

Il sistema di gestione dati del progetto EnvEurope per il monitoraggio ecologico a lungo termine nella rete LTER

Alessandro Oggioni (*), Paola Carrara (*), Tomas Kliment (**),
Johannes Peterseil (***), Herbert Schentz (***)

(*) Istituto per il Rilevamento Elettromagnetico dell'Ambiente (IREA - CNR),
Via Bassini 15, Milano (Italy), oggioni.a@irea.cnr.it e carrara.p@irea.cnr.it

(**) Istituto per le Scienze Marine (ISMAR - CNR), Arsenale - Tesa 104 Castello 2737/F Venezia (Italy),
tomas.kliment@gmail.com

(***) Ecosystem Research & Monitoring, Umweltbundesamt GmbH (UBA), Spittelauer Lände 5 Wien (Austria),
johannes.peterseil@umweltbundesamt.at e herbert.schentz@umweltbundesamt.at

Riassunto

Il progetto LIFE + EnvEurope (LIFE08 ENV/IT/000339) ha lo scopo di proporre, per la più estesa rete di ricerca ecologica a lungo termine europea (LTER Europe), un modello di alta qualità e un insieme di parametri comuni per il monitoraggio ambientale. Attualmente la rete LTER Europea conta più di 400 siti di ricerca, distribuiti in 22 paesi, in cui vengono raccolti dati ecologici in modo continuativo da almeno 10 anni in tre tipi di ecosistemi: terrestri, acquatici e marini.

Nell'ambito del progetto EnvEurope, il gruppo di gestione dei dati ha il compito di fornire strumenti e servizi per lo scambio di metadati e dati tra i partner che, in futuro, potranno essere utilizzati per l'intera rete europea. L'infrastruttura attualmente proposta, combina le pratiche adottate dalla Comunità LTER europea e internazionale, con le raccomandazioni europee (es. INSPIRE) e con gli standard internazionali (es. Open Geospatial Consortium - OGC).

Per realizzare un'architettura comune di scambio di dati, che garantisca un adeguato livello di interoperabilità nella rete sono state proposte alcune soluzioni:

- un profilo di metadati, per descrivere i dataset della comunità EnvEurope, sulla base dell'Ecological Metadata Language (EML);
- uno strumento di archiviazione, ricerca e gestione dei metadati come estensione del Drupal Ecological Information Management System (DEIMS), sviluppato per la rete LTER statunitense;
- un servizio di discovery per il rilevamento dei metadati da client remoti, basato su software opensource GeoNetwork e compatibile con il profilo di metadati proposto da INSPIRE;
- un GeoViewer per la geolocalizzazione dei siti e delle stazioni della rete;
- servizi web OGC, tipo Sensor Web e LinkedData, sono in fase di valutazione e di realizzazione.

Verranno presentate tutte queste soluzioni valutandone criticità, evoluzioni future, sinergie con altri progetti e buone pratiche.

Abstract

The last innovations in the information science have improved developing systems relating to the creation, collection, storage, processing, modelling, interpretation, display and dissemination of data and information focused on Environmental Science (Page & Wohlgemuth, 2010). International initiatives such as SEIS (Shared Environmental Information System), GMES (Global Monitoring Environmental and Security), GEOSS (Global Earth Observation System of Systems); projects like Humboldt, NatureSDIplus, BioFresh, GIGAS; consortia such as GBIF (Global Biodiversity Information Facility), LifeWatch, DataONE (Data Observation Network for Earth) and OGC (Open

Geo-spatial Consortium) and finally legal framework in Europe (INSPIRE - 2007/2/CE) and in United States (OMB Circular A-16 - 2002) stimulated effective implementation of an information technology innovations. A common feature consists of a development of an infrastructure, which facilitates discovery, evaluation and use of data, information and knowledge. Sharing of large datasets can establish a much deeper understanding for both nature and society, open up many new avenues of research or assist to policy-makers with relation to environmental policies. (AA.VV., 2011). Paper deals with issues related to an establishment of architecture for data exchange within the Long Term Ecological domain in Europe and propose solutions to resolve them in order to provide an interoperable system.

Introduzione

Secondo la più classica delle definizioni (Odum & Barret, 2005) l'ecologia è lo studio dei luoghi dove la vita si sviluppa e in particolare di tutti quei processi funzionali e quelle relazioni che avvengono tra gli organismi viventi e il loro ambiente. L'ecologia si occupa così di tutti i livelli di complessità nello studio dei viventi: popolazioni, comunità, ecosistemi, paesaggi e ecosfera; ma soprattutto delle interazioni tra tutti questi piani. Si può immaginare quindi una notevole complessità nei dati di tipo ecologico, per semplicità si preferisce però definirli come la combinazione di variabili ambientali (parametri chimici e fisici) e biologiche (presenza, abbondanza o densità di organismi). Le scale temporali e le tre dimensioni spaziali sono poi le vere scale di rappresentazione di questo tipo di dati (Harris, 1986).

L'interazione tra discipline e campi di indagine diversi (es. climatologia, geologia, oceanografia, idrologia, tassonomia) è quindi indispensabile per aumentare il grado di comprensione di qualunque sistema ecologico. L'ecologia ha quindi innata l'esigenza di sviluppare sistemi efficaci ed efficienti per archiviare, ricercare, accedere e integrare un'ampia gamma e grandi volumi di dati.

Tra la fine degli anni '90 e i primi anni 2000, a margine di iniziative di bioinformatica nate in seno alla U.S. Long-term Ecological Research Network (US-LTER) e al U.S. National Center for Ecological Analysis and Synthesis (NCEAS), viene coniato il termine 'ecoinformatica' (Kareiva, 2001; Brunt *et al.*, 2002). Già da questi primi passi emerge chiaramente come questa disciplina si debba occupare proprio della conservazione, del recupero e dell'integrazione dei dati ecologici. Infatti mentre la bioinformatica (Bisby, 2000) dà maggior enfasi alla risoluzione di problemi tassonomici soprattutto con un approccio molecolare (es. classificazioni, tassonomie e sinonimie), l'ecoinformatica ha il gradito compito di analizzare la struttura dei dati nelle diverse discipline, fornire strumenti informatici per la raccolta, l'archiviazione, la ricerca dei dati ecologici, in modo da combinarli e permetterne una analisi organica per comprendere, prevedere e contrastare i cambiamenti globali cui stiamo assistendo (Canhos *et al.*, 2004; Soberón & Peterson, 2004; Jones *et al.*, 2006; Bekker *et al.*, 2007; Guralnick & Hill 2009; Recknagel, 2011; Dengler *et al.*, 2011).

La rete LTER, oltre ad essere stata promotrice dello sviluppo dell'ecoinformatica, risulta essere anche un ottimo banco di prova per individuare modi efficaci per ricercare, accedere, integrare, conservare e analizzare in maniera accurata una vasta gamma e grandi volumi di dati per l'ambito ecologico (Reichman *et al.*, 2010; Michener *et al.*, 2011). Questo *network* ha una diffusione capillare nei 5 continenti (*International Long Term Ecological Research - ILTER*¹) ed è rappresentato in tutti i paesi europei (*European Long Term Ecosystem Research Network*²) con circa 400 siti di ricerca omogenei dal punto di vista dell'habitat o dell'uso del suolo. La rete europea, nata all'inizio degli anni '90, copre 4 tipologie di ambiente: marino, lacustre, forestale e fluviale; le sue attività si concentrano sulla evoluzione temporale dei processi ecosistemici di piccole dimensioni, per una migliorarne la comprensione in risposta a forzanti ambientali, sociali ed economiche. In questo contesto la rete europea, ma anche quella internazionale ha l'esigenza di iniziare un confronto *cross-domain* e a scala globale così da fornire un valore aggiunto ai dati a

¹ <http://www.ilternet.edu/>

² <http://www.lter-europe.net/>

lungo termine, per poter effettivamente migliorare la gestione ambientale dando un indirizzo alle politiche europee. Nella rete LTER a partire dal 2006, sotto l'auspicio del *Network of Excellence ALTER-Net*³ e sulla base delle infrastrutture esistenti, si iniziano a diffondere preziose serie di dati. E' però a partire dal 2010 con l'avvio del progetto EnvEurope (Mirtl & Krauze 2007; Mirtl, 2009; Mirtl, 2010) che si comprende in modo definitivo come sia necessario uno sforzo notevole di implementazione, informatizzazione e di uso di strumenti adeguati per fornire un *framework* che permette di condividere i dati seguendo le raccomandazioni di *Shared Environmental Information System* (SEIS) e sviluppando alcune componenti adatte per gli scopi del *Global Monitoring for Environment and Security* (GMES), che è uno degli scopi principali del progetto.

In questo lavoro saranno presentate le soluzioni fino ad ora implementate all'interno del progetto EnvEurope, le idee per sviluppi futuri, ma anche le interazioni con altre iniziative europee e internazionali legate agli aspetti di biodiversità (*Global Biodiversity Information Facility* - GBIF) o di ecologia (LTER International o US - LTER).

Contributo del progetto EnvEurope

Gli studi ecologici a lungo termine danno un effettivo contributo nel riconoscere e interpretare i cambiamenti ambientali (climatici, biologici, ecosistemici) se i dati sono raccolti su scala geografica ampia e con un approccio metodologico armonizzato. Il progetto Life+ EnvEurope⁴ (LIFE08 ENV/IT/000339) ha proprio lo scopo di armonizzare parametri, metodi e di gestire in maniera integrata i dati raccolti in 67 siti di ricerca delle reti LTER europea (figura 1). L'armonizzazione delle metodologie, combinata con il miglioramento della gestione e condivisione della grossa mole di dati esistenti permetteranno di ottenere una più profonda comprensione del funzionamento degli ecosistemi.

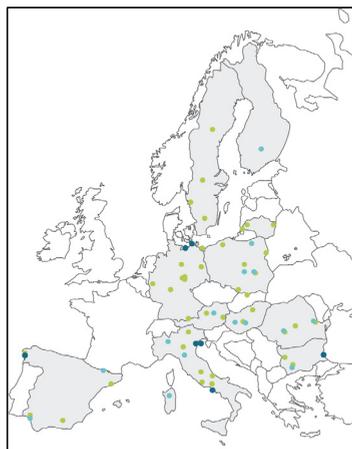


Figura 1. Distribuzione geografica dei 67 siti LTER coinvolti nel progetto EnvEurope, suddivisi per tipologia di ambiente: siti di acque continentali in azzurro, siti marini in blu e siti terrestri in verde.

In questo contesto progettuale l'azione di *Data Management* (Action 1) ha lo scopo di proporre una infrastruttura che permetta di conservare la grande mole di dati fino ad ora archiviati in modo autonomo nei vari siti, recuperandoli attraverso una armonizzazione delle informazioni che descrivono i dati (metadati) ma anche implementando strumenti per la condivisione i dati, così da sostenere le future attività di ricerca. L'attività così svolta servirà come banco di prova tecnologica e procedurale per tutta la rete LTER Europea.

³ <http://www.alter-net.info/>

⁴ <http://www.enveurope.eu/>

Il sistema di gestione dati in EnvEurope e i suoi componenti

La realizzazione di un'infrastruttura di gestione dei dati per una comunità, come quella ecologica, che ha lunghe e radicate tradizioni scientifiche ci ha costretti a procedere per passi successivi: valutando le pratiche per la conservazione di metadati e dati; riconoscendo le tipologie di dati presenti in ciascun dominio (lacustre, marino o terrestre); analizzando le soluzioni tecnologiche già adottate o sperimentate in altri progetti, realtà e iniziative; infine ricalibrando queste soluzioni secondo le esigenze indicate dai beneficiari del progetto. Tutto questo servirà ad ottenere soluzioni nell'ambito dell'ecoinformatica che possano essere ottimamente calibrate per la realtà della rete LTER, ma che permettano di ottenere una reale interoperabilità con altre infrastrutture dati o sistemi informativi trattanti dati ecologici o più genericamente dati ambientali.

Attraverso 2 questionari è stato possibile raccogliere informazioni relativamente alle pratiche adottate, dai singoli beneficiari, per la conservazione e la distribuzione dei metadati e dei dati nei siti e tra i siti. I risultati (Perteseil *et al.*, 2011) si possono considerare significativi della comunità: più del 85% dei siti di ricerca hanno risposto alle domande dei 2 questionari. Lo scenario emerso è però di una comunità poco abituata alla condivisione delle informazioni e dei dati attraverso i moderni strumenti dell'ecoinformatica. E' stato rilevato che:

- solo 23 siti su 67 adotta soluzioni standard per l'archiviazione dei metadati, utilizzando profili quali *DublinCore*, *Ecological Metadata Languages* (EML) e ISO19115;
- 21 siti hanno creato cataloghi di metadati usando MetaCat⁵ o utilizzato strumenti conformi allo standard CSW (*Catalogue Service of Web*) utilizzando GeoNetwork⁶;
- tutti i dati dei diversi beneficiari sono archiviati in fogli Excel che vengono ristrutturati di volta in volta secondo i bisogni progettuali;
- circa 80% dei beneficiari raccoglie l'informazione geografica insieme alla misura di campo e indica come fondamentale e necessaria la georeferenziazione del dato raccolto;
- solo il 27% dei siti utilizza *data portal* per la distribuzione di propri *dataset*, ma solo il 21% di questi provvede a distribuire i propri dati utilizzando servizi OGC (*Open Geospatial Consortium*) standard quali WFS (*Web Feature Service*), WCS (*Web Coverage Service*) o SWE (*Sensor Web Enablement*).

Da questi questionari è emerso che la tipologia prevalente di dati nella comunità EnvEurope sia di tipo osservativo, cioè misurazioni - quantità numeriche - di parametri che sono associati con caratteristiche temporali e luoghi geografici (Borgman *et al.*, 2007). Queste osservazioni si riferiscono poi a due comparti specifici, da un lato quello biologico (es. presenza di specie, elenchi tassonomici, abbondanze o densità di organismi, ecc.), dall'altro quello ambientale caratterizzato da parametri chimico-fisici (es. temperatura, salinità, trasparenza, torbidità, ecc.).

Prima di proporre qualunque soluzione è stata da noi effettuata un'analisi delle implementazioni, degli strumenti e degli standard adottati nelle altre reti LTER, esaminando anche le proposte in ambito europeo con particolare riguardo alle indicazioni date dalla Direttiva INSPIRE. Abbiamo così riconosciuto come nella rete US-LTER sono stati concepiti e adottati strumenti quali: MetaCat come database XML che permette l'archiviazione e il reperimento attraverso richieste sui metadati; Morpho come strumento di *data management* progettato appositamente per gli ecologi che permette l'inserimento dei dati, l'archiviazione sotto forma di *file system* e l'associazione con i relativi metadati; EML come standard di metadati studiato appositamente per le discipline ecologiche. Per la rete US-LTER nessun servizio web standard è attualmente adottato per la distribuzione dei dati, questi sono infatti distribuiti esclusivamente utilizzando fogli di calcolo o file testo.

Sono state poi analizzate le implementazioni adottate nella realtà europea, le normative comunitarie e gli standard internazionali per considerare soluzioni che permettano l'interoperabilità con i

⁵ <http://knb.ecoinformatics.org/>

⁶ <http://geonetwork-opensource.org/>

progetti, le iniziative e le realtà, più significative, attualmente esistenti. Sono stati da noi presi in esame iniziative internazionali come SEIS, GMES e GEOSS (*Global Earth Observation System of Systems*); progetti come Humboldt, NatureSDIplus, BioFresh, GIGAS; consorzi come GBIF, LifeWatch, DataOne (*Data Observation Network for Earth*); ma anche le indicazioni date dall'OGC e dal quadro giuridico Europeo (INSPIRE - 2007/2/CE). Ne emerso un quadro che potremmo dire omogeneo, indicando negli standard OGC le specifiche di condivisione e distribuzione dei dati di chimico-fisico e nelle soluzioni GBIF per la parte riguardante i dati biologici.

Sugli aspetti di data management però le situazioni adottate sono eterogenee sia a livello della rete LTER sia in senso globale, nel progetto EnvEurope si è così deciso di creare: un profilo di metadati per la comunità LTER Europa, uno strumento per la collezione, la ricerca e la visualizzazione dei metadati dei dataset e infine un GeoViewer per la visualizzazione su mappa dei siti di ricerca LTER.

Profilo di metadati per la comunità EnvEurope

La creazione di un profilo di metadati comune per la comunità EnvEurope e per quella LTER Europea è di fondamentale importanza per poter armonizzare le informazioni riguardanti i *dataset*. Questi infatti sono da considerarsi realmente condivisi solo quando sono raccolte informazioni che ne descrivano il contenuto, la modalità di raccolta, l'arco temporale coperto, la proprietà e la possibilità di utilizzo.

Nel progetto EnvEurope è stato usato come modello di riferimento per la descrizione dei dataset il linguaggio EML. Questa specifica di metadattazione è stata sviluppata per la descrizione di dataset ecologici, viene largamente usato come linguaggio di metadattazione in diverse realtà: dal 2008 per l'armonizzazione dei metadati nella rete ILTER, nel network US-LTER (Michener, 1997; Jones, 2001) e dal 2010 per descrivere i dati relativi alla biodiversità in ambito GBIF. Basato su un precedente lavoro effettuato per la società ecologica americana (*Ecological Society of America*, ESA), EML è implementato come una serie di schemi XML che possono essere usati in maniera modulare ed estendibile per dare una descrizione dei dati ecologici.

Visto l'ampio uso che si fa della specifica EML sia in ambito ILTER sia US-LTER, in EnvEurope è stato inevitabile utilizzare questo linguaggio come base per la creazione di un *Domain Specific Community Profile* (DSMP). Per EnvEurope il DSMP riflette però anche il risultato di richieste degli utenti coinvolti nel progetto, di una consultazione con esperti di metadattazione e di un confronto con le specifiche di metadattazione indicate dalla direttiva INSPIRE. Infatti se a livello internazionale e Nord americano le normative comunitarie non sono considerate, per un progetto finanziato dalla Comunità Europea e per una rete che è profondamente legata al continente europeo il considerare normative comunitarie diventa percorso obbligato. Il DSMP risultante e proposto per il progetto EnvEurope è così composto da 20 elementi di metadattazione (figura 2), esposti sotto forma di schema XML.

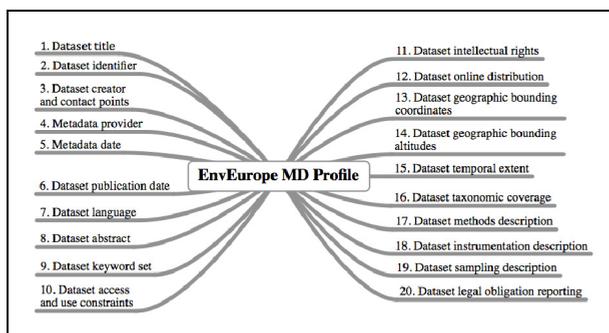


Figura 2. Elenco dei 20 elementi di metadattazione.

Nel confronto con le specifiche dettate dalla Direttiva INSPIRE abbiamo considerato i requisiti della stessa direttiva, è stato così creato un crosswalk tra il DSPM di progetto e il profilo di metadati di INSPIRE, basato su ISO 19115 (ISO, 2003). La prima parte del lavoro di confronto è consistito nella mappatura tra DSMP e *ISO Core* metadata, ma anche e soprattutto nella comparazione con i vincoli definiti nelle regole di implementazione di INSPIRE (INSPIRE, 2008). Questo lavoro, condotto in accordo con la metodologia descritta in Nogueras-Iso (2005) ha permesso di ottenere da un lato un profilo di metadati che è compatibile, avvalendosi di una tecnologia di traduzione XSLT (*eXtensible Stylesheet Language Transformations*) (Kliment, 2011), con il profilo di INSPIRE; dall'altro essendo implementato partendo da schemi EML è risultato interoperabile con altri network, consorzi o progetti che si avvalgono dello stesso linguaggio (es. GBIF, ILTER e US-LTER).

Drupal Ecological Information Management System (DEIMS) in ambito EnvEurope

DEIMS è l'acronimo utilizzato per indicare il sistema di gestione delle informazioni ecologiche nel *Content Management System* (CMS) Drupal. In EnvEurope tutte gli strumenti che riguardano l'archiviazione, la ricerca, la visualizzazione e l'analisi dei metadati sono stati implementati utilizzando lo stesso approccio utilizzato dalla rete US-LTER.

La collezione dei metadati per la descrizione dei dataset è, per il gruppo di data management di EnvEurope, il primo passo per la completa armonizzazione dei dati di progetto. In Drupal è stato così creato un portale *on-line*, denominato EnvEurope DEIMS⁷, per la compilazione dei metadati e uno strumento per la loro gestione. Nell'attuale versione, questo portale ha il compito fondamentale di permettere: ai soli partecipanti al progetto di poter archiviare i metadati che descrivono i dataset secondo modalità condivise; viene invece permesso anche a utenti non registrati di poter effettuare ricerche sui metadati per parole chiavi, per sito di ricerca LTER o per estensione temporale; infine è possibile visualizzarne il contenuto e analizzarlo in modo da poterne valutare la qualità e reale utilità.

Questa interfaccia è organizzata in 2 parti: una dedicata al completamento dei 20 campi di metadato e l'altra per la creazione di richieste per il recupero dei metadati (Fig. 3). I campi di metadatazione sull'interfaccia sono organizzati secondo la tipica struttura dei moduli della rete internet: caselle per il completamento a testo libero, finestre a scorrimento nel caso di campi a scelta multipla e caselle di spunta per scelte univoche. L'interfaccia creata per la ricerca all'interno dei metadati prevede che l'utente possa selezionare 5 diversi criteri: il nome del sito LTER Europe, le parole chiave presenti nel titolo o nell'*abstract*, l'estensione temporale del dataset, la lingua e la frequenza di campionamento. Questa stessa interfaccia permette poi di visualizzare il contenuto dei metadati attraverso un report in il metadato selezionato può essere esportato come formato XML, PDF o testo generico.

⁷ <http://enveurope.geocatalogue.ise.cnr.it/deims/>

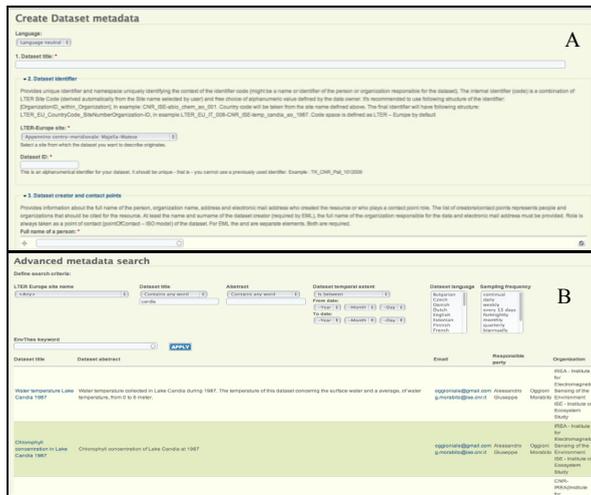


Figura 3. Interfaccia di EnvEurope DEIMS per l'inserimento dei metadati dei dati (A) e per la loro ricerca e visualizzazione (B).

Soluzioni basate su servizi web e il GeoViewer

Le informazioni georiferite sono tradizionalmente gestite per mezzo di Sistemi Informativi Geografici (GIS). L'avvento di internet ha permesso di generare e spostare queste applicazioni sul web. I Web-GIS solitamente sono applicazioni che forniscono servizi di mappatura attraverso il web, dove gli utenti possono controllare visivamente carte tematiche, gestire più strati sovrapposti, ed eseguire operazioni semplici, come pan e zoom. Inoltre, in queste applicazioni esiste sempre un archivio che gestisce e distribuisce i diversi layer che è strettamente associato ad un'interfaccia client, così repository diversi risultano gestiti da interfacce utente diverse in un rapporto di tipo uno a uno. Questo rapporto univoco è cambiato da quando sono stati sviluppati i geoservizi: servizi web che distribuiscono dati geografici e strumenti geografici attraverso la rete attraverso server. I servizi web sono gli elementi costitutivi delle cosiddette Infrastrutture di Dati Spaziali (SDI), le infrastrutture tecnologiche per condividere e utilizzare i dati spaziali in modo interoperabile. Secondo la più classica Architettura Service Oriented (SOA), adottata e raccomandata dalla Direttiva INSPIRE, da un lato i geoservizi possono servire i geodati a interfacce clienti diverse, dall'altro un solo client può accedere contemporaneamente a geoservizi diversi e distribuiti.

Nell'ambito del progetto EnvEurope, per la pubblicazione e la distribuzione dei dati ecologici, è stato proposto l'uso dei servizi supportati in ambito OGC e ampiamente utilizzati nel settore dei servizi Web per le informazioni geografiche. Alcuni degli OGC Web Services sono inclusi nelle raccomandazioni INSPIRE da utilizzare come servizi di rete (es. *Web Map Service* - WMS e *Web Feature Service* - WFS), altri come il Sensor Web sono raccomandati da GEOSS e sono stati adottati dal *World Wide Web Consortium* (W3C).

Come emerso dai questionari i dati che sono presenti in ambito EnvEurope sono principalmente osservazioni/misure di parametri, con associate caratteristiche spaziali (x, y, z) e una dimensione temporale (t). Solo in alcuni casi i dati possono essere rappresentati sotto forma di mappa: distribuzioni tematiche ottenute da osservazioni effettuate da sensori remoti, risultato di elaborazioni spaziali sui dati da una o più stazioni puntuali, oppure distribuzioni areali di specie. Pertanto servizi WMS e WFS probabilmente avranno una utilità limitata nell'ambito del dominio EnvEurope e LTER; servizi come il Sensor Web è invece uno standard più promettente, tanto più che in questo approccio i dati (osservazioni) non sono trattate come file ma direttamente come record di database, permettendo così un'analisi più flessibile e una gestione più semplice delle misure effettuate con alta granularità temporale.

In EnvEurope l'utilizzo di servizi web o di geoservizi per la pubblicazione dei dati, rappresenta un primo passo verso alcuni obiettivi definitivi di progetto: l'interoperabilità dei dati e la condivisione di archivi dati indipendenti e diffusi, dove le osservazioni raccolte da parte di beneficiari differenti possono essere archiviate e conservate, senza la necessità di impianti di stoccaggio centralizzato. Secondo questo indirizzo progettuale la realizzazione di un'applicazione client per visualizzare i dati serviti da servizi WMS e WFS, ma anche Sensor Web è di centrale importanza. Il progetto EnvEurope offre oggi un GeoViewer di progetto (figura 4) per la visualizzazione in un ambito geografico dei layer dei confini geografici dei siti LTER europei distribuiti come OGC Web Services (WMS e WFS). Il servizio di mappa consente inoltre, per ciascun sito LTER, di accedere direttamente alla ricerca dei metadati, di creare di metadati tramite gli strumenti precedentemente descritti e infine di raccogliere le informazioni geografiche dei diversi siti LTER. Infine dal GeoViewer è possibile aggiungere servizi WMS, funzionalità che consente agli altri utenti di caricare mappe pubblicate in altri progetti, iniziative, da altre istituzioni o autorità direttamente alla finestra EnvEurope GeoViewer.

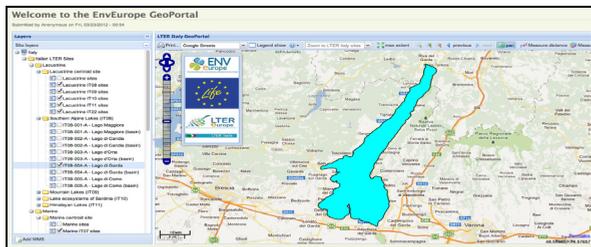


Figura 4. GeoViewer di progetto integrato nel sistema Drupal contenente anche gli strumenti per la metadatazione. Viene qui rappresentato a titolo di esempio il confine geografico del sito di ricerca LTER Lago di Garda.

Integrazione con altre iniziative

La completa interoperabilità tra iniziative, progetti e consorzi diversi è l'obiettivo principale in un'attività di gestione dati. La reale interoperabilità avviene però con l'utilizzo di standard condivisi o con l'adozione di strumenti (es. *broker*) (Nativi *et al.*, 2012) che permettano ai server o alle interfacce client di integrare informazioni o dati distribuiti anche utilizzando standard differenti.

In EnvEurope, visto il basso livello tecnologico dei beneficiari, si è preferito effettuare una analisi dell'esistente e adottare specifiche e geoservizi standard utilizzati da altre grandi realtà, progetti o iniziative. Così si è potuta realizzare la reale integrazione con tutta la rete LTER (ILTER e US-LTER) grazie all'utilizzo di identiche specifiche per la metadatazione, ma anche con tutte le infrastrutture che adottano le specifiche di metadatazione indicate dalla Direttiva INSPIRE o che utilizzano geoservizi OGC per la pubblicazione sul web dei dati.

Rispetto a queste altre iniziative hanno adottato approcci, specifiche e sviluppato strumenti per rispondere alle specifiche esigenze della comunità di riferimento. Realtà come GBIF, sicuramente il più comprensivo e ampio servizio per ricerche complesse su dati riguardanti la biodiversità, offre componenti differenti rispetto a quelli descritti fino ad ora. Questi componenti offrono facilitazioni per la pubblicazione, la ricerca, l'integrazione, il recupero e l'analisi di dati tassonomici che presentano per l'82% dei casi un riferimento geografico (Telenius, 2011). I dati di GBIF quindi sono di tipo geografico, si riferiscono alle occorrenze di diversità tassonomica, si recuperano e si interrogano per livelli sistematici differenti, infine presentano tutti conformità con gli standard DarwinCore o di *Access to Biological Collections Data (ABCD)* che ne facilitano la distribuzione includendo concetti riferiti specificatamente alla diversità biologica. Sono anche ricercabili e analizzabili attraverso l'uso di metadati codificati secondo la specifica EML. In GBIF la pubblicazione dei dati riguardanti la biodiversità avviene attraverso la rete ed è facilitata dalla piattaforma software IPT (*Integrated Publishing Toolkit*). IPT gestisce e permette di distribuire i

dati attraverso l'uso di una tripletta di file che riguardano: i metadati (EML), le liste e le occorrenze tassonomiche organizzate secondo *DataBase Management System* (DBMS) o file testo (es. CSV). L'integrazione tra EnvEurope e GBIF è stata da noi testata pubblicando un dataset riguardante occorrenze biologiche di specie di piante acquatiche lacustri, attraverso l'uso del software IPT. Sono stati importati i file di metadato creato attraverso l'uso dell'interfaccia EnvEurope DEIMS e le specifiche DSPM di progetto, le lista di specie e le occorrenze georiferite delle specie macrofite organizzate in un DataBase relazionale.

Sfide e opportunità

In un periodo come quello attuale dove si assiste a perdite di habitat, cambiamenti climatici di scala globale e rapido declino nella diversità biologica la domanda di informazioni, dati e misure armonizzate e distribuite da ricercatori per rispondere alle esigenze decisionali di politici o di amministratori pubblici è più che mai impellente (Telenius, 2011). L'ammontare delle informazioni e dei dati raccolti in questi anni in ambito ecologico e biologico è enorme (Pullin & Salafsky, 2010; Dengler *et al.*, 2011). Organizzare e realizzare infrastrutture dati robuste per la pubblicazione e la condivisione di dati non organizzati è ancora oggi una sfida. Le combinazioni tra le occorrenze geolocalizzate di specie e i dati ambientali raccolti con alta frequenza o granularità, però permettono opportunità analitiche che vanno oltre ai semplici dati grezzi. Se a questi dati aggiungiamo quelli raccolti in serie storiche più che decennali, con metodi armonizzati e parametri condivisi, vengono a prodursi occasioni di ricerca prima ad ora impensabili.

Il progetto EnvEurope, per conto della rete LTER Europea, si trova oggi in prima linea per far fronte a queste sfide e queste opportunità. Attualmente infatti l'infrastruttura per la condivisione dei dati e dei metadati è stata fino ad ora solo abbozzata e ancora molti nodi della rete, siti di ricerca nazionali e istituzioni sono ancora oggi arretrati; occorrerà nel prossimo futuro armonizzare la struttura dei dati e delocalizzare i servizi web in tutta la rete. In particolare andranno affrontate le sfide riguardanti:

1. la raccolta e la pubblicazione dei metadati, secondo il DSMP prodotto, in modo che si ottenga maggiore visibilità dei dati, dei parametri e dei metodi che vengono raccolti, misurati e adottati nella rete. Questo permetterebbe alla rete nel suo complesso, ma anche e soprattutto ai singoli siti di essere individuabili e coinvolgibili in progetti, pubblicazioni o iniziative;
2. l'adozione di servizi web che permettano la completa interoperabilità tra i dati della rete in modo che, adeguatamente controllati, questi dati siano disponibili in formati condivisi e che possano più facilmente essere utilizzati per analisi o ricerche;
3. la attuazione di strumenti e servizi da parte di tutti i siti di ricerca coinvolti nella rete LTER permetterebbe di distribuire le informazioni e i dati, così che queste non verrebbero centralizzati ma rimarrebbero archiviati nelle istituzioni che li hanno prodotti. Questo, oltre a migliorare la vitalità dei dati stessi, darebbe indipendenza ai singoli nodi.

La creazione di una rete così organizzata offrirebbe opportunità nella collaborazione con iniziative internazionali come GBIF, LifeWatch o DataOne che darebbero maggiori sviluppi e finanziamenti per nuove ricerche. Far interagire le informazioni e i dati provenienti da realtà come quelle citate, consentirebbe di creare quell'integrazione tra discipline che è tanto cara all'ecologia e permetterebbe di aumentare il grado di comprensione di qualunque sistema naturale più di quanto sia stato possibile fino ad ora.

Bibliografia

- AA.VV. 2011. Challenges and opportunities. Introduction. *Science*, 331(6018): 692.
- Bekker, R.M., H. Bruelheide & K. Woods. 2007. Longterm datasets: From descriptive to predictive data using ecoinformatics. *Journal of Vegetation Science*, 18: 457-570.
- Bisby, F.A. 2000. The quiet revolution: biodiversity informatics and the internet. *Science*, 289: 2309-2312.
- Borgman, C., J. Wallis & N. Enyedy. 2007. Little science confronts the data deluge: Habitat ecology, embedded sensor networks, and digital libraries. *International Journal on Digital Libraries*, 7(1), 17-30.

- Brunt, J.W., P. McCartney, K. Baker & S.G. Stafford. 2002. The future of ecoinformatics in long term ecological research. In: Callaos, N. (ed.) Proceedings of the 6th World Multiconference on Systemics, Cybernetics and Informatics: SCI 2002; July 14–18, 2002 Orlando, Florida, USA. Volume 7: Information systems development, II. pp. 367–372. International Institute of Informatics and Systemics, Orlando, FL, US.
- Canhos, V.P., S. Souza, R. Giovanni & D.A.L. Canhos. 2004. Global biodiversity informatics: setting the scene for a “new world” of ecological modelling. *Biodiversity Informatics*, 1: 1-13.
- Dengler, J., J. Ewald, I. Kühn, & R.K. Peet. (2011). Ecoinformatics and global change - an overdue liaison. *Journal of Vegetation Science*, 22(4), 577-581.
- Guralnick, R. & A. Hill. 2009. Biodiversity informatics: automated approaches for documenting global biodiversity patterns and processes. *Bioinformatics*, 25: 421-428.
- Harris G. 1986. Phytoplankton ecology: structure, function, and fluctuation. Chapman and Hall. 384 pp.
- INSPIRE. 2008. Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards Metadata.
- ISO. 2003. 19115:2003 - Geographic information - Metadata. ISO, Switzerland, 2003.
- Jones, M.B., C. Berkley, J. Bojilova & M. Schildauer. 2001. Managing Scientific Metadata. *IEEE Internet Computing*, 3: 59-68.
- Kareiva, P. 2001. Ecoinformatics: facilitating access to existing data sets. *Trends in Ecology and Evolution*, 16: 266.
- Kliment, T. & A. Oggioni. 2011. Bringing Eco-Biological Metadata to the INSPIRE metainformation world. GI2011-X-border-SDI/GDI-Symposium, 2011, Bad Schandau, Germany.
- Michener, W.K., J.W. Brunt, J.J. Helly, T.B. Kirchner & S.G. Stafford. 1997. Non geospatial Metadata for the Ecological Sciences. *Ecological Applications*, 7(1): 330-342.
- Michener, W.K., J. Porter, M. Servilla & K. Vanderbilt. 2011. Long term ecological research and information management. *Ecological Informatics*, 6(1), 13-24.
- Mirtl, M. & K. Krauze. 2007. Developing a new strategy for environmental re-search, monitoring and management: The European Long-Term Ecological Research Network’s (LTER-Europe) role and perspectives, in: Chmielewski, T.J. (ed.), Nature conservation management - From idea to practical results. ALTER-Net. Lublin-Lodz-Hesinki-Aarhus, ISBN 83-87414-98-0. pp. 36-52.
- Mirtl, M., M. Boamrane, L. Braat, E. Furman, K. Krauze, M. Frenzel, V. Gaube, E. Groner, A. Hester, S. Klotz, W. Los, I. Mautz, J. Peterseil, A. Richter, H. Schentz, K. Schleidt, M. Schmid, A. Sier, J. Stadler, R. Uhel, M. Wildenberg, S. Zacharias. 2009. LTER-Europe Design and Implementation Report - Enabling “Next Generation Ecological Science”: Report on the design and implementation phase of LTER-Europe un-der ALTER-Net & management plan 2009/2010. Umweltbundesamt (Federal Environment Agency Austria). Vienna. 220 pp. ISBN 978-3-99004-031-7.
- Mirtl, M. 2010. Introducing the next generation of ecosystem research in Europe: LTER-Europe’s multi-functional and multi-scale approach, in: Müller, F., C. Baessler, H. Schubert, & S. Klotz (eds), Long-term ecological research: between theory and application. Springer, Dordrecht. 456 pp. ISBN: 978-90-481-8781-2.
- Nativi, S., C. Massimo & J. Pearlman. 2012. The Brokering Approach for Multidisciplinary Interoperability: A Position Paper. *International Journal of Spatial Data Infrastructures Research*, 7: 1-15.
- Nogueras-Iso, J., F. Zarazaga-Soria, P.R. Muro-Medrano. 2005. Geographic Information Metadata for Spatial Data Infrastructures: Resources, Interoperability and Information Retrieval. Springer.
- Odum E. & G. Barrett. 2005. Fundamentals of ecology. Thomson Brooks/Cole. 598 pp.
- Page B. & V. Wohlgemuth. 2010. Advances in environmental informatics: Integration of discrete event simulation methodology with ecological material flow analysis for modelling eco-efficient systems. International Conference on Ecological Informatics and Ecosystem Conservation (ISEIS 2010), Procedia Environmental Sciences, 2: 696-705.
- Peterseil, J., T. Kliment, A. Oggioni, H. Schentz & P. Carrara. 2011. Report on data store in DB. 45 pp.
- Pullin, A.S. & N. Salafky. 2010. Save the Whales? Save the Rainforest? Save the Data! *Conservation Biology*, 24(4): 915-917. DOI: 10.1111/j.1523-1739.2010.01537.x
- Recknagel, F. 2011. Ecological informatics: a discipline in the making. *Ecological Informatics*, 6: 1-3.
- Reichman, O.J., M.B. Jones & M.P. Schildauer. 2011. Challenges and opportunities of open data in ecology. *Science*, 331 (6018): 703-705.
- Soberón, J. & T. Peterson. 2004. Biodiversity informatics: managing and applying primary biodiversity data. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 359: 689-698.
- Telenius, A. 2011. Biodiversity information goes public: GBIF at your service. *Nordic Journal of Botany*, 29: 378-381, doi: 10.1111/j.1756-1051.2011.01167.x