

Archivi Lucene-Like o Blind

Si dicono archivi “Lucene-Like” o archivi ciechi (da cui “Blind”) quegli archivi che non hanno l'intenzione di rappresentare un XML nel senso stretto del termine ma che consentono ad eXtraWay di agire nei confronti di un client come un server Lucene. In questi archivi c'è la massima libertà di esprimere unità informative e chiavi senza far ricorso al file di configurazione d'archivio.

La [procedura per la loro creazione](#) differisce lievemente dalla creazione di un archivio standard ma garantisce che l'archivio prodotto abbia già tutta la configurazione necessaria allo scopo.

In particolare, quindi, tali archivi saranno degli archivi “ciechi” ovvero archivi che automatizzano la gestione delle unità informative, delle relative chiavi e dei loro alias.

Le principali caratteristiche sono le seguenti.

- Nel file di configurazione non si citano le unità informative. Esse vengono desunte dal comando di salvataggio dei documenti (parametri `pne` e `pnce` che non sono più facoltativi) ovvero dagli XML stessi (tutti gli elementi figli diretti dell'elemento radice del XML sono considerate unità informative);
- Il file di configurazione non richiede più l'espressione di un alias. Se si fa uso di un alias non espresso esso viene sostituito d'ufficio dall'espressione “XML,/*/<alias>”;

Inoltre, per non essere tenuti alla dichiarazione delle caratteristiche di chiave, oltre ai default previsti dal server per elementi ed attributi¹⁾ è stata introdotta un'ulteriore modalità che consente di stabilire in che modo realizzare la chiave per un determinato elemento.

Nota: L'uso di questa modalità va compiuto con la massima attenzione in quanto il server demanda il controllo della configurazione dei propri vocabolari all'applicazione che lo pilota e quindi si corre seriamente il rischio che si creino delle incongruenze che sarebbe poi molto complesso sanare.

Si prenda un qualsiasi elemento `eLem`. Esso può essere corredato dai seguenti attributi, tutti facenti uso del namespace `xw`.

Attributo	Descrizione
xw:key_style	Equivale all'attributo <code>key_style</code> utilizzato nel file di configurazione d'archivio per indicare il tipo di chiave che si intende creare. Assume quindi i valori <code>single</code> , <code>one</code> , <code>one-md5</code> , <code>one-md5cs</code> , <code>multi</code> , <code>double</code> , <code>double-md5</code> , <code>double-md5cs</code> , <code>skip</code> , <code>spatial</code> . Il valore <code>'auto'</code> non ha senso in quanto corrisponde all'omissione dell'attributo.
xw:value_type	Equivale all'attributo <code>value_type</code> utilizzato nel file di configurazione d'archivio per indicare il tipo di valore che alimenterà il vocabolario. Assume quindi i valori <code>date</code> , <code>num</code> , <code>file</code> , <code>alnum</code> ed <code>alfa</code> che è comunque il default e può essere omesso.
xw:store	Nell'accezione <i>Lucene</i> indica, se pari a <code>false</code> , che del contenuto del campo si intendono fare gli indici ma che il suo contenuto non va registrato nel documento. <i>Il parametro è stato introdotto ma, attualmente, avendo un sistema di indici del tutto differente da quello Lucene e prevedendo la modifica che in Lucene non esiste, esso viene ignorato.</i>
xw:instance	Equivale all'attributo <code>instance</code> utilizzato nel file di configurazione d'archivio per indicare se l'elemento ripetibile può entrare a far parte di ricerche per adiacenza di gruppo ovvero bacini di pescaggio. Avendo un default negativo, si indica solitamente solo se persistono le condizioni di <i>Instance</i> con valore positivo.
xw: analyzer	Previsto dalla logica Lucene ma attualmente ignorato.

Con questa tecnica, è possibile dichiarare un qualsiasi elemento come data o numerico ovvero indicare se debba essere sottoposto ad indicizzazione mono valore, tutte condizioni che imporrebbero l'uso del file di configurazione d'archivio.

Attenzione: L'uso delle chiavi *auto-dichiarate* non è limitato ai soli archivi “Lucene Like” o “Blind”. Esse possono essere utilizzate teoricamente anche in archivi canonici ma in tal caso ci sono aspetti di performance e congruenza degli indici che non sono stati presi in esame.

¹⁾

Gli elementi vengono indicizzati come multi valore mentre gli attributi come mono valore.