

Ordinamento ricerche

Composizione delle regole di ordinamento delle selezioni

- [Composizione delle regole di ordinamento delle selezioni](#)
 - [Introduzione](#)
 - [I modificatori](#)

Introduzione

In primo luogo vediamo cosa si intenda per ordinamento e cosa siano gli oggetti ordinati. Ad essere ordinate sono solitamente delle unità informative ma in alcuni casi particolari si può assumere che siano dei sottinsiemi di queste unità ad essere ordinati.

L'ordinamento viene effettuato sulla base di valori contenuti in attributi o elementi delle unità. A tali componenti possono essere applicati di volta in volta una serie di modificatori che vedremo di seguito.

La sintassi dell'ordinamento deriva da quella precedentemente concepita per il programma HighWay Server che basa il tutto sull'identificatore del campo sul cui contenuto effettuare l'ordinamento. In eXtraWay il concetto di Campo viene naturalmente sostituito dal concetto di Canale che è appunto riconducibile al contenuto degli attributi ed elementi dell'unità informativa.

In generale, comunque, la sintassi è rappresentata da

```
NomeCampo[(eventuali modificatori)] [, altro nome campo coi suoi modificatori]
```

Come detto il concetto di NomeCampo andrebbe sostituito da un concetto di Canale che è di fatto rappresentato dal percorso XML che identifica la porzione di unità informativa che ci interessa sfruttare. Se in HighWay il nome del campo era sufficiente, in eXtraWay tale nome campo può assumere solo i valori NRECORD, UD e XML. Il primo rappresenta il numero logico dell'unità informativa (che tipicamente corrisponde al suo numero fisico), il secondo rappresenta il campo che ospita informazioni di servizio ed il terzo è quello che contiene la parte dell'unità informativa vera e propria corrispondente al frammento XML identificato dal catalogo. I valori dei percorsi XML validi entro tali Campi può essere estratto dalla struttura dei canali rilevati in indicizzazione da parte del server e tornata al Client all'atto del caricamento dell'archivio. A tale nome campo va quindi **sempre** associato un modificatore, xpart, quello che rappresenta la porzione dell'XML di nostro interesse. Il percorso XML indicato può contenere degli asterischi per indicare che l'oggetto in esame ha una posizione non nota o comunque non sempre uguale. In fine, il nome del campo porta ad un ordinamento crescente se esso viene espresso in caratteri maiuscoli (è sufficiente il primo) ed un ordinamento inverso, quindi decrescente, se espresso in caratteri minuscoli.

Ogni componente dell'ordinamento sarà separata dalla successiva da una virgola. Le componenti vengono considerate nell'ordine in cui sono state espresse, da sinistra verso destra: a valore uguale di un particolare componente si passa alla valutazione del successivo e così via. Si noti che due unità informative che contengano, per i Canali espressi, **esattamente lo stesso valore**, avranno lo stesso valore nell'ordinamento ma, se pure adiacenti, potranno essere indicate in un qualsiasi ordine. In pratica, ordinando più volte la stessa selezione, la posizione entro di essa di unità informative equivalenti potrebbe essere invertito.

La regola di ordinamento viene indicata al server in due modi:

- **Configurazione di Default:** Un archivio può essere configurato per avere un particolare ordinamento predefinito. Esso trova posto nel file nomearchivio.conf.xml in un elemento profile che ha due attributi. Il primo attributo, type, deve assumere il valore sort.default. Il secondo, value, assume il valore della regola espressa. A tale profilo ne può essere associato un'altro di avente tipo sort.defaultlimit il cui valore deve rappresentare il numero di documenti oltre il quale l'ordinamento, considerato troppo oneroso, non viene effettuato.

Esempio

```
<profile type="sort.defaultrule" value="XML(xpart:/civile/@datdec:d), XML(xpart:/civile/@numdec:n)"/>
<profile type="sort.defaultlimit" value="100"/>
```

- **Comandi Specifici:** Nell'esecuzione di ricerche oppure, ovviamente, di semplici ordinamenti, la regola di ordinamento viene inviata al server direttamente entro il comando di ricerca o ordinamento.

I modificatori

Scopo dei modificatori è quello di dichiarare con che modalità eseguire l'ordinamento della componente identificata o di unire tra loro componenti da combinare.

I modificatori vengono espressi immediatamente dopo il NomeCampo, senza alcuno spazio di separazione, e sono racchiusi tra parentesi tonde. Se all'interno delle parentesi tonde un valore può contenere a sua volta una tonda, in particolare una tonda chiusa, essa va escapata per mezzo del carattere '\'. L'ordine dei modificatori non è particolarmente significativo se pure si suggerisce di rifarsi se possibile all'ordine in cui essi sono presentati in questa documentazione. Anch'essi non sono separati da spazi ed in pratica è bene che non ci siano spazi lungo tutta la sintassi del Canale di ordinamento.

Vediamo nel dettaglio la loro sintassi ed il loro significato:

- **xpart** Indica di voler ordinare secondo il valore della parte del frammento XML identificabile per mezzo del percorso indicato. Al termine del percorso indicato è anche possibile indicare se l'ordinamento del valore estratto debba essere effettuato considerandolo alfanumerico (default) oppure numerico o data. Per impostare questo ulteriore parametro, dopo il percorso XML, dovremo indicare :d oppure :D per indicare che esso è di tipo 'data' ed indicare :n oppure :N per indicare che esso è di tipo numerico.

Avremo ad esempio: XML(xpart:/civile/@datdec:d), XML(xpart:/civile/@numdec:n)

Una simile sintassi indica che si vuole prendere la componente del frammento XML nota come /civile/@datdec, interpretandola come 'data' e di seguito la componente /civile/@numdec interpretandola come numerico. L'interpretazione di un componente sotto forma di 'data' comporta il tentativo di ricondurre il suo contenuto, qualsiasi sia la sua forma, al formato YYYYMMDD.

Per le parti la cui posizione non sia nota o non sia sempre certa si possono indicare dei percorsi con asterischi come ad esempio: XML(xpart:/NIR/asterisco/dataDoc/@norm:d)

- **part** Indica che del componente appena identificato si vuol prendere solo un sottinsieme del suo contenuto. La forma di questo modificatore è: (part: offset:size). L'offset, indicato in base 0, esprime il punto di partenza e size quanti caratteri prendere in considerazione. Se offset è un valore negativo si assume di partire dal fondo del valore, quindi da destra, anziché dall'inizio (al quale corrisponde l'offset 0) ovvero da sinistra. La lettura avviene sempre da sinistra verso destra quindi velendo acquisire gli ultimi 4 caratteri avremo una forma del tipo:(part:-4:4)
- **e_i_w** Sta per Empty Is Worst. Indica che, indipendentemente dall'ordine crescente o decrescente richiesto per questo Canale, se il valore risulta essere vuoto esso sarà messo in coda e quindi considerato al pari del più grande valore in caso di ordine crescente o del più piccolo in caso di ordine decrescente. Se corredato dall'ulteriore modificatore **absolute** esso indica che la posizione peggiore viene assunta anche nei confronti dei Canali meno prioritari, quelli indicati a destra del canale corrente nella regola di ordinamento. La forma è quindi (e_i_w: absolute)
- **break** Richiede che il Canale indicato funga da spartiacque per identificare le unità informative ridondanti. Sono considerate ridondanti le unità informative che siano uguali, in ordinamento, per tutti i Canali precedenti quello attuale ed il Canale stesso. Nella selezione finale le unità informative "spartiacque" saranno opportunamente identificati per scopi futuri. Se si applica anche l'ulteriore modificatore **skip** la selezione ordinata finale conterrà esclusivamente le unità informative non ridondanti, quindi la prima unità informativa avente le caratteristiche richieste mentre le altre equivalenti saranno ignorate. La forma sarà quindi (break) quando si vuole identificare lo spartiacque senza impattare sull'esito dell'ordinamento ovvero (break:skip) se vogliamo che la selezione ordinata finale contenga una sola unità informativa avente uguali valori nei canali sino ad ora espressi. A dire il vero, indicare altri canali a destra del canale che presenta questo modificatore non ha molto senso se non per l'elezione della prima unità documentaria nell'ordinamento e quindi quella che verrà eletta.
- **join** Unisce, nell'ordinamento, due Canali attigui. Se nella sequenza di Canali si intende ordinare per la somma dei valori presenti in due di essi o per il valore che può trovarsi in due Canali alternativi e solitamente non compresenti nell'unità informativa. Si può quindi operare per somma o per sostituzione. Quest'ultimo è il comportamento di default ma è sempre possibile indicarlo per evitare dubbi aggiungendo i modificatori add o alt (per somma e sostituzione).

Vediamo quindi due casi:

XML(xpart:/arc/@datdec:d), XML(xpart:/arc/@datdep:d)(join:alt)

Con quest'indicazione si richiede di ordinare per il valore dell'attributo datdec. Se tale valore è assente o semplicemente vuoto, il suo posto viene preso dal valore di datdep.

XML(xpart:/arc/@datdec:d)(part:0,4), XML(xpart:/arc/@sez:d)(join:add)

Con quest'indicazione si richiede di ordinare per i primi 4 caratteri del valore dell'attributo datdec una volta interpretato come data, quindi per l'anno in esso esistente. A tale valore si vuole sommare il contenuto dell'attributo sez. L'ordinamento avrà luogo per un valore solo, come se agli 'n' Canali indicati corrispondesse effettivamente solo uno di essi. Va da se che ciò differisce dall'ordinamento per canali successivi, in particolare nel caso alt

- **instance** Richiede che il Canale indicato, se presente più volte nell'unità informativa comporti la presenza nell'ordinamento della stessa unità più volte, tante quante sono le ripetizioni del Canale indicato. In questo caso il risultato dell'ordinamento presenterà un numero di unità informative maggiore o uguale alla selezione originale. Il risultato consente, con apposite modalità di interrogazione, di conoscere non solo l'unità informativa selezionata ma l'indice al suo interno della ripetizione del Canale indicato.

Esisteva anche il modificatore **title** che è divenuto obsoleto in eXtraWay. Se pure ancora supportato per compatibilità col passato non lo si documenta perché non ritenuto necessario.