

	 3D INFORMATICA	1/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

EXTRAWAY MANUALE TECNICO

	 3D INFORMATICA Extraway technical reference	2/93
03/08/2003		Ver. 0.0.4

1. PRESENTAZIONE 4

2. ARCHITETTURA DEL SISTEMA E FUNZIONAMENTO GENERALE..... 6

3. EXTRAWAY 7

3.1 MAPPATURE 8

3.2 INDICI..... 9

3.3 GERARCHIA DEI NODI PRIMARI 12

3.4 EQUIVALENZA DEGLI INDICI 13

3.5 ELEMENTI TRASPARENTI 14

3.6 DOPPIA INDICIZZAZIONE 15

3.7 CONCATENAZIONE..... 16

3.8 PERSONALIZZAZIONE DEGLI INDICI 17

3.9 THESAURI 18

3.10 ANALISI SPETTRALE 19

3.11 LINGUAGGIO NATURALE 20

3.12 INDICIZZAZIONE 21

3.13 OPZIONI DI RICERCA..... 22

3.13.1 UTILIZZO DELLA STOPLIST 22

3.14 COMPOSIZIONE DELLE FRASI DI RICERCA 24

3.14.1 AND E OR 24

3.14.2 ADJ 25

3.14.3 ESTENSIONE AL CONTORNO SEMANTICO 25

3.14.4 FRASE RACCHIUSA TRA APICI 25

3.14.5 RICERCA SULLE DATE 26

3.14.6 RAFFINAMENTO 26

3.15 AUTOMATISMI..... 27

4. ORGANIZZAZIONE DEI FILE..... 28

4.1 FILE TEMPORANEI 30

5. IL FRAMEWORK 31

6. IMPOSTAZIONI DI CONFIGURAZIONE DI APPLICAZIONE. 33

6.1 REGOLE DI COMPOSIZIONE TITOLI..... 43

7. IMPOSTAZIONI DI CONFIGURAZIONE DI ARCHIVIO. 47

	 3D INFORMATICA Extraway technical reference	3/93
03/08/2003		Ver. 0.0.4

8. GESTIONE DEGLI INDICI.....49

9. ATTIVITÀ DI LOGGING.....51

10. PROGETTAZIONE E CREAZIONE DI UN'APPLICAZIONE54

10.1 CREAZIONE DELLA DATA TYPE DEFINITION (DTD).....55

10.2 JAVA INTERFACE56

10.2.1 RIFERIMENTI PER SVILUPPATORI.....56

10.3 JAVA SERVER WEB TOMCAT80

10.4 TRASFORMAZIONE XML IN HTML81

10.5 ESEMPIO CONVERSIONE ARCHIVIO WEB DOCWAY IN FORMATO XML82

10.5.1 ESPORTAZIONE IN FORMATO XML82

10.5.2 VERSO UNA NUOVA DTD82

10.5.3 TRASFORMAZIONE XML VERSO NUOVA DTD.....83

10.5.4 CREAZIONE DI UN NUOVO ARCHIVIO.....84

10.5.5 CONFIGURAZIONE STRUTTURA ARCHIVIO.....86

10.5.6 CATALOGAZIONE NUOVO ARCHIVIO87

10.6 PERSONALIZZARE L'INTERFACCIA.....89

10.6.1 CONFIGURAZIONE DEL SYSTEM.....89

10.6.2 CONFIGURAZIONE PROPRERTIES DI TOMCAT89

10.6.3 PERSONALIZZAZIONE DTD SUI VERBI.....90

10.6.4 PERSONALIZZARE STILE E TEMPLATE HTML.....90

11. OPERAZIONI AMMINISTRATIVE.....91

12. IMPOSTAZIONI DI CONFIGURAZIONE MODULO XIPS.....92

13 INSTALLAZIONE93

	 3D INFORMATICA	4/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

1. Presentazione

ExtraWay è un motore di ricerca per database in formato XML nativo in grado di memorizzare, indicizzare e ricercare documenti in formato XML nativo.

ExtraWay *consente* indicizzazione testuale, ricerche con operatori booleani, probabilistiche, per somiglianza e per vicinanza di termini;
la definizione della struttura degli indici è completamente *personalizzabile*;
permette la definizione di vocabolari lessicografici, thesauri multilivello e l'analisi spettrale dei termini sia sull'intera base dati che relativamente all'insieme dei documenti in visione;
permette inoltre la visione gerarchica degli oggetti XML memorizzati anche per grandi quantità di elementi;
dispone inoltre di un ampio set di comandi per l'individuazione e la consultazione di porzioni di documenti, sia per grandi insiemi di documenti che per singole unità documentali;
mette inoltre a disposizione moduli aggiuntivi che abilitano all'uso di linguaggio naturale.

ExtraWay *espone inoltre* moduli e funzionalità (framework) per lo sviluppo di applicazioni web XML mediante una struttura multilivello.

ExtraWay offre un servizio multitasking per piattaforme Windows, Linux e Solaris, mentre il framework espone le proprie classi per la connessione, le interrogazioni, un esteso insieme di comandi per la manipolazione dei risultati.

Sono necessarie solo semplici personalizzazioni necessarie a definire i comportamenti e le diversità che contraddistinguono le applicazioni.

Il livello di personalizzazione si riferisce ai file XSL di stile sviluppati appositamente per le singole applicazioni ed alla definizione del relativo filtro ISAPI in grado di convertire, all'occorrenza, il risultato XML in un più comune formato HTML sulla base dei relativi template.

L'utilizzo della tecnologia XML è ideale per la gestione di banche dati di grandi dimensioni, che prevedono di mantenere il proprio formato nel lungo periodo, in quanto XML, oltre ad essere una nuova tecnologia dai promettenti sviluppi futuri, è una tecnologia proposta e diffusa dal W3C come lo standard web di prossima generazione.

La catalogazione delle informazioni avviene mediante una struttura di system dai "nomi parlanti" nella quale vengono registrate le risorse del motore, degli oggetti generici e tipici delle singole applicazioni.

Questo rende più intuitiva la manutenzione e la comprensione generale del meccanismo di funzionamento.

In questa stessa sezione di disco sono archiviate le impostazioni necessarie ad

	 3D INFORMATICA	5/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Extraway per il proprio funzionamento e per la gestione delle banche dati; in appositi file di impostazione sono disponibili e configurabili i comportamenti, le locazioni su disco e quant'altro serve per il corretto funzionamento di tutto l'ambiente.

	 3D INFORMATICA	6/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

2. Architettura del sistema e funzionamento generale

Il funzionamento dell'ambiente prevede l'uso e la comunicazione di diversi moduli. Intervengono la tecnologia Java, ISAPI, server Web, moduli di connessione e smistamento lato server e di conversione verso il formato di output.

Il client, da un'interfaccia web, esegue una richiesta mediante una URL.

Il comando viene intercettato dal server Web sul lato server, che smista la richiesta al filtro ISAPI definito nella URL e nella configurazione del server WEB (IIS, Apache, ...) che trasferirà la richiesta al server web TOMCAT, oppure si può giungere direttamente al server web Tomcat (questo dipende dalla forma della URL).

Sarà comunque il server Web Tomcat che contatterà le classi Java predisposte al confezionamento della request() verso il server Extraway.

Extraway interpreta la richiesta, la elabora e la ritorna al server web Tomcat.

Java predisporrà la response() che, se la richiesta proviene da ISAPI, gli sarà inviata senza apportare trasformazioni, altrimenti confezionerà un output XML, se il browser del client è Microsoft Internet Explorer 6.0 o superiore, oppure un output HTML in caso il browser del client sia diverso da IE 6.0.

Diversamente sarà il modulo ISAPI a predisporre un HTML di risposta al client tramite i template html e i fogli di stile preimpostati.

	 3D INFORMATICA	7/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3. Extraway

Extraway, oltre ad offrire le proprie funzionalità di server per l'accesso ai dati, espone le proprie caratteristiche di Information Retrieval di grande utilità quando si accede a banche dati di notevoli dimensioni.

Di particolare rilevanza:

- § indicizzazione del contenuto dei campi chiave esposto come un vocabolario lessicale (utile per agevolare la ricerca dei documenti);
- § indicizzazione delle singole parole, dell'intera frase (terminata da un CR) o di entrambe (doppia indicizzazione) del contenuto dei campi chiave;
- § generazione e gestione dei vocabolari vincolati e dei thesauri gerarchici (per imporre l'inserimento di dati ai soli termini consentiti);
- § gestione gerarchica degli oggetti xml;
- § modalità di gestione delle ricerche (opzioni di ricerca) che permettono di includere o escludere modalità quali singolare/plurale, maschile/femminile; di prendere in considerazione la vicinanza dei termini indicati in frasi di ricerca, ecc...;
- § estensioni di ricerca e vocabolari al linguaggio naturale (mediante l'uso di particolari thesauri che consentono estensioni verso termini legati alla semantica dei vocaboli utilizzati ed ai concetti che ogni termine identifica in un eventuale apposito albero gerarchico);
- § vocabolari in analisi spettrale (che filtrano il loro contenuto ai termini dei soli documenti appartenenti alla selezione corrente);
- § gestione comune o estesa o concatenata dei vocabolari (comune a più campi o estesa verso il contenuto di altri campi o concatenata al vocabolario di altri campi);
- § estensione delle funzionalità mediante librerie di terzi che compiono operazioni extra per conto di Extraway (vedi librerie Expert System);

Per ExtraWay un database è un insieme di :

- ? Mappature
- ? Indici
- ? Regole di indicizzazione
- ? Regole di processamento

	 3D INFORMATICA	8/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.1 Mappature

In una mappa sono contenuti i riferimenti verso i documenti XML, la loro dimensione, gli autori, la data di ultima modifica, la suddivisione del documento in "nodi primari" ed in rami "potabili" cioè di quei nodi che possono appartenere ad uno o più nodi primari, il cui significato non perde di validità anche se distinto dal nodo primario di appartenenza.

I **nodi primari** sono quelle parti di documenti XML che il creatore del database ha concepito per essere gli elementi tipici del risultato di una ricerca, come le entità di un database relazionale o le righe di tabelle principali.

La caratteristica dei nodi primari è che possono essere assemblati e comunque inseriti nel medesimo file consentendo ad **ExtraWay** di tracciarli quali frammenti di un ampio file di dati XML allo scopo di ottimizzarne il rintracciamento.

Potrebbe essere ad esempio il risultato di una migrazione da un database proprietario, nella quale si tratta di trasferire parecchie migliaia di record verso uno o più file debitamente conformati, raggruppando i dati per annata o per cartelle di destinazione o in altro modo in base alle preferenze dell'amministratore di **ExtraWay**.

Sulla base della conformazione prescelta **ExtraWay** provvederà ad estrarre, assemblare documenti XML, o frammenti di essi, sulla base delle richieste.

I rami "potabili" ("prunable branch") assumono particolare importanza nell'ambito della visualizzazione degli esiti delle ricerche : **ExtraWay** è in grado di eliminare tutte le porzioni "potabili" dei documenti XML che non assumono rilevanza, nelle quali non si individuano riferimenti verso i termini ricercati.

Questo comportamento è maggiormente performante quando la conformazione della struttura dei documenti XML (DTD) è appositamente progettata; e quando questo avviene consente al client di ricevere piccoli file nei quali sono riferiti i rami "potati" che potranno successivamente essere richiesti a completamento.

	 3D INFORMATICA Extraway technical reference	9/93
03/08/2003		Ver. 0.0.4

3.2 Indici

Informazioni di XPath

Gli indici sono puntatori verso ogni singolo termine presente sui documenti XML che compongono i dati.

Data la struttura gerarchica dei documenti XML, gli indici risultano molto più espressivi che rispetto a quelli di un database relazionale.

Ogni singola informazione non solo è indicizzata sulla base del nome della zona o del paragrafo che la contiene, ma anche in base al percorso che la lega alla radice di tutti i documenti XML della base dati.

Supponendo una collezione di documenti XML, la loro indicizzazione è strutturata in un file system in modo che se un file è contenuto gerarchicamente in una sottodirectory, questo è anche logicamente contenuto nel file XML padre che si trova a monte.

Per esempio:

File: /root.xml

```
<?xml version="1.0" encoding="UTF-16"?>
<ead>
  <dsc type="analyticcover">
    <c level="fonds">
      <unittitle>ARCHIVE</unittitle>
      <physdesc>100.000 units</physdesc>
    </c>
  </dsc>
</ead>
```

File: /ARCHIVE/sequence.xml

```
<?xml version="1.0" encoding="UTF-16"?>
<dsc type="analyticcover">
  <c level="fonds">
    <unittitle>XIV-XV centuries</unittitle>
    <physdesc>10.000 units</physdesc>
  </c>
</dsc>
```

File: /ARCHIVE/ XIV-XV centuries/sequence.xml

```
<?xml version="1.0" encoding="UTF-16"?>
<dsc type="analyticcover">
  <c level="fonds">
    <unittitle>XIV century</unittitle>
    <physdesc>4.500 unità</physdesc>
  </c>
</dsc>
```

In questo modo è possibile referenziare, mediante gli indici, il valore di "4.500 unità" attraverso l'intero percorso del DOM XML

	 3D INFORMATICA Extraway technical reference	10/93
03/08/2003		Ver. 0.0.4

/ead/archdesc/dsc/c/dsc/c/dsc/c/physdesc

con l'espressione Xpath

/ead/archdesc/dsc/c/dsc/c/dsc/c/[physdesc='4.500 unità']

o più semplicemente

//C[physdesc='4.500 unità']

I medesimi dati possono essere organizzati anche in un unico file che li racchiude tutti.
Ad esempio:

File: /root.xml

```
<?xml version="1.0" encoding="UTF-16"?>
<ead>
  <dsc type="analyticcover">
    <c level="fonds">
      <unittitle>ARCHIVE</unittitle>
      <physdesc>100.000 units</physdesc>
      ...
    <dsc type="analyticcover">
      <c level="fonds">
        <unittitle>XIV-XV centuries</unittitle>
        <physdesc>10.000 units</physdesc>
        ...
      <dsc type="analyticcover">
        <c level="fonds">
          <unittitle>XIV century</unittitle>
          <physdesc>4.500 unità</physdesc>
        </c>
      </dsc>
    </c>
  </dsc>
</ead>
```

Anche se questa modalità è raramente efficiente su grandi basi dati.

Un'ulteriore opportunità, che consente di mantenere la strutturazione di system delle varie parti di documento XML, consiste nel fare uso dell'istruzione "include" per "inserire" dati dislocati su diversi file all'interno del file XML principale che comunque all'atto dell'utilizzo riporterà alla condizione del file unico.

File: /root.xml

```
<?xml version="1.0" encoding="UTF-16"?>
<ead>
  <dsc type="analyticcover">
    <c level="fonds">
```

	 3D INFORMATICA	11/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

```

<unittitle>ARCHIVE</unittitle>
<physdesc>100.000 units</physdesc>
<!entity & ARCHIVE system "ARCHIVE/sequence.xml" >
  &ARCHIVE;
</c>
</dsc>
</ead>

```

Tutti questi modelli di gestione dei dati possono essere applicati e gestiti ed indicizzati da ExtraWay.

Gli archivi che possono avere testi in lingua straniera ove si presentano frequentemente caratteri con modificatori (accenti, umlaut...) dovrebbero essere tutti ri-compattati. Il compattamento, però, non ha luogo se già compattati a meno che non si intervenga sul file .stat.xml (valore 4 dello status d'archivio).

	 3D INFORMATICA	12/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.3 Gerarchia dei nodi primari

Indipendentemente dal modello di indicizzazione utilizzato, i documenti XML possono avere molti livelli di nodi primari, sia condivisi che collegati, sia che si tratti di piccoli file organizzati gerarchicamente in sottodirectory i cui nomi corrispondono a chiavi uniche. ExtraWay fornisce comunque una vista gerarchica concettuale dei nodi primari e supporta un set di comandi per:

- ? individuare una lista di tutti i nodi primari ad un determinato livello di profondità o indipendentemente dal livello gerarchico di appartenenza;
- ? individuare una lista di tutti i nodi primari indipendentemente dal livello gerarchico;
- ? individuare il percorso di un nodo primario a partire dalla radice;
- ? tornare un limitato insieme di nodi primari precedenti o successivi il nodo primario corrente;
- ? individuare la cima o la coda della gerarchia.

Per ogni nodo primario ExtraWay può tornare una determinata composizione di elementi o attributi tipicamente rappresentati dal nome o dal titolo.

Questo insieme di comandi consente ad un modulo client di gestire una vista gerarchica indipendentemente dalla locazione fisica dei dati.

	 3D INFORMATICA	13/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.4 Equivalenza degli indici

Relativamente alla costruzione degli indici, ExtraWay espone alcuni comportamenti che a volte situazioni reali richiedono.

Si tratta della possibilità di convergere più elementi in uno stesso indice. Questa necessità è utile quando vi sono più elementi di alto livello che nascono dallo stesso percorso differenziandosi leggermente.

Per impostare questa equivalenza ExtraWay consente di definire lo stesso indice per elementi individuati da espressioni Xpath diverse.

Esempio di dichiarazione:

```
<key name="XML,/civile/cla/tim" key_alias="XML,/civile/corpo"/>
```

```
<key name="XML,/civile/cla/tif" key_alias="XML,/civile/corpo"/>
```

consentendo di unificare in "XML,/civile/corpo" l'indice comune.

	 3D INFORMATICA	14/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.5 Elementi trasparenti

Agli effetti dell'indicizzazione alcuni elementi possono essere dichiarati ininfluenti, risultando trasparenti all'indicizzatore consentendo così ad ExtraWay di navigare fra gli elementi come se questi non esistessero.

Proprietà Information retrieval

Per definire un indice è necessario indicare:

? un nome;

? un titolo;

? un tipo di dato (stringa, numero, data);

? un tipo di unità :

parola: ogni termine separato da un separatore dichiarato in una lista di termini non indicizzabili diviene un'entità dell'indice;

stringa intera: ogni intera stringa, composta da una o più parole, diviene un elemento dell'indice.

Ogni entità indicizzata contiene:

? le informazioni relative al proprio posizionamento rispetto l'intero database XML;

? il numero delle occorrenze, cioè il numero delle volte che compare, rispetto all'elemento XML a cui l'indice fa riferimento.

Le proprietà di Information Retrieval di ExtraWay sono basate proprio sull'uso degli indici e sui riferimenti che questi mantengono verso la base dati.

	 3D INFORMATICA	15/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.6 Doppia indicizzazione

A volte può essere utile indicizzare un elemento XML in base ad entrambe le tipologie unitarie (singole parole o stringhe intere).

E' ad esempio il caso degli elementi XML che contengono il titolo di un documento o per elementi che raccolgono nome e cognome di persone.

	 3D INFORMATICA	16/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.7 Concatenazione

ExtraWay consente di definire indici composti dal contenuto concatenato di altri elementi XML. Questa condizione di ridondanza sugli indici spesso consente di evitare il medesimo problema relativo ai dati, offrendo però una migliore prestazione in termini di ricerca.

L'individuazione di dati contenuti in indici concatenati può essere richiesta anche tramite una ricerca che implica l'uso di operatori booleani, a scapito delle prestazioni di ExtraWay che deve fondere ed interpolare le ricerche su più elementi XML.

Per esempio:

```
"type='DPR' and year='2000' and number='445' and article='3' and comma='2'"
```

in luogo di

```
"reference='DPR 2000 445 3 2'"
```

I valori che vanno a comporre la concatenazione possono essere separati da spazio o da un stringa tipo; è possibile inoltre ordinare la sequenza dei valori degli elementi XML che la compongono.

	 3D INFORMATICA	17/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.8 Personalizzazione degli indici

ExtraWay espone la possibilità di personalizzare, mediante script creati appositamente, la costruzione degli indici.

Questi script saranno richiamati automaticamente all'atto dell'indicizzazione di quei particolari elementi XML.

	 3D INFORMATICA	18/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.9 Thesauri

E' possibile definire l'uso di tesauri su indicati elementi XML, cioè di vocabolari multi livello, gerarchici tipicamente utilizzati per definire classificazioni o vocabolari controllati.

E' possibile condividere questi tesauri a più database.

Possono essere file XML nei quali sono indicate le relazioni logiche, di dominio e di relazione fra i termini membri.

	 3D INFORMATICA	19/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.10 Analisi spettrale

ExtraWay è grado di confrontare le frequenze dei singoli termini con il corrente insieme di termini definito dall'insieme dei documenti della ricerca, fornendo così una indicazione di massima della rilevanza dei termini individuati della ricerca rispetto alla loro presenza sulla base dati.

	 3D INFORMATICA	20/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.11 Linguaggio naturale

Si tratta di un modulo aggiuntivo che permette di analizzare specifici elementi da un punto di vista morfologico, logico, sintattico e semantico esprimendo il risultato in termini di nodi concettuali.

Ogni nodo individuato è rappresentato da un insieme di frasi.

	 3D INFORMATICA	21/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.12 Indicizzazione

Il processo di indicizzazione può essere invocato manualmente, periodicamente o al succedersi di determinati eventi quali la modifica della collezione dei file XML che compongono la base dati XML.

Il processo di indicizzazione elabora i file verificandone prima il corretto formato, quindi creando o aggiornando l'indice.

In una seconda fase saranno costruiti i riferimenti alla base dati.

La fase finale aggiorna i file veri e propri.

Le proprietà di indicizzazione sono memorizzate in un file <database>.rules.xml, compilato tramite un editor dal creatore del database stesso.

Valori di default, assunti se non diversamente specificato:

- ? tutti gli elementi XML sono indicizzati per singole parole;
- ? tutti gli attributi XML sono indicizzati come stringhe intere;
- ? i nodi primari sono radice dei documenti.

	 3D INFORMATICA Extraway technical reference	22/93
03/08/2003		Ver. 0.0.4

3.13 Opzioni di ricerca

ExtraWay consente di impostare alcune caratteristiche di funzionamento della ricerca. E' possibile indicare alcune regole di comportamento da utilizzare all'atto dell'individuazione delle chiavi di ricerca sui documenti da selezionare.

In riguardo ai termini da individuare si potrà ordinare ad ExtraWay di individuarli :

Esempio: frase nel campo Testo: "repubblica presidente"

Documenti d'esempio:

documento 1: campo Testo: " il presidente della società ha visitato la filiale della repubblica Ceca ..."

documento 2: campo Testo: "la repubblica ha un presidente nominato dalle parti ..."

- in qualunque posizione
saranno individuati entrambi i documenti in quanto le parole "presidente" e "repubblica" compaiono entrambe, ma non è necessario che compaiano nello stesso ordine e vicinanza espressi dall'operatore che esegue la ricerca.
- a distanza di "x" parole
indicando che le parole inserite nella frase di ricerca debbano comparire entrambe, ma che fra le stesse non compaiano più di "x" parole non digitate dall'utente;
se poniamo una distanza pari a 2, il primo documento indicato non sarà trovato in quanto fra presidente e repubblica si trovano altre 7 parole compresi articoli e preposizioni)
- nello stesso ordine:
indica che le parole inserite dall'utente devono comparire nei documenti individuati, nello stesso ordine in cui sono state digitate dall'utente;
in questo caso il primo documento non sarà individuato in quanto "presidente" viene prima di "repubblica"
- almeno un termine presente:
sta ad indicare che i termini "presidente" e "repubblica" sono poste in OR anziché come prevede il default in AND;
in questo caso saranno individuati entrambi i documenti;
- ignora articoli e preposizioni:
significa che dalla frase digitata dall'utente saranno scartati tutti gli articoli e le preposizioni e saranno poste in ricerca solo le parole restanti

3.13.1 Utilizzo della stoplist

La stoplist è la lista di parole che si ritiene non debbano estrarre a far parte della frase di ricerca in quanto non rappresentano il contenuto dei dati ed inoltre appesantiscono ExtraWay a causa della loro frequenza.

	 3D INFORMATICA	23/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Si tratta degli articoli, delle preposizioni, dei singoli numeri delle singole lettere dell'alfabeto e quant'altro l'amministratore ritenga fuorviante essere inserito nella frase di ricerca.

Di norma questo file prende il suffisso .stp ed è dichiarato in Primary_node/key/stop_list di <applicazione>.conf.xml.

Il file di stoplist predefinito si chiama italian.stp.

L'uso della stoplist modifica le regole impostate con le opzioni di ricerca, soprattutto se l'impostazione sulla presenza di termini indica una distanza fra le parole; nel conteggio delle parole che definiscono la distanza dei termini posti in ricerca andranno scartate le parole di stoplist e quindi la distanza andrà calcolata sulle rimanenti.

Il mancato utilizzo della stoplist inserisce nel conteggio della distanza delle parole anche articoli, preposizioni e quant'altro.

Il risultato sarà certamente diverso per la medesima ricerca che faccia uso della stoplist rispetto ad una ricerca che la comprenda.

	 3D INFORMATICA Extraway technical reference	24/93
03/08/2003		Ver. 0.0.4

3.14 Composizione delle frasi di ricerca

3.14.1 AND e OR

La sintassi di ricerca prevede che più termini possano essere ricercati nell'ambito del medesimo canale di ricerca combinandoli con gli operatori booleani AND ed OR.

Questi operatori hanno ordine di valutazione diverso: AND ha priorità superiore ad OR.

Per esempio :

Esempio 1

[campo] = valore1 OR valore2 OR valore3

sarà interpretata come

[campo]=valore1 OR [campo]=valore2 OR [campo]=valore3

Esempio 2

[campo1]=valore1_1 AND [campo2]=valore2_1 OR valore2_2

sarà interpretata come

[campo1]=valore1_1 and [campo2]=valore2_1 or [campo2]=valore2_2

("AND" interessa il rapporto tra campo1 e solo il primo valore cercato nel campo2; La condizione di "OR" si applica al rapporto tra l'esito del primo "AND" ed il secondo valore cercato sul campo2.

Questo è ben diverso da quanto sembrava in virtù del diverso ordine di valutazione degli operatori booleani che privilegiano l'AND a discapito dell'OR.

Esempio 3

[campo1]=valore1_1 and [campo2]=valore2_1 or valore2_2

sarà interpretato come

([campo1]=valore1_1 and [campo2]=valore2_1) or (valore2_2)

Con l'uso delle parentesi è invece possibile impostare l'ordine di valutazione delle varie parti di cui è composta la ricerca.

Per Interpretare prima i valori espressi in OR ed il loro risultato con la parte in AND scrivere

([campo1]=valore1_1) and ([campo2]=valore2_1 or valore2_2)

Nuova tipologia di estensione rappresentata da un elenco di termini separati da virgole.

La presenza di virgole nelle chiavi deve essere "escapata" con il backslash.

[nomecampo]={1,3,6,8,9,4,11}

equivale a

	 3D INFORMATICA Extraway technical reference	25/93
03/08/2003		Ver. 0.0.4

([nomecampo]=1 OR [nomecampo]=3 OR [nomecampo]=6 OR [nomecampo]=8 OR [nomecampo]=9 OR [nomecampo]=4 OR [nomecampo]=11)

...ma si avvantaggia del fatto che la combinazione in OR di questi elementi espressi con il nuovo metodo può essere utilizzata come uno dei due estrami di un'adiacenza.

3.14.2 ADJ

La ricerca in ADJ consente di individuare i termini indicati in adiacenza tra loro; ma per il suo corretto funzionamento è necessario che i canali di ricerca interessati siano dichiarati di gruppo mediante "Instance=yes".

⁰⁰⁰⁷⁴²⁰Se nell'indicazione degli attrlist da porre in ADJ non viene indicato l'element di riferimento, ExtraWay risale il nodo fino alla determinazione dell'element padre al fine di bacino di pescaggio della ricerca.

3.14.3 Estensione al contorno semantico

Tramite le librerie di espansione di Expert System è possibile estendere la ricerca al contorno semantico dei singoli termini della frase di ricerca.

Questa estensione viene applicata automaticamente, se non specificato diversamente.

L'invio di una frase di ricerca alle librerie Expert System genera l'individuazione di termini o frasi (detti synset).

Ad ogni synset individuato sono "legati" altri synset semanticamente correlati.

La frase di ricerca che ne deriverà porrà in OR i synset originari individuati dalla frase di ricerca originale, ognuno dei quali sarà legato in AND con l'insieme dei propri synset semanticamente correlati.

3.14.4 Frase racchiusa tra apici

Per default gli 'attributes' vengono indicizzati come canali di ricerca monovalore mentre gli 'elements' vengono trattati come multi valore.

Per questi ultimi e per tutti i canali di ricerca soggetti ad indicizzazione monovalore l'interpretazione della frase di ricerca si comporta in questo modo :

	 3D INFORMATICA	26/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

se il testo da ricercare viene espresso tra doppi apici, la parte in essi contenuta viene ricercata senza applicarvi nessuna delle estensioni eventualmente previste e cercando di rilevare nel campo in esame esattamente quanto richiesto; quindi, cercare una sequenza di termini tra doppi apici comporta:

- a) L'esclusione di qualsiasi forma di estensione, compresa l'estensione maschile e femminile. Viene concessa l'estensione per Wild Cards (se preceduti dal carattere "backslash" si richiede che la "wild card" venga ricercata come chiave e non interpretata come motivo di estensione
- b) L'eliminazione della stop list.
- c) Valuazione della distanza di adiacenza ad 1 ed in ordine. In questo modo la frase ricercata verrà rispettata esattamente nella sua forma.
- d) Scarto dei caratteri fraintendibili come operatori di ricerca e scarto degli operatori logici interpretati come chiavi di ricerca. La presenza di caratteri come l'uguale o di sequenze come "non" o "e" che potrebbero essere trattati come operatori di ricerca viene saltata. I caratteri vengono ignorati e gli operatori logici vengono trattati come chiavi.

3.14.5 Ricerca sulle date

Sono consentite anche le wild cards purchè non combinate per ottenere intervalli di tempo; sono ammessi anche i punti interrogativi.

Si richiede il formato YYYYMMDD;

Esempio:

*1971 : richiede tutti di documenti, di qualsiasi anno, del settantunesimo giorno del diciannovesimo mese

1970*01 o 1970??01

documenti del primo del mese di qualsiasi mese del 1970

3.14.6 Raffinamento

Per ottenere il raffinamento di una precedente selezione, utilizzare, nell'ambito della frase di ricerca, la sintassi

[?SEL]=<nome completo file di selezione>

a cui applicare la restante parte della frase di ricerca, questo migliora i tempi di risposta, l'accuratezza del risultato e l'applicazione del reverse sulle chiavi di ricerca individuate.

	 3D INFORMATICA	27/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

3.15 Automatismi

ExtraWay alimenta autonomamente i campi inerenti data ed ora di inserimento e modifica corredati anche dal codice dell'utente modificatore.

All'atto dell'inserimento anche la modifica viene valorizzata con lo stesso identico valore.

In caso venga, per un errore di battitura o per un errore generico, indicato in frase di ricerca il nome di un canale di ricerca inesistente, ExtraWay si limita ad emettere un warning nel proprio log, in quanto, non essendovi nomi di canali di ricerca obbligatori, il server genera l'indicizzazione di default di ogni canale di ricerca che incontrato e quindi l'individuazione di un nome non ancora indicizzato può significare che l'indice di quel canale di ricerca non è ancora stata eseguita o che la costituzione dell'indice per qual canale di ricerca non ha individuato alcuna chiave.

Alla luce di questi fatti ExtraWay non può definire essere un errore l'individuazione di un nome di canale di ricerca non ancora indicizzato o con indice inesistente.

	 3D INFORMATICA	28/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

4. Organizzazione dei file

L'installazione di ExtraWay, indipendente dalla struttura di una tipica applicazione XML, genera una struttura di directory che contiene le parti necessarie al motore:

XW (directory principale di installazione)

- è bin : contiene i programmi (i server e le estensioni necessarie);
- è conf : contiene i file di configurazione;
- è logs : contiene i file di log redatti dal server XWLS ed eventualmente da ESServer (Expert System Server);
- è collect : contiene le raccolte suddivise in sotto-directory assegnate ai singoli utenti; i file contenuti in questa cartella sono invece le raccolte pubbliche;
- è tmp : in questa directory saranno scritti i file temporanei, se non dichiarata la variabile TEMP o TMP
- è db : contiene gli archivi non dichiarati nei files di configurazione
- è context : contiene metainformazioni quali documentazione, riferimenti a release ecc...

L'installazione di ExtraWay generalmente si inserisce nel contesto più ampio di una struttura idonea a contenere le applicazioni e quant'altro necessario per il funzionamento dell'ambiente XML.

La proposta 3D Informatica per una funzionale struttura di directory parte da una directory principale XWAY:

XWAY (directory principale ambiente di sviluppo XML)

- è apps: (contenitore principale della parte strutturale delle singole applicazioni)
 - è singole applicazioni :
 - è db : contiene il database dell'applicazione indicata nella directory superiore;
 - è bin : contiene le parti funzionali tipiche di ogni applicazione, quali ad esempio la parte isapi che consente l'accesso all'applicazione;
- è acl : questa directory (Access Control List) contiene un'applicazione di supporto per la gestione ed il controllo degli accessi;
- è xw : la struttura di directory sopradescritta che contiene ExtraWay;
- è java : questa è la directory principale che contiene il motore web, il layout delle applicazioni, le classi java di comunicazione con ExtraWay
- è TOMCAT : la directory del web server all'interno della quale si sviluppa la parte web delle applicazioni :
 - è Webapps : directory contenitore della applicazioni web XML :
 - è Xway (ex public) : questa è la directory principale delle applicazioni web :
 - è Web-inf : contiene le classi java che consentono la comunicazione con ExtraWay, la generazione del layout, l'iterazione utente – motore ecc..;
 - è Formatter : contiene la parte di layout generica;

	 3D INFORMATICA	29/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- è Application : contiene il layout delle singole applicazioni;
- è Singole directory per le singole applicazioni :
 - è Engine : personalizzazione del funzionamento delle singole applicazioni;
 - è Formatter : personalizzazione del layout delle singole applicazioni;
- è Base : implementazione di ACL (Access Control List).

	 3D INFORMATICA Extraway technical reference	30/93
03/08/2003		Ver. 0.0.4

4.1 File temporanei

La cancellazione dei file temporanei avviene ad ogni "TmpEraseTimeOut" ed elimina i file la cui data di ultimo accesso è superiore a "TmpEraseGap".

Se il numero di file contenuto dalla directory "hwtemp" si avvicina ai 65535 file o se lo spazio disponibile sul disco è inferiore al 5% si rimuovono file temporanei sempre più "recenti" mantenendo, al limite, quelli entro gli ultimi 5 minuti.

Questo compito del master viene svolto anche dai server figli in corrispondenza della propria uscita.

Su macchine ove il tempo di attesa tra una verifica e l'altra è alto ma i server si aprono e chiudono frequentemente, il test viene fatto altrettanto frequentemente, in corrispondenza ad ogni chiusura del server.

	 3D INFORMATICA	31/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

5. Il Framework

Extraway espone un framework, cioè un insieme di pacchetti Java e moduli di comunicazione, che si occupa delle principali funzionalità di risolvere le richieste tipiche di un'applicazione web per l'accesso e la consultazione di un archivio.

La sua architettura è composta da :

- ? Java Virtual Machine (rel 1.3 o superiore);
- ? Web Server Java Tomcat (Jakarta o altri);
- ? Generic.jsp e Error.jsp (Java Server Page) che si occupano della comunicazione fra le classi java ed il lato server ExtraWay;
- ? <nome_applicazione>.jar (pacchetto di classi java: Java Archive) che rappresentano le classi java che necessitano di essere personalizzate in base alle specifiche richieste delle singole applicazioni web;
- ? Generic-apps.jar (pacchetto di classi java: Java Archive) che sono il cuore del framework, cioè sono quelle classi che si occupano di gestire le funzionalità generiche esposte da ExtraWay per le applicazioni web;
- ? Hj.jar (pacchetto di classi java: Java Archive) che si occupa delle comunicazioni con ExtraWay.

Il ritorno tipico di una richiesta di ricerca dati è un involucro XML che riferisce un file di stile XSL definito in Generic-apps.jar come ad esempio:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<?xml-stylesheet type="text/xsl"
      href="../formatter/factory/query.xsl"?>
<response verbo="query" xverb="" dt="costit"
  uri="/xway/application/nif/engine/nif.jsp" host="localhost"
  port="-1" db="costit" user="Lettore" selid="3se3cdbeb691f3"
  qord="" qadj="-1"
  sele="0 3 ,20298, costit ;1 8
  3se3cdbeb691f3,193,A%5btesto%5d%3dadozione;1"
  pos="0" count="193" pageCount="10" keyCount="32" keypath=""
  keylist="">
  <!-- precedenti ricerche effettuate
        e relativi estremi di riferimento -->
  <ricerca liv="0" open="1" tdata="3" riferimento="" totale="20298"
    uid="" descr=" costit ">
    <ricerca liv="1" open="0" tdata="8" riferimento="3se3cdbeb691f3"
      totale="193" uid="A" descr="[testo]=adozione">
    </ricerca>
  </ricerca>
</response>
```

Il livello di personalizzazione a cui deve essere adattata un'applicazione riguarda certamente i file jar suddetti, nonché i file xsl che rappresentano lo stile del layout e del funzionamento

	 3D INFORMATICA	32/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

generale.

I file di stile XSL faranno riferimento a file HTML generici sui quali andranno applicati gli stili XSL ed i dati ritornati dal file XML di cui sopra un esempio.

Il livello di personalizzazione si estende all'uso degli stili CSS a cui chiaramente faranno riferimento i file HTML che l'utente consulerà sul proprio browser.

Il framework si occuperà di :

- § gestire i propri file di configurazione e di schema, quelli che compongono l'applicazione e quelli delle parti comuni (per le funzionalità di livello più generico);
- § creazione e gestione degli indici (creazione, visualizzazione ed aggiornamento);
- § creazione di catalogo (leggi importazione di dati in formato xml e relativa creazione di mappa ad oggetti degli ELEMENT xml);
- § generazione dei titoli (composti dal contenuto di determinati campi e proposti come lista di collegamenti ai documenti estratti dalla selezione);
- § creazione delle chiavi xml di un determinato catalogo;
- § connessione fra il modulo java ed il server Extraway;
- § conversione dell'xml tramite xsl allo scopo di produrre un output html;

I file generati e/o gestiti da Extraway sono platform-independent.

Con l'installazione di Extraway, su piattaforma Linux, possono essere installati una serie di script che consentono lo spegnimento e l'attivazione dei servizi ExtraWay e dipendenti.

Tipicamente con ExtraWay si installa anche il servizio XWLS (ExtraWay Log Server) dipendente ed il servizio XIPS (ExtraWay Inter Process Server).

Per l'uso dei relativi script è necessaria la presenza del modulo HwAdmin.

	 3D INFORMATICA Extraway technical reference	33/93
03/08/2003		Ver. 0.0.4

6. Impostazioni di configurazione di applicazione.

Per ogni applicazione, composta da uno o più database o archivi, esiste, ed è obbligatorio, un file che prende il nome dell'applicazione + ".conf.xml".

In questo file sono dichiarati i comportamenti che Extraway assume nei suoi confronti; vengono definiti i campi chiave dei vari archivi e le relative modalità di indicizzazione, l'uso e la definizione di moduli esterni quali librerie dinamiche, definizione del thesaurus sulle singole chiavi di ricerca, definizione e comportamento da assumere sui nodi primari, le selezioni e i comportamenti del server a fronte di rami xml da nascondere o visualizzare.

Si premette che la tecnologia XML è in veloce crescita e suscettibile di modifiche nel breve periodo.

A tale scopo si fornisce un riferimento di confronto per successivi aggiornamenti: i dati sotto descritti fanno riferimento al modulo "File di configurazione applicazioni Extraway" ver. 002.001.003 della scrivente.

Di seguito viene fornita una lista delle opzioni disponibili in questo file.

Legenda e Glossario:

- q riferimenti ad oggetti "ELEMENT" saranno MAIUSCOLI;
- q riferimenti ad oggetti "attlist" saranno minuscoli.
- q "Canale di ricerca" : si tratta della trasposizione xml di quello che nei comuni database si definisce campo; cioè è una unità di allocazione dati.
- q "Rami potabili" : element o attlist che, all'atto del caricamento del documento, possono essere richiesti o potati al fine di mostrare solo i dati in contesto alla visualizzazione richiesta.
- q File di stoplist : elenco dei termini da escludere durante l'elaborazione della frase di ricerca da parte del server e la visualizzazione del vocabolario (se non diversamente specificato).
- q Nella definizione degli element vengono utilizzati caratteri marcatori per definire l'obbligatorietà degli element o la quantità di presenze che questi possono avere nel file xml al fine della validazione del formato sulla base, appunto delle definizioni espresse nella DTD:
- q Se non è indicato alcun carattere, l'element si intende obbligatorio ed univoco;
- q "*" indica che l'element può comparire 0 o n volte;
- q "?" indica che l'element può comparire 0 o 1 volta;
- q "+" indica che l'element può comparire 1 o più volte;

	 3D INFORMATICA	34/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- ✓ **XWAY_CFG** : (root element) ELEMENT contenitore principale;
- ∅ **cfg_ver** : indicazione della versione di riferimento per questa DTD;
- ∅ **PRIMARY_NODE (+)** : consente di indicare ad Extraway l'identificazione di un nodo primario da un file XML;
 - § **ud_name** : nome dell'oggetto xml da identificare come nodo primario nella gerarchia DOM esposta;
 - § **ud_container** : nome dell'oggetto xml contenitore (padre) del nodo primario da definire; per xml contenenti più nodi primari;
 - § **UNIQUE_RULE (?)** : definisce la regola di univocità tramite la quale verificare l'univocità dei documenti all'atto dell'inserimento / modifica;
 - **search_rule** : indicare, con la sintassi sql-like esposta da Extraway, ed escludendo il carattere di "=" ed il valore, la frase di ricerca che deve essere scatenata per verificare l'univocità del documento in fase di salvataggio;
[XML,/merito/@pd] and [XML,/merito/@mecc]
 - § **KEY (*)** : definizione della modalità di indicizzazione dei singoli canali di ricerca;

Se erroneamente indico due elementi 'key' che fanno riferimento allo stesso canale di ricerca, ovvero allo stesso percorso di chiave, il primo indicato verrà interpretato ed ogni altra duplicazione verrà ignorata.

 - **name** : nome completo del canale di ricerca espresso mediante la sua posizione gerarchica nel DOM, oppure, in modo indipendente (dalla sua posizione assoluta);
 - **value_type** : tipologia del contenuto del canale di ricerca; valori consentiti :
 - ◆ **alfa** : set di caratteri alfanumerici;
 - ◆ **num** : set di caratteri numerici;
 - ◆ **date** : data in formato Highway (aaaammgg). Se il canale di ricerca viene indicato come "date" il server cercherà di interpretare la data opportunamente, convertendola negli indici nel formato indicato, indipendentemente dalla sua forma effettiva. Se la data viene espressa come gg/mm/aaaa o gg/mm/aa o usando altri separatori o sequenze di caratteri per esprimere il mese, verrà compiuto il tentativo di ricondurla al formato "interno" aaaammgg.
Per un formato data interpretabile si richiede esclusivamente che l'anno sia uno dei due estremi, come il giorno, e possibilmente sia espresso in 4 cifre. I separatori possono essere caratteri di vario genere purchè non numerici e lo spazio si può usare se e solo se la data viene indicata tra doppi apici.
 - ◆ **file** : memorizza il percorsi dei file che appartengono al catalogo di un dato archivio;
default = alfa;
 - **key_style** : identifica la modalità di indicizzazione del contenuto del canale di ricerca; valori consentiti :
 - ◆ **auto** : stile "one" per oggetti attlist e "multi" per oggetti element;
 - ◆ **one** : indica l'indicizzazione dell'intero contenuto come singola chiave;
 - ◆ **multi** : indica l'indicizzazione dei singoli termini;
 - ◆ **double** : rappresenta l'applicazione di entrambi gli stili "one" e "multi";
 - ◆ **skip** : indica la richiesta di non indicizzare il presente canale di ricerca;
default = auto;
 - **path_style** : identifica il comportamento da assumere in fase di indicizzazione; gli oggetti

	 3D INFORMATICA	35/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

annidati possono essere indicizzati nell'oggetto container oppure si può ignorarne l'esistenza con lo scopo di uniformare path diversi;

- ◆ auto : oggetto standard, produce un proprio canale di ricerca;
- ◆ container : indica che il contenuto di quest'oggetto e di tutti gli elementi in esso nidificati vengono fatti convergere nel canale di ricerca assegnato all'oggetto stesso. Qualora agli elementi in esso nidificati venisse dedicata esplicita dichiarazione di un canale di ricerca, essi non verrebbero inseriti in questo oggetto contenitore;
- ◆ transparent : indica ad ExtraWay che il contenuto di questo oggetto deve essere considerato appartenente all'elemento "a monte" come se il presente elemento non fosse mai stato incontrato; appare quindi come trasparente agli effetti dell'indicizzazione perchè non produce un proprio canale di ricerca;

default = auto;

- **stop_list** : indica il nome del file di stoplist da utilizzare per questo canale di ricerca. La stoplist viene utilizzata esclusivamente in fase di selezione e presentazione degli indici (visualizzazione vocabolario) e non di costituzione degli indici stessi ed anche in quelle fasi si può richiedere espressamente di inibirne l'uso;
default : per element -> italian.stp (nella directory /conf);
-> null (nessuna stoplist prevista);

E' consentito fare riferimento ad un altro file .stp da porre nella directory d'archivio, ma non potrà assumere il nome "italian.stp"
per attrlist -> nessuna stoplist prevista;

- **attr_collate** : indica, per un dato element, come utilizzare i suoi attrlist per produrre speciali chiavi di indicizzazione;
Pur non esprimendo una obbligatorietà, a livello di DTD, sui valori consentiti per questo attrlist, ExtraWay elabora solo i valori noti sotto indicati :
 - ◆ no : nessuna chiave di concatenamento;
 - ◆ spaced : la varie parti della "regola di composizione" vengono separate da spazio;
 - ◆ catenate : la varie parti della "regola di composizione" vengono concatenate senza spazi separatori;

oppure un certo numero (max 4) di regole di concatenazione personalizzate.

Sintassi :

- ◆ #<nomeregola>:<regola di composizione>
- ◆ max 4 regole

Simboli sintassi :

- ◆ # : inizio nome della regola
- ◆ \$: racchiude parti da sostituire
- ◆ @ : racchiude i nomi dei canali di ricerca a cui il loro contenuto andrà sostituito

La regola può contenere modificatori che permettono una maggiore elaborazione dei dati da concatenare.

Questi modificatori sono identificati dal carattere "|" (pipe) da porre dopo il nome del canale di ricerca (che a sua volta è racchiuso fra "@"):

La sintassi prevista è :

- ◆ d : indica che il valore del canale di ricerca deve essere trattato come un campo data e convertito nel formato "aaaammdd";
- ◆ , (virgola) carattere separatore;
- ◆ p : (part_of) che utilizza questa sintassi:
 - Ø offset : dal quale partire (base 0); se è negativo si intende la fine del valore

	 3D INFORMATICA	36/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- ∅ - : (separatore)
- ∅ dimensione da utilizzare
- ◆ , (virgola) carattere separatore;
- ◆ f : (formattazione) per il riempimento (fill) che prevede questa sintassi :
 - ∅ n : trattare il dato come numerico;
 - ∅ s : trattare il dato come stringa;
 - ∅ <n> : numero dei caratteri di riempimento (se il valore è negativo, il riempimento avviene a sinistra, se positivo a destra);

Esempio:

attr_collate="#<name>.\$@date/d,p0-4,fn-4"

- **key_alias** : rappresenta l'indicazione di uniformare l'indicizzazione di questo canale di ricerca a quello indicato in questo attlist e che di essi si ottenga un unico canale di ricerca alimentato con i dati provenienti da entrambe;
- **key_also**: permette di alimentare anche i vocabolari di altri canali di ricerca con le chiavi di questo oggetto xml;
- § **instance** : identificativo di raggruppamento per effettuare un ulteriore raffinamento nelle ricerche ristretto ai canali di ricerca con lo stesso "instance". Consente di eseguire selezioni nelle quali indicare (esplicitamente o implicitamente) che le adiacenze devono essere risolte nell'ambito dell'elemento indicato come "instance". In questo modo non è la semplice prossimità a soddisfare la condizione di adiacenza bensì la compresenza nell'ambito di una parte isolata dell'intero documento. Questo si presta sia per isolare vere e proprie istanze (ripetizioni di elementi ripetibili) sia quelli che vengono definiti "bacini di pescaggio".; valori consentiti :
 - ? yes : abilita la costituzione della chiave speciale di "instance" che viene usata per convalidare adiacenze e bacini di pescaggio.
 - ? no : condizione standard, l'elemento non rappresenta un bacino di pescaggio.
default = no;
- § **expert_system** : indica l'intervento delle librerie "expert_system " per questo canale di ricerca. Comporta l'uso di Expert System nella costituzione degli indici e quindi la possibilità di farne uso in selezione. Per tale scopo viene creato un apposito canale di ricerca "speciale" accodando al percorso di chiave indicato il postfisso "/#es"; valori consentiti :
 - ? yes : applicare librerie di "expert system";
 - ? no : non applicare librerie di "expert system";
default = no;
- § **extra_cat** : indica che i termini indicizzati per questo canale di ricerca devono essere elaborati anche dalle funzioni specificate in "addon_so";
 - ? yes : applicazione funzionalità di indicizzazione di librerie esterne;
 - ? no : non applicare funzionalità di indicizzazione di librerie esterne;
default = no;
- § **serial_id** : indica il nome logico del seriale;
- § **FILE_LOCATION (?)** : detta le regole secondo le quali creare il file xml che contiene i dati salvati con operazioni di editing dell'archivio; valori consentiti :
 - **mode** : valori consentiti :
 - ◆ single : un file per ogni documento salvato in directory dal nome incrementale;
 - ◆ day : un file al giorno, per tutti i documenti salvati in quella giornata, in una directory dal nome identificativo "<annomese>";

	 3D INFORMATICA	37/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- ◆ month : un file ogni mese, per tutti i documenti salvati in quell'arco di tempo, in una directory dal nome identificativo "<anno>";
- ◆ year : un file ogni anno, per tutti i documenti salvati entro l'anno solare;
- ◆ rule : indica che la regola di composizione del nome file deve essere ricercata nell'attributo "rule" che altrimenti viene ignorato;
sono consentite più regole di locazione dei file che saranno processate da ExtraWay in sequenza fino alla prima determinazione di una regola valida, oppure fino all'ultima regola espressa;

default = single;

- **dir** : modificatore di nome directory (suffisso da anteporre al nome delle directory generate automaticamente da Extraway);
- **rule** : questo attributo contiene un valore rappresentato da una particolare sintassi (molto simile a quella di "attr_collate")
 - ◆ \$: racchiude parti in cui saranno presenti delle componenti fisse da aggiungere se e solo se l'intera parte tra dollari acquisisce significato.
 - ◆ @ : (compresa in una parte racchiusa fra "\$") racchiude parti in cui deve apparire il nome della parte da acquisire che, a questo punto ed a differenza della composizione delle chiavi, deve essere espressa come percorso XML completo;
Esempio:
rule="\$@XML,/NIR//pubblicazione/@norm|p0-6,fn-6@\$" />
il nome del file verrà composto con il contenuto del campo XML,/NIR//pubblicazione/@norm, dal quale si assumeranno i primi 6 caratteri (da sinistra) e che verrà formattato ad un numerico di 6 (aggiungendo zeri a destra) qualora il valore acquisito fosse più breve.
 - ◆ ; : separatore delle varie rule che possono essere memorizzate in questo attlist;

esempio

<file_location

```
mode="rule" rule="$@XML,/doc/storia/in_fascicolo/@codice@$;$@XML,/doc/@nrecord|p0-8@/$doc" move_always="yes"/>
```

- **move_always** : indica ad ExtraWay che la locazione su file del primary_node può essere modificata a seguito di modifiche sui dati ed in base alle rule previste anche se la corrente locazione potrebbe contenere comunque il nodo modificato;
 - ◆ yes : la locazione su disco del primary_node può essere modificata anche se la corrente locazione lo potrebbe contenere;
 - ◆ no : la locazione su disco del primary_node non può essere modificata;

default : no;

§ **PRUNABLE_BRANCH (*)** : (rami potabili) consente di definire aree del documento (oggetti xml) che possano essere estratte o escluse dalla visualizzazione del documento;

- **element_name** : nome dell'oggetto xml che rappresenta l'oggetto potabile;
- **list** : permette di indicare l'estensione dell'azione di potatura :
 - ◆ yes : tutti gli elementi "element_name" nidificati nello stesso elemento padre rappresentano un'unica area di potatura. Tale area avrà inizio in corrispondenza del primo "element_name" incontrato e termine alla fine dell'ultimo;

	 3D INFORMATICA	38/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- ◆ no : singolarmente sui singoli "element_name";
default = no;
- **depth** : indicazione del livello su cui avviene la potatura (in caso di oggetti xml nidificati). Questo ha senso quando si incontra un elemento potabile all'interno di un altro elemento potabile. Dal momento che solo uno di essi può essere scelto come elemento di potatura (le aree di potatura non possono essere nidificate), quest'indicazione porterà alla selezione dell'elemento contenitore o di quello contenuto per la scelta dell'area da eleggere:
 - ◆ inner : potatura al livello più interno;
 - ◆ outer : potatura al livello più esterno;
default = outer;
- **itself** : indica se comprendere o escludere il tag dell'"element_name" dall'operazione di potatura;
 - ◆ yes : potatura che comprende l'element su cui avviene;
 - ◆ no : potatura che esclude l'element su cui avviene;
default = yes;
- § **TITLES (?)** : indica se per questo archivio debbano essere precalcolati i relativi titoli
 - **fixtitle** : precalcolare i titoli
 - ◆ on : i titoli saranno precalcolati;
 - ◆ off : i titoli non saranno precalcolati;
default : off;
- ∅ **UNIQUE_RULE (?)** : definisce la regola di univocità tramite la quale verificare l'univocità dei documenti all'atto dell'inserimento / modifica;
 - § **search_rule** : indicare, con la sintassi sql-like esposta da Extraway, ed escludendo il carattere di "=" ed il valore, la frase di ricerca che deve essere scatenata per verificare l'univocità del documento in fase di salvataggio;
[XML,/merito/@pd] and [XML,/merito/@mecc]
- ∅ **SERIALS (?)** :
 - § **id** : nome logico assegnato
 - § **mask** : è la maschera di valorizzazione composta da caratteri '#' per indicare i valori numerici (0-9) e '%' per indicare i valori alfabetici (A-Z);
 - § **prefix** : indica se il campo è seriale o è un repertorio;
valori consentiti :
 - yes : il campo è seriale;
 - no : il campo è un repertorio;
default : no;
 - § **empty** : consente di lasciare il seriale vuoto;
valori consentiti :
 - yes : il campo può essere lasciato vuoto;
 - no : il campo deve essere valorizzato;
default : no
 - § **skip_test** : consente la valorizzazione del campo anche con valori che non corrispondano alla maschera indicata;
valori consentiti :
 - yes : sono consentiti valori diversi da quelli di mask;
 - no : non sono consentiti valori diversi da quelli di mask;

	 3D INFORMATICA	39/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

default : no

Ø **FILE_LOCATION (?)** : detta le regole secondo le quali creare il file xml che contiene i dati salvati con operazioni di editing dell'archivio; valori consentiti :

§ **mode** : valori consentiti :

- single : un file per ogni documento salvato in directory dal nome incrementale;
- day : un file al giorno, per tutti i documenti salvati in quella giornata, in una directory dal nome identificativo "<annomese>";
- month : un file ogni mese, per tutti i documenti salvati in quell'arco di tempo, in una directory dal nome identificativo "<anno>";
- year : un file ogni anno, per tutti i documenti salvati entro l'anno solare;
- rule : indica che la regola di composizione del nome file deve essere ricercata nell'attributo "rule" che altrimenti viene ignorato;

default = single;

§ **dir** : modificatore di nome directory (suffisso da anteporre al nome delle directory generate automaticamente da Extraway);

§ **rule** : questo attributo contiene un valore rappresentato da una particolare sintassi (molto simile a quella di "attr_collate")

- \$: racchiude parti in cui saranno presenti delle componenti fisse da aggiungere se e solo se l'intera parte tra dollari acquisisce significato.
- @ : (compresa in una parte racchiusa fra "\$") racchiude parti in cui deve apparire il nome della parte da acquisire che, a questo punto ed a differenza della composizione delle chiavi, deve essere espressa come percorso XML completo;

Esempio:

```
rule="$@XML,/NIR//pubblicazione/@norm|p0-6,fn-6@$" />
```

il nome del file verrà composto con il contenuto del campo XML,/NIR//pubblicazione/@norm, dal quale si assumeranno i primi 6 caratteri (da sinistra) e che verrà formattato ad un numerico di 6 (aggiungendo zeri a destra) qualora il valore acquisito fosse più breve.

- **move_always** : indica ad ExtraWay che la locazione su file del primary_node può essere modificata a seguito di modifiche sui dati ed in base alle rule previste anche se la corrente locazione potrebbe contenere comunque il nodo modificato;
 - ◆ yes : la locazione su disco del primary_node può essere modificata anche se la corrente locazione lo potrebbe contenere;
 - ◆ no : la locazione su disco del primary_node non può essere modificata;

default : no;

Ø **THESAURUS (?)** : definisce il thesaurus d'archivio;

§ **TH_FILE (+)** : per definire informazioni relative al file contenitore dei dati del thesaurus;

- **th_filename** : nome del file di thesaurus completo di percorso assoluto o relativo rispetto all'archivio;
- **th_type** : definisce il suffisso, e quindi lo scopo del file di thesaurus;
 - ◆ ths : thesauro gerarchico;
 - ◆ thv : thesauro vincolato;

default = ths;

- **TH_LINK (+)** : definisce il rapporto tra i canali di ricerca indicati nel precedente element "key" ed i corrispondenti canali del Thesaurus per sfruttare il thesaurus

	 3D INFORMATICA	40/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

in selezione o presentazione dei vocabolari

- ◆ **xw_key** : rappresenta il nome del "canale di ricerca" identificato da "KEY/name";
- ◆ **th_key** : rappresenta il nome della chiave di thesaurus corrispondente (indicate nel file .dsp nella sezione [KeyLink] con la sintassi :
<nomechiave>=<valore univoco che identifica la chiave in base 1>
Ø dove <nomechiave>== <th_key>);
- Ø **HIGHLIGHT (?)** : indica i nomi da assegnare alle "processing instruction" (le procedure che Extraway utilizza per compiere determinate elaborazioni) previste per l'evidenziazione delle chiavi di ricerca;
 - § **open_tag_name** : indica il tag (indicatore di processing instruction) dal quale far partire l'evidenziazione all'interno di un element;
default: xw-sr
 - § **close_tag_name** : indica il tag (indicatore di processing instruction) con il quale terminare l'evidenziazione iniziata con "open_tag_name" all'interno di un element;
default: xw-er
 - § **attr_tag_name** : indica il tag (indicatore di processing instruction) che identifica il canale di ricerca da evidenziare;
default: xw-ar <nomeattributo>
 - § **skip_tag_name**: indica i rami potabili non ritornati al client;
default: xw-sk
 - § **unskip_tag_name** : indica i rami potabili ritornati al client;
default: xw-nsk
 - § **elected_element_tag_name** : indica i rami potabili, tornati al client, che contengono chiavi da evidenziare;
default: xw-ee
 - § **unelected_element_tag_name** : indica i rami potabili, non tornati al client, che contengono chiavi da evidenziare;
default: xw-ue
 - § **open_correct_name** : indica il tag (indicatore di processing instruction) dal quale ha inizio una parte di testo non riconosciuta dal correttore ortografico;
 - § **close_correct_name** : indica il tag (indicatore di processing instruction) con il quale termina la parte di testo non riconosciuta dal correttore ortografico;
 - § **attr_correct_name** : indica il tag (indicatore di processing instruction) che identifica l'attributo il cui contenuto non è riconosciuto dal correttore ortografico;
- Ø **HW_FIELDS** : nomi dei macro canali di ricerca che ospiteranno i dati (retaggio di Highway); attualmente obbligatorio ed in futuro presunto obsoleto;
 - § **SEARCH_ALIAS (*)** : consente di definire alias per i nomi dei canali di ricerca;
 - **search_name** : nome logico;
è consentito l'uso di wild cards per la definizione di un canale di ricerca "variabile", cioè che può essere definito con parti del nome costituite dinamicamente:
Esempio
<search_alias search_name="PD" search_key="XML,/NIR/*/vigenza/@id"/>
Non è consentito definire più search_name con lo stesso valore, anche se la relativa

	 3D INFORMATICA Extraway technical reference	41/93
03/08/2003		Ver. 0.0.4

search_key è diversa; è consentita una eccezione che riguarda il valore “@” per il quale sono invece consentite molteplici dichiarazioni.

- search_key : il valore indicato in “KEY/name”;
- § DSC_RULES (*) : una sorta di lookup, eseguita da Extraway, all’atto della restituzione del documento (si tratta di oggetti xml che contengono valori di riferimento da sostituire con valori esterni al database);
 - key_name : nome del canale del canale di ricerca (come definito in “key/name”);
 - key_rule : regola di trasformazione che converte il valore contenuto nell’oggetto xml, nel valore previsto in fase di visualizzazione;
- ∅ TITLES (?) : indica se per questo archivio debbano essere precalcolati i relativi titoli
 - § fixtitles : precalcolare i titoli
 - on : i titoli saranno precalcolati;
 - off : i titoli non saranno precalcolati;
 - default : off;
- ∅ ADDON_SO (?) : indica una eventuale libreria dinamica da utilizzare per l’archivio;
 - § path : indica il percorso di system nel quale individuare la DLL (path completo o parziale rispetto alla directory che ospita l’archivio);
 - § cat_init : nome delle funzione, esposta dalla DLL indicata in “path”, da invocare prima di iniziare un processo di indicizzazione;
 - § cat_add : funzione, esposta dalla DLL indicata in “path”, da invocare per l’oggetto “key/extra_cat” all’atto dell’indicizzazione di “key”;
 - § cat_get : funzione, esposta dalla DLL indicata in “path”, da invocare al termine dell’indicizzazione di ogni singolo documento;
 - § cat_end : nome della funzione, esposta dalla DLL indicata in “path”, da invocare al termine di un processo di indicizzazione;
 - § load_doc_filter : funzione da evocare al caricamento di ogni documento;
 - § search_translator : funzione che può essere evocata per trasformare una frase di selezione generica nella forma specifica per l’archivio in corso (utilizzato con le librerie Expert System);
- ∅ EXPERT_SYSTEM_EXTENSION (?) : per indicare se debbano essere utilizzate le librerie di estensione di Expert System;
 - § active = <value>;
 - ? true : abilita l’utilizzazione delle librerie di Expert System;
 - ? false : disabilita l’utilizzazione delle librerie di Expert System;
 - default = false;
- ∅ PROFILE (*) : indicazioni di profilo (tipicamente chiave=valore utilizzate nei file di impostazione) per definire comportamenti o valori da memorizzare ed utilizzare durante le elaborazioni di ExtraWay;
 - § type : rappresenta, nella trasposizione verso i file di impostazione, la chiave a cui assegnare “value”
 - § value : è, nell’esempio dei file di impostazione, il valore da assegnare alla chiave “type”.

Fra le chiavi definite da questo Element già previste troviamo:

Type=“Century” value=“<number>”

Per l’interpretazione del valore inserito in un campo data il server assume

	 3D INFORMATICA	42/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

che :

a) Se la data è espressa con l'anno intero (4 caratteri.) non si cambia nulla e si assume l'anno come viene dato.

b) Se la data è espressa in 2 cifre ci si assume che appartenga al ventesimo secolo quindi dal 1900 al 1999 salvo diversa implicazione.

b.1) Se c'è un'impostazione del secolo uguale o maggiore di 0 si assume che gli anni da quello indicato sino al '99' siano da assegnare al 1900 e quelli da '00' fino al limite escluso sono da assegnarsi al ventunesimo secolo.

b.2) Se c'è un'impostazione del secolo minore di 0 si assume che gli anni da 0 all'anno attuale più il valore assoluto della cifra espressa meno uno appartengano al ventunesimo secolo, gli altri sino a '99' al 1900.

Certamente un esempio chiarirà le cose.

Type="EsInitDir" value="path completo o parziale rispetto la directory di XW nel quale individuare le librerie di Expert System"

Type="EsLogFile" value="path completo o parziale rispetto la directory di XW nel quale le librerie di Expert System scrivono il proprio file di log"

Type="title" value="<campi e stringhe che compongono il titolo>"

Default = UD,/xw/@UdType "-" NRECORD

nome dell'elemento che identifica il nodo primario, un '-' separatore e l'NRecord del documento.

Type="title.separ" value="<lista caratteri separatori del titolo>"

	 3D INFORMATICA	43/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

6.1 Regole di composizione titoli

Il value del type=title può contenere una composizione di diciture (racchiuse fra ") e nomi di canali di ricerca, rivolta alla generazione di una stringa che rappresenti il documento nella vista sintetica (vista titoli).

I nomi dei canali di ricerca devono essere separati gli uni dagli altri da almeno uno spazio e non ha importanza il fatto che siano maiuscoli o minuscoli;

Le diciture, per le quali viene assicurato il mantenimento della forma maiuscola o minuscola in cui sono state indicate, devono essere separate le une dalle altre da almeno uno spazio; devono apparire racchiuse tra ";

E' sconsigliata una composizione del titolo che inizi o termini con una stringa fra "; al fine di non incorrere in una errata interpretazione della regola di composizione;

Dimensione massima consentita : 128 caratteri; se la stringa supera questa dimensione ExtraWay prende solo i primi 128 caratteri.

(<start[, num.caratteri]>) : questo formato può essere posto in coda ad un nome di canale di ricerca per indicare ad ExtraWay di prendere solo una parte del valore restituito;

start (in base 0) indica da quale carattere far partire la porzione di stringa da inserire nel titolo;

[num.caratteri] (opzionale) : indica di quanti caratteri deve essere composta la stringa da inserire nel titolo; in sua assenza ExtraWay preleva, a partire da start, l'intero valore del canale di ricerca.

'^' : l'utilizzo di questo carattere, prima o dopo il nome di un canale di ricerca, indica ad ExtraWay di eliminare eventuali spazi in testa o in coda al contenuto del canale di ricerca (viene perso in fase di composizione del titolo).

'~' : se le diciture che si inseriscono si intende condizionarle alla presenza di un valore nel canale di ricerca che le segue, devono contenere come ultimo carattere la tilde (~) che viene persa in fase di composizione del titolo ma che informa l'apposita funzione di evitare la dicitura se il campo che segue non esiste o è vuoto; lo stesso se si intende condizionare la dicitura al canale di ricerca che precede, la tilde (~) deve apparire in testa alla dicitura.

'/#' : in coda al nome dell'element, aggiunge in sequenza l'elenco degli suoi attributi separati da spazio

	 3D INFORMATICA Extraway technical reference	44/93
03/08/2003		Ver. 0.0.4

(first)[(x,y)] : aggiunge solo il primo dei suoi attributi; [eventualmente per la sottostringa indicata]

'/##' : come per il precedente parametro '/#' anche in questo caso viene fatta richiesta al server di tornare la lista degli attlist di un determinato element; in questo caso però i valori di ritorno non saranno i soli valori degli attlist separati da virgola, ma si tratterà di una lista di attlist, con i relativi nomi, il carattere uguale ed i valori racchiusi fra virgolette di chiusura.

(full)(x,y) : indica che il successivo riferimento di estrazione parziale, in caso di element o attrlist ripetitivi, si riferisce alla concatenazione delle iterazioni.

Per esempio

XML,/snintegrale/prv(full)(0,128)

Se XML,/snintegrale/prv è un element ripetuto più volte

(<prv>primo valore</prv><newline/>

<prv>secondo valore</prv><newline/>

....)

con questa opzione entreranno a fare parte del titolo i primi 128 caratteri della stringa determinata dalla concatenazione di tutti gli element consecutivi presenti in <snintegrale>;

senza il parametro "full" entrano a fare parte del titolo i primi 128 caratteri di ogni elemento <prv>.

Type="codepage" value="<valore della codepage>"

Default = imposta una tabella di conversione codepage di default:

Tabella valori di default di codepage

•	f	Æ	AE
Š	S	Ç	C
Œ	OE	È	E
Ž	Z	È	E
Š	s	Ê	E
Œ	oe	Ë	E
Ž	z	Ì	I
Ÿ	Y	Í	I
ı	i	Î	I
ç	c	Ï	I
À	A	Đ	D
Á	A	Ñ	N
Â	A	Ò	O
Ã	A	Ó	O
Ä	A	Ô	O
Å	A	Õ	O

	 3D INFORMATICA	45/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Ö	O	Ë	e
Ø	O	Ì	i
Ù	U	Í	i
Ú	U	Î	i
Û	U	Ï	i
Ü	U	Ð	d
Ý	Y	Ñ	n
Þ	TH	Ò	o
ß	SS	Ó	o
À	a	Ô	o
Á	a	Õ	o
Â	a	Ö	o
Ã	a	Ø	o
Ä	a	Ù	u
Å	a	Ú	u
Æ	ae	Û	u
Ç	c	Ü	u
È	e	Ý	y
É	e	Þ	th
Ê	e	ÿ	y

Type="**attach.maxperdir**" value="<num. massimo dei documenti per directory>"

Type="**attach.substext**" value="<estensioni dei file per i quali è consentita la sostituzione senza crearne uno nuovo;nessun default; separati da spazi.
Asterisco = tutte le estensioni sostituibili>"

Type="**attach.aliasext**" value="< estensione files trattati come file di testo;nessun default; separati da spazi >"

Type="**attach.idhome**" value="< percorsi alternativi per gli allegati caratterizzati da ID;separati da virgole;>"

Sintassi:

prima parte obbligatoria: nome del direttorio.

[seconda parte]: quadra aperta + numero minore dell'ID + trattino + numero maggiore dell'ID + quadra chiusa.

[terza parte]: label da mostrare in caso di unità removibile, racchiusa tra quadre.

Type="**attach.varhome**" value="<percorsi alternativi per gli allegati non caratterizzati dall'ID; separati da virgole; Ogni valore è un nome del direttorio.>"

	 3D INFORMATICA Extraway technical reference	46/93
03/08/2003		Ver. 0.0.4

Type="**separs.default**" value="<elenco caratteri, tra 1 e 255 da considerarsi separatori>"

Type="**catalog.strict**" value="<true>"

prevede di riconoscere come nodi primari tutti quelli espessamente configurati nel file <nomearchivio>.conf.xml ma anche tutti quelli non equivalentemente configurati purchè si tratti di elementi di primo livello entro il root element dei files.XML. Default=false;

Type="**check_time**" value="<valore numerico>"

rappresenta il tempo concesso ad un utente che ha posto il documento in stato di check-in (aperto il documento in stato di modifica) per il rilascio dello stesso.

Type="**sort.defaultrule**" value="<nome_campo_case_sensitive>"

l'indicazione della direzione del sort dell'ordinamento dipende dal case della prima lettera che compone il nome dell'element o dell'attlist che viene indicato;

maiuscolo : crescente;

minuscolo : decrescente (default);

Type="**sort.defaultlimit**" value="<valore numerico>"

indica il limite di documenti di cui è composta la selezione che subirà il processo di ordinamento o che non sarà ordinato;

non sono ammessi valori negativi;

questo limite viene applicato solo all'ordinamento di default;

In assenza di un'indicazione di ordinamento di default, le selezioni non subiscono alcun'ordinamento.

In presenza di una regola di ordinamento di default ed in assenza di diversa dichiarazione, quella di default viene applicata nei limiti eventualmente configurati.

In presenza di una regola di ordinamento esplicitata dal chiamante ogni altra impostazione decade e la regola viene applicata senza limitazioni.

Type="**TmpEraseTimeout**" value="<valore numerico >"

esprime l'intervallo di tempo che intercorre fra gli interventi di cancellazione dei file temporanei.

Type="**TmpEraseGap**" value="<valore numerico >"

Indica il tempo intercorso, in minuti, dalla data di ultimo accesso, perchè un file temporaneo possa essere cancellato.

	 3D INFORMATICA Extraway technical reference	47/93
03/08/2003		Ver. 0.0.4

7. Impostazioni di configurazione di archivio.

Per ogni archivio esiste, ed è automaticamente generato tramite i tools di ExtraWay, un file che prende il nome dell'archivio + ".stat.xml".

In questo file sono dichiarate informazioni generali sulla configurazione della struttura e dello stato dell'archivio, quali ad esempio lo stato di eventuale lock o compressione, indicazioni sui seriali e CRC, lo stato di indicizzazione, sui documenti cancellati ecc... .

Si premette che la tecnologia XML è in veloce crescita e suscettibile di modifiche nel breve periodo.

A tale scopo si fornisce un riferimento di confronto per successivi aggiornamenti: i dati sotto descritti fanno riferimento al modulo "File di configurazione archivi Extraway" ver. 000.000.002.001 della scrivente.

Di seguito viene fornita una lista delle opzioni disponibili in questo file.

Legenda e Glossario:

- q riferimenti ad oggetti "ELEMENT" saranno MAIUSCOLI;
- q riferimenti ad oggetti "attlist" saranno minuscoli.
- q "Canale di ricerca" : si tratta della trasposizione xml di quello che nei comuni database si definisce campo; cioè è una unità di allocazione dati.
- q "Rami potabili" : element o attlist che, all'atto del caricamento del documento, possono essere richiesti o potati al fine di mostrare solo i dati in contesto alla visualizzazione richiesta.
- q File di stoplist : elenco dei termini da escludere durante l'elaborazione della frase di ricerca da parte del server e la visualizzazione del vocabolario (se non diversamente specificato).
- q Nella definizione degli element vengono utilizzati caratteri marcatori per definire l'obbligatorietà degli element o la quantità di presenze che questi possono avere nel file xml al fine della validazione del formato sulla base, appunto delle definizioni espresse nella DTD:
- q Se non è indicato alcun carattere, l'element si intende obbligatorio ed univoco;
- q "*" indica che l'element può comparire 0 o n volte;
- q "?" indica che l'element può comparire 0 o 1 volta;
- q "+" indica che l'element può comparire 1 o più volte;

? XWAY_ARC : (root element) ELEMENT contenitore principale, in parte retaggi HighWay;

? attr : Riservato. Indica con valore esadecimale le caratteristiche d'archivio;

	 3D INFORMATICA	48/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

- ? **l_access** : Livello minimo di accesso all'archivio (0-255);
- ? **l_erase** : Livello minimo di cancellabilità dell'archivio (0-255);
- ? **l_config** : Livello minimo di configurabilità dell'archivio (0-255);
- ? **docs** : indicazione del numero totale dei documenti presenti nell'archivio;
- ? **erased** : indicazione della presenza e della quantità dei documenti cancellati. La rimozione dei documenti avviene in modo "fisico" ed il documento cancellato non è più recuperabile;
- ? **indexed** : indicazione dello stato di indicizzazione dell'archivio. Indica per quanti documenti (docs) sono stati effettivamente calcolati gli indici. Se questo valore non corrisponde al numero documenti in archivio alcune operazioni (come gli import) non possono aver luogo correttamente;
- ? **keys** : numero totale delle chiavi che compongono la totalità degli indici dell'archivio;
- ? **main_serial** : valore seriale (progressivo) principale utilizzato per il conteggio dei files nei quali trovano posto i nodi primari salvati. Viene usato quanto non si indica nulla nel elemento "file_location" del file di configurazione d'archivio o se la regola indicata è "single";
- ? **timestamp** : indica, con un formato data ed ora (comprensiva di millisecondi), il momento in cui è stata eseguita l'ultima operazione avente un impatto modificante sull'archivio;
- ? **SERIAL (?)** : definizione dei canali di ricerca con contenuto seriale:
 - ? **serial_name** : nome del canale di ricerca seriale;
 - ? **serial_mask** : modalità di valorizzazione dell'attributo "value" ("#" per indicare un valore incrementale numerico; "&" per un valore incrementale alfanumerico; ogni altro carattere viene considerato una costante)
 - ? **serial_value** : indicazione dell'ultimo valore seriale utilizzato in nel canale di ricerca "serial_name";
 - ? **skip test** : indica al server di evitare i test sul contenuto dei seriali quando questi risultano già valorizzati;
default : no.
- ? **DATA (?)** : definizione di particolari impostazioni o comportamenti (implementazione di coppie chiave=valore come definiti nei file di impostazione .ini)
 - ? **dtype** : nome della chiave a cui associare il valore indicato in "dval";
 - ? **dval** : valore da assegnare alla chiave dichiarata in "dtype";
- ? **CRC** : Controllore della validità del presente file.
 - ? **global** : rappresenta un check sum sulle componenti "sensibili" del presente file così da inibire l'intervento manuale in correzione. Ogni intervento sul contenuto del presente file deve essere effettuato in regime di manutenzione e deve riportare il file ad uno stato "corretto". Questo check sum consente al server di verificare che il file non sia stato manomesso.

	 3D INFORMATICA	49/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

8. Gestione degli indici

La gestione degli indici, in fase di progettazione o di riconfigurazione dell'applicazione - se necessaria - subisce la personalizzazione richiesta dal livello di interesse che il progettista rileva per l'individuazione dei dati; in altre parole si tratta di definire quali campi indicizzare, se fare uso di indici concatenati, ad indicizzazione semplice o doppi, se fare uso di thesaurus. La configurazione di <myapps>.conf.xml permette di accedere a queste funzionalità.

In questa fase è possibile indicare ad Extraway anche eventuali riferimenti di uguaglianza fra gli ELEMENT dell'applicazione; questo probabilmente genera ridondanza, ma consente di giungere ad un determinato elemento da più punti della gerarchia DOM (Document Object Model).

	 3D INFORMATICA	50/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

	 3D INFORMATICA	51/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

9. Attività di logging

ExtraWay esegue sempre, se non diversamente configurato, il solo controllo di concorrenza d'accesso e la stesura del "Journal" per gli eventi salienti.

ExtraWay esegue, solo se richiesto, la registrazione degli eventi salienti, verifiche e controlli, per eventuali "roll back" dei documenti modificati o cancellati.

Il file di configurazione si chiama xreg.conf.xml e si trova nel direttorio /conf.

Il formato interno previsto è:

```
<?xml version="1.0" encoding="iso-8859-1"?>
<!DOCTYPE xreg_cfg SYSTEM "http://www.3di.it/extraway/xreg_20030110.dtd">
<xreg_cfg>
<global active="true"
allow_collision="false"
packed="true"/>
</xreg_cfg>
```

active=true : attiva le funzionalità del servizio XReg di compilazione del registro (default=false);

allow_collision=false : non consente l'accesso contemporaneo dello stesso utente da ip address diversi (default=false).

Packed=true : indica se il file di registro debba essere registrato in forma compressa (default=true).

Il server non viene configurato per alimentare il registro.

Operazioni registrate : nuovo archivio, nuova mappa nodi primari, rigenerazione titoli, compattamento dati ed indici, cancellazione e generazione indici

Il server viene configurato per alimentare il registro.

Operazioni registrate : nuovo archivio, nuova mappa nodi primari, rigenerazione titoli, compattamento dati ed indici, cancellazione e generazione indici corredate di utente, archivio, ip address, tipo di operazione ed altri dati significativi, operazioni di check in ed out di documento, inserimento, modifica e cancellazione(vengono anche salvate le precedenti versioni del documento)

L'alimentazione del registro viene fatta direttamente dal server.

I log creati/alimentati, sono :

Log standard: xw.log (ciclico e rimosso ciclicamente dagli automatismi).

Log Journal: xw<YYYYMM>.log alimentato con registrazioni mirate ad indicare l'esecuzione di nuovo archivio, nuova mappa nodi primari, rigenerazione titoli, compattamento dati ed indici, cancellazione e generazione indici.

	 3D INFORMATICA	52/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

I log sono mensili e non rimosso dagli automatismi.

Registro: xreg<YYYYMM>.log : registra in record marcati con un apposito checksum le informazioni su operazioni registrate : nuovo archivio, nuova mappa nodi primari, rigenerazione titoli, compattamento dati ed indici, cancellazione e generazione indici correate di utente, archivio, ip address, tipo di operazione ed altri dati significativi, operazioni di check in ed out di documento, inserimento, modifica e cancellazione(vengono anche salvate le precedenti versioni del documento) arricchite da ulteriori dati.

La consultazione può avere luogo tramite un apposito modulo "xreg.exe" che funge da reader. Non è ciclico, ne rimosso dagli automatismi.

I files possono essere letti da xreg anche mentre sono in corso di alimentazione da parte del sever.

	 3D INFORMATICA	53/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

	 3D INFORMATICA	54/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10. Progettazione e creazione di un'applicazione

Dopo avere creato la nuova struttura di system per la nuova applicazione (per semplicità si consiglia di copiare una precedente applicazione ed intervenire cancellando o rinominando i file contenuti) è necessario pensare alla DTD che conterrà i dati.

	 3D INFORMATICA	55/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10.1 Creazione della Data Type Definition (DTD)

Se già si possiede una base dati, procedere all'esportazione in formato XML.

Il file che si ottiene rappresenta la mera trasposizione dei nomi dei campi del database in omonimi tag xml.

Il progettista si presume che a questo punto intenda conformare la struttura di memorizzazione del vecchio database verso le nuove funzionalità che questa tecnologia offre, creando una DTD ex-novo.

Questo passaggio può risultare ostico a causa della novità con la quale ci si trova a trattare, dalla complessità della sintassi per le DTD e dalla pur sempre presente complessità di una struttura di ricezione dati.

Potrebbe essere utile creare un esempio di xml che contenga qualche record; alcuni strumenti a disposizione su internet (www.w3c.org) permettono di automatizzare la creazione della DTD, fornendo a questi strumenti, un file xml d'esempio (questi strumenti funzionano correttamente solo se li si fornisce di un esempio xml piuttosto corposo dal quale sia possibile individuare le proprietà dei singoli ELEMENT).

Validare la nuova DTD mediante appositi strumenti (vedi i siti di Microsoft, W3C, ecc...).

	 3D INFORMATICA Extraway technical reference	56/93
03/08/2003		Ver. 0.0.4

10.2 Java Interface

Java si occupa della comunicazione fra Extraway e il client (oppure con il modulo ISAPI, che a sua volta si occuperà di restituire la risposta al client).

Per le normali operazioni previste da un archivio standard, le classi a cui fare riferimento sono già disponibili.

Potrebbe essere necessario apporvi modifiche nel qual caso si debbano personalizzare nuove funzionalità non previste.

Il framework prevede inoltre anche una serie di classi per le operazioni di manutenzione e di utilità generale (vedi sezione "Framework").

10.2.1 Riferimenti per sviluppatori

Nrecord: 14078

Obiettivo: Modificare le funzioni di caricamento documento in modo che siano in grado di caricare da un documento soggetto ad aree da sbiancare non semplicemente una o tutte le aree sbiancabili ma un ben preciso elenco di esse.

Descrizione: Mentre uno dei parametri da passare al caricamento documento per ottenere un particolare prunable branch era...

(elp:<numero>)

... dove <numero> rappresenta una delle aree precedentemente identificate e tornate o meno al chiamante (aree skippate o non skippate), ora il comando cambia forma prevedendo che si possano indicare più parti da tornare con separandole con delle virgole come nell'esempio che segue...

(elp:1,3,4,5,8,9)

Le cifre devono essere espresse in forma decimale.

Nrecord: 13680

Obiettivo : La ricerca [alias]="§" non ha avuto successo nonostante la dichiarazione nel conf.xml della chiave corrispondente con instance="yes" e la presenza del corrispondente elemento nei documenti.

Descrizione : Si ricorda che le chiavi corrispondenti ai caratteri 0xA6 e 0xA7 (il primo è un "pipe" diviso in due parti mentre il secondo è il simbolo solitamente attribuito ai paragrafi) sono chiavi speciali che il server introduce in alcuni vocabolari.

La prima di tali chiavi serve ad identificare gli estremi dei bacini di pescaggio ed in generale le "istanze" (da intendersi come ricorrenze di elementi) di un dato elemento e quindi dei suoi attributi. Viene applicata a tutti i canali di ricerca che hanno nel file <nomearchivio>.conf.xml l'indicazione "instance="yes".

La seconda chiave, invece, viene introdotta dal server per identificare la fine delle "frasi" nel testo e poter compiere le adiacenze in frase.

	 3D INFORMATICA	57/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Nrecord: 12727

Obiettivo : Mettere a disposizione un Comando di riapertura archivio che forzi la riapertura a tutte le istanze del server. Se un'istanza è impegnata in un'esecuzione che consuma tempo il comando deve tornare, se non complicato, preferibilmente:

0: OK tutte l'hanno riaperto,

1: non tutte le istanze hanno riaperto,

2: nessuna istanza l'ha riaperto

Descrizione : Solo eXtraWay.

Files di configurazione:

Nell'intenzione di rimuovere completamente l'uso del file <nomearchivio>.ini a favore del <nomearchivio>.conf.xml è stata abolita la sezione "codepage" dal file .ini a favore di una equivalente configurazione presso il file .conf.xml

La combinazione che precedentemente era data dal nome della sezione e dal nome della voce ora deve essere sostituita da una sola voce univoca e nel file .conf.xml trova posto nell'elemento 'profile' come nell'esempio che segue:

```
<xway_cfg>
<profile type="codepage" value="valore della codepage"/>
</xway_cfg>
```

Quanto indicato come "valore della codepage" non è che una stringa di valori che rappresentano il carattere da sostituire a quello rielvato tra 0x20 e 0xff, separati da una virgola. Per i caratteri che non si intende modificare, il valore numerico da impostare sarà 0. Di seguito un esempio per modificare i primi caratteri

```
<xway_cfg>
<profile type="codepage" value="0,161,0,169,163"/>
</xway_cfg>
```

L'esempio non sostituisce lo spazio (0x20), sostituisce il punto esclamativo (0x21) con il carattere decimale 161 (ovvero 0xA1) che è il punto esclamativo rovesciato, poi non sostituisce i doppi apici (0x22), sostituisce il diesis (0x23) con il simbolo di copyright (decimale 169 ovvero 0xA9) e sostituisce il dollaro (0x24) con il simbolo della lira (decimale 163 ovvero 0xA3).

Oltre a quanto detto, comunque, il server adotta un default diverso dal precedente. Mentre in precedenza le conversioni da code page non venivano mai effettuate se non era presente la sezione apposita nel file ini, ora l'eXtraWay server adotta una code page di default le cui conversioni sono visibili per mezzo del file allegato.

Comando "InterServer"

Implementato un metodo per diffondere ai server in esecuzione, master e children, un comando solitamente di natura amministrativa.

Nella fattispecie sono stati concepiti i comandi di reload della configurazione del server e quello di reload della configurazione d'archivio. Ogni operazione che viene diffusa può essere richiesta con esecuzione immediata oppure differita.

	 3D INFORMATICA	58/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Vediamo i dettagli. Il comando è il numero '7'. I suoi bits sono '1' per attendere una risposta dal server, '2' per diffondere a tutti i children lo stesso comando e '4' per diffondere a tutti i children ed anche al master il comando. Nel caso in cui si usi il bit '1' si deve impostare un attributo "timeout" che indica i secondi da attendere prima di restituire la risposta. Se si intende inviare il comando ad un server ben preciso si deve indicare il suo "pid" con un omonimo attributo. In caso di omissione di questo attributo, il comando viene inviato al server attuale (se i bit 2 e 4 sono assenti) o inviato a tutti i server (compreso il master se il bit è 4). Il comando da distribuire deve essere presente in un attributo "num". "timeout" è obbligatorio solo se si usa il bit '1', "pid" è facoltativo nel 99% dei casi. Se si indica il bit '1' ed il suo timeout la diffusione del comando viene stimolata così che tutti i server se ne accorgano al volo, altrimenti ogni server child se ne accorgerà all'atto della prossima operazione che lo coinvolge. Di conseguenza, per l'esempio dell'apertura archivio o della configurazione del server, non è indispensabile attendere un responso e forzare l'operazione perchè avvenga immediatamente. Il server che esegue il comando, comunque, lo porta a termine immediatamente indipendentemente dal bit '1'.

L'attributo "num" può assumere il valore '1' ad indicare che si richiede il reload della configurazione del server o '2' per indicare che si intende richiedere il reload della configurazione dell'archivio attuale.

Esempi:

Riapertura su tutti i server dell'archivio, subito, timeout 30 secondi.

```
<?xml version="1.0" encoding="windows1252"?>
<cmd c="7" bits="3" timeout="30" num="2"></cmd>
```

Riapertura su tutti i server dell'archivio, subito in locale, gli altri al primo accesso.

```
<?xml version="1.0" encoding="windows1252"?>
<cmd c="7" bits="2" num="2"></cmd>
```

Reload della configurazione del server su tutti i server, master compreso, subito in locale, gli altri al primo accesso.

```
<?xml version="1.0" encoding="windows1252"?>
<cmd c="7" bits="4" num="1"></cmd>
```

I server, ora, intercettano anche i sintomi di sistema che si usano, solitamente, proprio per provocare il reload della configurazione di un programma. La versione Unix compie il reload di tale configurazione anche se riceve il signal SIGHUP. Se viene ricevuto da un child, esso sarà il solo a processarlo, mentre se inviato al master, questo lo diffonderà a tutti i propri figli. Il server Win32, per contro, riceve un Windows Message che lo informa che sono cambiati i settaggi di sistema. In tale frangente, se non è stato il server stesso a stimolare questo comportamento, i child non si accorgono di nulla in corrispondenza al comando di sistema che però raggiunge sempre anche il master che provvederà a dar loro "la scossa". In pratica, è sempre il master che causa il reload di tutti i figli se lo stimolo viene dal sistema ma su Unix è anche possibile manualmente intervenire su un ben preciso child.

Nrecord: 0006812

Obiettivo : Dopo aver implementato le funzioni che consentono il ritorno parziale di una

	 3D INFORMATICA	59/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

porzione arbitraria del documento consentire di specificare la porzione desiderata.

Descrizione : Il server eXtraWay può indicare nel documento tornato delle processing instruction atte ad identificare i punti in cui sono presenti i termini oggetto di selezione, le parti del documento che non sono state intenzionalmente tornate e quelle che sono tornate.

I tags di default sono:

<?xw-sr>parola in reverse<?xw-er?>, per il reverse di un termine.

<?xw-ar nomeattributo?>, per indicare il reverse di un attributo dell'elemento precedente.

<?xw-sk numero?>, per indicare che l'area numerata non è stata tornata.

<?xw-nsk numero?>, per indicare che l'area numerata non è stata saltata bensì tornata.

<?xw-ee numero?>, per indicare che l'elemento numerato è stato tornato.

<?xw-ue numero?>, per indicare che l'elemento numerato non è stato tornato.

Queste ultime 4 voci servono a dare al chiamante tutta la numerazione delle aree o degli elementi ma identificando quelli tornati da quelli non tornati.

Nrecord: 0012756

Obiettivo : attualmente esiste un comando in formato 'xml' che permette di impostare a valido un archivio (@c pari a 5 e @bits pari a 9).

ritengo utile disporre anche di un comando che rende non valido un archivio validato

esempio: un amministratore mette a 'valido' l'archivio 'civile'; dopo un paio di secondi si rende conto che ha fatto un errore e l'archivio che doveva validare era 'penale'; a questo punto gli è facile validare 'penale', ma come invalidare 'civile' per evitare che sia propagato? Una reindicizzazione o una rigenerazione dei titoli possono essere efficaci ma mi sembrano eccessive.

Descrizione : Al pari del comando amministrativo per validare un archivio...

```
<cmd c="5" bits="9"/>
```

... esiste ora un comando inverso per devalidarlo...

```
<cmd c="5" bits="11"/>
```

... che ovviamente può essere espresso sia in decimale che in esadecimale con una qualsiasi delle seguenti forme '0xb', 'xb' e 'hb' con eventuali zeri prima della 'b' e con qualsiasi case.

Nrecord: 0013555

Obiettivo : Sviluppare una funzione del server che, dato un documento ed un elenco di regole, sottoponga il documento ad un parsing equivalente a quello effettuato per compiere highlight e riconosca (da un dizionario italiano, da altro dizionario, dal vocabolario dell'archivio o da una particolare relazione di thesaurus) i termini che non sono validi. Il documento deve quindi essere tornato al chiamante opportunamente decorato con processing instruction in grado di mostrare quali termini non siano stati riconosciuti come validi. Il documento può essere assoggettato a check-out.

Implementare quindi un comando per la manutenzione di questi dizionari, con la possibilità di inserire o rimuovere un termine o di esportare solo i termini che hanno un certo "data record" dato per assunto che quelli del nostro "know-how" abbiano un "data record" riservato.

In fase di salvataggio documenti provvedere alla rimozione di tutte le processing instruction

	 3D INFORMATICA	60/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

introdotte per qualsiasi motivo dal server.

Descrizione : ATTENZIONE: Parte delle indicazioni a seguire sono state modificate con la scheda [0014992](#).

Alimentazione del file di dizionario.

Il file di dizionario di default, inteso per la lingua italiana, si chiama "italian.lng" e si trova nella directory di configurazione assieme al file italian.stp che funge da stop list di default.

In tale file possono trovare posto sia chiavi che 3D ha stabilito essere appartenenti al set di base (data record 0x00010000) sia quelle che i singoli utenti decideranno di aggiungere in separata sede, durante l'arricchimento di questo dizionario dovuto all'esperienza fatta sul campo dai vari utenti. A tali chiavi viene assegnato un diverso data record pari al codice dell'utente che ha compiuto l'inserimento così che in seguito si possa sepre recuperare il solo elenco delle chiavi intenzionalmente inserite dagli utenti.

Per l'alimentazione del dizionario esiste un nuovo comando 0x9 che prevede i seguenti bits: 1 -> crea il file indicato che deve essere assente. Se usato con file esistente provoca errore. 2 -> aggiungi voce/i, indica che l'operazione amministrativa è appunto di aggiunta di una o più parole nel dizionario. 4 -> Arc relative, Indica che i percorsi dei files coinvolti sono relativi alla directory dell'archivio e non alla directory di configurazione di Xw. 8 -> Fix word, indica che la parola inserita deve essere considerata tra le parole che 3D ha scelto per la base di conoscenza, non dovrebbe essere messo a disposizione degli utenti.

Inoltre la funzione prevede che possa essere indicato il dizionario da alimentare nell'attributo "fn". Se assente si assume, appunto, il dizionario di default (italian.lng nella directory di configurazione) indipendentemente dal bit 4. Per indicare invece al fonte da cui compiere l'alimentazione si può indicare un attributo "key" per una singola chiave o un attributo "lfn" per il file che contiene la lista di termini. Se indicato key, lfn viene ignorato, se key manca, lfn è obbligatorio o si produce errore.

Esempi:

```
<cmd c="9" bit="11" lfn="..\tmp/italiano.lst"/>
```

Richiede di alimentare il dizionario di default (italian.lng) con il contenuto del file italiano.lst. Il file italian.lng non deve esistere perchè viene espresso il bit 1 e le chiavi inserite vengono considerate "standard 3D". Il file contenente le chiavi si trova in una directory "tmp" parallela a quella di configurazione.

```
<cmd c="9" bit="10" lfn="/temp/hwtemp/altre.lst"/>
```

Richiede l'inserimento di altre chiavi provenienti da una lista diversa

. Il file di destinazione è lo stesso di prima (ma questa volta esiste, non c'è il bit 1) ed il file si trova nel percorso assoluto indicato (ma nello stesso disco dove si trovano i files di configurazione di xw).

```
<cmd c="9" bit="10" key="universita"/>
```

Come nel caso precedente ma questa volta la stringa viene correttamente indicata. La presenza di un apice finale comporta che se il carattere precedente è una vocale nel dizionario verrà messa la parola con il carattere accentato.

```
<cmd c="9" bit="7" fn="civile.lng" lfn="..\diz/lingua.lst"/>
```

Crea ex novo il file "civile.lng" nella directory dell'archivio (bit 4). Il file non doveva esistere. Vi

	 3D INFORMATICA	61/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

introduce le parole presenti nel file lingua presente in una directory parallela a quella dell'archivio e chiamata "diz".

In tutti i casi l'encoding del comando indica l'encoding della chiave che si esprime in "key" ovvero, allo stesso modo, l'encoding delle chiavi espresse nel file indicato in "lfn". Detto questo se ne suggerisce l'uso in "WinLatin1" che è il formato in cui il server pone le chiavi nel dizionario.

 Uso del validatore.

Il validatore è da intendersi come una funzionalità che controlla il contenuto di parte di un documento XML dato o caricato dall'archivio. per compiere questo controllo si avvale di componenti esterni (thesaurus, dizionario) ai quali fa riferimento in funzione di particolari sintassi. Questo validatore "singolo" può anche essere inteso come loop di validazione che tornerà il primo documento disponibile dall'elenco espresso che soddisfa la condizione di non validità.

Per trattare questi temi prendiamo inizialmente in esame il validatore singolo e discuteremo solo in seguito il validatore multiplo.

Il comando di validazione è il comando '10' ovvero 0x0a. prevede, nell'attributo "bits", solo due valori alternativi, 1 e 2. Il valore 1 indica la scelta della validazione di un singolo documento, il valore 2 indica che si intende operare la validazione di un certo elenco di documenti che vedremo in seguito.

Assumendo quindi che il bit sia 1, il comando è composto dalle regole di validazione e dal documento ovvero dal riferimento ad esso.

Nel caso si faccia riferimento ad un documento dell'archivio senza passarlo nel comando, deve essere presente l'attributo 'd' per mezzo del quale indicare il numero fisico del documento.

Le regole di validazione verranno spiegate più avanti.

Esempio: Sto editando un documento e decido di passarlo per il validatore, mando quindi regole e documento al server.

```
<cmd c="xa" bits="1">
```

```
<rul>
```

```
testo|Ing|italian.Ing
```

```
titolo|Ing|italian.Ing
```

```
grande\ voce|th|CL
```

```
</rul>
```

```
<bst>
```

```
< civile>
```

```
... contenuto del documento.
```

```
</ civile>
```

```
</ bst>
```

```
</ cmd>
```

Se invece di mandare un documento devo richiedere che si elabori con le stesse regole un documento esistente, il comando sarà lievemente diverso e privo dell'elemento 'bst' (Byte

	 3D INFORMATICA	62/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Stream) contenente il documento.

```
<cmd c="xa" bits="1" d="1023">
```

```
<rul>
```

```
testo|Ing|italian.Ing
```

```
titolo|Ing|italian.Ing
```

```
grande\ voce|th|CL
```

```
</rul>
```

```
</cmd>
```

Se l'esito dell'operazione è positivo ma senza rilevare nel documento nulla di irregolare, il documento non viene tornato. In caso contrario esso viene tornato all'interno di un elemento 'bst' e "decorato" da una serie di processing instruction, sostanzialmente simili a quelle usate per il reverse e dislocate nello stesso modo, ma il cui valore (esiste un default ma può anche essere personalizzato) è appena diverso. I valori di default sono <?xw-sc?> e <?xw-ec?> per start e end di un'area da evidenziare negli elementi e <?xw-ac?> per un attributo.

Nella spiegazione delle regole di concatenazione verranno dati anche altri dettagli.

Per l'elaborazione del controllo di validazione su un insieme di documenti, per contro, si userà il bit 2 e si dovrà indicare non tanto il documento da elaborare quanto gli estremi per identificare i documenti interessati. Questi estremi possono essere rappresentati da un nome di selezione ovvero da due cifre separate da un asterisco nell'attributo "sel". Inoltre, se l'operazione viene reiterata, si può indicare il punto di partenza con l'attributo "num".

Vediamo un esempio:

Chiedo la validazione della grande voce nei documenti tra il 3000 ed il 4000 entrambe comprese:

```
<cmd c="xa" sel="3000*4000" bits="2">
```

```
<rul>
```

```
grande\ voce|th|CL
```

```
</rul>
```

```
</cmd>
```

ovvero, validazione dei documenti di una selezione

```
<cmd c="xa" sel="3se92043ac.tmp" bits="2">
```

```
<rul>
```

```
grande\ voce|th|CL
```

```
</rul>
```

```
</cmd>
```

In questo caso, il ritorno non si limiterà a tornare il documento in un elemento 'bst' ma tornerà anche uno o due details (elementi dtl). Viene sempre tornato un dtl avente un attributo "dtype" con valore "ndoc" ed un valore "dval" che rappresenta il numero fisico del documento in esame. Se poi si è indicata una selezione viene tornato un ulteriore elemento "dtl" il cui attributo "dtype" sarà "num" ed il cui "dval" sarà la posizione raggiunta nella selezione.

In seguito, quando si vorrà passare al successivo, si potrà indicare il punto da raggiungere all'interno della selezione come "ennesimo" documento.

Se la precedente richiesta avesse tornato un documento fisico qualsiasi indicando che la

	 3D INFORMATICA	63/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

posizione raggiunta nella selezione era 123, il comando successivo sarebbe:

```
<cmd c="xa" sel="3se92043ac.tmp" bits="2" num="124">
```

```
<rul>
```

```
grande\\ voce|th|CL
```

```
</rul>
```

```
</cmd>
```

per passare al documento successivo. Al termine del ciclo il comando da esito positivo senza tornare, come detto in precedenza, il documento nel suo elemento 'bst'.

Vediamo ora le regole di validazione. Nell'elemento 'rul' si ritiene siano presenti una o più regole separate tra loro dai comuni separatori XML (spazi, riporti a capo, tabulatori). In ogni regola, nelle quali gli eventuali spazi presenti dovranno essere preceduti dal carattere di escaping '\', si trovano 3 componenti, a loro volta separate dal pipe (|). La prima componente è il canale di ricerca, quindi la parte del documento, da sottoporre a validazione. Essa può essere tanto un percorso di chiave valido quanto un alias preso dal <nomearchivio>.conf.xml. La seconda parte indica la modalità di validazione e la terza indica una stringa che assume un particolare significato a seconda della seconda componente.

Le modalità attualmente sviluppate sono 3: "Ing", "arclng" e "th". Le prime due sono sostanzialmente equivalenti e richiedono che il campo, trattato in modo testuale, venga filtrato tramite un dizionario. Nel caso della regola "Ing" la terza componente indica appunto il dizionario da usare (default italian.Ing). I percorsi relativi saranno risolti partendo dalla directory dei files di configurazione di eXtraWay. Nel comando "arclng" il comportamento è identico ma i percorsi relativi verranno risolti a partire dalla directory che ospita l'archivio. In fine, la regola "th" è quella che convalida il contenuto "intero" dell'elemento o attributo passando verificandone la presenza nel thesaurus secondo quanto indicato nella terza componente.

Per questa particolare regola, al terza componente è a sua volta frammentata usando la virgola. La prima parte, se presente, deve essere un codice di relazione così come elencati nel file .DSP. La seconda componente, se presente, è la chiave di partenza. Si potranno in futuro ampliare le modalità di comportamento di questa regola.

Vediamo ora alcuni esempi

```
testo|Ing|italian.Ing
```

Richiede che in contenuto del campo "testo" ovvero del percorso XML corrispondente a tale alias, venga filtrato con il correttore ortografico avvalendosi del file italian.Ing che si trova nella directory di configurazione dei programmi.

```
titolo|arclng|Ing/italian.Ing
```

Richiede che il contenuto del campo con alias "titolo" venga filtrato con il correttore ortografico usando un dizionario che si chiama "italian.Ing" dislocato in una directory "Ing" all'interno della directory che ospita l'archivio stesso.

```
grande\ voce|th|
```

```
piccola\ voce|th|CL
```

Richiede che il campo corrispondente all'alias "grande voce" e "piccola voca" (notare il carattere di escape per non farsi ingannare dallo spazio nel nome dell'alias) venga validato con il thesaurus. Nel canale del thesaurus collegato ad ogni campo indicato si cercherà una

	 3D INFORMATICA	64/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

qualsiasi chiave che metta in relazione il contenuto del campo stesso con altre chiavi. Per "grande voce", non avendo indicato alcuna relazione nella terza componente, qualsiasi chiave esistente sarà soddisfacente mentre per piccola voce si richiede, quanto meno, che esista una relazione che lega il contenuto del campo con un altro termine con la relazione CL.

classificazione_bid|th|NT,radice

Avendo indicato due componenti nella terza parte si richiede che il contenuto del campo avente alias "classificazione_bid" sia risolto tramite il thesaurus. In questo caso non ci si accontenterà di una chiave qualsiasi ma si comporrà esattamente la chiave che parte da "radice" usa la relazione NT (ne dominio generico) ed ha come altro estremo il contenuto del campo indicato. Se esiste la chiave esatta, il campo è valido.

Veniamo ora alla decorazione: Come per le processing instruction inerenti il reverse dei termini nei documenti, queste possono essere personalizzate con le voci "open_correct_name", "close_correct_name" e "attr_correct_name" nella sezione "highlight" del file di configurazione d'archivio. (Vds. versione [2.0.2.0](#)).

Il loro valore di default, comunque, è <?xw-sc?>, <?xw-ec?> e <?xw-ac?>. Quest'ultimo viene corredato dall'identificatore dell'attributo e del contenuto da evidenziare separato da due punti. Nel caso delle regole "Ing" e "arclng" il sistema cerca anche di dare suggerimenti sui valori presenti nel dizionario che potrebbero essere sostituiti al valore non riconosciuto. Per ottenere questi valori si fissano i primi 3 caratteri e sui restanti si valuta la somiglianza accettando un errore su 5.

Un attributo non validato potrebbe avere il seguente aspetto

...

```
<cla bigvox="000" litvox="000"><?xw-ac @bigvox:000?>
```

...

ad indicare che la bigvox indicata non è presente nel thesaurus. Se invece si trattasse del trattamento tramite dizionario, potremmo avere una condizione come la seguente sempre applicata ad un attributo.

```
<civile ...
```

```
dis="(Rigetta, Trib. Torno, 9 febbraio 2000)."
```

```
codpar="174"><?xw-ac @dis:Trib?><?xw-ac @dis:Torno:torino,tornio?>
```

...

suggerendo, dopo il primo termine che è quello da porre in evidenza i due suggerimenti rilevati dal dizionario.

La stessa cosa su un elemento avrebbe un aspetto diverso come segue

...estromessi dall'attività lavorativa e <?xw-sc prossimo,prossima?>prossimi<?xw-ec ?> al trattamento pensionistico di anzianità ...

In questo esempio i suggerimenti vengono inseriti nella processing instruction d'apertura.

Nota integrativa sulle regole di controllo ortografico.

Il comando di controllo ortografico e validazione valori da thesaurus ed affini può essere inviato anche senza l'elemento "rul". In tal caso il server applica le regole di default. Queste devono essere state indicate nel <nomearchivio>.conf.xml in una voce di profilo il cui tipo è

	 3D INFORMATICA	65/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

"spellchk.rule" ed il cui valore deve essere sostanzialmente equivalente a quanto detto precedentemente per l'elemento "rul".

Esempio:

...

```
<profile type="spellchk.rule" value="testo|Ing|italian.Ing titolo|Ing|italian.Ing grande\
voce|th|CL"/>
```

...

Nuovi paramtri del comando.

In precedenza il comando di validazione di un documento o di una lista di documenti non prevedeva che il documento stesso venisse assoggettato a check-out. Essendo questa la primaria esigenza ora la funzione si regola di conseguenza.

In altri termini: la funzione non torna il documento se non c'è motivo ovvero se il documento non presenta irregolarità. In tal caso il documento viene sempre assoggettato a check-out tornando il numero fisico (attributo 'd') ed il suo check-sum (attributo 'chk'). Specificando il bit 4, invece il documento viene tornato anche se non presenta alcuna irregolarità. Se e solo se il documento tornato non presenta irregolarità si può richiedere di non compiere il check-out di tale documento con il bit 8.

Immaginiamo quindi la condizinoe tipo, quella in cui si sottopone al server un elenco di documenti da analizzare (bits="2"). Ci verrà tornato il primo documento irregolare e sarà bloccato. Se non vogliamo intervenire dovremo sbloccarlo. Le altre combinazioni hanno maggiormente senso se si carica un solo documento e non un elenco perchè prevedono che il documento venga tornato sempre. Diciamo quindi che bits passa da '2' a '1' e poi aggiungiamo gli altri vbit. Per bits="5" la funzione torna il documento comunque, irregolare o meno, e sarà bloccato e quindi da sbloccare. La combinazione bits="9" non ha senso visto che il bit 8 assume significato solo se c'è il bit 4. Con bits="13" invece, il documento viene tornato e sarà bloccato se e solo se presenta irregolarità, altrimenti sarà sbloccato. Ovviamente si possono usare le combinazioni 6, 10 (priva di significato come la 9) e 14 su liste di documenti ma verrà sempre tornato il primo documento disponibile data la posizione di partenza nella lista.

Aggiunto ulteriore bit 0x10 al comando di spellchk da applicarsi nel caso di una scansione di lista documenti. Indicando questo bit la scansione avviene a partire dal punto indicato ma procedendo in senso inverso.

Apportato ampliamento al comportamento di aggiunta di uno o più termini ad un dizionario. Se si è a conoscenza del nome del dizionario o si ritiene di esplicitare quello di default il server cercherà su disco il file indicato (con percorso relativo o meno a seconda dei bits usati) e cercherà di alimentarlo. Se non trova il file indicato assume che il valore impostato sia un percorso XML di un elemento o attributo XML riconducibile, se possibile, ad una chiave (o addirittura può essere l'alias stesso usato in ricerca). Il server cercherà quindi di tradurre la stringa indicata per risalire al dizionario corrispondente alla regola che comprende appunto quella componente XML.

	 3D INFORMATICA	66/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Esempio:

```
<cmd c="9" bits="2" fn="XML,/civile/cla/tim" key="pippo"/>
```

```
<cmd c="9" bits="2" fn="titolo" key="pippo"/>
```

Entrambe i comandi richiedono di inserire la chiave "pippo" nel dizionario che viene fatto corrispondere al percorso XML indicato ovvero alla chiave di ricerca indicata. Si noti, per altro, che il percorso indicato può anche essere un percorso che confluisce in un canale di ricerca completamente diverso come proprio il caso di /civile/cla/tim del nostro esempio.

Nrecord: 0014992

Obiettivo : Restituire nella PI che decora i termini non riconosciuti l'informazione relativa alla possibilità di inserimento in dizionario del termine (ovvero segnalare la presenza di regole di verifica ambigue per la costruzione della relazione utilizzata per la verifica nel th).

Descrizione : A differenza di quanto [precedentemente sviluppato](#) le processing instruction sono lievemente cambiate. Anche la P.I. usata negli elementi ora acquisisce il "due punti" che è sempre presente per poter prevedere un altro "due punti" che indica dove si trova l'istruzione di inseribilità.

La forma attuale è:

- Per gli elementi:

```
<?xw-sc :eventuali suggerimenti separati da virgole:inseribilità?>termine<?xw-ec?>
```

- Per gli attributi:

```
<?xw-ac @attributo:eventuali suggerimenti separati da virgole:inseribilità?>
```

In breve, quindi, tra il primo ed il secondo simbolo ':' ci sono i suggerimenti (eventualmente stringa vuota) e dopo il secondo simbolo ':', prima della fine della P.I. (?>) c'è la stringa "ins" se inseribile e "noins" se non inseribile perchè regolato da norma abigua.

Nrecord: 0009018

Obiettivo : Tutti gli archivi.

Per ognuno dei canali di ricerca inclusi nel riquadro "Parole" di ogni archivio, la ricerca m pla (con più termine) deve considerare l'utilizzo ulti dell'operatore logico "AND" tra i termini esca al posto dell'operatore "ADJ" attualmente utilizzato.

Tipo di intervento richiesto:

Nuova Funzionalità (NF)

Descrizione : Le informazioni che possono essere inviate al server, durante le ricerche, in materia di adiacenze sono:

a) La distanza generica, positiva o negativa;

b) La distanza specifica per un dato attributo, anch'essa positiva o negativa;

In precedenza si applicava la distanza stabilita per il campo specificato ed in sua assenza quella generica. Questo comportamento vale ancora e gli deve essere sommato un comportamento ulteriore. L'indicazione della distanza '0' comporta che per il campo in esame (caso AdjDist:0) ovvero per l'intera frase di selezione (caso distanza generica uguale a 0) gli operatori sottintesi (quelli semplicemente omessi e non espressamente indicati come "ADJ")

	 3D INFORMATICA	67/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

siano trattati come AND e non come ADJ. In questo caso, vale prima di tutto la qualifica specifica del campo e poi il valore generico quindi si può:

- a) Richiedere adiacenza a distanza X su un campo e distanza Y su tuttigli altri, con ADJ come operatore sottinteso;
- b) Richiedere AND sottinteso su un campo e distanza Y su tutti gli altri con ADJ sottinteso;
- c) Richiedere ADJ sottinteso su un campo a distanza X e AND sottinteso su tutti gli altri
- d) Richiedere AND sottinteso su tutti i campi.

Esempio:

Distanza generica 4

[Testo|(AdjDist:0)]=progesso civile and [NOTE]=antichi romani

Comporta che tra processo e civile si consideri un AND e tra antichi e romani si instauri un'adiacenza a distanza 4 ordinata.

Inversamente...

Distanza generica 0

[Testo|(AdjDist:-3)]=progesso civile and [NOTE]=antichi romani

Comporta che tra processo e civile si consideri un'adiacenza disordinata a distanza 3 e tra antichi e romani si imponga un operatore AND.

Nrecord: 0015298

Obiettivo : Prevedere la possibilità di chiedere al server la composizione di un file "delta" non legato ad una particolare data bensì di natura "tombale" che comprenda cioè tutti i files d'archivio, tutti gli XML e tutti i files accessori del caso (ad esempio gli allegati).

Descrizione : Aggiunto a quanto detto nella scheda indicata, qualora il contenuto del timestamp che si indica al server fosse la costante "full", il server provvederà a calcolare non tanto il "delta" ovvero l'elenco dei soli files (d'archivio, XML contenenti i nodi primari, gli allegati) modificati dalla data indicata bensì un file cumulativo comprendente tutti i files dell'archivio. In buona sostanza è come se il test del "delta" venisse effettuato con un timestamp fissato ad un anno, mese e giorno assolutamente precedente la primigena creazione dell'archivio.

Nrecord: 0015392

Obiettivo : Archivio Civile.

Cerco i documenti aventi anno compreso tra il 1999 ed il 2002 e numero tra 9 ed 11 nei riferimenti normativi. Isolo la sentenza 7479 del 1996 e giocherello con conformi e vedi di un solo livello. Una di queste operazioni mi da un elenco di 4 documenti che porto nell'albero della selezione. Da essi ne scarto due e mi rimangono:

SENT 5279 del 09/04/2001

SENT 12611 del 16/12/1998

Se raffino questi due per [rif.anno]=1995 trovo due documenti

Se raffino questi due per [anno decisione]=2001 trovo un documento.

Se metto quest'ultima selezione in "and not" con i due documenti, dovrei trovarne uno e non ne trovo alcuno.

	 3D INFORMATICA	68/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Descrizione : Nelle selezioni/raccolte ove si aggiungono documenti non è possibile mantenere correttamente traccia degli indici delle parole che hanno condotto alla selezione perché si introduce un nuovo record a scombinare quanto registrato. Ne consegue che una simile operazione annulla la catena riferimenti della selezione in esame e successive selezioni che la coinvolgano non sapranno porre in reverse i termini selezionati con la selezione originaria.

Nrecord: 0010121

Obiettivo : Come da capitolato è necessario garantire ad es. uno script che esporta le mappe (UDP/UDD), il cross reference, le raccolte e le ricerche frequenti

L'esportazione delle chiavi degli allarmi sarà opera di nif32 (massimoM)

Descrizione : Implementato nuovo comando (8) che consente di esportare informazioni proprie di un archivio e non solo.

La parte "bits" del documento può presentare il bit 0x80000000 per indicare che si vuole ottenere l'export direttamente nel messaggio di risposta, altrimenti ci si deve avvalere dall'attributo "fn" per indicare in quale file si vuole l'esportazione. In assenza anche di quest'impostazione il server provvederà a creare un file temporaneo (volatile) con l'esportazione richiesta.

Escluso il bit indicato, l'attributo "bits" viene utilizzato per indicare cosa si voglia esportare secondo la seguente tabella:

- 1: mappa documenti dell'archivio in esame
- 2: documenti in una raccolta pubblica
- 3: documenti in una raccolta privata
- 4: chiavi del vocabolario
- 5: chiavi del thesaurus

Vediamo i casi singolarmente:

Nel caso 1, si possono indicare il primo documento fisico dal quale esportare la mappa (attributo "num") e l'ultimo documento da esportare (attributo "num2"). In assenza degli estremi il minore è il primo documento dell'archivio ed il maggiore è l'ultimo. Per valori al di fuori di questi limiti (prima del primo e dopo l'ultimo) si assumono il primo e l'ultimo. Se il valore dell'estremo "destra" è inferiore all'estremo "sinistra" si provoca errore.

Nel caso 2 e 3 si esportano i documenti appartenenti alla raccolta identificabile con il nome indicato nell'attributo "lfn". Per ogni documento in raccolta se ne esporta l'archivio di corrispondente, il suo numero fisico e l'nrecord.

Per il caso 4 si esportano le chiavi con le relative frequenze. Si deve indicare nell'attributo "id" la chiave che si intende esportare, sotto forma di alias valido o di percorso XML completo. Si possono inoltre indicare con gli attributi "num" e "num2" gli estremi, maggiore e minore, della frequenza che vogliamo venga esportata (valori di default 0 e numero documenti in archivio) oppure con gli attributi "key" e "key2" la prima ed ultima chiave che si intende esportare (valori di default, tutte le chiavi del vocabolario). In fine, indicando un attributo "sel" si può richiedere che la selezione indicata sia fonte di analisi spettrale delle chiavi da esportare.

	 3D INFORMATICA	69/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

In testa ad ogni file di esportazione ci sono una serie di elementi "global" che danno informazioni su quanto esportato e su come l'esportazione abbia avuto luogo, di seguito ci sono gli elementi esportati.

Vediamo ora alcuni esempi:

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

```
<cmd c="8" bits="1" fn="c:/temp/hwtemp/export.xml" num2="x10">
```

```
</cmd>
```

Esporta dall'archivio in esame i documenti a partire dal primo sino al 16° (Vds. Allegato 1)

```
<?xml version="1.0" encoding="windows1252"?>
```

```
<cmd c="8" bits="2" fn="c:/temp/hwtemp/collect.xml" lfn="penale-edita">
```

```
</cmd>
```

Esporta una raccolta pubblica chiamata "penale-edita". (Vds. Allegato 2)

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

```
<cmd c="8" bits="4" id="titolo" num="5" num2="25" key="A" key2="C">
```

```
</cmd>
```

Esporta le chiavi tra 'A' e 'C' aventi frequenza compresa tra 5 e 25 documenti dal vocabolario del canale di ricerca avente alias "titolo" (Vds. Allegato 3)

Nrecord: 0008311

Obiettivo : Nonostante l'importazione con rispetto dei direttori e delle relazioni basate sugli "nrecord" che da essi derivano, la costituzione del catalogo dell'archivio asiri ha perso delle informazioni inerenti le relazioni. Sotto al documento primario (nrecord 77) si hanno due rami, il primo dei quali è la numerazione nera (nrecord 40275). Il primo documento al suo interno è la documentazione presso l'archivio centrale dello stato (nrecord 81) che sappiamo avere dei figli ma la cui gerarchia si interrompe subito.

Descrizione : Per la costituzione del catalogo da documenti XML organizzati in direttori erano stati previsti i seguenti parametri:

LffFile: Nome del file o del direttorio dal quale importare.

LffParentElementId: nome elemento per mezzo del quale identificare il nome del direttorio che qualifica i figli del documento attuale.

LffParentAttributeId: nome dell'attribute per mezzo del quale identificare il nome del direttorio che qualifica i figli del documento attuale.

LffZeroFill: numero dei caratteri complessivi da raggiungere aggiungendo zeri a sinistra del nome del direttorio per renderlo equivalente alla dimensione del campo da controllare (tipicamente NRECORD).

Questi parametri, in particolare il secondo ed il terzo, consentono di identificare il valore nel file XML che corrisponde al nome del direttorio dei documenti figli. Così come concepito si presta per determinare il valore di un element esclusivo o di un attribute in un element esclusivo.

Avendo incontrato, per contro, un caso diverso in cui lo stesso element può apparire più volte ed ogni volta acquisisce un senso diverso in funzione del valore di un particolare attribute, ecco che è stato inserito il parametro 'LffParentAttributeValue' per indicare il valore che qualifica l'attribute che, a sua volta qualifica l'elemento.

	 3D INFORMATICA	70/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

In sintesi avremo:

- LffParentElementId compilato, LffParentAttributeld e LffParentAttributeValue non compilati: si richiede di riconoscere il contenuto dell'element indicato confidando nel fatto che sia univoco in tutto il nodo primario;
- LffParentElementId e LffParentAttributeld compiati, LffParentAttributeValue non compilato: si richiede di riconoscere il contenuto dell'attribute indicato internamente all'element indicato confidando nel fatto che sia univoco in tutto il nodo primario;
- Tutti e tre i parametri compilati: si richiede di riconosere il contenuto dell'element indicato in quella istanza di esso riconoscibile per la presenza dell'attribute indicato avente il valore indicato.

Nrecord: 0014536

Obiettivo : Se tento di rimuovere un documento il motore restituisce "operazione non effettuata o interrotta".

Archivio xdocwaystruttur

Andare in visualizzazione di un gruppo e premere pulsante cancella.

URL: <http://www2:8090/xway/application/xdocway/engine/xdocway.jsp?db=xdocwaystruttur>

Descrizione : Esempio cancellazione documento: si richiede la rimozione del sesto documento dell'archivio in esame previo suo check-out.

```
<?xml version="1.0" encoding="windows1252"?>
```

```
<cmd c="3" bits="0" d="6" chk="68EA7F275E9FE7FBF39DB029EF0D6020">
```

```
</cmd>
```

Nrecord: 0009753

Obiettivo : Alla luce di quanto sviluppato per gli obiettivi [0005121](#) e [0007159](#), implementare funzionalità analoghe per l'eliminazione di un nodo primario.

Descrizione : I documenti rimossi figurano nel file XML che li ospitava precedentemente sotto forma di processing instruction. Al pari della p.i. che viene usata per i documenti modificati si compone della chiave che identifica l'operazione, il numero documento in esadecimale ed il timestamp corrispondente all'operazione di cancellazione. Per le modifiche si aggiunge anche la URN del file di nuova destinazione.

Modifica:

```
<?xw-moved 0x1234 20030120174544.123
```

```
c:\xw\apps\nif\db\civile\civile\000000\000001.xml?>
```

Cancellazione:

```
<?xw_erased 0x1234 20030120174650.483?>
```

Nrecord: 0013054

Obiettivo : Gestire la presenza di un prefisso non significativo nella interpretazione degli id per la produzione delle relazioni fra doc in catalogazione.

Descrizione : In aggiunta alle precedenti voci di configurazione dinamica utilizzabili per il comando di costituzione del catalogo, ora si può introdurre una voce...

	 3D INFORMATICA	71/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

(LffIgnorePrefix:<prefisso>)

...per imporre che il prefisso, appunto, di un ID che deve avere un primo carattere non numerico possa essere ignorato all'atto del riconoscimento degli estremi che qualificano i rapporti di gerarchia.

Nrecord: 0007630

Obiettivo : Creare le relazioni tra nodi primari ed il loro padre in virtù del nome del direttorio in cui si trova il file corrente.

Tale funzione deve essere fatta come opzione successiva all'indicizzazione. L'attributo o l'elemento che designano il nome del sottodirettorio viene specificato o su linea di comando o con un default nell'archivio.ini.

(In effetti propongo che il file che contiene le regole cioè l'attuale archivio.xml si chiami di default archivio.rules.xml es. lexs.rules.xml, è meno fuorviante e così ha un nome preciso nella nostra terminologia).

Questo obiettivo è successivo all'esportazione XML in gerarchia di sottodirettori

Descrizione : Il serve acquisisce una nuova funzionalità senza perdere la compatibilità col passato.

In precedenza il comando 172 (0xAC) prevedeva che in Point.rec ci fosse il nome del file o il percorso dal quale comporre il catalogo dell'archivio eXtraWay. Ora, con il server indicato, questo parametro può essere il nome secco (ed allora si segue il comportamento canonico) un una combinazione di parametri articolati ove solo il primo è indispensabile.

Ecco l'elenco dei parametri:

(LffFile:<nomefile o percorso dal quale importare>)

(LffParentElementId:<nome elemento chiave univoca>)

(LffParentAttributId:<nome attributo eventuale chiave univoca>)

(LffZeroFill:<numero caratteri a '0' di fill>)

Mentre 'LffFile' è l'equivalente del precedente nome file nel quale i separatori di direttorio E' BENE che siano indicati come '/' per non confondere il server, 'LffParentElementId' e 'LffParentAttributId' servono ad identificare la chiave univoca come valore di un certo elemento o di un suo attributo. Se si parla del valore dell'elemento esso deve essere indicato subito dopo il carattere '>' che chiude il tag di apertura dell'elemento stesso. Il valore di 'LffZeroFill' invece ci dice quanto è estesa la chiave univoca se numerica in quanto il nome dei direttori che ospitano i files relazionati potrebbero essere privi degli zeri.

Esempio:

(LffFile:f:/highway/apps/asiri/asiri/)(LffParentElementId:nrecord)(LffZeroFill:10)

Nrecord: 0011304

Obiettivo : Alla luce della scheda 0011302 è bene provvedere, in fase di inserimento o modifica di un documento, alla costruzione o ricostruzione del suo titolo. Modificare i comandi esistenti perchè possano compiere quest'operazione direttamente.

Descrizione : Come detto, la generazione dei titoli in modo automatico avviene se mnel file <nomearchivio>.conf.xml trova posto una nuova voce di configurazione che ha il seguente

	 3D INFORMATICA Extraway technical reference	72/93
03/08/2003		Ver. 0.0.4

aspetto.

```
<titles fixtitle="on"/>
```

Questo si applica sia per la comune generazione del catalogo che per il salvataggio dei documenti. In quest'ultima fase viene comunque ricostruito il titolo dei documenti modificati che fossero stati precedentemente corredati da un titolo o mandando come parametro di salvataggio il valore 256 (0x100) come nell'esempio seguente:

```
<?xml version="1.0" encoding="iso-8859-1"?>
```

```
<cmd c="3" bits="256" d="0" pnce="civile" pne="listof_civile" ....>
```

```
</cmd>
```

Nrecord: 0014484

Obiettivo : 1) Occorrerebbero funzioni da parte del motore per restituire le informazioni riguardanti i seriali, data una chiave.

2) (Per Fabio) hj.jar dovrebbe fornire metodi per tali comandi del motore.

Descrizione : Il comando teso ad acquisire dal server lo stato dei canali di ricerca identificati per un dato archivio (0xAC-0x01) è stato arricchito come segue. L'elemento che descrive un canale di ricerca può ora avere 3 attributi in più che sono:

serial_id: Nome identificativo del seriale associato a quel canale di ricerca.

serial_mask: Maschera di valorizzazione del seriale in esame.

serial_flags: Valore esadecimale che identifica le caratteristiche del seriale secondo questi valori:

0x00000001: Repertorio o Seriale Protocollare

0x00000002: Seriale per il quale si ammette il valore vuoto

0x00000004: Seriale per il quale non si compie il test di validità del contenuto.

Nrecord: 0012285

Obiettivo : Ho effettuato le seguenti operazioni:

Caricamento della pagina di ricerca sull'archivio COSTSN alla URL

<http://www3/xway/application/nif/engine/hcprot.dll?host=mbussetti&db=costsn>

ho cercato la parola "casa" nell'intero testo (397 documenti trovati);

dall'eleco dei titoli ho visualizzato il documento numero 11 (il primo della seconda pagina); l'ho aperto in modifica e ho inserito il mio nome nel campo "giudice". Ho cliccato il pulsante "Salva documento" ma in visualizzazione mi è riapparso il documento di prima, cioè senza la modifica apportata.

La stessa cosa fatta precedentemente su www3 (cioè senza specificare nell'URL host=mbussetti) aveva avuto esito positivo.

Descrizione : Dalla versione attuale del server, se il chiamante dimentica di indicare il nodo contenitore, il server cerca di acquisire quella di default per il nodo primario indicato direttamente dal <nomearchivio>.conf.xml.

Inoltre, sempre da questa versione, quando il server torna un errore da un comando XML, esso non viene tornato solo come componente del messaggio ma come vero e proprio codice

	 3D INFORMATICA	73/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

d'errore.

Nrecord: 0011943

Obiettivo : Per la replicazione nell'applicativo 'nif' si presenta la seguente necessita': un comando (focalizzato su un archivio) che, partendo da un timestamp comunicato dal client, restituisca (a) il timestamp corrispondente all'ultima modifica dell'archivio e, se questo e' maggiore del timestamp ricevuto, (b) un elenco dei file modificati nell'intervallo tra i due timestamp (sia file di archivio che file documento). In alternativa a (b) si potrebbe prevedere la restituzione di (c) un file ('tar'? 'tar.gz'? 'tar.Z'? 'tar.bz'? 'zip?') con tutti e soli i file che dovrebbero essere elencati in (b); questo potrebbe forse essere utile se il client non ha accesso al filesystem su cui risiede l'archivio (risiede su un'altra macchina? gira con un utente che non ha i permessi per raggiungere i file in questione?).

Potrebbe anche essere opportuno prevedere una versione del messaggio che si limita a restituire il timestamp attuale (forse senza neppure ricevere il timestamp dal client) senza preparare elenco di file o file piu' o meno compresso con i medesimi (comandi di tipo 'status' del client).

Descrizione : Vedansi la configurazione di prova e la scheda [0015298](#)

Nrecord: 0010974

Obiettivo : l'operatore '<>' (diverso da) sembrerebbe non funzionare correttamente.

Verificato sull'archivio 'snciv' dell'applicativo 'nif':

- 'snciv' presenta 84 documenti;
- l'analisi spettrale, sull'intero archivio, di "[anno decisione]" (campo "anno" nella maschera di ricerca sintetica) mostra 2 documenti per il 1997, 68 documenti per il 1998 e 1 documento per il 1999 (per 13 documenti sembrerebbe non indicato tale valore);
- la ricerca "[anno decisione]=1998" trova (come mi aspetto) 68 documenti;
- la ricerca "[anno decisione]<>1998" trova invece 84 documenti (l'intero archivio) mentre me ne aspetterei solo 16.

En passant, non sarebbe male fare in modo (se cosi' non e' gia') che la ricerca dell'operatore '<>' sia diretta e non sequenziale.

Descrizione : Implementato un metodo per esprimere ranges di valori nei quali gli estremi non siano da considerarsi validi. Se in precedenza esprimere un range era possibile solo nella forma

{eventuale valore minore|eventuale valore maggiore}

con almeno uno dei due estremi espressi la ricerca era l'equivalente di

>= eventuale estremo minore e <= eventuale estremo maggiore

ora si può esprimere la stessa sintassi sostituendo il carattere '|' col carattere '^' che quindi cambia l'espresione in

{eventuale valore minore^eventuale valore maggiore}

che viene interpretata come

> eventuale estremo minore e < eventuale estremo maggiore

	 3D INFORMATICA	74/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Chiaramente, per gli estremi non espressi, in entrambe i casi si raggiunge il massimo o minimo valore previsto dal vocabolario.

Nrecord: 0011622

Obiettivo : Dopo aver completato la segnalazione 0011621 accettare nelle frasi di ricerca il formato [?COL]="nomeraccolta" al pari del preesistente [?SEL] così da evitare al client il fastidio, di volta in volta, di doverse la andare a calcolare.

Descrizione : Dalla versione corrente, oltre ad esplicitare nella frase di selezione una selezione precedente indicando il nome di campo fittizio [?SEL] è possibile ottenere un risultato equivalente usando direttamente una raccolta. Il campo fittizio si chiama [?COL] (Da collezione o collection) ed eventualmente tra doppi apici il nome della raccolta. Dal nome della raccolta il server risale al file di raccolta su disco e lo usa al pari di una vera e propria selezione.

Dal momento che esistoo raccolte private e raccolte pubbliche l'indicazione del nome della raccolta puro e semplice comporta l'uso di raccolte pubbliche mentre se il primo carattere del nome della raccolta indicato è un più ('+', carattere illegale nei nomi di raccolta), allora la raccolta indicata si intende essere una delle raccolte private dell'utente.

Nrecord: 0012242

Obiettivo : Implementare un comando che, data una raccolta privata, possa copiarla in altra raccolta privata o in una raccolta pubblica provvedendo alla sovrascrittura se necessario.

Descrizione : L'operazione di copia di una raccolta fallisce in tutti i casi in cui:

- a) Manca la raccolta d'origine o è stata erroneamente indicata per pubblica una raccolta privata e vice versa;
 - b) La raccolta di destinazione esiste già ma non è stato dato il comando di "overwrite";
 - c) La raccolta di destinazione è pubblica ed esiste già ma, nonostante il comando di sovrascrittura, appartiene ad un altro autore quindi non è possibile modificarla;
- ...ed in generale gli errori di lettura o scrittura sul sipporto

Nrecord: 0008109

Obiettivo : Rilevata anomalia in fase di indicizzazione in presenza di un ENTITY contenuta in un ATTRIBUTE e risolta prima della relativa indicizzazione!

Descrizione : Implementato un nuovo parametro di caricamento dei documenti eXtraWay che può essere usato al pari dei comandi atti a stimolare i vari casi di highlighting.

Il parametro è 'lu' (Load Unconverted), ha per default il valore 0 e può essere impostato ad 1 per forzare il server a non convertire i documenti caricati in formato WinLatin1.

Non si deve fare nulla di particolare nel caso della reindicizzazione perchè viene impostato internamente al server in quei casi.

Nrecord: 0005121

Obiettivo : Gestire comando di append nuovo file XML o di notifica esistenza nuovo file XML con un nome dato.

	 3D INFORMATICA Extraway technical reference	75/93
03/08/2003		Ver. 0.0.4

Accettare il nome del file se non esiste ancora altrimenti appendere un suffisso che lo renda univoco

Descrizione : Configurando l'elemento file_location del file di configurazione degli archivi eXtraWay è possibile indicare come comporre i nomi di percorso e di file XML ove verranno salvati i documenti proprietari.

Vds. documentazione allegata.

Nrecord: 0012635

Obiettivo : esigenze legate all'applicativo 'nif', renderebbero desiderabile lo sviluppo di comandi presso il server eXtraWay applicabili a una lista di archivi e non a un solo archivio. esempi concreti: il comando per ottenere il timeStamp (c='4', bits='1'), i comandi per ottenere lo stato di lock o per modificare i medesimi (c='5', bits='3', '4' e '5'), il futuro comando per ottenere lo stato relativo alla propagabilità (anche quello per la modifica degli stessi?).

Descrizione : Le informazioni elencate di seguito hanno senso nella manualistica per sviluppatori.

Il server espone, da questa versione in poi, un comando supplementare in forma XML, come tutti quelli più moderni.

Il comando 6 (attualmente non prevede bits) consente l'esecuzione di più comandi in sequenza. Nell'implementazione attuale non si considera ancora di poter utilizzare l'output del precedente comando come input di quello successivo ma il comando è stato concepito per mantenere aperta questa possibilità.

La sintassi del comando consta nell'inserimento all'interno del tag di comandi di tutti gli altri comandi, sequenzialmente elencati, e separati da una particolare processing instruction atta ad identificarli.

Parimenti la risposta sarà composta da un tag "rsp" che a sua volta conterrà tanti tags "rsp" separati tra loro da una corrispondente processing instruction.

La processing instruction, oltre ad indicare dove finisce un comando ed inizia il successivo, ci indica anche a quale archivio fare riferimento da quel comando in poi.

Più chiaro di mille parole è l'esempio che segue. Immaginiamo di voler richiedere alcune informazioni generiche su alcuni archivi e di voler evitare di lanciare una sequenza di comandi complessa. Il comando potrebbe essere il seguente:

```
<?xml version="1.0" encoding="windows1252"?>
```

```
<cmd c="6" bits="0">
```

```
<cmd c="5" bits="5"></cmd>
```

```
<?xw-cmd?>
```

```
<cmd c="4" bits="1"></cmd>
```

```
<?xw-cmd arc:penale?>
```

```
<cmd c="5" bits="5"/>
```

```
<?xw-cmd?>
```

```
<cmd c="4" bits="1"></cmd>
```

```
<?xw-cmd arc:merito?>
```

```
<cmd c="5" bits="5"/>
```

	 3D INFORMATICA	76/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

```
<?xw-cmd?>
```

```
<cmd c="4" bits="1"></cmd>
```

```
</cmd>
```

Come si può notare l'encoding viene espresso una volta sola per tutte e vale per l'intero comando articolato. All'interno del comando se ne trovano altri rappresentati da tag da terminare con un tag di chiusura o auto terminati. Tra ogni comando appare la processing instruction `<?xw-cmd?>` che può contenere al suo interno la dichiarazione di un archivio. Nel nostro esempio si presuppone che il comando inviato interessi l'archivio civile, inizialmente, al quale viene richiesto lo stato di locking (5-5), Subito dopo si richiede il time stamp (4-1) e di seguito si cambia archivio passando a "penale" al quale vengono fatte le stesse richieste per poi passare a "merito" con gli stessi comandi.

La risposta avrà l'aspetto seguente:

```
<?xml version="1.0" encoding="windows-1252"?>
```

```
<rsp ack="1" e="0">
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="lckr" dval="off"/>
```

```
<dtl dtype="lcka" dval="off"/>
```

```
</rsp>
```

```
<?xw-rsp?>
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="info" dval="20030128163930.265"/>
```

```
</rsp>
```

```
<?xw-rsp?>
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="lckr" dval="off"/>
```

```
<dtl dtype="lcka" dval="off"/>
```

```
</rsp>
```

```
<?xw-rsp?>
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="info" dval="19000101000000.000"/>
```

```
</rsp>
```

```
<?xw-rsp?>
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="lckr" dval="off"/>
```

```
<dtl dtype="lcka" dval="off"/>
```

```
</rsp>
```

```
<?xw-rsp?>
```

```
<rsp ack="1" e="0">
```

```
<dtl dtype="info" dval="19000101000000.000"/>
```

```
</rsp>
```

```
</rsp>
```

Anche in questo caso l'encoding vale per l'intera risposta. Per ogni comando si ha una risposta

	 3D INFORMATICA	77/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

ma qualora un comando non fosse andato a buon fine sarebbe possibile anche l'interruzione di questo ciclo con un ritorno di risposte inferiore alla richiesta. In tal caso, comunque, la risposta "globale" o l'ultima di quelle interne sarebbe negativa.

Nella sintassi non è richiesta la processing instruction dopo l'ultimo comando e prima della chiusura del comando contenitore così come nella risposta non viene tornata la processing instruction dopo l'ultima risposta.

Nrecord: 0014972

Obiettivo : Attivare, parametrizzabile, il test sulle stringhe che compongono il thesaurus in modo case insensitive invece del solito case sensitive così da accettare un termine indipendentemente dall'accezione con il quale è stato stilato.

Descrizione : Nella dichiarazione delle regole del controllore ortografico la regola nota come "th" comporta il test in modo case insensitive a partire da questa versione del server. Così facendo il test più "lasco" da maggior probabilità di trovare termini che differiscono solo per il case.

Data la natura del Thesaurus e dell'ordine che alle chiavi viene dato, qualora il test richieda l'indicazione di una chiave di partenza (solitamente "radice") essa non viene può essere sottoposta al controllo case insensitive (altrimenti la scansione del thesaurus potrebbe divenire ciclica e lunghissima) e deve quindi essere espresa in modo corretto rispetto al case indicato. Per effettuare il test in modo "case sensitive", come avveniva prima e come avviene solitamente nell'accesso al thesaurus, la regola deve essere cambiata in "th_cs".

In sintesi: Nel caso di regola "th_cs" indipendentemente dalla presenza della componente "radice", la chiave viene verificata "as is" e se manca nella forma prevista si assume che non esiste. Nel caso di regola "th" in assenza di radice la chiave viene verificata in modo insensitive (non esistendo una seconda parte ci si accontenta di sapere se la chiave esiste) mentre se esiste una radice essa viene verificata nel thesaurus come "case sensitive" ma la validità della chiave, effettuata confrontando la seconda metà della stessa avviene in modo "case insensitive".

Nrecord: 0014855

Obiettivo : gestire nel comando di inserimento in dizionario la scelta del canale opportuno (dizionario o thesaurus) sulla base del parametro che qualifica il percorso xml e della rispettiva regola espressa nel conf

è opportuno in caso di errore distinguere il caso di regola ambigua nella verifica da thesaurus va inoltre gestito il caso di chiave completa (proporrei come default nel caso di key_style="double" ma comunque configurabile nella regola)

Descrizione : Implementato comando di inserimento nel thesaurus di chiavi utilizzate dal controllore ortografico (verificatore di presenza in elenchi del thesaurus) in modo che possa operare appunto sul thesaurus e non semplicemente sul dizionario di un campo. Il comando è una estensione di quello precedente richiedendo sempre il "file name" che però rappresenta il canale del thesaurus di nostro interesse e non un nome di file ed il parametro "key" che rappresenta appunto la chiave.

	 3D INFORMATICA	78/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Il comando, quindi, avrà un aspetto somigliante al seguente:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
```

```
<cmd c="9" bits="2" fn="presidente" key="prova"></cmd>
```

Il comando '9' richiede l'amministrazione dello spell checker, il bit 2 indica la richiesta di inserire una chiave nella fonte corrispondente al canale noto come "presidente". Nel file di configurazione a tale canale (che in questo frangente è espresso come alias di ricerca ma corrisponde ad una chiave ben precisa, si dovrà associare un canale del thesaurus, riportato nel DSP. Parimenti nel file di configurazione dovrà essere citata nella regola di spell checking una regola che coinvolga tale canale e che descriva sia la relazione da usare sia la radice della chiave da creare. Il server provvederà a creare la chiave che unisce la radice al termine indicato nel parametro "key" unendoli con la relazione espressa nella regola. Parimenti, se esiste nel DSP la dichiarazione di regola inversa, il server costituirà anche quella.

Se la regola è ambigua e non può condurre alla corretta costituzione di una chiave di relazione, la funzione sviluppata torna "NO_FIELD" (887). Tra gli altri errori significativi c'è "NO_PATH" (882) che indica la mancanza o incongruenza di uno dei parametri in ingresso o l'impossibilità, da essi, di giungere alla configurazione d'archivio richiesta.

Nrecord: 0013362

Obiettivo : TUTTI gli Archivi. Quando si effettua una ricerca in linguaggio naturale, il sistema restituisce una finestra con i termini oggetto della ricerca più dei sottoelementi. Si richiede di affiancare, a sinistra di ogni termine, il totale dei documenti restituiti dalla ricerca. Ad esempio termine di ricerca= furto gioielli e selezione di "concetti" - il sistema restituisce i due termini principali più i loro derivati. Si chiede di gestire la risposta del sistema come indicato nell'esempio che segue (i valori tra parentesi sono a titolo di esempio)

+ (198) (furto)

(21) (crimine, reato, delitto)

(7) (razzia, ruberia)

(31) (furto aggravato)

.....

+ (110) Oggetto ornamentale di metallo prezioso lavorato e spesso (monile, gioiello, prezioso, bijou, gioia)

(35) (guarnitura, finitura, abbellimento, guarnizione, rifinitura, ornamento, finimento)

(5) (corona)

(19)(collana, vezzo)

.....

Descrizione : All'esistente comando '1' atto ad ottenere le informazioni inerenti la disambiguazione di una stringa tramite Expert System, si possono ora aggiungere alcuni parametri supplementari per ottenere altri dati.

Se si aggiunge il bit 0x20, si richiede oltre a quanto ottenuto sino ad ora, la frequenza delle chiavi (synset) identificate presso il vocabolario di un dato canale di ricerca espresso dall'attributo 'key', espresso come alias o percorso di chiave. In questo caso, l'elemento 'esinfo'

	 3D INFORMATICA	79/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

tornato precedentemente può comprendere un attributo in più "freq" che indica la frequenza per il synset indicato. Se poi si vuole sottoporre quanto appena indicato ad una ulteriore restrizione data dall'analisi spettrale su una selezione, allora si deve aggiungere anche il bit 0x40 ed indicare nell'attributo 'sel' l'id della selezione effettuata.

Esempio:

```
<?xml version="1.0" encoding="windows1252"?>
<cmd c="1" bits="x2b" key="probabilistica">
<txt>comperare una borsa</txt>
</cmd>
```

Torna la disambiguazione completa di termini alternativi e contorno sematico con indicazione anche delle frequenze (Vds. esempio 1 allegato)

Lo stesso esempio di cui sopra che avesse x6b come 'bits' ed un attributo 'sel' ad indicare una selezione effettuata compierebbe la stessa operazione ma sotto forma di analisi spettrale condizionata ad una selezione.

	 3D INFORMATICA	80/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10.3 Java server web Tomcat

Il server web per java potrebbe, nel caso si siano modificate le classi predisposte dal framework, necessitare di una nuova configurazione.

Comunque questo server dovrà essere configurato per la connessione ad Extraway e per la comunicazione con il modulo ISAPI.

Se la richiesta del client giunge direttamente a Tomcat, allora tramite Extraway sarà confezionata la risposta e convertita in formato html o xml (dipende dal browser del client).

Se sarà il modulo ISAPI a inviare la richiesta a Tomcat, allora la risposta dell'elaborazione dovrà tornare ad Highway Connector (modulo ISAPI).

Per il corretto funzionamento di Tomcat server web è necessario configurare il file `\highway\java\tomcat\it.highwaytech.apps.nif.properties` nel quale dichiarare i percorsi locali e remoti ai file XSL di stile e DTD di schema ("localDTDPath", "remoteDTDPath", "localStylesheetPath" e "remoteStylesheetPath").

	 3D INFORMATICA	81/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10.4 Trasformazione XML in HTML

Il framework già propone fogli di stile e template HTML per la conversione verso gli stati interattivi più comuni, proponendo un'interfaccia grafica preimpostata.

Tipicamente questi file (xsl, html e di impostazione) dovranno essere rimaneggiati al fine di ottenere la grafica prevista per l'applicazione che si vuole sviluppare.

La trasformazione dell'xml avviene tramite identificazione degli "element" sui template html; un id appositamente conformato indicherà all'xsl come trattare l'xml per il confezionamento dell'html.

	 3D INFORMATICA	82/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10.5 Esempio conversione archivio web Docway in formato XML

In questo esempio, che sarà maggiormente esplicativo per i lettori che conoscono l'applicazione Docway di 3D Informatica, percorre le fasi salienti dei primi passi verso la conversione XML di un archivio esistente.

Essendo l'applicazione Docway composta da più database, si precisa che le prime operazioni che saranno eseguite riguarderanno la conversione del database delle strutture.

10.5.1 Esportazione in formato XML

Tramite uno strumento che permetta di esportare i dati in formato xml (p.e. con hwout.bat - o .cmd o .sh – di 3D Informatica) estrarre i dati presenti sull'archivio da convertire.

Il risultato è un file ascii che sostituisce i nomi dei campi del db di Docway in tag xml omonimi, fra i quali ne racchiude il contenuto.

Il file ottenuto contiene :

- § la "declaration" nella quale compare la dichiarazione del set di caratteri consentiti ("encoding=utf-8");
- § un ELEMENT root di nome "export";
- § i tag principali rappresentano le tabelle di origine;
- § per ogni ELEMENT sono elencate le varie proprietà e caratteristiche con la sintassi prevista.

Al termine delle dichiarazioni degli ELEMENT, sono indicate le ENTITY che permettono l'uso di caratteri speciali, diversamente non supportati da XML; quindi compaiono i dati esportati. In questo caso si parla di DTD interna, in quanto è esplicitata direttamente e non riferita mediante una declaration verso un file esterno.

Fino a questo punto non si può parlare di conversione, in quanto la struttura xml dovrà essere progettata al fine di definire la base dati dal punto di vista di questa nuova tecnologia. E' quindi necessario risolvere l'esportazione verso una ben precisa DTD che potrebbe avere ben poco in comune con il vecchio database; e quindi riconfigurare l'esportazione al fine di ottenere un file di dati da catalogare con Extraway al fine di creare il nuovo database (che per xml possiamo definire come uno schema di oggetti vincolati in base alla gerarchia che avremo impostato con la nuova DTD).

10.5.2 Verso una nuova DTD

	 3D INFORMATICA	83/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Come sopra citato la preparazione di una nuova DTD passa dalla costruzione di un file xml d'esempio, dal quale si potrà più facilmente definire il nuovo schema.

Stralcio di una struttura xml d'esempio:

```
<list_of_struttur>
  <record>
    <struttura_interna cod_amm="xxx" cod_aoo="xxx"
cod_uff="xxxxx" cod_responsabile="xxx" ...>
      <nome>xxxxxxxx</nome>
      <indirizzo comune="xxx" prov="xx" cap="xxxxxx" ...
</indirizzo>
      <email addr="xxxx" />
      <telefono num="xxxx" tipo="xxx"/>
      <telefono num="xxxx" tipo="xxx"/>
      ...
      <sito_web url="xxxxxxxxxx" />
      <sito_web url="xxxxxxxxxx" />
      ...
      <competenze>xxxxxxxx</competenze>
      <note>xxxxxxxx xxxxx</note>
      <creazione data="xxxxxxxx" operatore="xxxx"
matricola="xxx" />
      <modifica data="20020520" operatore="Paolo"
matricola="1" />
    </struttura_interna>
  </record>
  <record>
    ...
  </record>
</list_of_struttur>
```

Sulla base di questo esempio, mediante l'uso di strumenti appositi o manualmente, si provvede a creare una nuova DTD.

10.5.3 Trasformazione XML verso nuova DTD

L'esportazione ottenuta con "hwout" ha fornito un export xml che traspone i nomi dei campi in ELEMENT.

Ora si vuole convertire quegli ELEMENT, ma soprattutto il loro contenuto, verso la nuova struttura dichiarata nella nuova DTD.

La tecnologia xml ci giunge in aiuto con l'ausilio degli stili XSL.

Tramite idoneo XSL, e apposito strumento (xmllint, saxon, xalan, ecc...) è possibile ottenere un nuovo file dati in formato xml che rispetta la struttura DTD definita.

La progettazione prevede una prima definizione degli oggetti di livello elevato, per poi giungere ad oggetti di basso livello, detti nodi primari.

Xsl rappresenta un vero e proprio linguaggio di programmazione che permette di interpretare il file xml a cui è stato applicato (mediante una conforme dichiarazione) e di generare un file di

	 3D INFORMATICA Extraway technical reference	84/93
03/08/2003		Ver. 0.0.4

output opportunamente elaborato.

Nel nostro caso di tratterà del file dati definitivo, a cui abbiamo applicato la nuova DTD.

Stralci di codice del file xsl di conversione:

Carica l'intera struttura DOM dell'xml:

```
<xsl:template match="*">
</xsl:template>
```

Riscrive l'xml

```
<xsl:template match="comuni" mode="table_root">
  <xsl:element name="record">
    <xsl:element name="comune">
      <xsl:attribute name="nome"><xsl:value-of select="comune"/></xsl:attribute>
      <xsl:attribute name="cap"><xsl:value-of select="cap"/></xsl:attribute>
      <xsl:attribute name="prov"><xsl:value-of select="provincia"/></xsl:attribute>
      <xsl:attribute name="nazione"><xsl:value-of select="nazione"/></xsl:attribute>
      <xsl:apply-templates select="istat"/>
      <!-- Info creazione/modifica -->
      <xsl:call-template name="InfoInsMod"/>
    </xsl:element>
  </xsl:element>
</xsl:template>
```

Individuando "Comuni" a partire da root (table_root) e creando :

```
<record><comune nome="xxx" cap="xxx" prov="xxx" nazione="xxxx" /></record>
prelevando i valori da
<export>
  <comune>xxx</comune>
  <cap>xxx</cap>
  <provincia>xxx</provincia>
  <nazione>xxx</nazione>
</export>
```

10.5.4 Creazione di un nuovo archivio

Abbiamo a questo punto i dati in formato xml e la relativa DTD come la nostra analisi ha previsto debba essere.

Prepariamo la struttura di system per accogliere il nostro nuovo archivio.

Il sistema più semplice consiste nel copiarne una esistente e di eseguire le opportune rinomine e cancellazioni.

	 3D INFORMATICA	85/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

Nella directory "highway\apps" si trovano i database delle applicazioni: copiare una directory presente e rinominarla con il nome dell'applicazione che si sta creando.

All'interno di questa nuova directory, crearne un'altra con lo stesso nome, nella quale andremo a mettere il file xml da catalogare e la relativa DTD.

Il database che ospiterà i dati deve essere creato in "highway\apps\<myapps>": prendere una directory "null" e copiare il suo contenuto in <myapps>, cioè il database vuoto contenuto in "null" e rinominarlo con il nome dell'archivio dell'applicazione.

I file che compongono il database sono :

§ <myapps>.stat.xml

(rappresenta il file di definizione dell'archivio, nel quale sono memorizzate informazioni generali sul livello di accessibilità, configurabilità e cancellazione, informazioni statistiche sui documenti totali, indicizzati, cancellati, informazioni di configurazione del seriale obbligatorio NRECORD);

Equivale al file sts dell'ambiente Highway

§ <myapps>.udp

Unità documentaria point...

Equivale al file pnt dell'ambiente Highway

§ <myapps>.udd

Unità documentaria dynamic...

Equivale al file pad dell'ambiente Highway

§ <myapps>.idx, <myapps>.vcb, <myapps>.ref

invariati.

§ <myapps>.ths, <myapps>.thv

invariati.

§ <myapps>.ini

invariato

§ <myapps>.conf.xml

(Vedi. file di configurazione applicazioni Extraway)

§ <myapps>.ent

elencare quelle entity che il server deve riconoscere e decodificare.

Elencare tutte le entity effettivamente traducibili con un carattere compreso tra 1 e 255

I file contenitori dei titoli (<myapps>.tit e <myapps>.tip) sono stati aboliti.

I file <myapps>.ser, <myapps>.rlz, <myapps>.rll mantengono la loro estensione ed il significato.

Per esigenze di cross-platform, tutti i file che sono rimasti invariati rispetto all'ambiente Highway, hanno modificato il loro formato interno perdendo compatibilità con il passato.

	 3D INFORMATICA Extraway technical reference	86/93
03/08/2003		Ver. 0.0.4

Dobbiamo ancora impostare le caratteristiche d'archivio dei singoli campi che permettono di indicare i campi chiave, le caratteristiche delle indicizzazioni su di essi ecc..

Il file da configurare si chiamerà "<myapps>.conf.xml" (se non è presente, prelevarlo da un'altra directory e rinominarlo).

In questo file vengono dichiarati le parti articolate accessorie (p.e. thesaurus gerarchici o vincolati) che permettano di estendere la ricerca a parti vincolate o referenziate in livelli di contesto gerarchico adiacente.

Altro file di configurazione necessario è <<myapps>.stat.xml >, che contiene informazioni relative all'archivio (in quanto contenitore generale) quali documenti totali, indicizzati, ecc. validate con una firma presente all'interno dello stesso file (crc global).

Ad ogni variazione dell'archivio, questa chiave viene generata ex-novo.

Accertarsi della presenza del file di stoplist (italiano.stp), che contiene l'elenco delle parole da scartare sulla frase di ricerca (p.e. congiunzioni, articoli, ecc..).

La descrizione delle impostazioni da dichiarare in questo file esulano dagli argomenti trattati in questo manuale.

Ancora una cosa resta da fare prima di catalogare la nuova applicazione: bisogna indicare ad Extraway che esiste: nella directory "highway\conf" dobbiamo intervenire sul file xw.ini dichiarando, nella sezione "[Archivi]", una chiave che rappresenta un alias per la nostra applicazione con valore il suo percorso locale:

Esempio:

```
[Archivi]
myapps=c:\highway\apps\myapps\db\myapps\myapps.sts
```

Ora Extraway sarà in grado di individuare la nostra applicazione.

10.5.5 Configurazione struttura archivio

Questa operazione deve essere eseguita sul file <myapps>.conf.xml.

Il formato utilizzato per la strutturazione delle informazioni è l'XML, ma ricorda molto la forma dei file di impostazione, dove le sezioni sono qui rappresentate dagli ELEMENT principali.

<xway_cfg/hw_fields> rappresenta i nomi delle chiavi di ricerca - <search_key> - (i valori racchiusi fra parentesi quadre) con la relativa indicazione di su quali ELEMENT quella chiave interviene - <search_name> -; notare il formato di <search_key> che prevede un carattere che il nome inizi con un carattere "@".

<xway_cfg/primary_node> contiene i nomi degli ELEMENT che subiscono l'indicizzazione da parte di Extraway; rappresentano i nodi primari degli oggetti generati all'atto della catalogazione.

	 3D INFORMATICA Extraway technical reference	87/93
03/08/2003		Ver. 0.0.4

La sintassi individua il nome dell'ELEMENT, il tipo di dato contenuto (<value_type>) e la modalità di generazione degli indici (singola parola, frase, doppia indicizzazione o concatenazione, o ELEMENT di riferimento, ecc...).

Nell'ambito della regola di concatenazione degli ELEMENT che compongono la chiave per definire una regola di concatenamento indicheremo:

§ "@" : delimitatore del nome dell'ELEMENT;

§ "\$" : carattere di condizionamento.

Per le regole di concatenazione degli element, vedere relativa documentazione.

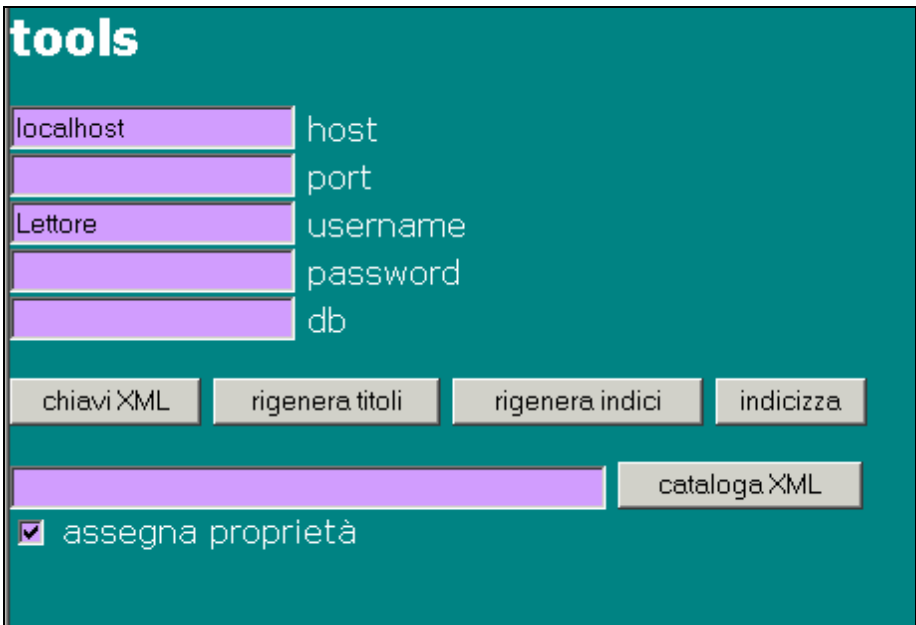
10.5.6 Catalogazione nuovo archivio

A questo punto è tutto pronto.

Dobbiamo importare i dati, che per XML si dice "catalogare", in quanto si tratta di una operazione di creazione ed allocazione degli oggetti definiti dai tipi ELEMENT del nostro file XML definitivo.

Ci viene in aiuto il framework predisposto da 3D Informatica che espone una serie di utility per l'esecuzione di operazioni amministrative.

All'indirizzo ":porta/<alias a highway>/engine/tools.jsp" sono disponibili i tools per quelle che sono le operazioni di manutenzione e mantenimento più comuni previste (l'uso della porta, a seguito del nome dell'host, nella url di accesso, permette la convivenza di più server web sullo stesso computer).



Maschera degli strumenti

Digitiamo la directory (path completo) nella quale è collocato il file xml da catalogare nella

	 3D INFORMATICA	88/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

casella di testo affiancata dal pulsante che dovremo poi premere: "cataloga XML".

Extraway a questo punto procede a creare il proprio schema, la propria allocazione gerarchica degli oggetti ELEMENT estratti dal file XML in input (con la vecchia tecnologia si potrebbe dire che sta importando i dati).

In questa fase Extraway si genera le mappe di individuazione dei file d'archivio (definisce aree, sottoaree, collezioni, oggetti; definisce l'unità informativa (l'informazione più piccola, unitaria)) e di applicazione; crea i file di indice ed espone le regole di indicizzazione per la validazione sintattica.

Dopo la generazione dello schema del nuovo archivio, Extraway costruisce automaticamente gli indici delle chiavi, così come dichiarato nel relativo file di <myapps>.conf.xml.

L'indicizzazione di un archivio può avvenire in varie fasi:

- manualmente, su richiesta dell'operatore;
- periodicamente, mediante automatismi a livello server;
- fornendo un file esterno ad Extraway;

Una nuova applicazione è stata creata, catalogata ed indicizzata.

Nessuna personalizzazione sull'interfaccia è ora disponibile; il server Tomcat non è in grado di individuarla.

Per giungere ad un minimo di funzionamento generale bisognerà avviare una procedura di configurazione dell'interfaccia e le relative impostazioni per il server TOMCAT.

E' importante specificare che è sempre possibile aggiungere nuovi ELEMENT allo schema dell'applicazione (resta sottinteso che a seguito di una modifica di schema debba avvenire una indicizzazione, completa o per la sola nuova parte).

	 3D INFORMATICA	89/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

10.6 Personalizzare l'interfaccia

La parte relativa alla grafica ed al comportamento dell'applicazione viene impostata nelle relative sottodirectory di "`highway\public`".

La comunicazione fra le classi java ed il client (tramite Tomcat) avviene grazie ad un piccolo modulo che riceve la richiesta dal client e la impacchetta per Extraway; quindi riceve la risposta da Extraway e la impacchetta per il client.

Questo file, "`<myapp>.jsp`", è memorizzato in "`highway\public\application\<myapp>engine`" è uguale per tutte le applicazioni (almeno fino a che qualche funzionalità particolare non ne richiede la modifica) e quindi deve essere copiato da una locazione in cui sia presente e rinominato in "`<myapp>.jsp`".

10.6.1 Configurazione del system

La parte relativa ai comportamenti ed all'interfaccia dell'applicazione è collocata nella directory "`highway\public`"; si dettagliano alcune locazioni principali:

- § `highway\public\application\<myapp>engine`: moduli che ricevono ed elaborano le richieste del client e generano l'output da ritornare al browser;
- § `highway\public\application\<myapp>engine\template`: DTD di validazione personalizzate sui verbi (leggi azioni) che vengono utilizzate per la composizione e validazione dell'output da inviare al client;
- § `highway\public\application\<myapp>formatter`: stili e template html per la composizione dell'interfaccia di output;
- § `highway\public\application\<myapp>formatter\factory`: xls di stile per la generazione dell'interfaccia client;
- § `highway\public\application\<myapp>formatter\html`: template html che forniranno l'interfaccia, opportunamente formattata dagli stile xsl, verso il client.

10.6.2 Configurazione properties di Tomcat

Tomcat, per il suo corretto funzionamento e per l'individuazione delle applicazioni, deve essere correttamente configurato.

Il file su cui intervenire si trova in "`highway\javaltomcat`" e si chiama "`it.highwaytech.apps.<myapp>.properties`".

Questo file imposta i comportamenti, le caratteristiche e le locazioni dell'applicazione `<myapp>`.

Qui vengono impostati i percorsi per l'individuazione dei file di stile e di schema di `<myapp>`

	 3D INFORMATICA	90/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

sia mediante notazione su disco locale (e relativa LAN) che mediante notazione remota (indirizzo internet):

- § localDTDPath;
- § localStylesheetPath;
- § remoteDTDPath;
- § remoteStylesheetPath.

Siccome <myapp> potrebbe essere composta da più database, in questo file andranno indicati i riferimenti ad ognuno di essi, sullo stile dei file di impostazione (ini).

Per esempio:

```
#<database1> si identifica con la sezione di ini [<database1>]
```

All'interno di ognuno di questi alias di database saranno indicati i percorsi agli stili da applicare relativamente ai verbi (leggi azioni) a cui sono assoggettati i database dalle richieste dell'utente.

Per esempio:

```
xsl.query.<database1>=<path parziale rispetto all'applicazione>\<stile>.xsl
xsl.showdoc.<database1>=<path parziale rispetto all'applicazione>\<stile>.xsl
...
```

10.6.3 Personalizzazione DTD sui verbi

In base al contesto in cui vengono mostrati i dati sulle richieste dell'utente (verbi) è necessario, a scopo di validazione, creare apposite DTD che possono espandere la DTD del database per maggiormente descrivere i dati.

10.6.4 Personalizzare stile e template html

A questo punto si rende necessario entrare nel codice dei file di stile (xsl) e dei template html per costruire l'interfaccia utente personalizzata sul database.

Dal framework, tipicamente da una precedente applicazione, si copiano i file esistenti e si procede con una prima fase di rinomina dei file e dei nomi degli ELEMENT in esso contenuti, e quindi della grafica esposta nei template html.

	 3D INFORMATICA	91/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

11. Operazioni amministrative

All'indirizzo "":porta/<alias a highway>/engine/tools.jsp", come dichiarato in "Catalogazione nuovo archivio", sono rese disponibili alcune delle più comuni operazioni amministrative per la manutenzione degli archivi.

Le operazioni, al momento della stesura della presente documentazione, preconfezionate previste sono:

- § visualizzazione chiavi xml;
- § generazione titoli (stringa composta da una serie di ELEMENT che identifica il contenuto del documento);
- § rigenerazione indici (re-indicizzazione del database);
- § indicizzazione di sole parti aggiuntive;
- § catalogazione xml (riconoscimento degli oggetti, dei relativi percorsi, e conseguente indicizzazione);
- § casella di spunta che permette di assegnare la proprietà del file indicato nella casella di testo ad Extraway (che lo marcherà con un timestamp in modo da riconoscerlo e riconoscere l'ultimo intervento su di esso effettuato).

	 3D INFORMATICA Extraway technical reference	92/93
03/08/2003		Ver. 0.0.4

12. Impostazioni di configurazione modulo XIPS

Extraway Inter Process Server (XIPS) cioè il modulo di intercomunicazione fra server Extraway.

Installato con Extraway si occupa di mettere in comunicazioni i vari server Extraway dislocati su diverse macchine, attraverso il server XIPS della macchina alla quale accedere.

✓ **Xips_cfg** : elemento radice;

Ø **c_class** : classe di riferimento di indirizzo IP con il quale il server XIPS comunica utilizzare la forma "xxx.xxx.xxx"

§ **mask**

Ø **server**

§ **pool_id** : identifica il quarto elemento dell'indirizzo IP definito in c_class

Xips.conf.xml può essere usato indipendentemente su tutte le macchine indicate in "pool_id".

Ogni macchina, indipendentemente dal numero di IP address disponibili, troverà il proprio indicatore di appartenenza al pool cercando tra i propri indirizzi ip quello che corrisponde a 194.243.252.? (con una maschera calcolata automaticamente pari a 255.255.255.0).

Le macchine sapranno quindi scambiarsi i dati riconoscendosi con questa classe di indirizzi comune.

ExtraWay provvede ad eseguire una istanza del server Xips (se già in essere, quella sopravvive) e dialoga con questo server (o con il server Xips di altre postazioni) per la comunicazione tra server dello stesso Pool (più Host votati alla stessa missione).

	 3D INFORMATICA	93/93
03/08/2003	Extraway technical reference	Ver. 0.0.4

13 Installazione

L'installazione del server ExtraWay prevede:

? il modulo MSG.DAT;

? il file di impostazione XW.INI (in una directory "conf" parallela a quella di ExtraWay) molto simile ad Highway.ini;