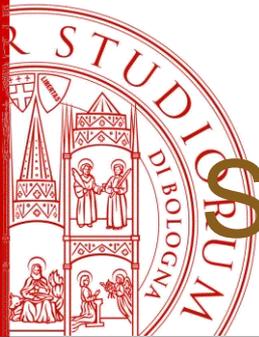


La Naming Convention di Akoma Ntoso

Fabio Vitali
+ *Monica Palmirani*

University of Bologna

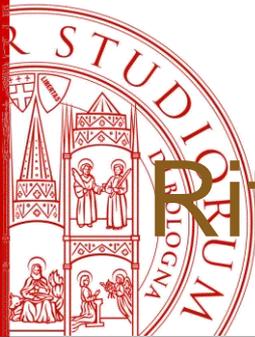


Scopo di questa presentazione

Introdurre gli ultimi sviluppi della Akoma Ntoso Naming Convention come meccanismo per indirizzare

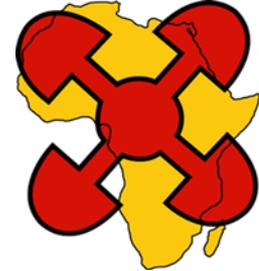
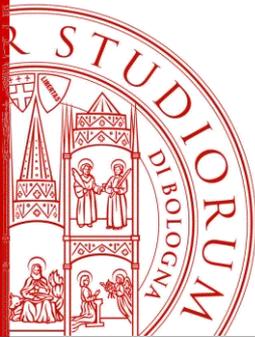
- Interi documenti
- Porzioni di documenti

... sia staticamente (ad una specifica versione) sia dinamicamente (a tutte le versioni o ad una versione decisa dinamicamente).



Riferimenti legali in Akoma Ntoso

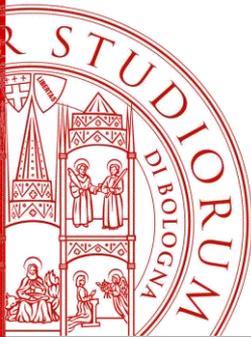
- I riferimenti legali sono raramente a documenti fisici, e più spesso a concettualizzazioni astratte di documenti: le leggi fanno riferimento a leggi, non a copie fisiche di leggi su un server da qualche parte.
- Questo è vero in particolare quando i documenti cambiano nel tempo
 - Un riferimento statico punta ad una versione congelata ad uno specifico momento nel tempo
 - Un riferimento dinamico è stato creato verso un documento che esisteva all'epoca in una specifica versione, ma ci si aspetta che esso si trasferisca alla versione più recente del documento destinazione, o più precisamente alla versione più appropriata a seconda del momento e del contesto in cui stiamo operando.



CEN Metalex

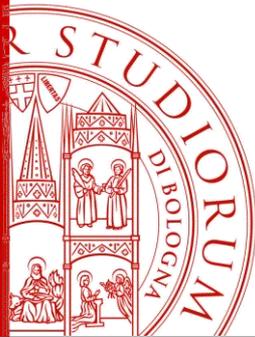
Per permettere un uso naturale degli URI, i nomi dovrebbero essere:

1. *Persistenti,*
2. *Globali,*
3. *Memorizzabili,*
4. *Significativi,*
5. *Indovinabili tra livelli,*
6. *Indovinabili tra classi di documenti,*
7. *Indovinabili tra componenti di documento.*



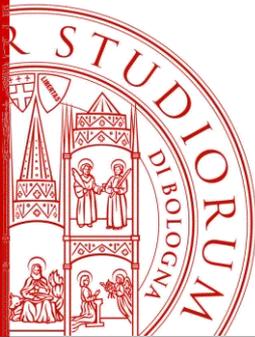
La Naming Convention di Akoma Ntoso

- Uno dei 4 documenti ufficiali in corso di approvazione dal Technical committee LegalDocML di OASIS
- Fornisce giustificazioni per la sintassi dei riferimenti legali
- Fornisce una sintassi per riferimenti normativi che usano gli URI (IRI)
- Fornisce una sintassi per riferimenti di frammenti usando gli attributi id



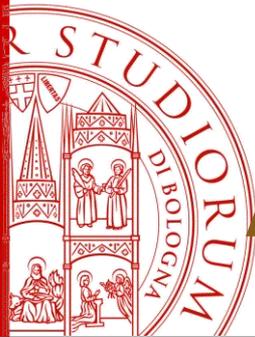
URI, IRI, URL

- Uniform Resource Identifiers (URIs) sono la sintassi usata nel World Wide Web come riferimenti a risorse (documenti e concetti, idee, cose) accessibile via web.
- Recentemente i requisiti dei parlanti di lingue diverse dall'inglese hanno portato a creare una versione degli URI che accettano caratteri non-Latini, chiamata International Resource Identifiers (IRIs)
- Uniform Resource Locators (URL) sono URI e IRI che specificano il server e la locazione dentro al server dove è disponibile un documento. Non è appropriato per documenti Akoma Ntoso.



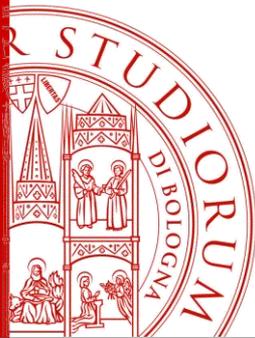
IRI assoluti, IRI references

- UN IRI assoluto non ha parti mancanti
 - E' completamente indipendente dal contenuto
 - E.g.: <http://www.governo.it/Governo/Costituzione/principi.html>
- Un IRI reference omette una o più parti a sinistra (a partire dall'inizio)
 - Le parti mancanti vengono fornite dal contesto locale (chiamato base)
 - E.g.: [Costituzione/principi.html](#)
- IRI reference globali omettono solo protocollo e host, ma hanno tutte le altre parti dopo la prima /
 - Permette a host arbitrari di fornire la risoluzione
 - E.g.: [/Governo/Costituzione/principi.html](#)



Akoma Ntoso NC: documenti

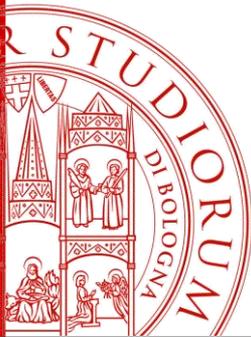
- `[protocol]://[resolver]/[workfeatures]/[expressionfeatures]/[manifestationfeatures]`
- Tutti i riferimenti a documenti sono O ad un work, O ad una expression, O ad una manifestation
 - Akoma Ntoso non tratta Items
 - Un riferimento di tipo Expression si ottiene aggiungendo informazioni ad un riferimento di tipo Work
 - Un riferimento di tipo Manifestation si ottiene aggiungendo informazioni ad un riferimento Expression
 - Tutte le informazioni sono organizzate in **features** separate da caratteri (separatori) che disambiguano la stringa



Work level features

[/akn/it/act/decreto/ministerotesoro/2013-02-22/97](#)

- Stringa fissa **/akn/**
- Country o subdivision (ISO 3166-1 or ISO 3166-2)
 - es.: **it** (Italia), **it-er** (Emilia-Romagna)
- Type (uno dei document type di Akn)
 - **act, bill, debateRecord**, etc.
- Document subtype
 - Una stringa usata nella tradizione locale (e.g.: **decreto**)
- Emanating actor
 - Se non ovvio e stabilito dal tipo di documento
- Creation date (YYYY-MM-DD or YYYY)
- Numero or titolo o feature di disambiguazione

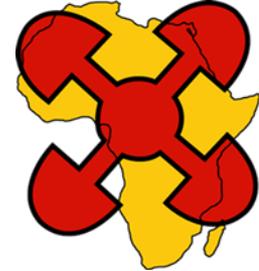


Expression level features

/akn/it/act/decreto/ministerotesoro/2013-02-22/97
ita@2014-11-04

- human language (ISO 639-2 alpha-3)
- Il carattere “@” + identificatore di versione
 - Se è una data, in sintassi YYYY-MM-DD
 - Altrimenti una stringa (e.g. il nome dello stadio di drafting di un disegno di legge)
- Il carattere ":" + view date
- Informazioni di authoring the contenuto
 - E.g. : il curatore del documento che ha creato il contenuto

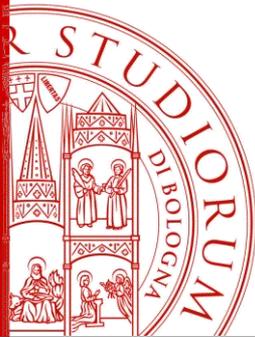
Tutti i riferimenti expression-level DEBBONO avere o il carattere @ o il carattere : Se non c'è una data, si assume che la data sia "oggi".



Manifestation level features

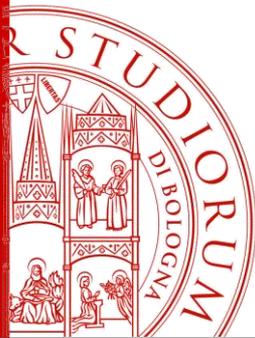
/akn/it/act/decreto/ministerotesoro/2013-02-22/97
ita@2014-11-04/***IPZS.xml***

- L'autore del markup (opzionale)
- La data del markup (opzionale)
- Il carattere “.” + specifica del formato



Identificatore di frammento

- Solo nei documenti XML!
- Specificato negli attributi **eld** e **wld** dei documenti Akoma Ntoso
 - **eld**: associato a questa espressione. Può cambiare in una nuova espressione (ad es. renumbering)
 - **wld**: associato al work. Non può mai cambiare neanche dopo pesanti modifiche in nuove espressioni. Usato solo se diverso all'**eld**.

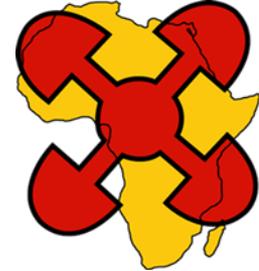
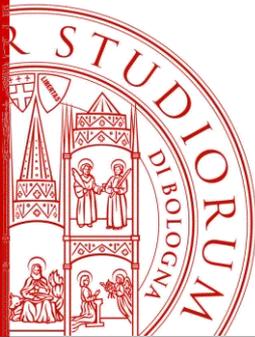


La sintassi degli identificatori di fragment

[prefix “___”] element_ref [“_”number]

- **Prefix**: una stringa (a volte vuota) che fornisce unicità al resto
- **Element_ref**: l'identificatore dell'elemento (di solito il suo nome)
- **Number**: una rappresentazione (a volte vuota) della numerazione dell'elemento.

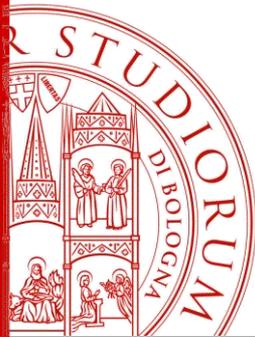
E.g.: **art_12__cl_3**



Prefisso

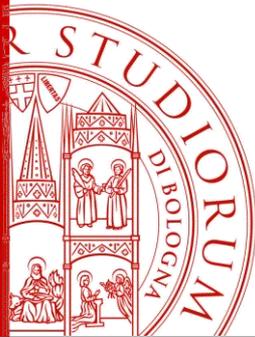
Fornisce unicità all'elemento descritto

- Elementi unici del documento ***non richiedono*** prefisso
 - E.g.: preface, body, meta, etc.
- Elementi numerati globalmente ***non richiedono*** prefissi
 - E.g.: article 137
- Elementi numerati localmente ***richiedono*** prefisso
 - E.g.: sezione 2 del Tomo I è diverso dalla sezione 2 del Tomo II
- Elementi non numerati ***richiedono*** prefisso
 - E.g.: terza riga della prima tabella dell'Annesso I
 - E.g.: il quarto eol (end of line) dopo il terzo eop (end of page)



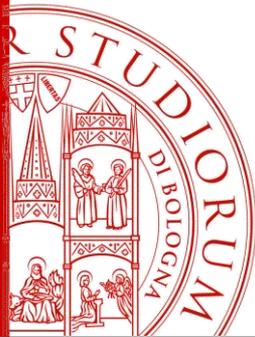
Element_ref

- Di solito, il nome dell'elemento
 - Troppi elementi per creare abbreviazioni per tutti.
- Una piccola lista di elementi usa invece abbreviazioni
 - E.g.: art per article, sect per section, cl per clause



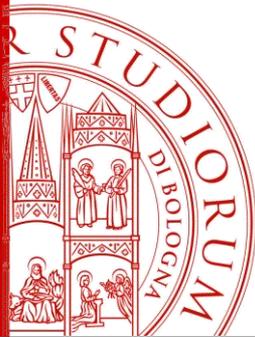
Numero

- **Elementi globalmente e localmente unici:** nessun numero
 - E.g.: c'è solo una prefazione (globalmente unico)
 - E.g.: c'è solo un num in un contenitore gerarchico (localmente unico)
- **Elementi numerati esplicitamente:** una rappresentazione del numero come usato nel testo
 - Rimuovendo tutta la punteggiatura e caratteri ovvi
 - E.g.: "art. 12 bis" -> **12bis**
- **Elementi non numerati:** stabiliti contando gli elementi precedenti e espressi sempre con numeri arabi
 - E.g.: terza riga della prima tabella dell'Annex I => **annex_1__table_1__tr_3**
 - E.g.: il quarto eol dopo il terzo eop => **eop_3__eol_4**



Conclusioni

- La naming convention fornisce un ambiente di lavoro per i link ipertestuali tra documenti.
 - Usabile nella navigazione (per documenti collegati)
 - Ma anche nel consolidamento (costruire nuove versioni combinando pezzi del documento originale e modificativo)
- Dare nomi precisi alle cose è estremamente importante per fare funzionare queste feature in maniera corretta e automatica.
- Gli errori impediranno servizi automatici e aumenteranno le esigenze di controlli e aggiustamenti manuali.



AKN resolver

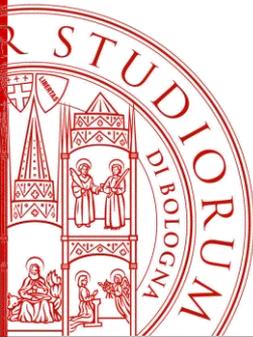
- Un servizio server-side per la risoluzione (e il suggerimento di documenti collegati) basato sulla naming convention.
- Una architettura federata di performer e risolutori che svolgono in maniera coordinata compiti diversi.



Architettura del resolver

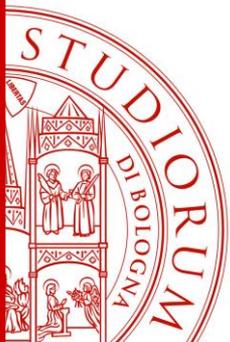
- I due moduli del resolver AKN resolver:
 - Il *performer* riceve la richiesta, ne fa parsing creando una lista di feature, identifica e chiama il resolver
 - Il *resolver* riceve una feature list, e genera la lista di best match e suggerimenti
 - Il *performer* poi esegue la azione richiesta





La lista di resolvers è la *federazione* dei performers

- Nella misura in cui usano gli stessi resolver, i vari performer sono intercambiabili
- Esiste una lista di resolver disponibili. Tutti i performer che accedono a questa lista useranno gli stessi resolver e otterranno gli stessi risultati
- E' possibile allora realizzare performer specializzati che supportino sintassi diverse e che ritorneranno sempre gli stessi risultati



Demo

<http://akresolver.cs.unibo.it/>