerte zur Schweizer Flora- Ber.Geobot.Inst.ETH. 64:64-207, Zurich.
- LUNEVA NATALIA N_, IRENA G. CHUKHINA & ELIZAVETA G. LEBEDEVA, 2000 An information retrieval system for VIR's herbarium, St. Petersburg, Russia, as a tool for cultivated plant research. *Genetic Resources and Crop Evolution* 47: 147–152.
- MACDOUGALL A.S., J.A. LOO, S.R. CLAYDEN, J.G. GOLTZ, H.R. 1998 Defining conservation priorities for plant taxa in southeastern New Brunswick, Canada using herbarium records. *Biological Conservation* 86 325±338
- OLSVIG-WHITTAKER L. & BERENDSOHN W.G., 2000 Computerizing and networking biological collection data , Pp. 5-12 in: Berendsohn, W. G. (ed.), *Resource Identification for a Biological Collection Information Service in Europe (BioCISE)*. - Botanic Garden and Botanical Museum Berlin-Dahlem, Dept. of Biodiversity Informatics.
- PANFILI E., PESCIAIOLI P., VENANZONI R., 2004 - AnArchive: un software per la gestione di dati d'erbario, floristici e vegetazionali direttamente da internet. *Inform. Bot. Ital.*, 36(1):110-112.
- PIGNATTI S. 1982 - *Flora d'Italia*. Edagricola, Bologna.
- PONDER W.F., CARTER G.A., FLEMONS P., CHAPMAN R.R., 2001. Evaluation of Museum Collection Data for Use in Biodiversity Assessment. *Conservation Biology* 15 (3), 648-657
- SHAFFER H.B., FISHER R.N., DAVIDSON C., 1998. The role of natural history collections in documenting species declines. *TREE* 13, 27-30
- SIGNORE O., 1995. Calcolatori e cataloghi. Ricerche e documenti per il Catalogo dei Beni Culturali, 2. Regione Toscana – Giunta Regionale
- SOBERON J., 1999. Linking biodiversity information sources. *TREE* 14, 291
- TUTIN T.G., HEYWOOD V.H., MOORE D.M. VALENTINE D.H., WALTERS S.M., WEBB. D.A. 1976 - *Flora Europaea*. Cambridge University Press.

WYSE JACKSON D., 1997. International Transfer Format for Botanic Garden Plant Records ITF2.
Botanic Gardens Conservation International

ISTITUZIONI E PROGETTI EUROPEI

BGCI BOTANIC GARDEN CONSERVATION INTERNATIONAL: <http://www.bgci.org.uk/>

INDEX HERBARIORUM: <http://www.nybg.org/bsci/ih/>

IUCN The World Conservation Union: <http://www.iucn.org>

PROGETTO BIOCASE: <http://www.biocase.org/>

PROGETTO BIOCISE: <http://www.bgbm.org/BioCise/>

PROGETTO CDEFD: <http://www.bgbm.org/CDEFD/CollectionModel/cdefd.htm>

PROGETTO SIMES: <http://www.ercim.org/simes/>

MODELLI DI INFORMAZIONE

ASC (1993): An information model for Biological Collections (Draft) Report of the Biological Collections Data Standards Workshop, August 18-24, 1992. Association of Systematic Collections, Committee on Computerization and Networking.
<gopher://kaw.keil.ukans.edu:70/11/standards/asc>

ASC/CNC (1997): A Library of Information Models (Schemas) for Biological Collections. The Association of Systematic Collections Computerization and Networking Committee.
http://gizmo.lbl.gov/DM_TOOLS/OPM/BCSL/BCSL.html

BERENDSOHN, W. G. (1994): IOPI World Vascular Plant Checklist. A CASE Model of Checklist System Data. [Draft version 6 in:] Wilson, K. (ed.): Global Plant Checklist project plan, version 1.2. International Organization for Plant Information (IOPI), Sidney
<http://www.bgbm.org/IOPI/IOPIModel73/7301root.htm>.

BERENDSOHN, W. G. (1997): A taxonomic information model for botanical *databases*: the IOPI model. *Taxon* 46:283-309.

BERENDSOHN, W. G., ANAGNOSTOPOULOS, A., HAGEDORN, G., JAKUPOVIC, J., NIMIS, P.L. & VALDÉS, B. (1996): The CDEFD Information Model for Biological Collections.
<http://www.bgbm.org/CDEFD/CollectionModel/cdefd.htm>

BERENDSOHN, W. G., ANAGNOSTOPOULOS, A., HAGEDORN, G., JAKUPOVIC, J., NIMIS, P. L., VALDÉS, B., GÜNTSCH, A., PANKHURST, R. J. & WHITE, R. J. (1999): A comprehensive reference model for biological collections and surveys. *Taxon* 48: 511-562.
<http://www.bgbm.org/biodivinf/docs/CollectionModel/>

BERENDSOHN, W. G., ANAGNOSTOPOULOS, A., JAKUPOVIC, J., NIMIS, P. L. & VALDÉS, B. 1996. A Framework for Biological Information Models. In: Valdés, B. & Silvestre, S. (ed.), Proceedings of the VIII OPTIMA meeting. *Lagascalia* 19:667-672. <http://www.bgbm.org/cdefd/objects/>

BERENDSOHN, W. G., GREILHUBER, J., ANAGNOSTOPOULOS, A., BEDINI, G., JAKUPOVIC, J., NIMIS, P.L. & VALDÉS, B. (1996): A comprehensive datamodel for karyological *databases*. *Plant Syst. Evol.* 205:85-98.

BISBY, F. (1995): Plant Names in Botanical *Databases*. Plant Taxonomic *Database* Standards No. 3, International Working Group on Taxonomic *Databases* for Plant Sciences (TDWG). Hunt Institute for Botanical Documentation, Pittsburgh

ELANKOVAN, S., BERENDSOHN, W. G. & MEYER, H. (1996): Person Teams in the CDEFD and IOPI Models: An Implementation. <http://www.bgbm.org/CDEFD/PersonTeams/Title.htm>.

MEYER, H., BERENDSOHN, W. G. & ELANKOVAN, S. (1996): Taxonomic Identification Within the CDEFD Information Model. <http://www.bgbm.org/cdefd/identifications/title.htm>.

Missouri Botanical Garden (1996): Missouri Botanical Garden Research Data Model. Tropicos II ER Diagram. <http://mobot.mobot.org/W3T/Search/vast.html>

- Smithsonian Institution (1994): Logical Data Model for Museum Collections Transaction Management. Collections and Research Information System (CRIS) Development Program, National Museum of Natural History, Smithsonian Institution.
 gopher://nmnhgoph.si.edu:70/11/.compute/.cris/.logicaldm.
- SPNHC (1994): Guidelines for the Care of Natural History Collections Part VI. Definitions [of general English terms in natural history collections]. Society for the Preservation of Natural History Collections. <http://palimpsest.stanford.edu/byorg/spnhc/spnhc1.html>
- WILSON, P. D. (1993): Missouri Botanical Garden Research Data Model. Release 4.0. Prepared for the Research and Steering Committee. Missouri Botanical Garden (unpublished).
 gopher://cissus.mobot.org:70/11/pub

SISTEMI DEDICATI

- ANONYMOUS (1997): ITIS - Integrated Taxonomic Information System. Documentation Menu: ITIS *Database* - Physical Model. <http://www.itis.usda.gov/plantproj/itis/document.html>
- COLWELL, R. K. (1997): The Biota Data Model (Biota Manual, Appendix A).
<http://viceroy.eeb.uconn.edu/BiotaPages/DataModel.html>
- D.L. FILER 2001. Databases and herbarium efficiency BioNET- International News
<http://www.bionet-intl.org/html/whatsnew/newsletters/newsletters.htm>
- MINNIGERODE, M. D. (1998): Overview of Tracy, Datamodel.
<http://www.csdl.tamu.edu/FLORA/input/overview.html>
- MIP (1996): SMASCH [Specimen Management System for California Herbaria] Documentation. Museum Informatics Project (MIP), Information Systems and Technology, University of California, Berkeley. http://www.mip.berkeley.edu/www_apps/smasch/
- New York Botanical Garden (1998): CASSIA [Collection And Specimen System for Information and Analysis], Introduction. <http://www.nybg.org/bsci/cass/spec.html>
- SHATTUCK, S., FITZSIMMONS, N. & BAIRD, D. (1999): Technical Development - the BioLink Datamodel. <http://www.ento.csiro.au/biolink/index.html>
- SPECIFY 1999- Biodiversity and Biological Collection Management, Beta 2.0 Draft Users Guide. Appendix A: Table names, Appendix B: Field names. <http://www.usobi.org/specify/Docs.html>

STANDARD DI INTERSCAMBIO DI DATI

- ALLKIN, B. & WHITE, R. (1991): XDF. A language for the definition and exchange of biological data sets. Royal Botanic Garden, Kew.
- CONN, B.J. (1996) (ed.): HISPID3. Herbarium Information Standards and Protocols for Interchange of Data. Version 3. Royal Botanic Gardens Sydney. <http://www.rbgsyd.gov.au/hiscom/>
- CONN, B.J. (1997): BarCode Technology for Herbarium and Museum Biological Collections, Technical Report. <http://www.rbgsyd.gov.au/HISCOM/barcode/barcode.html>
- CROFT, J. R. (ed.) (1989): HISPID - Herbarium Information Standards and Protocols for Interchange of Data [version 1]. Australian National Botanic Gardens, Canberra.
- Flora of Texas Consortium (1997): Data Specification, Common data exchange format - Core Data. <http://www.csdl.tamu.edu/FLORA/ftc/ftcffld4.htm>
 gopher://muse.bio.cornell.edu/00/standards/hispid/hispid91.std,
<http://www.anbg.gov.au/projects/hispid/hispid-contents.html>.
- IUCN/WWF (1987): The International Transfer Format (ITF) for Botanic Garden Plant Records. Plant Taxonomic *Database* Standards No. 1. Hunt Institute for Botanical Documentation, Pittsburgh.
- WHALEN, A. (ed) (1993): HISPID - Herbarium Information Standards and Protocols for Interchange of Data [version 2]. National Herbarium of New South Wales, Sydney.

WYSE JACKSON, D. (compiler) (1997): International Transfer Format for Botanic Garden Plant Records (version 2.00 draft 3.2.). Botanic Gardens Conservation International, Richmond.
<http://www.rbgekew.org.uk/BGCI/itf2-d31.doc>

STANDARD NOMENCLATURALI

- BRIDSON, G.D.R. & SMITH, E. R. (1991): *Botanico-Periodicum-Huntianum/supplementum*. Hunt Institute for Botanical Documentation, Pittsburgh..
- BRUMMIT, R.K. & DAVIES, R.A. (ed.) (1998+). Authors of plant names, pilot project for the Plant Name Project. Royal Botanic Gardens Kew, Harvard University Herbaria & Australian National Herbarium. http://pnp.huh.harvard.edu/authors/servlet/query_author.html.
- COOK, E. M. (1995): Economic Botany Data Collection Standard. Prepared for the International Working Group on Taxonomic *Databases* for Plant Sciences (TDWG). Royal Botanic Gardens, Kew.
- Di GREGORIO, A. & JANSEN, L. J. M. (1996): *FAO Land Cover Classification: A Dichotomous, Modular-Hierarchical Approach*.
<http://www.fao.org/WAICENT/FAOINFO/SUSTDEV/EIdirect/EIre0019.htm>
- GREUTER W., J. MCNEILL, R. BARRIE, H.-M. BURDET, V. DEMOULIN, T. S. FILGUERIAS, D. H. NICOLSON, P. C. SILVA, J. E. SKOG, P. TREHANE, N. J. TURLAND, D. L. HAWKSWORTH (EDITORS & COMPILERS) (2000): *International Code of Botanical Nomenclature (Saint Louis Code)* adopted by the Sixteenth International Botanical Congress St. Louis, Missouri, July - August 1999. *Regnum Vegetabile* 138. Koeltz Scientific Books, Königstein.
- GREUTER, W., BARRIE, F., BURDET, H.M., CHALONER, W.G., DEMOULIN, V., HAWKSWORTH, D.L., JORGENSEN, P.M., NICOLSON, D.H., SILVA, P.C., TREHANE, P. & MCNEILL, J. (ED.) (1994): *International Code of Botanical Nomenclature (Tokyo Code)*. *Regnum Vegetabile* 131. Koeltz Scientific Books, Königstein. <http://www.bgbm.org/iapt/nomenclature/code/default.htm>
- HEYWOOD, C. A., HEYWOOD, V. H. & WYSE JACKSON, P. S. (1990): *International Directory of Botanical Gardens*. Koeltz, Koenigstein.
- HOLLIS, S. & BRUMMITT, R. (1992): *World Geographical Scheme for Recording Plant Distributions*. *Plant Taxonomic Database Standards No. 2*, International Working Group on Taxonomic *Databases* for Plant Sciences (TDWG). Hunt Institute for Botanical Documentation, Pittsburgh <http://www.bgbm.org/TDWG/geo/default.htm>
- HOLMGREN, P. K., HOLMGREN, N. H. & BARNETT, L. C. (1990): *Index Herbariorum*, Pt. 1: The Herbaria of the World (ed. 8). *Regnum Vegetabile* 120. <http://www.nybg.org/bsci/ih/ih.html>
- IPNI The International Plant Names Index <http://www.ipni.org/index.html>
- ISO (1988): *Codes for the representation of names of countries*. Third edition; ISO 3166: 1988 Aug 15; 53 p. [gopher://muse.bio.cornell.edu:70/00/standards/iso/iso-3166](http://muse.bio.cornell.edu:70/00/standards/iso/iso-3166).
- ISO (1994): *Codes from ISO 3166*. Updated by the RIPE Network Coordination Centre, in coordination with the ISO 3166 Maintenance Agency at DIN Berlin. <http://www.chemie.fu-berlin.de/outerspace/mirror-packages/german/iso-3166.html>
- ISO (1997): *Some Codes from ISO 3166*. Updated by the RIPE Network Coordination Centre. Source: ISO 3166 Maintenance Agency. <http://www.chemie.fu-berlin.de/outerspace/mirror-packages/german/iso-3166.html>
- ISO 1997. ISO 3166-1. *Codes for the representation of names of countries and their subdivisions - Part 1: Country codes*. Berlin.
- ISO 1998. ISO 3166-2. *Codes for the representation of names of countries and their subdivisions - Part 2: Country subdivision code*. Berlin.
- ISO 1998. ISO 639-2. *Codes for the representation of names of languages - Part 2: Alpha-3 code*. Vienna.
- ISO 1999. ISO 3166-3. *Codes for the representation of names of countries and their subdivisions - Part 3: Code for formerly used names of countries*. Berlin.

- IUCN (1994): IUCN Red List Categories. Prepared by the IUCN Species Survival Commission. As approved by the 40th meeting of the IUCN Council, Gland, Switzerland, 30 November 1994.
- LAWRENCE et al. 1968. *Botanico-periodicum-huntianum*. Hunt Institute for Botanical Documentation, Pittsburgh.
- Nature Conservancy (1997): The National Vegetation Classification System: List of Types. <http://consci.tnc.org/library/pubs/class/vol2/4list.pdf>
- STAFLEU, F. A. & COWAN, R. S. (1976-): Taxonomic literature, ed. 2 and its Supplements. Reg. Veg. 125 etc.
- Takhtajan, A. (1986): *Floristic Regions of the World*. University of California Press. Bishen Singh, Dehra Dun.
- TREHANE, P., BRICKELL, C.D., BAUM, B.R., HETTERSCHIED, W.L.A., LESLIE, A.C., MCNEILL, J., SPONGBERG, S.A. & VRUGTMAN, F. (1995): *International Code of Nomenclature for Cultivated Plants*. Quarterjack Publishing, Wimborne.
- WCMC (1995): *Plant Occurrence and Status Scheme, a Standard for Recording the Relationship between a Plant and a Place*. WCMC, Cambridge <http://plants.usda.gov>

STANDARD DI METADATA

- ANONYMOUS 1996: Dublin Core Metadata Element Set: Reference Description. OCLC Online Computer Library Center, Inc. http://purl.org/metadata/dublin_core_elements
- CROSSLEY, D. 1996: ERIN Standards for Spatial Information. http://www.erin.gov.au/gis/gis_standards.html
- FGDC 1994: Content standards for digital spatial metadata (June 8 draft). Federal Geographic Data Committee. Washington, D.C. <http://geology.usgs.gov/tools/metadata/standard/metadata.html>
- FGDC 1997: FGDC-STD-005 Vegetation Classification and Information Standards. Federal Geographic Data Committee, Vegetation Subcommittee. Federal Geographic Data Committee Secretariat, Reston <http://www.fgdc.gov/standards/documents/standards/vegetation/>.
- FGDC 1998: FGDC-STD-001-1998. Content standard for digital geospatial metadata (revised June 1998). Federal Geographic Data Committee. Washington, D.C. <http://www.fgdc.gov/metadata/csdgm/>
- FGDC Biological Data Working Group and USGS Biological Resources Division 1999: Content Standard for Digital Geospatial Metadata. Part 1: Biological Data Profile. Federal Geographic Data Committee, Washington D.C. <http://www.fgdc.gov/standards/documents/standards/biodata/biodatap.html>
- MDA 2000: SPECTRUM: The UK Museum Documentation Standard. London, UK.