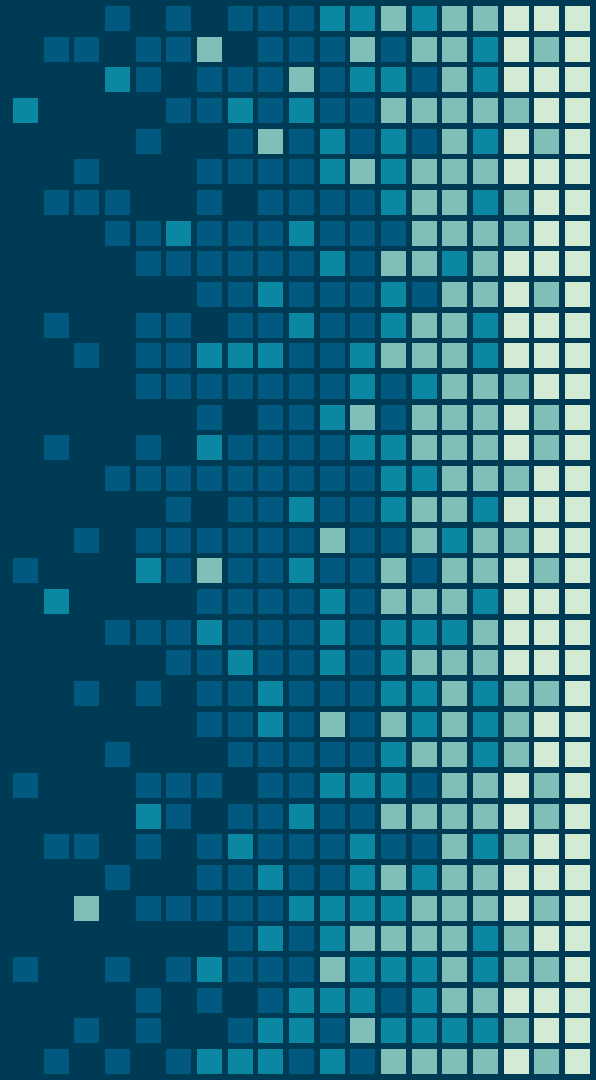


Codifica di Testi


Il progetto di codifica

Chiara Di Pietro
chiaradipietrocdt@gmail.com





Definire il modello di codifica

- Individuare le caratteristiche del documento che si intendono marcare (compresa la sua struttura)
 - Queste potranno variare da testo a testo:
 - **testo letterario in prosa:** caratteristiche stilistiche, apparato critico, interventi editoriali, nomi, etc.
 - **testo letterario in poesia:** caratteristiche stilistiche, apparato critico, interventi editoriali, metro, figure retoriche, etc.
 - **documento storico:** forme non standard/di interesse linguistico, nomi, date, etc.
 - **testo contemporaneo:** analisi morfologica, fonologica, etc.
 - **collezione di testi:** ricerche di vario tipo su corpus
- 

Definire lo schema di codifica – 1

- A meno che non si opti per la versione P5 Lite usata fin qui (**sconsigliato**), sarà indispensabile generare una versione dello schema di codifica che si è definito.
- Lo strumento da usare a questo scopo è Roma:
 - sul sito TEI: <http://www.tei-c.org/Roma/>
 - in caso di problemi con Roma scaricare template già pronte (**non saranno però accettati progetti con template TEI All**)
- Caldamente raccomandato di sperimentare più volte prima di provare a produrre la versione finale
- Salvare **sempre** la versione finale (eventualmente anche quelle di prova) in formato ODD XML oltre a DTD e schema (opzione Save customization)
- Produrre anche la documentazione

Definire lo schema di codifica – 2


- Selezionare i moduli TEI (o anche i singoli elementi) necessari per codificare le caratteristiche desiderate:
 - metadati → moduli di base ([header](#))
 - struttura del testo → moduli di base ([textstructure](#))
 - caratteristiche generali → moduli di base ([core](#))
 - testi poetici → modulo [verse](#)
 - testi teatrali → modulo [drama](#)
 - trascrizioni/edizioni → modulo [transcr](#)
 - analisi di vario tipo → modulo [analysis](#)
 - corpora linguistici → modulo [corpus](#)
- Predisporre una mappa dettagliata di elementi e attributi (+ valori)

Definire lo schema di codifica – 3

- Potrebbe essere necessario definire uno o più nuovi elementi.
- Questi potrebbero anche essere semplicemente *syntactic sugar* per elementi TEI:
 - ad esempio si potrebbe decidere di creare un nuovo elemento **<sv>** invece di usare **<seg type="semiverso">**
- I nuovi elementi possono essere definiti al momento di generare lo schema di codifica usando Roma.
 - Possibile introdurre anche nuovi attributi, ma solo all'interno delle classi esistenti.
- Preparare una documentazione da inserire al momento in cui si produrrà lo schema di codifica.



Definire la modalità di presentazione

- La modalità di presentazione deve essere progettata e preparata in base al tipo di testo codificato e alle caratteristiche che sono state marcate
 - La visualizzazione può cercare di replicare l'aspetto dell'originale:
 - edizione (semi)diplomatica
 - testo con impaginazione grafica (giornale, rivista)
 - Oppure può (o deve) scegliere un altro approccio:
 - edizione critica con note e apparato
 - trascrizione di parlato
 - corpus di testi
- 

Codifica del testo: come procedere

- Procedere attraverso una serie di fasi distinte:
 - codifica di una parte ridotta per testare lo schema di codifica (DTD o schema RelaxNG)
 - preparazione della struttura gerarchica
 - codifica di tutto il testo
 - verifica dello schema di codifica
 - eventuali variazioni da apportare allo schema di codifica
 - correzione del testo
 - inserimento metadati → importante l'<encodingDesc> per la descrizione del tipo di codifica adottato!!!
 - verifica finale

Codifica del testo: risorse

- Editor con validazione continua (XML Copy Editor, Oxygen, etc.).
- Documentazione dello schema di codifica a portata di mano (lista elementi, attributi e relativi valori),
- Finestra del navigatore aperta sulle norme TEI:
 - versione HTML consultabile sul WWW sul sito del consorzio TEI (<http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html>);
 - pacchetti compressi con Zip scaricabili dal sito TEI su Sourceforge (<http://sourceforge.net/projects/tei/>);

Fogli di stile: come procedere

- Anche in questo caso è bene suddividere il lavoro in fasi:
 - definire chiaramente l'aspetto della presentazione finale (eventualmente con schizzi su carta!);
 - definire eventuali obiettivi accessori (liste di nomi, etc.);
 - stabilire come verranno visualizzati i singoli elementi;
 - controllare la lista in modo da evitare sovrapposizioni;
 - scrivere il foglio di stile cominciando dallo scheletro (X)HTML, passare poi agli elementi strutturali e per ultimi quelli a livello di frase
 - verificare ogni singola regola, passare alla successiva solo se tutto va bene

Fogli di stile: risorse

- Editor con validazione continua (i fogli di stile XSL sono documenti XML validabili!)
- Documentazione per XSL:
 - specifiche e link: <http://www.w3.org/Style/XSL/>
 - guide e tutorial: [XSL Transformations](#) e [XSLT Tutorial](#)
- Documentazione per CSS:
 - specifiche e link: <http://www.w3.org/Style/CSS/>
 - guide e tutorial:
 - [CSS HTML.it](#), [Beginner's guide to CSS](#) e [Tableless layout HOWTO](#)



Fogli di stile: bibliografia minima

- Elliotte Rusty Harold, *XML 1.1 Bible*. Wiley, 2004
 - Michael H. Kay, *XSLT Programmer's Reference* 2nd Edition. Wrox, 2001.
 - Eric A. Meyer, *Cascading Style Sheets 2.0 Programmer's Reference*. Osborne, 2001.
 - NB: i due testi su XSLT e CSS sono *programmer's reference*, da usare come riferimento, ma non come introduzione e tutorial.
- 