

Musume

Nuova piattaforma di riconoscimento indirizzi

*Java e “Lucene” come strumenti open source
per il riconoscimento, la georeferenziazione e la geocodifica di indirizzi*

Version: 1.1.6 -

Date: 20/05/2013

Sommario

Il progetto “Musume”	3
Obiettivi - Open Source e territorio	3
L’architettura	3
Il motore di ricerca	3
Creare l’applicazione di riconoscimento con Lucene	4
L’Analizzatore (Analysis)	5
Il processo di indicizzazione (Index Tokens)	5
Ricerca i dati indicizzati (Search Index)	5
Visualizzare i dati ricercati	6
Aspetti Tecnici	7
La base dati territoriale	7
Tracciato File dei Dati Territoriali delle Località	8
Tracciato File dei Dati Territoriali delle Strade	9
Tracciato File dei Dati Territoriali dei Civici	10
La ricerca dei dati	13
Scenari possibili in output	14
Musume come normalizzatore interattivo	17
Musume come normalizzatore batch	20
APPENDICE “A”	33
APPENDICE “B”	42

Il progetto “Musume”

Obiettivi - Open Source e territorio

L’obiettivo di questo progetto è quello di realizzare uno strumento di riconoscimento di un indirizzo in una data località e restituirne i dati normalizzati, le informazioni geografiche (coordinate spaziali), le informazioni Istat (codifiche territoriali e statistiche) e informazioni postali (CAP e denominazioni).

Questo progetto, denominato “MUSUME” (“Figlia” in giapponese), nasce dall’esigenza di usufruire di uno strumento per l’uso territoriale potente e svincolato da architetture proprietarie costose e poco elastiche alle innovazioni e al supporto.

Infatti, come vedremo più avanti, gli strumenti che verranno utilizzati per la sua realizzazione sono **open source**.

L’architettura

L’architettura di *Musume* si suddivide in tre componenti principali :

La base dati territoriale

La normalizzazione, geocodifica e georeferenziazione con “Lucene”

L’interfaccia utente

Il motore di ricerca

Il motore di ricerca che verrà utilizzato per accedere ai dati sarà il **LUCENE** , il motore open source di “Apache Software Foundation”.

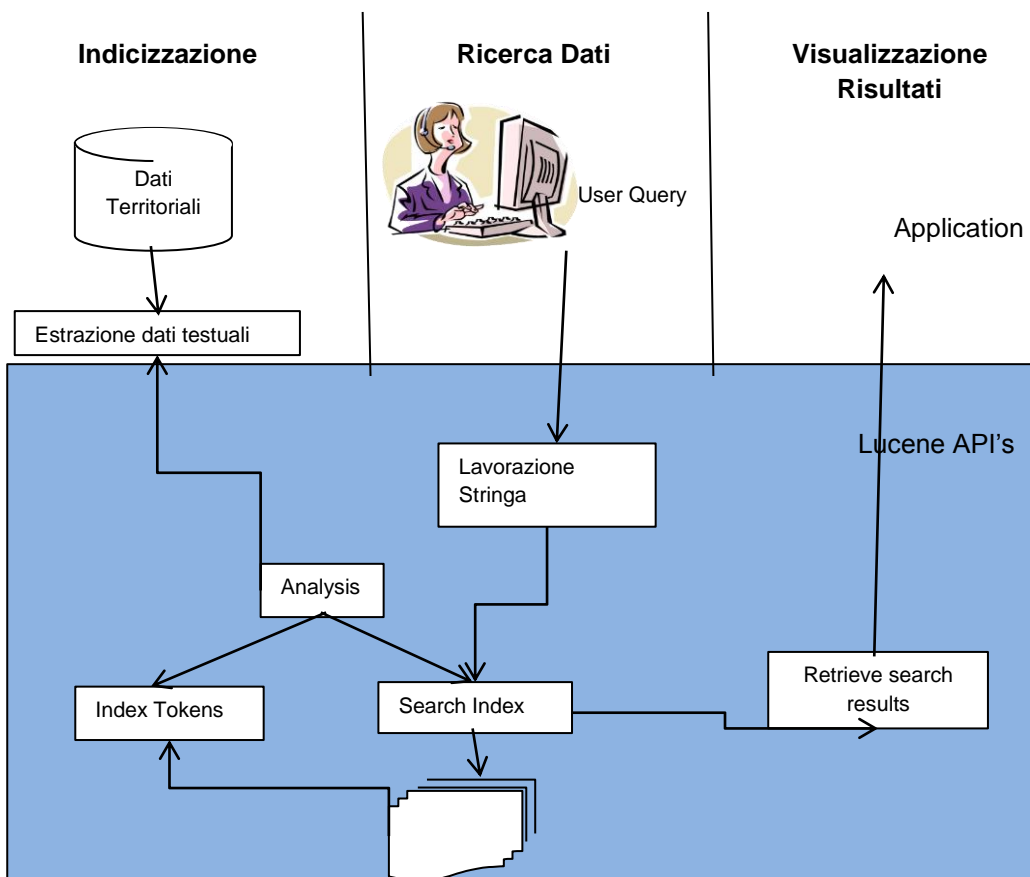
E’ una libreria per ricerche di tipo testuale, scalabile e ad alte prestazioni che permette di aggiungere indici e capacità di ricerca alle proprie applicazioni (viene utilizzato da Wikipedia, LinkedIn, Enciclopedia Britannica, FedEx, ecc.).

Creare l'applicazione di riconoscimento con Lucene

Come mostrato nella figura seguente ,costruire un'applicazione completa con il Lucene implica l'estrazione , l'analisi, l'indicizzazione, la ricerca e la visualizzazione dei risultati di ricerca. I dati territoriali estratti, convertiti e arricchiti verranno indicizzati con un'organizzazione ad hoc (che vedremo piu' avanti) in modo che nel risultato finale le informazioni ottenute siano quelle richieste;

quindi il disegno degli indici dei dati rappresenterà il nodo principale del progetto, perché permetterà alle logiche di ricerca di portare in output le informazioni volute nel modo voluto.

Di seguito una illustrazione sintetica del processo di Indicizzazione/ricerca/visualizzazione dei dati



L'Analizzatore (Analysis)

Il processo di analisi converte i dati testuali nelle unità fondamentali di ricerca, denominate "termini" (tokens) eliminando tutto ciò che rimane; infatti durante l'analisi i dati vengono sottoposti a diverse operazioni : l'estrazione dei termini, la rimozione delle parole comuni, lo scarto delle punteggiature, la conversione delle parole in minuscolo, ecc.

L'analisi viene effettuata sia durante le indicizzazioni e sia durante l'elaborazione delle query (query parsing).

Tutti i termini (tokens) estratti verranno aggiunti all'indice di Lucene.

In Lucene esistono diversi tipi di analizzatori (SimpleAnalyzer, StandardAnalyzer, StopAnalyzer, SnowballAnalyzer, ecc.) che differiscono in base alla tipologia di estrazione dei termini da indicizzare e dei filtri da applicare.

La fase di analisi è un passo fondamentale nella qualità finale del risultato e dev'essere appropriato alla tipologia di dati trattati, più avanti vedremo nel dettaglio questa fase.

Il processo di indicizzazione (Index Tokens)

L'indicizzazione è il processo di conversione dei dati di testo in un formato che ne consenta una rapida ricerca, il prodotto dell'indicizzazione è denominato indice (index).

Un esempio della logica di questo processo è l'indice che si trova alla fine di un libro: Questo indice punta al percorso di argomenti che compaiono nel libro.

Prima che nuovi dati vengano aggiunti all'indice, questi devono essere processati dall'analizzatore.

Ricerca i dati indicizzati (Search Index)

La ricerca è il processo che guarda nell'indice per le parole per poi accedere ai documenti che le contengono.

Il processo di ricerca avviene dopo la preparazione delle query , e data la natura della ricerca e dei dati trattati e' necessario personalizzarle e combinarle tra loro (in modo booleano).

Una FuzzyQuery permette di estrarre le informazioni di indirizzi scritti male, come "Via Gariboldi" anziche' "Via Garibaldi";

una PrefixQuery permette di estrarre i termini per radice; ecc.

Verra' implementata una ricerca per fonetismi, dove in alcuni casi di ambiguita' l'arbitro della scelta sara' proprio l'assonanza dei termine.

Visualizzare i dati ricercati

Il processo di ricerca restituira' un elenco di riferimenti classificati dei termini ricercati (uno o piu' candidati), con tutti gli attributi territoriali :

Coordinate geografiche del civico

Codifiche Istat

Codice avviamento postale

Diciture ufficiali di DUG e toponimi

L'interfaccia di interrogazione del software di riconoscimento verra' realizzata in Java, i parametri di input consteranno nelle canoniche informazioni territoriali di localita' , indirizzo, cap, e livello di precisione della ricerca (questo andra' ad incidere nella scelta della tipologia di ricerca da utilizzare nel motore).

Musume verra' esposto all'utente attraverso un servizio WEB di tipo http/SOAP, e' prevista anche una versione batch via Web Service e a console.

Aspetti Tecnici

La base dati territoriale

I dati da indicizzare saranno strutturati su due strati distinti, uno basato sui dati territoriali di Regione Toscana, con lettura da tabelle Sqlite, l'altro basato sui dati del fornitore , trasformati in archivi sequenziali.

I dati di Regione Toscana verranno estratti dalle tabelle dei comuni, strade, civici e cippi chilometrici.

I dati territoriali Valuelab trasformati in archivi sequenziali basati sulle coordinate geografiche dei singoli numeri civici.

Per questi ultimi la trasformazione degli shape files (a polilinea) in civici georeferenziabili (point) avverrà tramite una applicazione che effettuerà conversioni di tipo geometrico basandosi sulle informazioni contenute nelle matrici territoriali degli stessi file e negli attributi leggibili (quali civici, se pari o dispari, ecc.).

Ciascun civico verrà arricchito nei suoi attributi dai dati postali e Istat in modo da ricreare una sorta di documento completo per ciascun punto.

I dati territoriali ValueLab una volta pronti verranno scaricati in tre flussi sequenziali, mentre i dati di Regione Toscana verranno acceduti direttamente nel database durante il processo di indicizzazione.

Tracciato File dei Dati Territoriali delle Località

Keys	Fields	Descrizione	Trattamento Campo per indicizzazione
*	Comune ricerca	Testo nome comune utilizzato per ricerca	Index .Analyzed

*	provincia	Sigla provincia	Index .Not Analyzed
	Commune	Descrizione Comune	Index .Not Analyzed
	codiceregioneistat	Codice Regione Istat	No Index
	codiceprovinciaistat	Codice Provincia Istat	No Index
	codicecomuneistat	Codice Comune Istat	No Index
	Codicefrazione	Codice Frazione	No Index
*	cap	Codice di avviamento Postale (2010)	Index .Not Analyzed
	fonte	Fonte origine del dato (Istat, VL, RT)	No Index
	Xlocalita	Coordinata X del centroide di localita'	No Index
	Ylocalita	Coordinata Y del centroide di localita'	No Index

Tracciato File dei Dati Territoriali delle Strade

Keys	Fields	Descrizione	Trattamento Campo per indicizzazione
*	codistat	Codici Istat Completi (regione+provincial+comune)	Index .Not Analyzed
	codiceregioneistat	Codice Regione Istat	Index .Not Analyzed

	codiceprovinciaistat	Codice Provincia Istat	Index .Not Analyzed
	codicecomuneistat	Codice Comune Istat	Index .Not Analyzed
*	dug	D.U.G. denominazione urbanistica generalizzata	Index .Analyzed
*	toponimo	Nome strada	Index .Analyzed
*	Indirizzo Completo	DUG + Toponimo	Index .Analyzed
	Via Abbreviata	Nome Strada abbreviato	Index .Not Analyzed
	codicestrada	Codice identificativo della strada	Index .Not Analyzed
	macrodig	Famiglia di appartenenza DUG (1=vie e simili;2=piazza e simili;3=localita' e simili;4=salite e simili)	No Index
	Xstrada	Coordinata X	No Index
	Ystrada	Coordinata Y	No Index
	Latitudine	Latitudine	No Index
	Longitudine	Longitudine	No Index

Tracciato File dei Dati Territoriali dei Civici

Keys	Fields	Descrizione	Trattamento Campo per indicizzazione
*	codistat	Codici Istat Completi (regione+provincial+comune)	Index .Not Analyzed
*	codicestrada	Codice identificativo della strada	Index .Not Analyzed
	codiceregioneistat	Codice Regione Istat	Index .Not Analyzed
	codiceprovinciaistat	Codice Provincia Istat	Index .Not Analyzed
	codicecomuneistat	Codice Comune Istat	Index .Not Analyzed
	dug	D.U.G. denominazione urbanistica generalizzata	Index .Analyzed
	toponimo	Nome strada	Index .Analyzed
*	civico	Civico in formato numerico	Index .Not Analyzed
	Via abbreviata	Indirizzo Completo (dug + toponimo)	No Index
	sezioneistat	Codice Sezione Censuaria 2001 Istat	No Index
	Xsezione	Coordinata X centroide di sezione Istat utm 32	No Index

	Ysezione	Coordinata Y centroide di sezione Istat utm 32	No Index
	xsezioneGB	Coordinata X centroide di sezione Istat (GAUSS BOAGA)	No Index
	YsezioneGB	Coordinata Y centroide di sezione Istat (GAUSS BOAGA)	No Index
	Latitudine	Coordinata Geografica decimale civico (latitudine)	No Index
	longitudine	Coordinata Geografica decimale civico (longitudine)	No Index
	Xcivico	Coordinata X del civico utm 32	No Index
	ycivico	Coordinata Y del civico utm 32	No Index
	xgaussboaga	Coordinata X del civico (GAUSS BOAGA)	No Index
	ygaussboaga	Coordinata Y del civico (GAUSS BOAGA)	No Index
	cap	Codice di avviamento Postale (2010)	No Index
	Min Civico	Minimo numero civico riferito alla strada	No Index
	Max Civico	Massimo numero civico riferito alla strada	No Index
	Fonte	Fonte di origine del civico (VL=VALUELAB,RT=Regione Toscana)	No Index

I flussi creato vengono così indicizzati e analizzati tramite le classi

org.regione.toscana.musume.CreaIndiceLoc, **org.regione.toscana.musume.CreaIndiceStrada** e **org.regione.toscana.musume.CreaIndiceCiv** che utilizzano le API di Lucene per l'operazione.

L'analisi non viene effettuata su tutti i campi mono termine e numerici che costituiscono l'indice perché inutili da analizzare (la fase di analisi prepara i termini da memorizzare nell'indice scegliendo quelli necessari e scartando a seconda dell'analizzatore quelli non utili ai fini della ricerca).

Al momento l'analisi viene effettuata tramite lo StandardAnalyzer, l'analizzatore di default (e' in studio la possibilità di utilizzare un analizzatore più specifico per la lingua italiana, come ad esempio SnowballAnalyzer che si presta alle lingue europee).

I campi indicizzati sono quelli utili alla "normalizzazione", ossia :

- Indirizzo (dug + toponimo)
- Civico
- Comune
- Sigla provincia
- DUG (come alternativa all'indirizzo completo, da usare con il toponimo separato)
- Toponimo (da usare con DUG per ricerche alternative all'indirizzo completo)

La ricerca dei dati

La ricerca dei dati indicizzati avviene, a console, tramite l'utilizzo la classe **org.regione.toscana.musume.Geocode** dove si combinano diverse tipologie di queries in modo da rendere "intelligenti" gli accessi sugli indici e ottenere il risultato atteso.

Per la modalità Web Service interattivo viene chiamata la classe **org.regione.toscana.musume.GeocodeWS** che lavora su stringhe in input e in output .

Per la modalità Web Service batch viene chiamata la classe **org.regione.toscana.musume.MusumeBatch** che si occupa della gestione del file di input, dell'estrazione e lavorazione dei singoli indirizzi e della scrittura dell'output.

La ricerca avviene in tre fasi :

1. Accesso ai dati Località – viene effettuato il riconoscimento del comune e si estraggono i campi chiave e i vari attributi
2. Accesso ai dati Strada – viene effettuato il riconoscimento di dug e toponimo, con un filtro basato sui codici istat del comune trovato
3. Accesso ai dati dei Civici – Viene effettuato un filtro basandosi sui codici istat del comune, sul codice strada selezionato e sul civico uguale o più vicino al civico di input

Si distinguono 3 livelli di “normalizzazione” :

1. Ricerca per “termini esatti” , modalita' di ricerca con “term query” (certo)
2. Ricerca per “termini simili” , modalita' di ricerca con “fuzzy query” , (probabile)
3. Ricerca per “radice” , modalita' di ricerca con “prefix query”

I dati da utilizzare per la ricerca verranno scelti da un opportuno parametro che imposterà la fonte (Regione Toscana o Valuelab).

N.B. La scelta della fonte ricade nella scelta del tipo di informazioni che si vogliono ottenere in output, i dati Valuelab (fonte = 2) contengono informazioni di geocodifica Istat 2001 (cella e coordinate del centroide di sezione) che al momento non sono presenti nei dati di Regione Toscana (fonte = 1), d'altro canto i dati territoriali di quest'ultima sono molto piu' precisi di quelli Value Lab per via delle acquisizioni operate da Enti Locali anche in modo manuale, inoltre contengono localita' non presenti nelle cartografie standard.

Scenari possibili in output

Le stringhe di input vengono sottoposte ad una lavorazione specifica del trattamento territoriale, in modo da formare le query di accesso ai dati in modo ottimale.

La lavorazione consiste in una serie di algoritmi , come ad esempio la separazione della dug dal toponimo, o del civico dall'indirizzo, oppure trattare parole speciali come il presso (c/o), ecc.

Viene effettuata una prima ricerca sull'indice delle località per estrapolare i codici istat del comune, la tabella contiene tutti i comuni e tutte le frazioni della Toscana.

Il risultato della prima ricerca può restituire tre risultati :

- comune non trovato (output di indirizzo non riconosciuto)
- più comuni trovati (verrà proposta la lista dei candidati)
- comune trovato (verranno estratti i codici istat per la ricerca successiva)

In caso di unico candidato trovato nella ricerca dei comuni , verrà effettuata l'accesso all'indice delle strade, i risultati possibili saranno :

- strada non trovata (output con i dati della sola località)
- più strade trovate (verrà proposta la lista dei candidati)
- strada trovata (verrà estratto il codice strada per la ricerca successiva)

In caso di unico candidato nella ricerca delle strade verra' effettuato l'accesso all'indice dei civici con i codici istat del comune e il codice strada di appartenenza, i risultati possibili saranno :

- Civico trovato (output con tutti gli attributi geografici e istat)
- Civico non trovato, ma presenti altri civici sulla stessa strada (output con i dati del civico piu' vicino)
- Nessun civico trovato per la strada (output con i soli dati della strada e del comune)

Dei parametri di input e' possibile omettere alcune delle variabili in modo da allargare i candidati ottenuti, ad esempio e' possibile omettere la DUG o il civico per ottenere rispettivamente tutte le stesse toponomastiche e tutti i civici.

Il livello "certo" si ottiene utilizzando la seguente struttura nelle query di ricerca :

DUG : TermQuery su ogni termine che la compone , tutte in "should" (equivalente all'"OR" booleano)

Toponimo : TermQuery su ogni termine, tutte in "must" ("AND" booleano)

Civico : TermQuery in "Should"

Comune : TermQuery in "Must"

Sigla Provincia : TermQuery in "Must"

Il livello "probabile" :

DUG : FuzzyQuery (con similarita' 0.4f) su ogni termine che la compone , tutte in "should"

Toponimo : FuzzyQuery (con similarita' 0.4f) su ogni termine che la compone , tutte in "Must"

Civico : TermQuery in "Should"

Comune : FuzzyQuery (con similarita' 0.6f) su ogni termine che la compone , tutte in "Must"

Sigla Provincia : TermQuery in "Must"

La ricerca per radice :

DUG : PrefixQuery su ogni termine che la compone , tutte in “must”

Toponimo : PrefixQuery su ogni termine, tutte in “must”

Civico : PrefixQuery in “Should”

Comune : PrefixQuery in “Must”

Sigla Provincia : PrefixQuery in “Must”

Del risultato ottenuto verra' analizzato lo score attribuito da Lucene in modo da applicare un primo filtro al risultato che potrebbe presentare piu' candidati possibili, lo score piu' alto rappresenta il candidato corretto, piu' score uguali delle ambiguita', a valle dell'analisi dello score, il risultato ottenuto dalla ricerca verra' analizzato tramite algoritmo fonetico METAPHONE (codec open source di conversione del suono di pronuncia in stringa) e tramite Soundex (altro codec simile al metaphone) in modo da eliminare dal range di risultati le strade e le localita' che sicuramente non appartengono alla richiesta iniziale ma che potrebbero essere presenti per via della ricerca a similitudine in fuzzy logic.

Prima di arrivare ad effettuare la ricerca, le stringhe di input vengono trattate in modo tale da essere utilizzate nel miglior modo possibile considerando gli elementi che compongono un indirizzo, ossia COMUNE, DUG, TOPONIMO e CIVICO.

Per questo scopo gli input verranno analizzati da un programma che si occupera' esclusivamente della separazione DUG / TOPONIMO, poi da un altro programma che estrarra' il civico (parte numerica e non).

Nella preparazione delle stringhe si controllano i nomi di persona (da un database con oltre 30.000 record), le preposizioni , i numeri romani , ecc. per fare in modo da dare il giusto peso ad ogni singolo termine nella ricerca del dato stradale.

Ad esempio, nello scrivere “Via di Novoli” le parole importanti sono “Via” e “Novoli”, dove la prima e' DUG e la seconda toponimo, quest'ultima verra' considerata in AND booleano mentre la preposizione “DI” contera' poco nella determinazione del risultato in quanto OR booleano.

Musume come normalizzatore interattivo

La normalizzazione degli indirizzi in modalità interattiva verrà effettuata tramite web service sincrono, dove il motore nelle operazioni di riconoscimento sarà costituito dalla classe :

org.regione.toscana.musume.GeocodeWS

mentre la struttura dei messaggi XML di invio e ricezione sono strutturati secondo l'XML schema (XSD) visibile in appendice "B".

il WSDL del servizio è visualizzabile in appendice "A".

Un esempio di interrogazione via browser del servizio è'

<http://webtrial.rete.toscana.it/normaws/Musume.jws?method=richiesta&algoritmo=1&provincia=fi&comune=firenze&indirizzo=via+roma+124&cap=00000&fonte=2>

che produrrebbe a video il seguente SOAP envelope :

```
<?xml version="1.0" encoding="ISO-8859-1" ?>
_ <soapenv:Envelope xmlns:soapenv="http://schemas.xmlsoap.org/soap/envelope/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
_ <soapenv:Body>
_ <richiestaResponse
  soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/">
<richiestaReturn href="#id0" />
  </richiestaResponse>
_ <multiRef id="id0" soapenc:root="0"
  soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  xsi:type="ns1:RispostaNormalizzataType"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
  xmlns:ns1="http://tipi.musume.toscana.regione.org">
<choiceValue href="#id1" />
<indirizzoAmbiguo xsi:type="ns1:IndirizzoAmbiguo" xsi:nil="true" />
<indirizzoRiconosciuto href="#id1" />
<indirizzoScartato xsi:type="ns1:IndirizzoScartato" xsi:nil="true" />
```

```

<localitaAmbigua xsi:type="ns1:LocalitaAmbigua" xsi:nil="true" />
<localitaNormalizzata xsi:type="ns1:LocalitaNormalizzata" xsi:nil="true" />
<localitaScartata xsi:type="ns1:LocalitaScartata" xsi:nil="true" />
<tipoRispostaNorm xsi:type="xsd:string">0</tipoRispostaNorm>

</multiRef>

_ <multiRef id="id1" soapenc:root="0"
  soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  xsi:type="ns2:IndirizzoRiconosciuto"
  xmlns:ns2="http://tipi.musume.toscana.regione.org"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/">

<datiGeoreferenziazioneInd href="#id2" />
<datiNormalizzazioneInd href="#id3" />
<descrizione xsi:type="xsd:string">Indirizzo Riconosciuto</descrizione>

</multiRef>

_ <multiRef id="id2" soapenc:root="0"
  soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  xsi:type="ns3:DatiGeoreferenziazioneInd"
  xmlns:ns3="http://tipi.musume.toscana.regione.org"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/">

<XCivicoPuntualeGB xsi:type="xsd:string">1681470</XCivicoPuntualeGB>
<XIstat xsi:type="xsd:string">1681470</XIstat>
<YCivicoPuntualeGB xsi:type="xsd:string">4849123</YCivicoPuntualeGB>
<YIstat xsi:type="xsd:string">4849123</YIstat>
<latitudine xsi:type="xsd:string">43.77279217</latitudine>
<longitudine xsi:type="xsd:string">11.25446322</longitudine>

</multiRef>

_ <multiRef id="id3" soapenc:root="0"
  soapenv:encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  xsi:type="ns4:DatiNormalizzazioneInd"
  xmlns:ns4="http://tipi.musume.toscana.regione.org"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/">

<cap xsi:type="xsd:string">50100</cap>
<civico xsi:type="xsd:string">124</civico>

```

```
<codiceComuneIstat xsi:type="xsd:string">017</codiceComuneIstat>
<codiceProvinciaIstat xsi:type="xsd:string">048</codiceProvinciaIstat>
<codiceRegioneIstat xsi:type="xsd:string">09</codiceRegioneIstat>
<codiceToponimo xsi:type="xsd:string">0017550198757</codiceToponimo>
<dug xsi:type="xsd:string">VIA</dug>
<fonte xsi:type="xsd:string">ISTAT</fonte>
<localita xsi:type="xsd:string">FIRENZE</localita>
<provincia xsi:type="xsd:string">FI</provincia>
<sezioneIstat xsi:type="xsd:string">0000172</sezioneIstat>
<toponimo xsi:type="xsd:string">ROMA</toponimo>
</multiRef>
</soapenv:Body>
</soapenv:Envelope>
```

Lo stesso servizio e' richiamabile tramite la realizzazione di un client specifico.

Musume come normalizzatore batch

La normalizzazione batch degli indirizzi verra' effettuata tramite web service asincrono, dove il motore nelle operazioni di riconoscimento sara' costituito dalla classe :

org.regione.toscana.musume.MusumeBatch

mentre l'interfaccia verra' gestita da una classe java che si occupera' della lettura del file di input, dell'estrazione dei campi interessati alla normalizzazione e della creazione dell'output.

L'elaborazione batch utilizzerà i dati cartografici di Regione Toscana, o in alternative i dati cartografici ValueLab (l'utilizzo dell'una o dell'altra sara' evidenziato dal tag "fonte" in output).

Gli input e gli output saranno dinamici nella forma, e sara' l'utilizzatore a sceglierne il tipo dalla finestra del client , ad esempio :

Sistemi Informativi Georeferenziati

NORMALIZZAZIONE BATCH : preparazione ed invio dati











utente abilitato: invio

Selezionare il file da inviare	F:\musume\Musume.1. <input type="button" value="Sfoglia..."/>	Inserire un identificativo per la richiesta (max 15 caratteri)	<input type="text" value="test"/>	
Selezionare il tipo di input	<input type="text" value="File di Testo"/>	Selezionare il tipo di output	<input type="text" value="File di testo Standard"/>	Fonte Dati: <input type="text" value="Regione Toscana"/>

Regione Toscana - Settore Sistemi informativi documentari e servizi per l'accesso al pubblico

L'output si potra' monitorare, scaricare o consultare dalla pagina stile "file manager" che si presentera' ad invio avvenuto (**l'esecuzione sara' terminata quando il file con estensione log non avra' piu' la dimensione a 0 bytes**) :

Download Files Elaborati

Check all <input type="checkbox"/>	Nome File  	Type	Size   	Ultima Modifica  
<input checked="" type="checkbox"/>	 test_20120320112604.txt	file	104 KB	Mar 20, 2012 11:26:17 AM
<input checked="" type="checkbox"/>	 test_20120320112604_stat.log	file	1 KB	Mar 20, 2012 11:26:17 AM
<input checked="" type="checkbox"/>	 test_20120320112604SCA.txt	file	17 KB	Mar 20, 2012 11:26:17 AM
Azione sui files selezionati				
<input type="button" value="ZipDownload"/>				

Tramite i pulsanti ad **X** si potranno cancellare i files, tramite i checkbox selezionare quelli da zippare e poi farne il download.

I formati previsti sono :

Input

Sono previste tre tipologie di dati di input al Web Service Batch :

[XML Standard](#)

[File Ascii sequenziale](#)

XML Standard

Il file XML standard riporta i tag dei dati da normalizzare e georeferenziare senza avere informazioni aggiuntive; affinché possa essere riconosciuto dal Web Service è necessario che i tag dei dati abbiano dei nomi precisi, di seguito un esempio di XML Standard :

```
<?xml version="1.0" encoding="UTF-8"?>
```

```
<richiesta>
```

```
  <indirizzo>via roma 1</indirizzo>
```

```
  <comune>firenze</comune>
```

```
  <provincia>fi</provincia>
```

```
  <cap>50100</cap>
```

```
    <codice>1</codice>
```

```
</richiesta>
```

I tag richiesta, indirizzo, comune, provincia, cap, codice sono tutti obbligatori;

il tag codice dovrà riportare una chiave identificativa del record.

L'XML Standard può essere rappresentato da più tag **RICHIESTA**, che dovranno essere tutti tag principali che rappresenteranno ciascuno un record da trattare.

File ASCII Sequenziale

L'input di tipo Ascii sequenziale dovrà rispettare una struttura con una sequenza ben definita dei campi da normalizzare su ciascun record sequenziale, separati dal carattere PIPE ("|"), esempio :

VIA ROMA 1 | FIRENZE | FI | 53100 | 00001

La prima posizione è rappresentata dall'indirizzo completo, la seconda dal comune, la terza dalla sigla provincia, la quarta dal CAP e in ultima posizione il codice identificativo del record da trattare.

Output

Sono previste 4 tipologie di dati di output dal Web Service Batch :

XML Standard

File Ascii sequenziale

XML Cartografico

File Ascii sequenziale Cartografico

XML Standard

Il file XML Standard riporta i tag dei dati normalizzati e georeferenziati seguendo questa struttura (contenuta in un tag "record" per ciascun record elaborato) :

Tag name	Descrizione
chiave	Chiave che lega il record di input (e' anche presente come attributo del tag radice denominato "record")

indirizzo_input	Indirizzo di input
comune_input	Comune di input
provincia_input	Sigla provincia di input
cap_input	Cap di input
esito	Esito (0=Indirizzo Riconosciuto, 8=Localita' riconosciuta, 1=Indirizzo Scartato, 7=Indirizzo Ambiguo, 6=Localita' ambigua)
descrizione_esito	Descrizione Esito (Indirizzo Riconosciuto, Localita' riconosciuta, Indirizzo Scartato, Indirizzo Ambiguo, Localita' ambigua)
comune	Comune
provincia	Sigla provincia
cap	Cap
DUG	D.U.G.
Toponimo	Toponimo
Civico	Numero civico
Codice Regione Istat	Codice Regione Istat
Codice Provincia Istat	Codice Provincia Istat
Codice Comune Istat	Codice Comune Istat
Est GB	Coordinata X del civico (gauss boaga)
Nord GB	Coordinata Y del civico (gauss boaga)
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
Distanza Civico	La distanza del civico di input (se mancante e' 0) dal civico piu' vicino trovato in cartografia (considerando la parita')

Codice Sezione Istat	Codice sezione di censimento istat (2001)
X sezione	Coordinata x del centroide di sezione istat
Y sezione	Coordinata Y del centroide di sezione istat
Codice Toponimo	Campo identificativo della strada nel database territoriale di riferimento
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=ValueLab)
Codice strada	Il codice identificativo del toponimo nel database territoriale di riferimento

Esempio :

```

<?xml version="1.0" encoding="UTF-8"?>
<record chiave="1035">
  <chiave>1035</chiave>
  <indirizzo_input>Piazza IV Novembre</indirizzo_input>
  <comune_input>Sesto Fiorentino</comune_input>
  <provincia_input>FI</provincia_input>
  <cap_input> </cap_input>
  <esito>0</esito>
  <descrizione_esito>Indirizzo Riconosciuto</descrizione_esito>
  <comune>SESTO FIORENTINO</comune>
  <provincia>FIRENZE</provincia>
  <cap>50019</cap>
  <dug>PIAZZA</dug>
  <toponimo>IV NOVEMBRE</toponimo>
  <civico>0</civico>
  <codiceregioneistat>09</codiceregioneistat>
  <codiceprovinciaistat>048</codiceprovinciaistat>
  <codicecomuneistat>043</codicecomuneistat>
  <codicesezistat2001>0000007</codicesezistat2001>
  <xsezione_gb>1676884</xsezione_gb>
  <ysezione_gb>4855725</ysezione_gb>
  <xcivico_gb>1676887</xcivico_gb>
  <ycivico_gb>4855757</ycivico_gb>
  <latitudine>43.83358828</latitudine>
  <longitudine>11.19974964</longitudine>
  <diffcivico>4</diffcivico>
  <fonte>VL </fonte>
  <codice_strada>0007940119400</codice_strada>
</record>

```


File ASCII Sequenziale

L'output di tipo Ascii sequenziale sarà rappresentato da una struttura fissa, con un record di testata che identifica i campi e dai record formattati secondo una sequenza precisa, e si avranno tre files :

Record normalizzati

Record Scartati

Statistiche elaborative

Tracciato record normalizzati :

Nome Colonna	Descrizione
Esito	0=Indirizzo riconosciuto ,1=Indirizzo Scartato e localita' riconosciuta
Record di Input	Indirizzo, comune, sigla provincia, cap e campo chiave di input
Sigla Provincia	Comune
Comune	Sigla provincia
DUG	D.U.G.
Toponimo	Toponimo
Civico	Numero civico
Codice Regione Istat	Codice Regione Istat
Codice Provincia Istat	Codice Provincia Istat
Codice Comune Istat	Codice Comune Istat
CAP	CAP
Est GB	Coordinata X del civico (gauss boaga)

Nord GB	Coordinata Y del civico (gauss boaga)
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
Distanza Civico	La distanza del civico di input (se mancante e' 0) dal civico piu' vicino trovato in cartografia (considerando la parita')
Codice Sezione Istat	Codice sezione di censimento istat (2001)
X sezione	Coordinata x del centroide di sezione istat
Y sezione	Coordinata Y del centroide di sezione istat
Codice Chiave	Chiave identificativa del record di input
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=ValueLab)
Codice Toponimo	Campo identificativo della strada nel database territoriale di riferimento
Descrizione Esito	La descrizione dell'esito della normalizzazione

Esempio :

|0|VIALE SANTA VITTORIA, 52 |FOIANO DELLA CHIANA |AR|
|AR1A001007|AR|FOIANO DELLA CHIANA |VIA |S. VITTORIA
|52 |09|051|018|52045|1727113|4797103|43.29236801|11.79925003|-50
|0000022|1727113|4797103|VL |AR1A001007|0017550207036 |INDIRIZZO RICONOSCIUTO
|

Tracciato record scartati :

Nome Colonna	Descrizione
Esito	1=Indirizzo Ambiguo ,1=Localita' scartata, 1=Localita' ambigua , 1=Ambiguita' (elemento ambiguo figlio del record con Indirizzo Ambiguo o Localita' ambigua)

Record di Input	Indirizzo, comune, sigla provincia, cap e campo chiave di input
Sigla Provincia	Comune
Comune	Sigla provincia
DUG	D.U.G.
Toponimo	Toponimo
Civico	Numero civico
Codice Regione Istat	Codice Regione Istat
Codice Provincia Istat	Codice Provincia Istat
Codice Comune Istat	Codice Comune Istat
CAP	CAP
Coordinata X	Coordinata X del civico (gauss boaga)
Coordinata Y	Coordinata Y del civico (gauss boaga)
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
Distanza Civico	La distanza del civico di input (se mancante e' 0) dal civico piu' vicino trovato in cartografia (considerando la parita')
Codice Sezione Istat	Codice sezione di censimento istat (2001)
X sezione	Coordinata x del centroide di sezione istat
Y sezione	Coordinata Y del centroide di sezione istat
Codice Chiave	Chiave identificativa del record di input
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=Valuelab)
Codice Toponimo	Campo identificativo della strada nel database territoriale di riferimento
Descrizione Esito	La descrizione dell'esito della normalizzazione

es. :

```
|1|FRAZIONE RIGUTINO SUD 119 |AREZZO |AR|
|AR1A03400E| | | | | | | | | |VL |AR1A03400E| |INDIRIZZO
| | | | | | | | | | | | | | | | | | | | | | | |
AMBIGUO
```

Gli indirizzi ambigui, avranno consecutivamente nel file il grappolo delle ambiguità georeferenziate e normalizzate, in modo da facilitare il processo di recupero del dato.

N.B. E' stato applicato un filtro che NON riporta eventuali ambiguità non georeferenziate nel grappolo di possibili scelte.

Es. (riferito all'esempio precedente)

```
|1|LOCALITA' RIGUTINO SUD,119 |AREZZO
|AR| |AR1A03400E| | |LOCALITA' |RIGUTINO SUD
|119 | | | | |1732386|4806427|43.37461287|11.86811806|0
|0000658|1732678|4806041|VL |AR1A03400E|0005900195791 |AMBIGUITA'
|
```

Analogamente agli indirizzi, anche le localit  ambigue avranno i grappoli dei comuni trovati.

XML Cartografico

Il file XML Cartografico riporta i tag dei dati normalizzati e georeferenzati seguendo questa struttura (contenuta in un tag "record" per ciascun record elaborato) :

Tag name	Descrizione
chiave	Chiave che lega il record di input (e' anche presente come attributo del tag radice denominato "record")
indirizzo_input	Indirizzo di input
comune_input	Comune di input
provincia_input	Sigla provincia di input
cap_input	Cap di input
esito	Esito (0=Indirizzo Riconosciuto, 8=Localita' riconosciuta, 1=Indirizzo Scartato, 7=Indirizzo Ambiguo, 6=Localita' ambigua)
descrizione_esito	Descrizione Esito (Indirizzo Riconosciuto, Localita' riconosciuta, Indirizzo Scartato, Indirizzo Ambiguo, Localita' ambigua)
x	Coordinata X
y	Coordinata Y
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=ValueLab)
Codice strada	Il codice identificativo del toponimo nel database territoriale di riferimento

Esempio :

```

<?xml version="1.0" encoding="UTF-8" ?>
_ <record chiave="1035">
<chiave>1035</chiave>
<indirizzo_input>Piazza IV Novembre</indirizzo_input>
<comune_input>Sesto Fiorentino</comune_input>
<provincia_input>FI</provincia_input>
<cap_input />
<esito>0</esito>
<descrizione_esito>Indirizzo Riconosciuto</descrizione_esito>

```

```

<xcivico_gb>1676887</xcivico_gb>
<ycivico_gb>4855757</ycivico_gb>
<latitudine>43.83358828</latitudine>
<longitudine>11.19974964</longitudine>
<fonte>VL</fonte>
<codice_strada>0007940119400</codice_strada>
</record>

```

File ASCII Sequenziale Cartografico

L'output di tipo Ascii sequenziale Cartografico sarà rappresentato da una struttura fissa, con un record di testata che identifica i campi e dai record formattati secondo una sequenza precisa, e si avranno tre files :

Record normalizzati

Record Scartati

Statistiche elaborative

Tracciato record normalizzati :

Nome Colonna	Descrizione
Esito	0=Indirizzo riconosciuto ,1=Indirizzo Scartato e localita' riconosciuta
Record di Input	Indirizzo, comune, sigla provincia, cap e campo chiave di input
Coordinata X	Coordinata X del civico (gauss boaga)
Coordinata Y	Coordinata Y del civico (gauss boaga)
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=Valuelab)
Codice Chiave	Chiave identificativa del record di input

Codice Toponimo	Campo identificativo della strada nel database territoriale di riferimento
Descrizione Esito	La descrizione dell'esito della normalizzazione

Esempio :

```
|0|Piazza IV Novembre                               |Sesto
Fiorentino                                           |FI|      |1035
|1676887|4855757|43.83358828|11.19974964|VL      |1035      |0007940119400
|INDIRIZZO RICONOSCIUTO
```

Tracciato record scartati :

Nome Colonna	Descrizione
Esito	1=Indirizzo Ambiguo ,1=Localita' scartata, 1=Localita' ambigua , 1=Ambiguita' (elemento ambiguo figlio del record con Indirizzo Ambiguo o Localita' ambigua)
Record di Input	Indirizzo, comune, sigla provincia, cap e campo chiave di input
Coordinata X	Coordinata X del civico (gauss boaga)
Coordinata Y	Coordinata Y del civico (gauss boaga)
Latitudine	Latitudine (geografica decimale) del civico
Longitudine	Longitudine (geografica decimale) del civico
fonte	Fonte dati cartografici (RT=Regione Toscana; VL=Valuelab)
Codice Chiave	Chiave identificativa del record di input
Codice Toponimo	Campo identificativo della strada nel database territoriale di riferimento
Descrizione Esito	1=Indirizzo Ambiguo ,1=Localita' scartata, 1=Localita' ambigua , 1=Ambiguita' (elemento ambiguo figlio del record con Indirizzo Ambiguo o Localita' ambigua)

es. :

```
|1|Piazza IV Novembre                               |Sesto  
Fiorentino                                           |FI|      |1035      |      |      |      |VL  
|1035      |      |INDIRIZZO AMBIGUO
```

Gli indirizzi ambigui, avranno consecutivamente nel file il grappolo delle ambiguità già georeferenziate e normalizzate, in modo da facilitare il processo di recupero del dato.

```
|0000658|1732678|4806041|VL      |AR1A03400E|0005900195791      |AMBIGUITA'  
|
```

Analogamente agli indirizzi, anche le località ambigue avranno i grappoli dei comuni trovati.

Lo scopo finale è quello di fornire una totale georeferenziazione degli indirizzi, per fare questo il risultato delle coordinate sarà a scalare a partire dal civico, se non presente salire via via fino al centroide della località.

Gli step di georeferenziazione saranno (dal più preciso al meno preciso) :

- Civico
- Strada
- Sezione Istat
- Centroidi di località

Essendo un processo Asincrono, il risultato dell'elaborazione sarà disponibile dopo un tempo variabile a seconda del numero di record da trattare e dal formato da gestire.

A fine elaborazione, quando disponibili, i files dell'elaborazione saranno scaricabili da un'apposita finestra nel browser.

Le elaborazioni saranno sempre composti da un file dei record normalizzati, un file dei record scartati (con i dettagli dello scarto) e un file di log dove vengono riassunte le statistiche elaborative.

APPENDICE "A"

WSDL del servizio Musumebatch :

```
<?xml version="1.0" encoding="UTF-8" ?>
_ <wsdl:definitions targetNamespace="http://localhost/retenew/Musumebatch.jws"
  xmlns:apachesoap="http://xml.apache.org/xml-soap"
  xmlns:impl="http://localhost/retenew/Musumebatch.jws"
  xmlns:intf="http://localhost/retenew/Musumebatch.jws"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
  xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
  xmlns:wsdlsoap="http://schemas.xmlsoap.org/wsdl/soap/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  _ <!--
    WSDL created by Apache Axis version: 1.4
    Built on Apr 22, 2006 (06:55:48 PDT)
  -->
  <wsdl:message name="NomeprogettoResponse" />
  <wsdl:message name="TipoIOResponse" />
  _ <wsdl:message name="TipoIORequest">
  <wsdl:part name="tipoIO" type="xsd:string" />
  </wsdl:message>
  _ <wsdl:message name="FontedatiRequest">
  <wsdl:part name="fontedati" type="xsd:string" />
  </wsdl:message>
  _ <wsdl:message name="NomeprogettoRequest">
  <wsdl:part name="nomeprogetto" type="xsd:string" />
  </wsdl:message>
  _ <wsdl:message name="NomeFileinRequest">
  <wsdl:part name="nomefilein" type="xsd:string" />
  </wsdl:message>
  <wsdl:message name="NomeFileinResponse" />
  _ <wsdl:message name="elaboraRequest">
  <wsdl:part name="nomefilein" type="xsd:string" />
  <wsdl:part name="tipoIO" type="xsd:string" />
  <wsdl:part name="nomeprogetto" type="xsd:string" />
  <wsdl:part name="fontedati" type="xsd:string" />
  </wsdl:message>
  _ <wsdl:message name="elaboraResponse">
  <wsdl:part name="elaboraReturn" type="xsd:string" />
  </wsdl:message>
  <wsdl:message name="FontedatiResponse" />
  _ <wsdl:portType name="Musumebatch">
  _ <wsdl:operation name="elabora" parameterOrder="nomefilein tipoIO nomeprogetto
    fontedati">
  <wsdl:input message="impl:elaboraRequest" name="elaboraRequest" />
  <wsdl:output message="impl:elaboraResponse" name="elaboraResponse" />
  </wsdl:operation>
  _ <wsdl:operation name="NomeFilein" parameterOrder="nomefilein">
  <wsdl:input message="impl:NomeFileinRequest" name="NomeFileinRequest" />
```

```

<wsdl:output message="impl:NomeFileinResponse" name="NomeFileinResponse" />
  </wsdl:operation>
_ <wsdl:operation name="TipoIO" parameterOrder="tipoIO">
<wsdl:input message="impl:TipoIORequest" name="TipoIORequest" />
<wsdl:output message="impl:TipoIOResponse" name="TipoIOResponse" />
  </wsdl:operation>
_ <wsdl:operation name="Nomeprogetto" parameterOrder="nomeprogetto">
<wsdl:input message="impl:NomeprogettoRequest" name="NomeprogettoRequest" />
<wsdl:output message="impl:NomeprogettoResponse" name="NomeprogettoResponse" />
  </wsdl:operation>
_ <wsdl:operation name="Fontedati" parameterOrder="fontedati">
<wsdl:input message="impl:FontedatiRequest" name="FontedatiRequest" />
<wsdl:output message="impl:FontedatiResponse" name="FontedatiResponse" />
  </wsdl:operation>
</wsdl:portType>
_ <wsdl:binding name="MusumebatchSoapBinding" type="impl:Musumebatch">
<wsdlsoap:binding style="rpc" transport="http://schemas.xmlsoap.org/soap/http" />
_ <wsdl:operation name="elabora">
<wsdlsoap:operation soapAction="" />
_ <wsdl:input name="elaboraRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
_ <wsdl:output name="elaboraResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/retenew/Musumebatch.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
_ <wsdl:operation name="NomeFilein">
<wsdlsoap:operation soapAction="" />
_ <wsdl:input name="NomeFileinRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
_ <wsdl:output name="NomeFileinResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/retenew/Musumebatch.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
_ <wsdl:operation name="TipoIO">
<wsdlsoap:operation soapAction="" />
_ <wsdl:input name="TipoIORequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
_ <wsdl:output name="TipoIOResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/retenew/Musumebatch.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
_ <wsdl:operation name="Nomeprogetto">
<wsdlsoap:operation soapAction="" />
_ <wsdl:input name="NomeprogettoRequest">

```

```

<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="NomeprogettoResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/retenew/Musumebatch.jws" use="encoded" />
  </wsdl:output>
  </wsdl:operation>
- <wsdl:operation name="Fontedati">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="FontedatiRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="FontedatiResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/retenew/Musumebatch.jws" use="encoded" />
  </wsdl:output>
  </wsdl:operation>
  </wsdl:binding>
- <wsdl:service name="MusumebatchService">
- <wsdl:port binding="impl:MusumebatchSoapBinding" name="Musumebatch">
<wsdlsoap:address location="http://localhost/retenew/Musumebatch.jws" />
  </wsdl:port>
  </wsdl:service>
</wsdl:definitions>

```

WSDL del servizio Musume :

```

<?xml version="1.0" encoding="UTF-8" ?>
- <wsdl:definitions targetNamespace="http://localhost/musume/Musume.jws"
  xmlns:apacheSOAP="http://xml.apache.org/xml-soap"
  xmlns:impl="http://localhost/musume/Musume.jws"
  xmlns:intf="http://localhost/musume/Musume.jws"
  xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
  xmlns:tns1="http://tipi.musume.toscana.regione.org"
  xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/"
  xmlns:wsdlsoap="http://schemas.xmlsoap.org/wsdl/soap/"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  - <!--
    WSDL created by Apache Axis version: 1.3
    Built on Oct 05, 2005 (05:23:37 EDT)
  -->
- <wsdl:types>
- <schema targetNamespace="http://tipi.musume.toscana.regione.org"
  xmlns="http://www.w3.org/2001/XMLSchema">
<import namespace="http://schemas.xmlsoap.org/soap/encoding/" />
- <complexType name="AmbiguitaIndItem">
- <sequence>
<element name="civico" nillable="true" type="xsd:string" />
<element name="indirizzo" nillable="true" type="xsd:string" />
<element name="x" nillable="true" type="xsd:string" />

```

```

<element name="y" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="AmbiguitaInd">
- <sequence>
<element maxOccurs="unbounded" name="ambiguitaIndItem" nillable="true"
  type="tns1:AmbiguitaIndItem" />
<element name="ambiguitaIndItemCount" type="xsd:int" />
  </sequence>
  </complexType>
- <complexType name="IndirizzoAmbiguoType">
- <sequence>
<element name="ambiguitaInd" nillable="true" type="tns1:AmbiguitaInd" />
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="IndirizzoAmbiguo">
- <complexContent>
- <extension base="tns1:IndirizzoAmbiguoType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="DatiGeoreferenziazioneType">
- <sequence>
<element name="XCivicoPuntualeGB" nillable="true" type="xsd:string" />
<element name="XIstat" nillable="true" type="xsd:string" />
<element name="YCivicoPuntualeGB" nillable="true" type="xsd:string" />
<element name="YIstat" nillable="true" type="xsd:string" />
<element name="latitudine" nillable="true" type="xsd:string" />
<element name="longitudine" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="DatiGeoreferenziazioneInd">
- <complexContent>
- <extension base="tns1:DatiGeoreferenziazioneType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="DatiNormalizzazioneIndRiconosciutoType">
- <sequence>
<element name="cap" nillable="true" type="xsd:string" />
<element name="civico" nillable="true" type="xsd:string" />
<element name="codiceComuneIstat" nillable="true" type="xsd:string" />
<element name="codiceProvinciaIstat" nillable="true" type="xsd:string" />
<element name="codiceRegioneIstat" nillable="true" type="xsd:string" />
<element name="codiceToponimo" nillable="true" type="xsd:string" />
<element name="dug" nillable="true" type="xsd:string" />
<element name="fonte" nillable="true" type="xsd:string" />
<element name="localita" nillable="true" type="xsd:string" />
<element name="provincia" nillable="true" type="xsd:string" />
<element name="sezioneIstat" nillable="true" type="xsd:string" />

```

```

<element name="toponimo" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="DatiNormalizzazioneInd">
- <complexContent>
- <extension base="tns1:DatiNormalizzazioneIndRiconosciutoType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="IndirizzoRiconosciutoType">
- <sequence>
<element name="datiGeoreferenziazioneInd" nillable="true"
  type="tns1:DatiGeoreferenziazioneInd" />
<element name="datiNormalizzazioneInd" nillable="true"
  type="tns1:DatiNormalizzazioneInd" />
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="IndirizzoRiconosciuto">
- <complexContent>
- <extension base="tns1:IndirizzoRiconosciutoType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="IndirizzoScartatoType">
- <sequence>
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="IndirizzoScartato">
- <complexContent>
- <extension base="tns1:IndirizzoScartatoType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="AmbiguitaLocItem">
- <sequence>
<element name="localita" nillable="true" type="xsd:string" />
<element name="sigla" nillable="true" type="xsd:string" />
  </sequence>
  </complexType>
- <complexType name="AmbiguitaLoc">
- <sequence>
<element maxOccurs="unbounded" name="ambiguitaLocItem" nillable="true"
  type="tns1:AmbiguitaLocItem" />
<element name="ambiguitaLocItemCount" type="xsd:int" />
  </sequence>
  </complexType>
- <complexType name="LocalitaAmbiguaType">
- <sequence>

```

```

<element name="ambiguitaLoc" nillable="true" type="tns1:AmbiguitaLoc" />
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
</complexType>
- <complexType name="LocalitaAmbigua">
- <complexContent>
- <extension base="tns1:LocalitaAmbiguaType">
<sequence />
  </extension>
</complexContent>
</complexType>
- <complexType name="DatiNormalizzazioneLocRiconosciutaType">
- <sequence>
<element name="cap" nillable="true" type="xsd:string" />
<element name="codiceComuneIstat" nillable="true" type="xsd:string" />
<element name="codiceProvinciaIstat" nillable="true" type="xsd:string" />
<element name="codiceRegioneIstat" nillable="true" type="xsd:string" />
<element name="fonte" nillable="true" type="xsd:string" />
<element name="localita" nillable="true" type="xsd:string" />
<element name="provincia" nillable="true" type="xsd:string" />
<element name="sezioneIstat" nillable="true" type="xsd:string" />
<element name="XIstat" nillable="true" type="xsd:string" />
<element name="YIstat" nillable="true" type="xsd:string" />
  </sequence>
</complexType>
- <complexType name="DatiNormalizzazioneLoc">
- <complexContent>
- <extension base="tns1:DatiNormalizzazioneLocRiconosciutaType">
<sequence />
  </extension>
</complexContent>
</complexType>
- <complexType name="LocalitaNormalizzataType">
- <sequence>
<element name="datiNormalizzazioneLoc" nillable="true"
  type="tns1:DatiNormalizzazioneLoc" />
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
</complexType>
- <complexType name="LocalitaNormalizzata">
- <complexContent>
- <extension base="tns1:LocalitaNormalizzataType">
<sequence />
  </extension>
</complexContent>
</complexType>
- <complexType name="LocalitaScartataType">
- <sequence>
<element name="descrizione" nillable="true" type="xsd:string" />
  </sequence>
</complexType>
- <complexType name="LocalitaScartata">
- <complexContent>

```

```

- <extension base="tns1:LocalitaScartataType">
<sequence />
  </extension>
  </complexContent>
  </complexType>
- <complexType name="RispostaNormalizzataType">
- <sequence>
<element name="choiceValue" nillable="true" type="xsd:anyType" />
<element name="indirizzoAmbiguo" nillable="true" type="tns1:IndirizzoAmbiguo" />
<element name="indirizzoRiconosciuto" nillable="true" type="tns1:IndirizzoRiconosciuto"
  />
<element name="indirizzoScartato" nillable="true" type="tns1:IndirizzoScartato" />
<element name="localitaAmbigua" nillable="true" type="tns1:LocalitaAmbigua" />
<element name="localitaNormalizzata" nillable="true" type="tns1:LocalitaNormalizzata" />
<element name="localitaScartata" nillable="true" type="tns1:LocalitaScartata" />
<element name="tipoRispostaNorm" nillable="true" type="xsd:string" />
  </sequence>
</complexType>
</schema>
</wsdl:types>
- <wsdl:message name="CapRequest">
<wsdl:part name="cap" type="xsd:string" />
  </wsdl:message>
<wsdl:message name="FonteResponse" />
- <wsdl:message name="richiestaRequest">
<wsdl:part name="algoritmo" type="xsd:string" />
<wsdl:part name="indirizzo" type="xsd:string" />
<wsdl:part name="comune" type="xsd:string" />
<wsdl:part name="provincia" type="xsd:string" />
<wsdl:part name="cap" type="xsd:string" />
<wsdl:part name="fonte" type="xsd:string" />
  </wsdl:message>
<wsdl:message name="AlgoritmoResponse" />
<wsdl:message name="IndirizzoResponse" />
- <wsdl:message name="FonteRequest">
<wsdl:part name="fonte" type="xsd:string" />
  </wsdl:message>
<wsdl:message name="CapResponse" />
- <wsdl:message name="ComuneRequest">
<wsdl:part name="comune" type="xsd:string" />
  </wsdl:message>
<wsdl:message name="ProvinciaResponse" />
- <wsdl:message name="AlgoritmoRequest">
<wsdl:part name="algoritmo" type="xsd:string" />
  </wsdl:message>
- <wsdl:message name="ProvinciaRequest">
<wsdl:part name="provincia" type="xsd:string" />
  </wsdl:message>
- <wsdl:message name="richiestaResponse">
<wsdl:part name="richiestaReturn" type="tns1:RispostaNormalizzataType" />
  </wsdl:message>
<wsdl:message name="ComuneResponse" />
- <wsdl:message name="IndirizzoRequest">

```

```

<wsdl:part name="indirizzo" type="xsd:string" />
  </wsdl:message>
- <wsdl:portType name="Musume">
- <wsdl:operation name="Indirizzo" parameterOrder="indirizzo">
<wsdl:input message="impl:IndirizzoRequest" name="IndirizzoRequest" />
<wsdl:output message="impl:IndirizzoResponse" name="IndirizzoResponse" />
  </wsdl:operation>
- <wsdl:operation name="Provincia" parameterOrder="provincia">
<wsdl:input message="impl:ProvinciaRequest" name="ProvinciaRequest" />
<wsdl:output message="impl:ProvinciaResponse" name="ProvinciaResponse" />
  </wsdl:operation>
- <wsdl:operation name="Algoritmo" parameterOrder="algoritmo">
<wsdl:input message="impl:AlgoritmoRequest" name="AlgoritmoRequest" />
<wsdl:output message="impl:AlgoritmoResponse" name="AlgoritmoResponse" />
  </wsdl:operation>
- <wsdl:operation name="Comune" parameterOrder="comune">
<wsdl:input message="impl:ComuneRequest" name="ComuneRequest" />
<wsdl:output message="impl:ComuneResponse" name="ComuneResponse" />
  </wsdl:operation>
- <wsdl:operation name="richiesta" parameterOrder="algoritmo indirizzo comune provincia
  cap fonte">
<wsdl:input message="impl:richiestaRequest" name="richiestaRequest" />
<wsdl:output message="impl:richiestaResponse" name="richiestaResponse" />
  </wsdl:operation>
- <wsdl:operation name="Fonte" parameterOrder="fonte">
<wsdl:input message="impl:FonteRequest" name="FonteRequest" />
<wsdl:output message="impl:FonteResponse" name="FonteResponse" />
  </wsdl:operation>
- <wsdl:operation name="Cap" parameterOrder="cap">
<wsdl:input message="impl:CapRequest" name="CapRequest" />
<wsdl:output message="impl:CapResponse" name="CapResponse" />
  </wsdl:operation>
</wsdl:portType>
- <wsdl:binding name="MusumeSoapBinding" type="impl:Musume">
<wsdlsoap:binding style="rpc" transport="http://schemas.xmlsoap.org/soap/http" />
- <wsdl:operation name="Indirizzo">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="IndirizzoRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="IndirizzoResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="Provincia">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="ProvinciaRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="ProvinciaResponse">

```



```

<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="Algoritmo">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="AlgoritmoRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="AlgoritmoResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="Comune">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="ComuneRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="ComuneResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="richiesta">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="richiestaRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="richiestaResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="Fonte">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="FonteRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />
  </wsdl:input>
- <wsdl:output name="FonteResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
  </wsdl:output>
</wsdl:operation>
- <wsdl:operation name="Cap">
<wsdlsoap:operation soapAction="" />
- <wsdl:input name="CapRequest">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://DefaultNamespace" use="encoded" />

```

```

</wsdl:input>
- <wsdl:output name="CapResponse">
<wsdlsoap:body encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"
  namespace="http://localhost/musume/Musume.jws" use="encoded" />
</wsdl:output>
</wsdl:operation>
</wsdl:binding>
- <wsdl:service name="MusumeService">
- <wsdl:port binding="impl:MusumeSoapBinding" name="Musume">
<wsdlsoap:address location="http://localhost/musume/Musume.jws" />
</wsdl:port>
</wsdl:service>
</wsdl:definitions>

```

APPENDICE "B"

Schema XML dei messaggi di input/output del servizio web musume :

```

<?xml version="1.0" encoding="UTF-8" ?>
- <schema xmlns="http://www.w3.org/2001/XMLSchema" xmlns:tns="http://localhost/"
  targetNamespace="http://localhost/" elementFormDefault="qualified">
  - <!--
  =====
  Definizione dei tipi usati nella Richiesta del servizio
  =====
  -->
<element name="richiesta" type="tns:RichiestaDatiIndirizzoType" />
- <complexType name="RichiestaDatiIndirizzoType">
- <sequence>
- <element name="Algoritmo" type="string">
- <annotation>
<documentation>
1 - Codice corrispondente all'algoritmo di identificazione CERTA
2 - Codice corrispondente all'algoritmo di identificazione AFFIDABILE
3 - Codice corrispondente all'algoritmo di identificazione PROBABILE
</documentation>
  </annotation>
</element>
<element name="Indirizzo" type="string" />
<element name="Comune" type="string" />
<element name="Provincia" type="string" />
<element name="Cap" type="string" />
- <element name="Fonte" type="string">
- <annotation>
<documentation>
1 - Codice corrispondente all'uso dei dati di Regione Toscana
2 - Codice corrispondente all'uso dei dati di ValueLab
</documentation>
  </annotation>
</element>

```

```

</sequence>
</complexType>
  - <!--
    =====
      Definizione dei tipi usati nella Risposta del servizio
    =====
  -->
<element name="risposta" type="tns:RispostaNormalizzataType" />
- <complexType name="RispostaNormalizzataType">
- <choice>
<element name="IndirizzoRiconosciuto" type="tns:IndirizzoRiconosciutoType"
  minOccurs="0" maxOccurs="1" />
<element name="IndirizzoScartato" type="tns:IndirizzoScartatoType" minOccurs="0"
  maxOccurs="1" />
<element name="LocalitaScartata" type="tns:LocalitaScartataType" minOccurs="0"
  maxOccurs="1" />
<element name="LocalitaAmbigua" type="tns:LocalitaAmbiguaType" minOccurs="0"
  maxOccurs="1" />
<element name="IndirizzoAmbiguo" type="tns:IndirizzoAmbiguoType" minOccurs="0"
  maxOccurs="1" />
<element name="LocalitaNormalizzata" type="tns:LocalitaNormalizzataType"
  minOccurs="0" maxOccurs="1" />
  </choice>
- <attribute name="tipoRispostaNorm" type="string" use="required">
- <annotation>
<documentation>
Il codice '0' indica che l'indirizzo e' stato riconosciuto correttamente dal servizio.
Il codice '1' indica che l'indirizzo e' stato scartato dal servizio.
Il codice '2' indica che la localita' e' stata scartata dal servizio.
Il codice '3' indica che i parametri di input sono errati.
Il codice '6' indica il nome della localita' e' ambiguo.
Il codice '7' indica che l'indirizzo fornito al servizio e' ambiguo
Il codice '8' indica che il servizio fornisce solo i dati di normalizzazione e non quelli di
georeferenziazione.
Il codice '9' indica un errore nella lettura dei dati.</documentation>
  </annotation>
  </attribute>
</complexType>
- <complexType name="IndirizzoRiconosciutoType">
- <sequence>
<element name="Descrizione" type="string" fixed="Indirizzo Riconosciuto" />
<element name="DatiNormalizzazioneInd"
  type="tns:DatiNormalizzazioneIndRiconosciutoType" />
<element name="DatiGeoreferenziazioneInd" type="tns:DatiGeoreferenziazioneType" />
  </sequence>
</complexType>
- <complexType name="IndirizzoScartatoType">
- <sequence>
<element name="Descrizione" type="string" fixed="Indirizzo Scartato" />
  </sequence>
</complexType>
- <complexType name="LocalitaScartataType">
- <sequence>

```

```

<element name="Descrizione" type="string" fixed="Localita Scartata" />
  </sequence>
  </complexType>
- <complexType name="LocalitaAmbiguaType">
- <sequence>
<element name="Descrizione" type="string" fixed="Localita Ambigua" />
- <element name="AmbiguitaLoc">
- <complexType>
- <sequence minOccurs="2" maxOccurs="30">
<element name="Localita" type="string" />
<element name="Sigla" type="string" />
  </sequence>
  </complexType>
  </element>
  </sequence>
  </complexType>
- <complexType name="IndirizzoAmbiguoType">
- <sequence>
<element name="Descrizione" type="string" fixed="Indirizzo Ambiguo" />
- <element name="AmbiguitaInd">
- <complexType>
- <sequence minOccurs="2" maxOccurs="30">
<element name="Indirizzo" type="string" />
<element name="Civico" type="string" />
<element name="x" type="tns:SevenDigitStringType" />
<element name="y" type="tns:SevenDigitStringType" />
  </sequence>
  </complexType>
  </element>
  </sequence>
  </complexType>
- <complexType name="LocalitaNormalizzataType">
- <sequence>
<element name="Descrizione" type="string" fixed="Normalizzata la Localita" />
<element name="DatiNormalizzazioneLoc"
  type="tns:DatiNormalizzazioneLocRiconosciutaType" />
  </sequence>
  </complexType>
- <complexType name="DatiNormalizzazioneIndRiconosciutoType">
- <sequence>
<element name="Provincia" type="string" />
<element name="Localita" type="string" />
<element name="Cap" type="string" />
<element name="Dug" type="string" />
<element name="Toponimo" type="string" />
<element name="Civico" type="string" />
<element name="CodiceRegioneIstat" type="string" />
<element name="CodiceProvinciaIstat" type="string" />
<element name="CodiceComuneIstat" type="string" />
<element name="SezioneIstat" type="string" />
<element name="Fonte" type="string" />
<element name="CodiceToponimo" type="string" />
  </sequence>

```

```

    </complexType>
  - <complexType name="DatiNormalizzazioneLocRiconosciutaType">
  - <sequence>
    <element name="Provincia" type="string" />
    <element name="Localita" type="string" />
    <element name="Cap" type="string" />
    <element name="CodiceRegioneIstat" type="string" />
    <element name="CodiceProvinciaIstat" type="string" />
    <element name="CodiceComuneIstat" type="string" />
    <element name="SezioneIstat" type="string" />
    <element name="XSezioneIstat" type="tns:SevenDigitStringType" />
    <element name="YSezioneIstat" type="tns:SevenDigitStringType" />
    <element name="Fonte" type="string" />
  </sequence>
  </complexType>
  - <complexType name="DatiGeoreferenziazioneType">
  - <sequence>
    <element name="X-CivicoPuntuale-GB" type="tns:SevenDigitStringType" />
    <element name="Y-CivicoPuntuale-GB" type="tns:SevenDigitStringType" />
    <element name="X-Istat" type="tns:SevenDigitStringType" />
    <element name="Y-Istat" type="tns:SevenDigitStringType" />
    <element name="Latitudine" type="string" />
    <element name="Longitudine" type="string" />
  </sequence>
  </complexType>
  - <simpleType name="SevenDigitStringType">
  - <restriction base="string">
    <pattern value="[0-9]{7}" />
  </restriction>
  </simpleType>
</schema>

```