

# GeoHarP: la completa armonizzazione dei dati geografici

Cristoforo Abbattista, Vincenzo Barbieri

Planetek Italia s.r.l., Via Massaua 12, I-70123 Bari, Italia, [abbattista@planetek.it](mailto:abbattista@planetek.it) , [barbieri@planetek.it](mailto:barbieri@planetek.it)

## Sintesi

L'introduzione di sistemi WebGIS per la pubblicazione di dati cartografici ha reso possibile la consultazione di un numero crescente di archivi geografici. Questi archivi generalmente presentano caratteristiche disomogenee, che rendono difficile il loro utilizzo in modo sinergico (*overlay, query*). **GeoHarP (GEOgraphic HARmonizing Platform)** è un sistema web che consente l'integrazione e l'armonizzazione di dati geografici disponibili *on-line*, attraverso procedure di conversione "al volo". La piattaforma consente di accedere ad archivi cartografici remoti, serviti con tecnologie, sistemi di riferimento e proiezione differenti e di renderli disponibili alla consultazione in modo normalizzato e conforme alle richieste dell'utente. GeoHarP è in grado di normalizzare anche le caratteristiche semantiche dei dati provenienti da fonti differenti.

Le principali caratteristiche di GeoHarP sono:

- la disponibilità di algoritmi di conversione *on-line* e *on-the-fly*;
- la capacità di integrazione semantica tra diversi database;
- l'accesso trasparente e non invasivo verso i sistemi che pubblicano i dati.

GeoHarP permette di passare ad una piena interoperabilità tra sistemi differenti eliminando il lavoro di trasferimento dei dati geografici da un sistema ad un altro e, grazie ai meccanismi di conversione *on-line*, si evitano le laboriose procedure di *batch* di conversione dei dati (archiviati secondo differenti modalità). Il sistema consente la conversione al volo, da un sistema di proiezione all'altro (Gauss-Boaga, UTM, Geografico), senza risultare invasivo nei confronti dei server che rendono disponibili gli archivi.

## Abstract

The introduction of WebGIS systems distributing cartographic data is allowing the consultation of an increasing number of geographic archives. These archives are generally based on non homogeneous features and standards, making their synergic use a complex task (*overlay, query*).

**GeoHarP (GEOgraphic HARmonizing Platform)** is a web based system that allows integration and harmonization across geographic data available *on-line*, through its "*on-the-fly*" conversion procedures. The platform is conceived to access to remote cartographic archives, served by different technologies, reference systems and projection, and to make it ready to the consultation in compliance with standards and customer needs. GeoHarP allows also standardization of semantic characteristics of data coming from different sources.

The main features of GeoHarP are:

- the availability of conversion algorithms *on-line* and *on-the-fly*;
- the ability to realize a semantic integration between different databases;
- the transparent and not invasive access towards geographic data distributing systems.

GeoHarP allows a full interoperability across different systems, skipping the need of geographic data transfer from a system to another and, thanks to its *on-line* conversion procedure, avoiding the ponderous data batch conversion procedures (when data are archived with different system).

The system allows the conversion *on-the-fly*, from one projection to another (Gauss-Boaga, UTM, Geographic), without asking any intervention on those *server* that makes archives available.

### Introduzione e Scenario attuale dei sistemi distribuiti

Gli archivi di dati geografici vivono, oggi, un periodo di grande sviluppo ed uno dei principali obiettivi è garantire la massima condivisione degli stessi dati, in modo da rispondere alle necessità di ogni tipologia d'utenza. GeoHarP è una piattaforma sviluppata per il pieno raggiungimento di questo obiettivo.

Lo scenario attuale dei sistemi di archiviazione di dati GIS ed EO, vede coinvolti una serie numerosa di attori, istituzionali e non, cui si affianca un'altrettanto numerosa serie di sistemi, che ne permettono la pubblicazione. Le tecnologie maggiormente presenti sul mercato hanno i nomi di ESRI ArcIMS, Autodesk MapGuide, Intergraph Geomedia WebMap, MapInfo MapXtreme, UMN MapServer, ecc.

I fornitori di tecnologia, pur avendo la necessità di garantirsi una quota del mercato, sono ben consci dei requisiti di integrazione dei loro clienti, che molto spesso sono chiamati a gestire un substrato tecnologico molto eterogeneo. Ognuno di essi, infatti, ha recentemente introdotto una serie di componenti, che rendesse il proprio sistema conforme alle specifiche OGC (Open GeoSpatial Consortium), identificando in queste specifiche la risposta alle richieste di interoperabilità degli utenti dei sistemi GIS.

Un primo livello di integrazione dei sistemi è perciò garantito da questa scelta architettonica dei fornitori, il che restringe le problematiche dell'utente all'interoperabilità a livello di dato geografico. Un sistema che voglia rispondere a tali esigenze, deve analizzare e risolvere due criticità, insite nei dati cartografici:

- L'eterogeneità dei sistemi di riferimento del dato cartografico;
- L'assenza di uno standard per la codifica delle informazioni alfanumeriche collegate al dato geografico, ovvero l'assenza di una semantica comune.

La prima criticità trova le proprie origini in motivazioni di tipo storico e geografico, così che in Italia, ad esempio, sussistono tre diverse proiezioni cartografiche (UTM, Gauss-Boaga, Geografica) e tre differenti Datum (Roma40, ED50, WGS84), per ognuno dei due fusi in cui è diviso il territorio italiano. La loro combinazione da origine a nove differenti modi per rappresentare mappe e quasi tutti sono attualmente usati nei diversi ambiti amministrativi italiani. Se dovessimo allargare l'analisi al territorio europeo, la questione si complicherebbe ulteriormente. Un sistema che voglia garantire l'integrazione di dati così diversi, deve essere capace di convertirli in tempo reale (*on-the-fly*) senza richiedere che essi siano preventivamente allineati alle specifiche di progetto, attraverso laboriose procedure *batch*, che porterebbero ad una notevole perdita di tempo e denaro: due risorse notoriamente scarse. Tale conversione deve valere sia per dati vettoriali (temi puntuali, lineari, poligonali), che raster (ortofoto, mappe satellitari, carte tecniche, DEM, IGM), ed è proprio la conversione dei dati raster che costituisce la sfida maggiore.

La seconda criticità rappresenta la capacità di capire il significato delle informazioni memorizzate. Ogni database ha, infatti, un proprio differente modello di dati, che nasce dalle analisi e dalle necessità di progetto, così che il termine "strada" potrebbe avere differenti significati per ogni database coinvolto nell'integrazione: un indirizzo, o l'intera via comprensiva di larghezza,



**Figura 1: GeoHarP: Overview funzionale**

lunghezza, tipologia, ecc. C'è la necessità di avere un dizionario semantico, che regoli tale rappresentazione dell'informazione.

La piattaforma WebGIS GeoHarP è stata definita per rendere interoperabili sistemi differenti per:

- 1) Tecnologia;
- 2) Dislocazione fisica;
- 3) Sistema di riferimento del dato cartografico;
- 4) Caratteristiche semantiche del dato alfanumerico.

Inoltre il sistema è caratterizzato da due proprietà fondamentali:

- 1) è non invasivo rispetto ai server integrati;
- 2) tutto il processing avviene *on-the-fly* senza nessun tipo di *pre-processing*.

### L'architettura di GeoHarP

La piattaforma GeoHarP è costituita da quattro componenti di base:

- un **server GeoHarP**, che memorizza e gestisce le informazioni dei diversi sistemi da interconnettere;
- uno **user client GeoHarP** (*browser based*), che accede e visualizza le informazioni usando le informazioni ottenute dal server;
- un **administration client GeoHarP** (*browser based*), che configura e gestisce l'anagrafica dei server remoti, il catalogo cartografico virtuale e il dizionario semantico;
- un insieme di Internet Map Server (IMS) dislocati nella rete, detti **nodi GeoHarP**.

Essendo GeoHarP totalmente trasparente anche ai fornitori di dati cartografici, quindi non invasivo nei confronti dei propri nodi, essi non necessitano di installare alcun modulo software aggiuntivo per essere parte della rete GeoHarP. Questo è reso possibile dalle architetture *Open* con cui sono implementati i più diffusi IMS ed in particolare quelli per cui si è scelto di implementare il *driver* in GeoHarP.

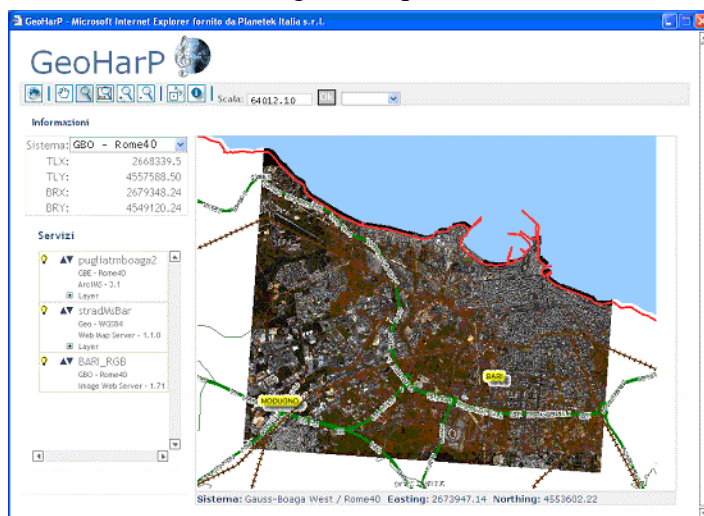
### Modulo Client

Il client GeoHarP in esecuzione sulla macchina client è un'applicazione WebGIS che, utilizzando un comune browser web, permette all'utente di scegliere tra una serie di cartografie, raster e vettoriali, presenti a catalogo. Indipendentemente da dove queste cartografie risiedono, dall'IMS che le fornisce, dal SRS (Spatial Reference System) con cui sono memorizzate, l'utente potrà visualizzarle e sovrapporle, per trarne le indicazioni di cui necessita.

Come già accennato, all'utente è totalmente trasparente:

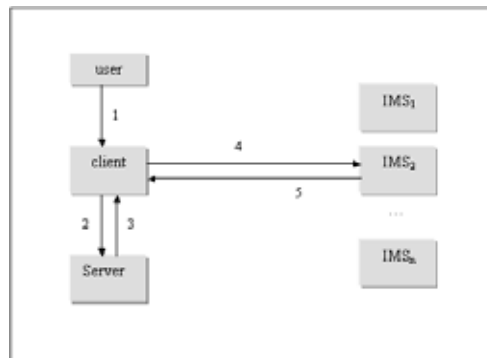
- qual'è il SRS nativo con cui l'informazione è memorizzata sul server cartografico (è l'utente a scegliere, di volta in volta, come vuole che l'informazione gli venga fornita);
- chi è il *Data Provider*;
- con quale tecnologia il dato viene fornito per la consultazione.

Attraverso il modulo client, l'utente è in grado di sfruttare tutte le funzionalità di consultazione dei dati cartografici, che sono quelle tipiche di applicazioni WebGIS come, *pan*, *zoom*, *identify*, cambio scala, vai a X/Y, visualizza X/Y, più la funzionalità innovativa di cambio SRS.



**Figura 2: Esempio di interfaccia WebGIS.  
La mappa raster di sfondo è un'immagine  
IKONOS ad 1 metro di risoluzione**

Nello stesso modulo risiede la logica necessaria all'integrazione geografica e semantica dell'informazione. Questo tipo di architettura (*thick client*) è stata ottenuta attraverso l'implementazione di un'*applet java*, per quello che riguarda il componente **mappa**, e logica *Javascript* per l'interazione col *browser*. L'applicazione garantisce, perciò, la possibilità di avere un server snello, cui non sono richieste grandi capacità prestazionali. Alcuni IMS offrono il cambio di SRS della mappa visualizzata rispetto al SRS con cui la mappa è memorizzata, a fronte di un'opportuna configurazione del server stesso da parte dell'amministratore. Si verifica però che le prestazioni di tali sistemi decadono molto velocemente al crescere del numero degli utenti connessi, mentre decentrando il processamento sui client non si è affetti da tali problemi. Nel contempo se si vuole sfruttare tale funzionalità, è possibile configurare l'anagrafica di ogni singolo nodo di GeoHarP affinché si sfruttino tutte le potenzialità offerte dal nodo.



**Figura 3: Schema di accesso ai dati**

Le informazioni che il *client* deve reperire sono: indirizzo IP del server, tecnologia dell'IMS, tipologia della cartografia, sistema di riferimento.

Attualmente GeoHarP è in grado di gestire le seguenti combinazioni Proiezione, Datum, Fuso:

Proiezione	Datum	Fusi
UTM	WGS84	32,33
UTM	ED50	32,33
Gauss-Boaga	Roma40	Est, Ovest
Geografica	WGS84	
Geografica	ED50	
Geografica	Roma40	

Le tecnologie gestite, ossia i protocolli di scambio informazione, sono:

- *ArcXML*: cartografia **vettoriale** servita da ArcIMS di ESRI;
- *WMS (Web Map Server) 1.1.1*: cartografia **vettoriale** servita da IMS conformi alle specifiche OGC;
- *ECWP (Enhanced Compression Wavelet Protocol)*: cartografia **raster** servita da Image Web Server di ER Mapper.



**Fig. 4: layer "autostrade" confrontato con l'ortofoto in due visioni con layer acceso e spento**

La combinazione delle scelte effettuate per tecnologia e SRS, racchiude gran parte dei Sistemi Informativi Territoriali attualmente esistenti in Italia. Pertanto, usando GeoHarP come piattaforma di interoperabilità, si è in grado di integrare gran parte dell'esistente patrimonio cartografico, senza alcuna necessità di migrazione del dato o della tecnologia.

L'utente non percepisce in nessun modo l'interazione con un'architettura distribuita e deve solamente scegliere in quale SRS, tra quelli disponibili, consultare il progetto creato, potendolo cambiare in ogni momento.

Una serie di procedure, interne al client WebGIS, effettua tutte le conversioni, *on the fly*, di cui il dato necessita. Tali procedure, basate sul componente *CConvert*, software per la conversione tra i sistemi di coordinate sviluppato da Planetek Italia, sono ottimizzate e tarate, inoltre, sulla risoluzione della mappa richiesta. Infatti un'elevata precisione degli algoritmi di conversione del dato può essere resa del tutto superflua dalla bassa risoluzione a schermo dell'immagine. Pertanto non è sensato processare i dati per ottenere precisioni dell'ordine di 10 metri, se la risoluzione a schermo dell'immagine è di 100 m per pixel.

Per ottenere una mappa, con un determinato *Bounding Box* da visualizzare a video attraverso uno specifico SRS (ad esempio UTM32-ED50), se la cartografia è memorizzata in Gauss-Boaga, bisogna:

- 1) identificare il rettangolo in Gauss-Boaga che contiene il rettangolo in UTM32-ED50;
- 2) richiedere la mappa, corrispondente a questo nuovo box, con la tecnologia propria del nodo GeoHarP;
- 3) Estrarre dalla mappa ottenuta in UTM-ED50 i pixel necessari alla costruzione della mappa in Gauss-Boaga per il rettangolo richiesto, applicando le opportune leggi di trasformazione (modello affine) e di interpolazione (*nearest value*).

Analoghi meccanismi si realizzano per eseguire query spaziali come, ad esempio, l'operazione di richiesta informazioni alfanumeriche di un tematismo in una determinata coordinata. In questa fase il client deve eseguire l'armonizzazione semantica richiesta dal sistema. Infatti il client deve:

- 1) convertire il punto da interrogare da UTM-ED50 a Gauss-Boaga;
- 2) richiedere le informazioni, corrispondenti a questo nuovo punto, con la tecnologia propria del nodo GeoHarP;
- 3) visualizzare le informazioni ottenute, armonizzandone la semantica degli attributi e normalizzandole secondo il dizionario memorizzato sul server GeoHarP.

Entrambe le operazioni descritte sono svolte per ogni nodo della rete che fornisce dati al progetto analizzato dall'utente, ma in entrambi i casi, questo livello di granularità è reso totalmente trasparente.

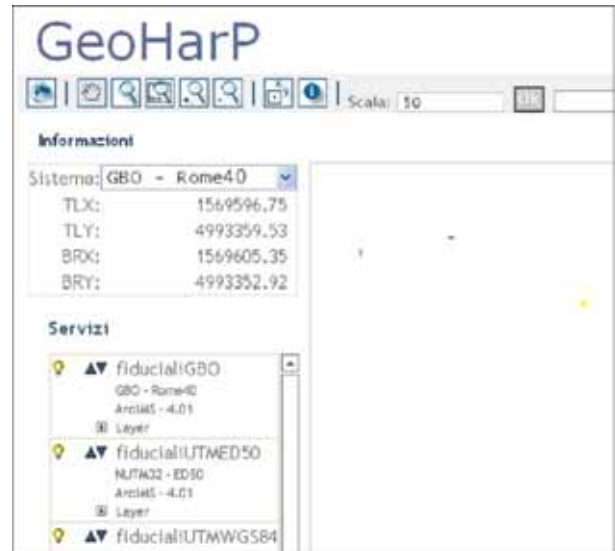
La validazione degli algoritmi di conversione implementati, è stata eseguita, analizzando i risultati ottenuti su cartografia vettoriale acquisita sul campo nei tre differenti Datum (ED50, WGS84, Roma40). Sono stati creati tre layer cartografici, ognuno nel proprio sistema di riferimento, con tre tematismi diversi: una barra orizzontale di 7 pixel, una verticale di sette pixel ed un cerchio pieno di diametro 7 pixel. Si è valutato con GeoHarP fino a che scala fosse evidente all'utente un errore di registrazione. Ognuno dei tre temi, di tipo puntuale, è stato pubblicato su tre server differenti, ed è stato reso consultabile attraverso un progetto GeoHarP. Si è verificato che fino alla scala 1:100 l'utente non è in grado di percepire l'errore di registrazione. Sotto tale scala si riscontrano gli inevitabili limiti degli algoritmi implementati.

Inoltre se si effettua l'interrogazione dei tre layer cartografici con lo strumento idoneo del client GeoHarP, indipendentemente dal SRS usato in consultazione, per tutti e tre i layer vengono riportate le informazioni corrette.





**Figura 5: Alla scala di 1:20.000 i punti sono ancora perfettamente sovrapposti**



**Fig. 6: Alla scala 1:50 si notano i tre punti non registrati con distanze intorno ai 2, 3 m**

### Modulo di Amministrazione

Il modulo di amministrazione risiede presso il server GeoHarP ed è utile per svolgere tutte le operazioni necessarie al popolamento del catalogo cartografico virtuale, da cui l'utente potrà selezionare i layer da consultare.

L'amministratore ha il compito di inserire nel sistema le seguenti informazioni:

- Internet Map Server remoti
  - o Indirizzo IP
  - o Tecnologia
- Cartografia remota
  - o Tipologia
  - o Scala di acquisizione
  - o Informazioni sul Copyright
  - o Sistema di riferimento nativo
- Dizionario semantico
  - o Aggiunta, modifica, eliminazione dei termini del proprio dizionario interno
  - o Gestione delle associazioni tra attributi delle mappe catalogate e termini del dizionario interno

Poiché l'architettura progettata fa esclusivo riferimento a cartografia servita da IMS remoti, l'amministratore può in ogni momento invocare la sincronizzazione delle proprie informazioni con quelle effettivamente residenti presso i server remoti, con uno strumento opportunamente predisposto, evitando così che il catalogo virtuale risulti obsoleto rispetto alle informazioni reali.

### Conclusioni

La piattaforma GeoHarP può essere funzionale ad un'ampia schiera di utenti e di applicazioni, soprattutto in ambiti dove il settore cartografico è presente da più anni, e ha visto l'avvicinarsi di tecnologie e metodologie differenti per la definizione del proprio repertorio cartografico. Un caso esemplare è fornito dalla pubblica amministrazione, dove vi sono anche esigenze di interoperabilità tra la cartografia fornita a diversi livelli, da quello comunale, fino a quello nazionale, al fine di raggiungere quegli obiettivi di cooperazione dettati dall'iniziativa INSPIRE. GeoHarP è concepito come modulo di sistemi come l'applicazione Cart@net Multiserver e Multistandard, sviluppato da Planetek Italia per interconnettere le informazioni geografiche della rete SINAnet dell'APAT.