

# **INTRODUZIONE ALL'INFORMATICA UMANISTICA**

## **Corso online di formazione per insegnanti**

# ***Creare il testo digitale***

## **Linguaggi, procedure e modelli**

**Francesca Tomasi**

Dipartimento di Filologia Classica e Italianistica

Università di Bologna

[francesca.tomasi@unibo.it](mailto:francesca.tomasi@unibo.it)

**DIGITAL HUMANITIES AND DIGITAL KNOWLEDGE (DHDK)**

International second cycle degree

<http://corsi.unibo.it/2Cycle/DigitalHumanitiesDigitalKnowledge>

**DIGITAL HUMANITIES ADVANCED RESEARCH CENTRE (/DH.arc)**

<http://dharc.unibo.it>

**PATRIMONIO CULTURALE NELL'ECOSISTEMA DIGITALE (CHede)**

Phd

<https://phd.unibo.it/chede/it>



# In sintesi, cosa vedremo?

1. Perché creo un **testo digitale**? Qual è lo scopo? Conservare a lungo termine, manipolare, interrogare, disseminare, rendere universalmente usabile e riutilizzabile?
2. Rapporto **testo/paratesto**. Descrivere un testo digitale: standard e modelli per il paratesto. Un esempio in ambiente Web: il Dublin Core (DC) e le sue possibili linearizzazioni. I **metadati**.
3. Creare un testo digitale: il ruolo del **linguaggio XML**. Realizzazione di un XML ben formato. La costruzione di un file XML valido con DC. Il ruolo degli Schemi di marcatura.
4. Usare XML per creare il testo digitale. Il passaggio dal paratesto alla marcatura del full-text. La funzione della **Text Encoding Initiative (TEI)** per header (paratesto) e body (testo).
5. Arricchire la marcatura. L'**identificazione univoca delle risorse digitali** e dei suoi elementi descrittivi.
6. Il processo di **trasformazione**: da XML a HTML. L'accesso via Web



# Un esempio. Il catalogo di libri digitali di un editore

Cosa ci serve per produrre un catalogo?

- ✓ La descrizione del libro (componenti paratestuali)
- ✓ L'accesso al testo pieno, possibilmente in diversi formati (PDF, HTML, XML, Epub, etc...)
- ✓ Delle categorie per la classificazione della risorsa ed il browsing della raccolta
- ✓ Degli strumenti di analisi, esplorazione o arricchimento

Ovvero: una serie di file, contenenti la descrizione del paratesto (metadati) e del testo (full-text), conservati in un qualche disco rigido, accessibili attraverso un'interfaccia.

Un esempio internazionale:

<https://www.openbookpublishers.com/product/483>

OpenBook Publishers Catalogue About Publish with us Libraries Sales Open Software Support us

**Digital Humanities**

**Digital Scholarly Editing**  
Theories and Practices

Matthew James Driscoll (editor) • Elena Pierazzo (editor)

READ ONLINE DOWNLOAD PRINTED EDITIONS

This volume presents the state of the art in digital scholarly editing. Drawing together the work of established and emerging researchers, it gives pause at a crucial moment in the history of technology in order to offer a sustained reflection on the practices involved in producing, editing and reading digital scholarly editions—and the theories that underpin them.

The unrelenting progress of computer technology has changed the nature of textual scholarship at the most fundamental level: the way editors and scholars work, the tools they use to do such work and the research questions they attempt to answer have all been affected. Each of the essays in Digital Scholarly Editing approaches these changes with a different methodological consideration in mind. Together, they make a compelling case for re-evaluating the foundation of the discipline—one that tests its assertions against manuscripts and printed works from across literary history, and the globe.

The sheer breadth of Digital Scholarly Editing, along with its successful integration of theory and practice, help redefine a rapidly-changing field, as its firm grounding and future-looking ambit ensure the work will be an indispensable starting point for further scholarship. This collection is essential reading for editors, scholars, students and readers who are invested in the future of textual scholarship and the digital humanities.

Endorsements Metrics Contents Contributors

**Endorsements**

*The quality of these essays is uniform and high; they represent the state of the art in this area. The essays range over all the important technical and intellectual debates in digital scholarly editing and provide an excellent introduction to the field as well as a report on where we are. There are some fine discussions of the thorny theoretical topics as well as contributions that discuss particular projects without falling into the tedious show-and-tell format: we always hear why something matters.*

**Professor Gabriel Egan**  
Director of the Centre for Textual Studies at De Montfort University

**Metrics**

Usage by Measure

Google Books Views  
OAPEN Downloads  
OBP Downloads  
OBP HTML Reader Sessions  
OBP PDF Reader Sessions  
Open Edition Downloads  
World Reader Users  
Retail Distributors Downloads

Measures Timeline Geographical Information Full Report

**Contents**

1. Introduction: Old Wine in New Bottles? (pp. 1–16)  
Matthew James Driscoll • Elena Pierazzo  
<https://doi.org/10.11647/obp.0095.01>

2. What is a Scholarly Digital Edition? (pp. 19–40)  
Patrick Sahle  
<https://doi.org/10.11647/obp.0095.02>

3. Modelling Digital Scholarly Editing: From Plato to Heraditus (pp. 41–58)  
Elena Pierazzo  
<https://doi.org/10.11647/obp.0095.03>

4. A Protocol for Scholarly Digital Editions? The Italian Point of View (pp. 59–82)  
Marina Buzzoni  
<https://doi.org/10.11647/obp.0095.04>

Book Series  
Digital Humanities Series vol. 4 •  
ISSN Print: 2054-2410 • ISSN Digital: 2054-2429

DOI  
<https://doi.org/10.11647/obp.0095>

License  
<http://creativecommons.org/licenses/by/4.0/>

Copyright  
Matthew James Driscoll; Elena Pierazzo; Copyright of each individual chapter is maintained by the author(s).

Published On  
2016-08-15

ISBN  
Paperback 978-1-78374-238-7  
Hardback 978-1-78374-239-4  
PDF 978-1-78374-240-0  
HTML 978-1-80064-514-1  
XML 978-1-78374-627-9  
EPUB 978-1-78374-241-7  
MOBI 978-1-78374-242-4

Language  
English

Print Length  
290 pages (xvi + 274)

Dimensions  
Paperback 156 x 16 x 234 mm  
(6.14" x 0.61" x 9.21")  
Hardback 156 x 17 x 234 mm  
(6.14" x 0.69" x 9.21")

Weight  
Paperback 912g (32.17oz)  
Hardback 1294g (45.64oz)

Media  
Illustrations 22

OCLC Number  
1030403495

LCN  
2019453274

BIC  
H: U; D

BISAC  
LT000000; COM087000; COM065000

LCC  
Z286.E43

Keywords  
Digital scholarly editing; digital humanities;  
textual scholarship; computer technology; theories;  
practices

Powered by Theth

- Browse by Categories**
- Anthropology, Archaeology and Religion
  - Art and Music
  - Cinema and Photography
  - Classics
  - Digital Humanities
  - Economics, Politics and Sociology
  - Education
  - Environmental Studies
  - Health
  - History and Biography
  - History of the Book
  - History of Science
  - Prizewinning Books
  - Innovative Publications
  - Accessibility
  - What People Say About Us
  - Media Room
  - Support us
  - Sponsor Us
  - Contact Us
  - Donate
  - Vacancies, Training & Volunteering
  - FAQs
- Partnerships**
- Publishing Services and Partnerships
  - Library Membership

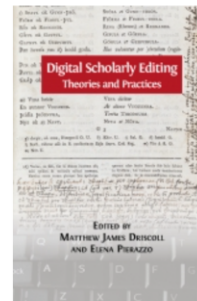
[Homepage](#) - [All Books](#) - [Digital Scholarly Editing: Theories and Practices](#)

## Digital Scholarly Editing: Theories and Practices

Matthew James Driscoll and Elena Pierazzo (eds)

<https://orcid.org/0000-0001-9662-1271>

Paperback	ISBN: 978-1-78374-238-7	£19.95	<a href="#">Add to Cart</a>
Hardback	ISBN: 978-1-78374-239-4	£30.95	<a href="#">Add to Cart</a>
PDF	ISBN: 978-1-78374-240-0	£0.00	<a href="#">Download</a>
epub	ISBN: 978-1-78374-241-7	£3.99	<a href="#">Add to Cart</a>
mobi	ISBN: 978-1-78374-242-4	£3.99	<a href="#">Add to Cart</a>
XML	ISBN: 978-1-78374-627-9	£0.00	<a href="#">Download</a>



*Digital Scholarly Editing: Theories and Practices*

Matthew James Driscoll and Elena Pierazzo (eds) | August 2016

290 | 22 colour illustrations | 6.14" x 9.21" (234 x 156 mm)

[Digital Humanities Series](#), vol. 4 | ISSN: 2054-2410 (Print); 2054-2429 (Online)

ISBN Paperback: 9781783742387

ISBN Hardback: 9781783742394

ISBN Digital (PDF): 9781783742400

ISBN Digital ebook (epub): 9781783742417

ISBN Digital ebook (mobi): 9781783742424

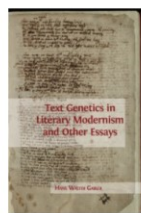
ISBN Digital (XML): 9781783746279

DOI: 10.11647/OBP.0095

BIC subject codes: H (Humanities), U (Computing and information technology), D (Literature and literary studies); BISAC:

EDU037000 (EDUCATION / Research), TEC000000 (TECHNOLOGY & ENGINEERING / General); OCLC Number: 1030403495.

### You may also be interested in:



**Categorie  
(soggetti)**

**Diversi formati  
del file  
scaricabili**

**Accesso diretto  
alla lettura**

**Metadati (dati  
catalografici/  
dati dell'oggetto  
digitale)**

**Libri correlati**



# Progetti nel settore delle DH

“At their simplest, digital projects can consist of a set files (**assets**) stored in an information architecture such as a database or file system (**structure**) where they can be accessed (**services**) and called by a browser (**use/display**).

The basic elements:

- a repository of files or digital **assets**,
- some kind of information architecture or **structure**,
- a suite of **services**,
- and a **display for user experience**.”



*DH101. INTRODUCTION TO DIGITAL HUMANITIES*  
By JOHANNA DRUCKER  
Concepts, Methods, and Tutorials for Students and Instructors, 2014,  
<http://dh101.humanities.ucla.edu/>



# **Le componenti paratestuali di una risorsa digitale, ovvero i metadati**



# Il ruolo dei metadati

**Descrizione controllata o normalizzata delle caratteristiche di un oggetto culturale (tipicamente digitale)**

**=**

**Scelta degli elementi descrittivi più appropriati sulla base di scopo dell'operazione e della tipologia di oggetto**

## **Vari modi per classificare i metadati:**

- Descrittivi, Amministrativi, Strutturali
- Contenuto, Struttura, Presentazione, Codifica
- Funzione, Obiettivo, Dominio, Comunità
- Set di Elementi, Vocabolari, Ontologie

For a complete list, see: Jenn Riley - *Seeing Standards: A Visualization of the Metadata Universe*. <http://jennriley.com/metadatamap/> Copyright 2009-2010





# Perché ci servono i metadati?

**Descrivere**

**Classificare**

**Categorizzare**

**Soggettare**

**Con lo scopo di:**

Cercare

Scoprire

Usare

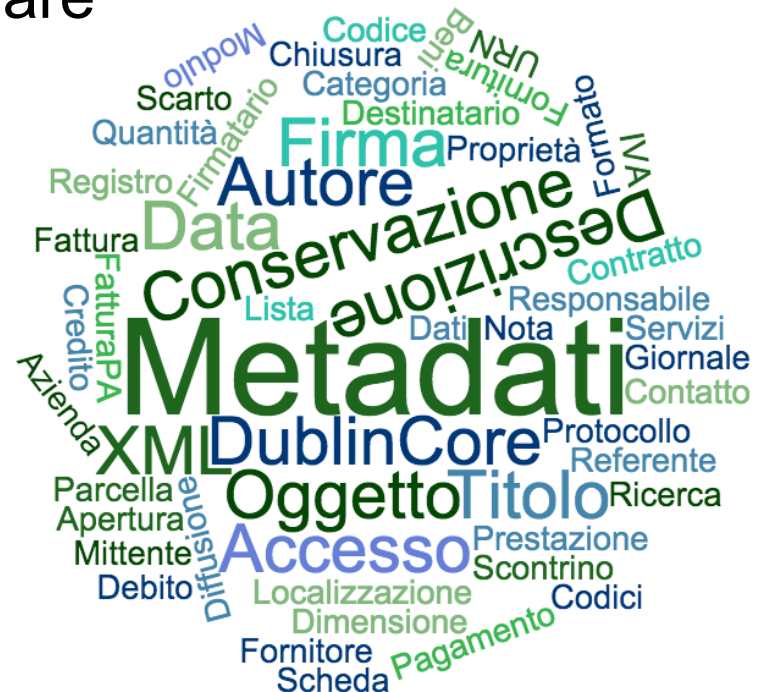
Scambiare

Ma anche:

Preservare

Documentare

(a livello amministrativo,  
tecnico e gestionale)





## SBN Catalogue

- [Basic search](#)
- [Advanced search](#)
- [Modern book](#)
- [Early printed book](#)
- [Music](#)
- [Graphic arts](#)
- [Cartography](#)
- [Authority](#)

## Other Catalogues

- [Available catalogues \(Z39.50\)](#)
- [Your catalogue](#)
- [Ancient catalogues](#)

## Services

- [Loan and reproduction](#)
- [SBN libraries](#)
- [Performed queries](#)
- [My bookmark](#)
- [Statistics](#)

Query: Any = alighieri tavoni (words in AND)

[← Back to results](#)

Format  [Hide locations](#) [Print](#) [Email](#)

Record: 3/6

- [Record Unimarc](#)
- [Download Unimarc](#)
- [Record Usmarc](#)
- [Download Marc21](#)
- [Export Endnote](#)
- [Export Refworks](#)
- [Citazioni](#)
- [Add favorites](#)
- [Permalink](#)

Alighieri, Dante

ISBD	<b>De vulgari eloquentia / Dante Alighieri ; a cura di Mirko Tavoni.</b> - Milano : Mondadori, 2017. - XCVIII, 423 p. ; 20 cm. - (Oscar classici ; 85). - [ISBN] 978-88-04-67586-0.
Livello bibliografico	Monografia
Tipo documento	Testo a stampa
Collezione	- <a href="#">Oscar classici</a> ; 85
Nomi	- <a href="#">Alighieri, Dante</a> <a href="#">scheda di autorità</a> - <a href="#">Tavoni, Mirko</a>
Classificazione Dewey	- <a href="#">450.1</a> (23.) LINGUA ITALIANA, SARDA, DALMATICA, ROMENA, LINGUE LADINE. FILOSOFIA E TEORIA DELL'ITALIANO
Lingua di pubblicazione	ITALIANO
Paese di pubblicazione	ITALIA
Codice identificativo	IT\ICCU\MIL\0929589

Le biblioteche aderenti al servizio ILL SBN, possono attivare la richiesta di prestito interbibliotecario e fornitura documenti per la propria utenza.

[Selezionare una o più biblioteche contraddistinte dal quadratino e poi cliccare su questo link.](#)

### Locations

- [B00563 UBOSB](#) Biblioteca Sala Borsa - Bologna - BO
- [CN0181 TO0HC](#) Biblioteca civica - Margarita - CN
- [MI0185 MILNB](#) Biblioteca nazionale Braidense - Milano - MI
- [MI1359 MILEL](#) Archivio della produzione editoriale della regione Lombardia - APE- RL - Milano - MI

Il record catalografico:

- ✓ Descrizione
- ✓ Classificazione
- ✓ Collegamento al record di autorità

[← Back to results](#)

[Print](#) [Email](#)



# Uno standard per i metadati: Il caso Dublin Core



# Dublin Core Metadata Element set

- ✓ E' nata in ambito OCLC la comunità del Dublin Core metadata che unisce produttori, autori e detentori dei diritti sulla documentazione in rete, al fine di fornire strumenti per l'accesso alle risorse digitali.
- ✓ Il consorzio di utilizzatori che si è costituito ha incominciato ad implementare una architettura per i metadati che venga incontro alle necessità dei venditori e dei produttori di informazioni.
- ✓ Il set minimo proposto nel dicembre 1996 è costituito da 15 elementi di base e si è esteso anche a sottoelementi o qualificatori: lo standard è in via di sviluppo ma il cosiddetto *core* dei 15 elementi della descrizione è rimasto stabile. Ciascun elemento è definito usando un set di 10 attributi.
- ✓ Nucleo di elementi essenziali ai fini della descrizione di qualsiasi materiale digitale (video, immagini, pagine Web, ecc.), nonché risorse fisiche, come libri o CD, e oggetti come opere d'arte, via rete telematica.

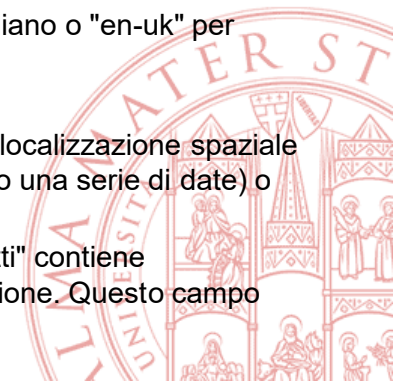
TRAD.IT: <https://www.iccu.sbn.it/it/normative-standard/linee-guida-per-la-digitalizzazione-e-metadati/dublin-core-metadata-element-set-versione-1.1-traduzione-italiana/index.html>

SITO UFFICIALE: <http://dublincore.org/> ; specifiche <http://purl.org/dc/elements/1.1/>



# Le 15 categorie del DC

1. **Titolo** (*Title*) - Nome dato alla risorsa. In particolare, un Titolo sarà un termine con il quale la risorsa è formalmente conosciuta.
2. **Autore** (*Creator*) - Entità che ha la responsabilità principale della produzione del contenuto della risorsa.
3. **Soggetto** (*Subject*) - Argomento principale della risorsa. In particolare un Soggetto può essere espresso da parole o frasi chiave, o da codici di classificazione che descrivono l'argomento della risorsa.
4. **Descrizione** (*Description*) - Spiegazione del contenuto della risorsa. Testo descrittivo libero che può includere un riasunto analitico, un indice, o una rappresentazione grafica del contenuto.
5. **Editore** (*Publisher*) - Entità responsabile della pubblicazione della risorsa. Esempi di Editore possono essere una persona, un'organizzazione o un servizio che si occupa di rendere disponibile la risorsa nella sua forma attuale.
6. **Autore di contributo subordinato** (*Contributor*) - Entità responsabile della produzione di un contributo al contenuto della risorsa. Esempi di Autore secondario includono una persona, un'organizzazione o un servizio che contribuiscono alla produzione della risorsa.
7. **Data** (*Date*) - Data associata ad un evento del ciclo di vita della risorsa. Normalmente la data è associata al momento di creazione o di disponibilità della risorsa e viene indicata attraverso una stringa di 8 caratteri nella forma YYYY-MM-DD, come definita nel profilo dello standard ISO 8601.
8. **Tipo** (*Type*) - Natura o genere del contenuto della risorsa. L'elemento "Tipo" include termini che descrivono categorie generali, funzioni, generi, o livelli di aggregazione per contenuto presi generalmente da un vocabolario controllato.
9. **Formato** (*Format*) - Manifestazione fisica o digitale della risorsa. Normalmente l'elemento "Formato" può includere il tipo di supporto o le dimensioni, ossia grandezza e durata, della risorsa.
10. **Identificatore** (*Identifier*) - Riferimento univoco alla risorsa nell'ambito di un dato contesto. Solitamente le risorse vengono identificate per mezzo di una sequenza di caratteri alfabetici o numerici secondo un sistema di identificazione formalmente definito. Esempi di tali sistemi di identificazione includono l'Uniform Resource Identifier (URI) (incluso l'Uniform Resource Locator (URL)), il Digital Object Identifier (DOI) e l'International Standard Book Number (ISBN).
11. **Fonte** (*Source*) - Riferimento a una risorsa dalla quale è derivata la risorsa in oggetto. La risorsa in questione potrebbe derivare, in tutto o in parte, da un'altra risorsa fonte.
12. **Lingua** (*Language*) - Lingua del contenuto intellettuale della risorsa. Per i valori dell'elemento Lingua si utilizza un codice di linguaggio, seguito opzionalmente da un codice di paese, entrambi su due caratteri. Ad esempio "it" per l'italiano o "en-uk" per l'inglese usato nel Regno Unito.
13. **Relazione** (*Relation*) - Riferimento ad una risorsa correlata.
14. **Copertura** (*Coverage*) - Estensione o scopo del contenuto della risorsa. Normalmente Copertura include la localizzazione spaziale (il nome o le coordinate geografiche di un luogo), il periodo temporale (l'indicazione di un periodo, una data o una serie di date) o una giurisdizione (ad esempio il nome di un'entità amministrativa).
15. **Gestione dei diritti** (*Rights*) - Informazione sui diritti esercitati sulla risorsa. Normalmente un elemento "Diritti" contiene un'indicazione sulla gestione dei diritti sulla risorsa, o un riferimento al servizio che fornisce questa informazione. Questo campo comprende gli Intellectual Property Rights, il copyright, e vari diritti di proprietà.



# DCMI - Metadata Terms

<https://www.dublincore.org/specifications/dublin-core/dcmi-terms/>

<b>Properties in the <i>/terms/</i> namespace:</b>	<a href="#">abstract</a> , <a href="#">accessRights</a> , <a href="#">accrualMethod</a> , <a href="#">accrualPeriodicity</a> , <a href="#">accrualPolicy</a> , <a href="#">alternative</a> , <a href="#">audience</a> , <a href="#">available</a> , <a href="#">bibliographicCitation</a> , <a href="#">conformsTo</a> , <a href="#">contributor</a> , <a href="#">coverage</a> , <a href="#">created</a> , <a href="#">creator</a> , <a href="#">date</a> , <a href="#">dateAccepted</a> , <a href="#">dateCopyrighted</a> , <a href="#">dateSubmitted</a> , <a href="#">description</a> , <a href="#">educationLevel</a> , <a href="#">extent</a> , <a href="#">format</a> , <a href="#">hasFormat</a> , <a href="#">hasPart</a> , <a href="#">hasVersion</a> , <a href="#">identifier</a> , <a href="#">instructionalMethod</a> , <a href="#">isFormatOf</a> , <a href="#">isPartOf</a> , <a href="#">isReferencedBy</a> , <a href="#">isReplacedBy</a> , <a href="#">isRequiredBy</a> , <a href="#">issued</a> , <a href="#">isVersionOf</a> , <a href="#">language</a> , <a href="#">license</a> , <a href="#">mediator</a> , <a href="#">medium</a> , <a href="#">modified</a> , <a href="#">provenance</a> , <a href="#">publisher</a> , <a href="#">references</a> , <a href="#">relation</a> , <a href="#">replaces</a> , <a href="#">requires</a> , <a href="#">rights</a> , <a href="#">rightsHolder</a> , <a href="#">source</a> , <a href="#">spatial</a> , <a href="#">subject</a> , <a href="#">tableOfContents</a> , <a href="#">temporal</a> , <a href="#">title</a> , <a href="#">type</a> , <a href="#">valid</a>
<b>Properties in the <i>/elements/1.1/</i> namespace:</b>	<a href="#">contributor</a> , <a href="#">coverage</a> , <a href="#">creator</a> , <a href="#">date</a> , <a href="#">description</a> , <a href="#">format</a> , <a href="#">identifier</a> , <a href="#">language</a> , <a href="#">publisher</a> , <a href="#">relation</a> , <a href="#">rights</a> , <a href="#">source</a> , <a href="#">subject</a> , <a href="#">title</a> , <a href="#">type</a>
<b>Vocabulary Encoding Schemes:</b>	<a href="#">DCMIType</a> , <a href="#">DDC</a> , <a href="#">IMT</a> , <a href="#">LCC</a> , <a href="#">LCSH</a> , <a href="#">MESH</a> , <a href="#">NLM</a> , <a href="#">TGN</a> , <a href="#">UDC</a>
<b>Syntax Encoding Schemes:</b>	<a href="#">Box</a> , <a href="#">ISO3166</a> , <a href="#">ISO639-2</a> , <a href="#">ISO639-3</a> , <a href="#">Period</a> , <a href="#">Point</a> , <a href="#">RFC1766</a> , <a href="#">RFC3066</a> , <a href="#">RFC4646</a> , <a href="#">RFC5646</a> , <a href="#">URI</a> , <a href="#">W3CDTF</a>
<b>Classes:</b>	<a href="#">Agent</a> , <a href="#">AgentClass</a> , <a href="#">BibliographicResource</a> , <a href="#">FileFormat</a> , <a href="#">Frequency</a> , <a href="#">Jurisdiction</a> , <a href="#">LicenseDocument</a> , <a href="#">LinguisticSystem</a> , <a href="#">Location</a> , <a href="#">LocationPeriodOrJurisdiction</a> , <a href="#">MediaType</a> , <a href="#">MediaTypeOrExtent</a> , <a href="#">MethodOfAccrual</a> , <a href="#">MethodOfInstruction</a> , <a href="#">PeriodOfTime</a> , <a href="#">PhysicalMedium</a> , <a href="#">PhysicalResource</a> , <a href="#">Policy</a> , <a href="#">ProvenanceStatement</a> , <a href="#">RightsStatement</a> , <a href="#">SizeOrDuration</a> , <a href="#">Standard</a>
<b>DCMI Type Vocabulary:</b>	<a href="#">Collection</a> , <a href="#">Dataset</a> , <a href="#">Event</a> , <a href="#">Image</a> , <a href="#">InteractiveResource</a> , <a href="#">MovingImage</a> , <a href="#">PhysicalObject</a> , <a href="#">Service</a> , <a href="#">Software</a> , <a href="#">Sound</a> , <a href="#">StillImage</a> , <a href="#">Text</a>
<b>Terms for vocabulary description:</b>	<a href="#">domainIncludes</a> , <a href="#">memberOf</a> , <a href="#">rangeIncludes</a> , <a href="#">VocabularyEncodingScheme</a>



# Esprimere contenuti con il DC

## Dublin Core Metadata Initiative



- The Dublin Core Metadata Initiative provides a common metadata standards for resources such as media, library books, etc.
- It defines standards for information such as:

Title	Format
Creator	Identifier
Subject	Source
Description	Language
Publisher	Relation
Contributor	Coverage
Date	Rights
Type	

- Resources can be described using:

- Text
- HTML
- XML
- RDF XML

### Sample Metadata

Format="video/mpeg; 5 minutes"  
Language="en"  
Publisher="Kats Online, LLC"  
Title="My Favorite Cat Video"  
Subject="Cats"  
Description="A short video of a black cat playing with string."



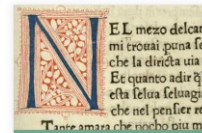
<http://dublincore.org>

# Alcuni esempi

- ✓ Raccolte digitali di testi pieni
  - Biblioteca Italiana, <http://www.bibliotecaitaliana.it/>
  - Oxford Text Archive,  
<https://ota.bodleian.ox.ac.uk/repository/xmlui/>
- ✓ Un sito Web
  - Vespasiano da Bisticci, Lettere,  
<http://vespasianodabisticciletters.unibo.it>
- ✓ Un repository istituzionale (una raccolta di prodotti scientifici)
  - Dspace - Università della Tuscia, <http://dspace.unitus.it/>







BiBit

Più di 1600 testi consultabili e liberamente scaricabili, codificati in XML-TEI, sono in edizione integrale e si fondano sulle più autorevoli edizioni di riferimento.



Incunaboli

La digitalizzazione più estesa possibile degli incunaboli volgari esistenti nelle biblioteche italiane e straniere.



Scrittori d'Italia

Riproduzione in formato digitale dell'omonima collana, fondata nel 1910 dall'editore Laterza di Bari sotto la direzione di Benedetto Croce.

## Titolo Della dissimulazione onesta

<b>Autore:</b>	Accetto, Torquato	<b>Pubblicazione:</b>	Roma: Biblioteca Italiana, 2003
<b>Genere:</b>	Trattati	<b>Periodo:</b>	600

### Descrizione fonte cartacea

<b>Autore:</b>	Accetto, Torquato	<b>Titolo:</b>	Della dissimulazione onesta
<b>Pubblicazione:</b>	Genova: Costa e Nolan, 1990	<b>Altra Responsabilità:</b>	Nigro, Salvatore Silvano; Manganelli, Giorgio
<b>Record SBN:</b>	IT\ICCU\VEA\0063161		

### Descrizione versione digitale

**Dimensione:** 72328 bytes

<http://www.bibliotecaitaliana.it/scheda/bibit001546>

### Links

[File XML](#)

[File METS](#)

[File MAG](#)

[Vai al testo](#)

```
▼ <mag:bib level="m">
  <dc:identifier/>
  <dc:identifier xsi:type="SBN">IT\ICCU\VEA\0063161</dc:identifier>
  <dc:title>Della dissimulazione onesta</dc:title>
  <dc:creator>Accetto, Torquato</dc:creator>
  <dc:publisher>Roma : Biblioteca Italiana</dc:publisher>
  <dc:subject>858.5 - Scritti miscellanei italiani. 1585-1748.</dc:subject>
  <dc:subject>Trattati</dc:subject>
  <dc:subject>600</dc:subject>
  <dc:date>2003</dc:date>
  <dc:type>text</dc:type>
  <dc:format>electronic - 72324</dc:format>
  <dc:source>Accetto, Torquato. Della dissimulazione onesta - Genova: Costa e Nolan, 1990</dc:source>
  <dc:source>IT\ICCU\VEA\0063161</dc:source>
  <dc:language>ita</dc:language>
</mag:bib>
```

Selected Klondike Short Stories

Please use the following text to cite this item or export to a predefined format:

BIBTEX

CMDI

London, Jack, 1876-1916, *Selected Klondike Short Stories*, Oxford Text Archive, <http://hdl.handle.net/20.500.12024/3044>.

Share:

f

t

g+

Oxford Text Archive

Authors

London, Jack, 1876-1916

Date of publication

1899

Type

Text

Language(s)

English

OTA identifier

ota:3044

Collection(s)

Core Collection

Show full item record

This item is **Publicly Available** and licensed under:  
Attribution-ShareAlike 3.0 Unported (CC-BY-SA 3.0)

Files for this item

Download all local files for this item (2.28 MB)

Name

3044.epub

Size

220.83 KB

Format

Unknown

Description

Version of the work for e-book readers in the EPUB format

Download file

Name

3044.html

Size

480.35 KB

Format

HTML

Description

Version of the work for web browsers

Download file

Preview

Name

3044.mobi

Size

814.96 KB

Format

Unknown

Description

Version of the work for e-book readers in the Mobipocket format

Download file

Name

3044.txt

Size

392.61 KB

Format

Text file

Description

Version of the work in plain text with all tags and formatting information removed

Download file

Preview

Name

3044.xml

Size

427.25 KB

Format

XML

Description

Version of the work in the original source TEI XML file

Download file

OXFORD

TEXT

ARCHIVE

Bodleian Libraries

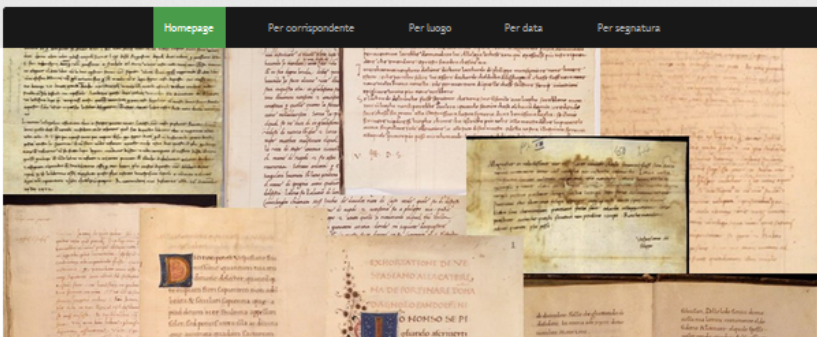
UNIVERSITY OF OXFORD

Browse	
> All of the Repository	
dc.contributor	Triggs, Jeffery North American Reading Project, Oxford University Press
dc.contributor.author	London, Jack, 1876-1916
dc.coverage.placeName	New York
dc.date.accessioned	2018-06-14
dc.date.accessioned	2019-07-04T10:31:09Z
dc.date.available	2019-07-04T10:31:09Z
dc.date.created	1899
dc.identifier	ota:3044
dc.identifier.citation	<a href="http://purl.ox.ac.uk/ota/3044">http://purl.ox.ac.uk/ota/3044</a>
dc.identifier.uri	<a href="http://hdl.handle.net/20.500.12024/3044">http://hdl.handle.net/20.500.12024/3044</a>
dc.description.abstract	Taken from the original type-setting tapes used by the Library of America: 1911-1918
dc.format.medium	Digital bitstream
dc.format.mimetype	text/xml
dc.language	English
dc.language.iso	eng
dc.publisher	University of Oxford
dc.relation.ispartof	Oxford Text Archive Core Collection
dc.relation.replaces	<a href="http://purl.ox.ac.uk/ota/1819">http://purl.ox.ac.uk/ota/1819</a>
dc.rights	Distributed by the University of Oxford under a Creative Commons Attribution-ShareAlike 3.0 Unported License
dc.rights.uri	<a href="http://creativecommons.org/licenses/by-sa/3.0/">http://creativecommons.org/licenses/by-sa/3.0/</a>
dc.rights.label	PUB
dc.subject.lcsh	Short stories, American -- 20th century
dc.title	Selected Klondike Short Stories
dc.type	Text
has.files	yes
branding	Oxford Text Archive
files.size	2392043
files.count	5
otaterms.date.range	1800-1899

<https://ota.bodleian.ox.ac.uk/repository/xmlui/handle/20.500.12024/3044?show=full>

# Vespasiano da Bisticci, Lettere

A digital edition



### Authorities

Ricerca guidata attraverso elenchi.

L'edizione è stata modificata al fine di estrarre: persona, codici, lessico, citazioni.

Authorities -

### Tavola sinottica

Tabella relazionale.

Una tabella che raccoglie: segnatura, mittente, destinatario, data cronica, data topica, fonte.

Tavola sinottica -

### Nota filologica

I principi della trascrizione.

Identificazione e descrizione di: segni della scrittura, segni dell'estensore, segni del supporto, segni dell'editore.

Nota filologica -

### I testimoni

Le testimonianze della raccolta.

La descrizione degli esemplari: codici manoscritti e di dedica, documenti d'archivio, testi a stampa.

I testimoni -

## Il corpus

- La raccolta
- I corrispondenti
- I manoscritti
- I copisti
- Le biblioteche

## Il progetto

*Vespasiano da Bisticci, Lettere*

Edizione digitale delle Lettere di Vespasiano da Bisticci, copista fiorentino vissuto nell'anno del XV secolo. Le lettere, inedite e ritrovate fino ad oggi rintracciate, sono navigabili attraverso un sistema di faccette tematiche (corrispondente, data, luogo, segnatura). L'edizione è accompagnata da strumenti filologici di orientamento e guida (Introduzione delle parole sottoforma di authorities, tavola sinottica, nota filologica, descrizione dei testimoni). Le lettere sono accompagnate da informazioni contestuali (la raccolta delle lettere, i corrispondenti, i manoscritti realizzati dalla scuola, i copisti di Vespasiano, le biblioteche prodotte) necessarie ad inserire i documenti in una prospettiva metatestuale in evoluzione e aggiornamento costante, anche in futura dimensione collaborativa.

## Guida all'edizione

### Il progetto

Il markup e l'ipertesto

### Il commento al testo

Le facce di annotazione

### La base di conoscenza

URI, RDF e ontologie

Centro di Risorsa per la Ricerca - Multimedia

L'attività di ricerca del Centro

Progetti finalizzati

Contact Information



Il progetto di edizione si inserisce nell'attività di ricerca del CRR-MM, il Centro di Risorsa per la Ricerca dell'Ateneo di Bologna.

### Produzione video

Ideazione, progettazione, ripresa, montaggio e post-produzione di audiovisivi.

### Creazione e gestione collezioni digitali

Gestione, conservazione, accessibilità nel tempo e citabilità delle collezioni multimediali archiviate.

Il CRR-MM realizza prodotti per la valorizzazione della ricerca d'Ateneo. Il Centro di Risorsa per la Ricerca Multimediale (CRR-MM) offre ai ricercatori e ai docenti dell'Università di Bologna servizi.

Attività di produzione, elaborazione, conservazione, diffusione e fruizione di materiale multimediale come collezioni di testi e immagini, di audiovisivi e registrazioni sonore, di animazioni.

Telefono: +39 051 209839  
Email: francesca.tomasini@unibo.it

Dipartimento di Filologia Classica e Italianistica  
Alma Mater Studiorum  
Università di Bologna  
Via Zamboni 31  
40126, Bologna

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

The Simple Dublin Core Ge... Vespasiano da Bisticci Letters... http://vespasianodabisticcilett...

view-source:http://vespasianodabisticciletters.unibo.it/

```
1 <!DOCTYPE HTML>
2 <!--
3 Citrus: A responsive HTML5 website template by HTML5Templates.com
4 Released for free under the Creative Commons Attribution 3.0 license (html5templates.com/license)
5 Visit http://html5templates.com for more great templates or follow us on Twitter @HTML5T
6 -->
7
8 <html>
9 <head>
10 <title>Vespasiano da Bisticci Letters. A digital edition</title>
11 <meta name="DC.Title" content="Vespasiano da Bisticci, Lettere. Edizione digitale">
12 <meta name="DC.Creator.PersonalName" content="Tomasini, Francesca">
13 <meta name="DC.Creator.PersonalName.Address" content="francesca.tomasini@unibo.it">
14 <meta name="DC.Subject" content="Markus Languages">
15 <meta name="DC.Subject" content="Edizione digitale">
16 <meta name="DC.Subject" content="XML/TEI">
17 <meta name="DC.Subject" scheme="BNCf" content="filologia umanistica">
18 <meta name="DC.Subject" scheme="BNCf" content="Letteratura del Quattrocento">
19 <meta name="DC.Description" content="Edizione digitale della raccolta di lettere del libraio fiorentino del Quattrocento Vespasiano da Bisticci">
20 <meta name="DC.Publisher" content="Universit  agrove; di Bologna, CRR-MM (Centro di Risorsa per la Ricerca - Multimediale)">
21 <meta name="DC.Date" scheme="ISO8601" content="2013-10-18">
22 <meta name="DC.Type" content="Text, Manuscript">
23 <meta name="DC.Type" content="Text, Manuscript">
24 <meta name="DC.Type" content="Text, Manuscript">
25 <meta name="DC.Format" scheme="IMT" content="text/html">
26 <meta name="DC.Identifier" content="http://vespasianodabisticciletters.unibo.it">
27 <meta name="DC.Language" content="it">
28 <meta name="DC.Date.X-MetadataLastModified" scheme="ISO8601" content="2013-10-28">
29 </head><noscript>
30 <link rel="stylesheet" href="css/Sgrid/core.css" />
31 <link rel="stylesheet" href="css/Sgrid/core-desktop.css" />
32 <link rel="stylesheet" href="css/Sgrid/core-1200px.css" />
33 <link rel="stylesheet" href="css/Sgrid/core-noscript.css" />
34 <link rel="stylesheet" href="css/style.css" />
35 <link rel="stylesheet" href="css/style-desktop.css" />
36 </noscript>
37 <script src="css/Sgrid/jquery.js"></script>
38 <script src="css/Sgrid/init.js?use=mobile,desktop,1000px&mobileUI=&mobileUI.theme=none"></script>
39 <!--[if IE 9]><link rel="stylesheet" href="css/style-ie9.css" /><![endif]>-->
40 <script>
41 (function(i,s,o,g,r,a,m){i['GoogleAnalyticsObject']=r;i[r]=i[r]||function(){
42 (i[r].q=i[r].q||[]).push(arguments)},i[r].l=1*new Date();a=s.createElement(o),
43 m=a.src=g;r=a.async=1;a.src=g;m.parentNode.insertBefore(a,m)
44 })(window,document,'script','//www.google-analytics.com/analytics.js','ga');
45
46 ga('create', 'UA-4212655-10', 'auto');
47 ga('send', 'pageview');
48 </script>
49 <body class="homepage">
50 <div id="header-wrapper">
51 <header id="header">
52 <div class="Sgrid-layout">
53 <div class="row">
54
55 <div class="12u" id="logo">
56 <!-- Logo -->
57 <h1>Vespasiano da Bisticci, Lettere</h1>
58
59 <p>A digital edition</p>
60 </div>
61 </div>
62 </div>
63 </div>
64 </div>
65 </div>
66 </div>
67 </div>
68 </div>
69 </div>
70 </div>
71 </div>
72 </div>
73 </div>
74 </div>
75 </div>
76 </div>
77 </div>
78 </div>
79 </div>
80 </div>
81 </div>
82 </div>
83 </div>
84 </div>
85 </div>
86 </div>
87 </div>
88 </div>
89 </div>
90 </div>
91 </div>
92 </div>
93 </div>
94 </div>
95 </div>
96 </div>
97 </div>
98 </div>
99 </div>
100 </div>
```



[Unitus DSpace](#) >  
[Dipartimento di Scienze Umanistiche della Comunicazione e del Turismo](#) >  
[DISUCOM - Archivio della produzione scientifica](#) >

Utilizza questo identificativo per citare o creare un link a questo documento: <http://hdl.handle.net/2067/2331>

**Titolo:** Annotating digital libraries and electronic editions in a collaborative and semantic perspective

**Autori:** [Barbera, Michele](#)  
[Meschini, Federico](#)  
[Morbidoni, Christian](#)  
[Tomasì, Francesca](#)

**Parole chiave:** Digital Libraries  
Electronic Editions  
OAC  
Semantic Annotations

**Data:** 2012

**Editore:** Springer-Verlag

**Citazione:** Barbera M., Meschini F., Morbidoni C., Tomasì F. 2012. Annotating digital libraries and electronic editions in a collaborative and semantic perspective. In: Digital Libraries and Archives, IRCDL2012. 8th Italian Research Conference on Digital Libraries (IRCDL). Bari. 9-10 February 2012. "Communications in computer and information science" 354 : 45-56

**Abstract:** The distinction between digital libraries and electronic editions is becoming more and more subtle. The practice of annotation represents a point of convergence of two only apparently separated worlds. The aim of this paper is to present a model of collaborative semantic annotation of texts (SemLib project), suggesting a system that finds in Semantic Web and Linked Data the solution technologies for enabling structured semantic annotation, also in the field of electronic editions in Digital H

[...more](#)

**URI:** <http://hdl.handle.net/2067/2331>

**ISSN:** 1865-0929

È visualizzato nelle collezioni: [DISUCOM - Archivio della produzione scientifica](#)

File in questo documento:

File	Descrizione	Dimensioni	Formato
<a href="#">paper_annotation_last.pdf</a>		218.75 kB	Adobe PDF <a href="#">Visualizza/apri</a>

Questo documento è protetto dal copyright originale

[Visualizza licenza](#)

[Visualizza tutti i metadati del documento](#)

[Suggerisci questo documento](#)

[View Statistics](#)

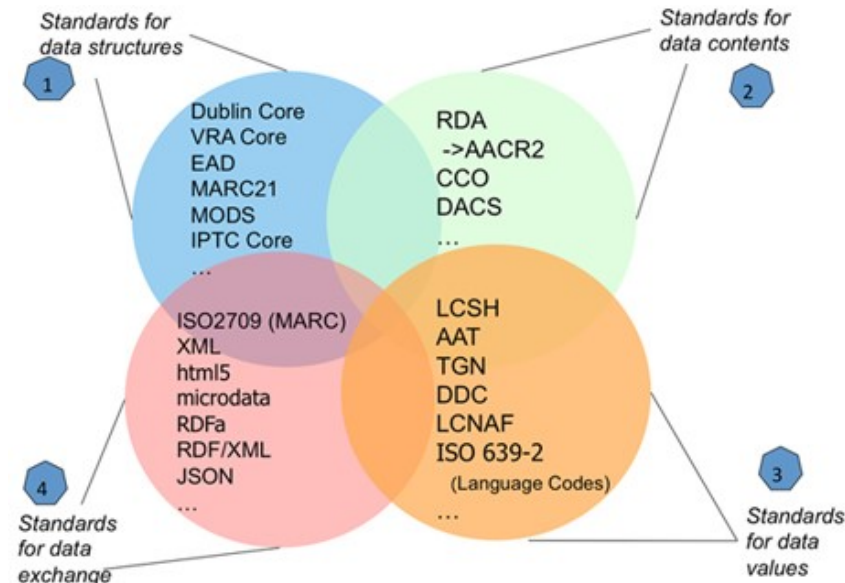
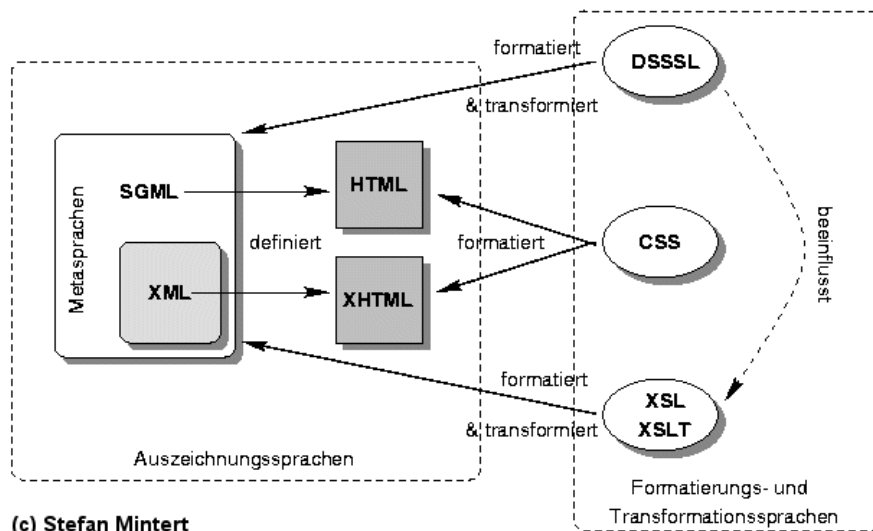
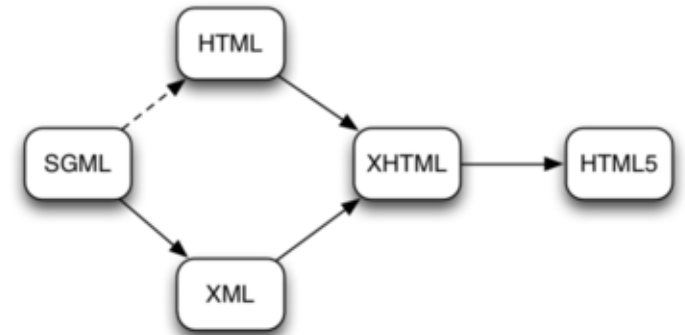
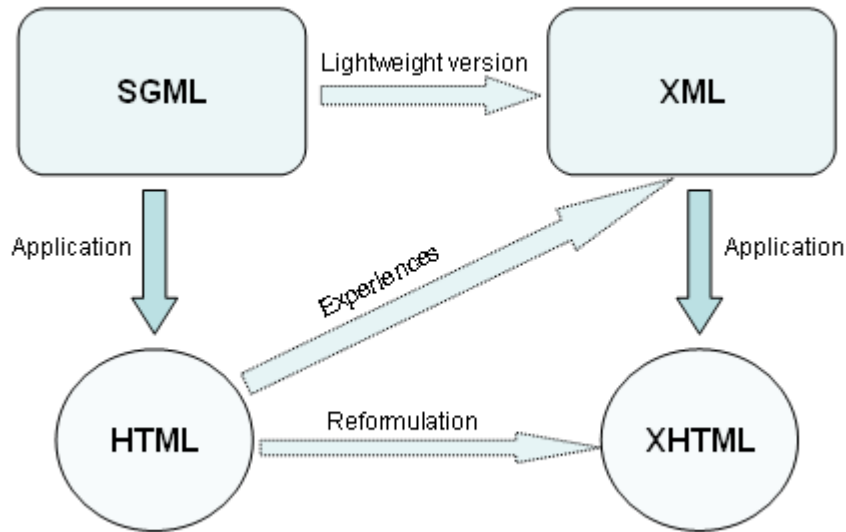
Tutti i documenti archiviati in DSpace sono protetti da copyright. Tutti i diritti riservati.



# I linguaggi di markup. XML



# Introduzione ai linguaggi di marcatura





- ✓ **Metalinguaggio** di markup, cioè un linguaggio che permette di definire “altri linguaggi” a seconda della sua applicazione.
- ✓ Standard ufficiale sviluppato dal W3C (World Wide Web Consortium) nel 1998: deriva da SGML quale suo sottoinsieme semplificato, ma ad oggi lo sostituisce. (<http://www.w3.org/XML>)
- ✓ Nasce con l'obiettivo di **rappresentare documenti** (e.g. un testo letterario) e/o **dati strutturati** (e.g. i riferimenti bibliografici) su supporto digitale.
- ✓ Un testo marcato in sintassi XML è detto **documento XML**: contiene sia il testo che i tag (anch'essi testo) utilizzati per descrivere le informazioni insite nel testo; è “**leggibile**” dall'utente senza l'utilizzo di software specifici (i.e. con qualsiasi editor di testo). E' infatti **indipendente** da qualsiasi software e hardware.





# Concetti chiave di XML

- ✓ Documenti validi e documenti ben formati
- ✓ Il concetto di tipo di documento: DTD e Schemi. I namespace
- ✓ La struttura gerarchica ad albero della descrizione
- ✓ I sistemi di validazione (il browser)
- ✓ Il ruolo degli standard per la preservazione, l'interoperabilità e il riuso
- ✓ Struttura logica vs layout

Ultima Recommendation del W3c:

<https://www.w3.org/TR/2008/REC-xml-20081126/>

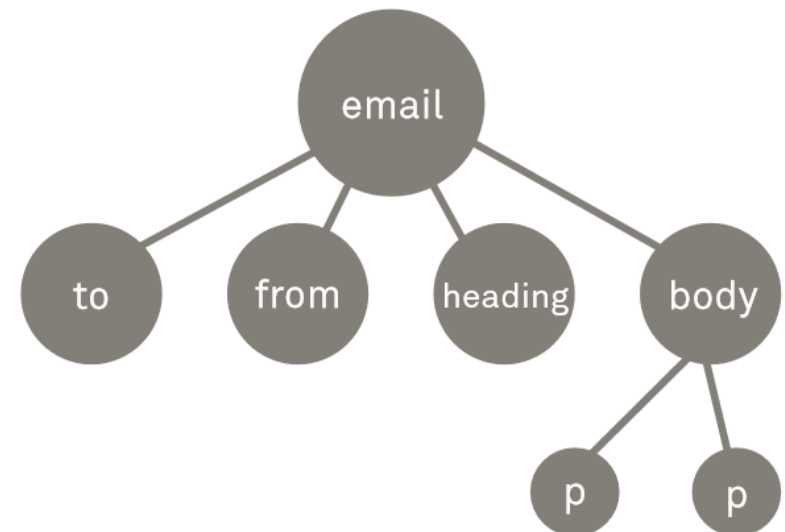


# struttura di un documento XML

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

```
<email>  
  <to>Mark</to>  
  <from>Veronika</from>  
  <heading>Reminder</heading>  
  <body>  
    <p>Don't forget me this weekend!</p>  
    <p>We have to do homeworks.</p>  
  </body>  
</email>
```

ogni documento XML (.xml) è caratterizzato da una **struttura gerarchica**, ad albero, composta da **nodi** (o **elementi**)



# sintassi di un documento XML

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

```
<email to="Mark" from="Veronika" sign="no">  
  <heading>Reminder</heading>  
  <body>  
    <p n='1'>Don't forget me this weekend!</p>  
    <p n='2'>We have to do homeworks.</p>  
  </body>  
</email>
```

N.B. non tutte le informazioni devono essere espresse come elementi

gli **attributi** sono una forma alternativa, a volte più efficiente, per descrivere gli stessi contenuti

si parla di attributo (@to) e **valore dell'attributo** ('Mark')

un attributo si riferisce all'elemento in cui è contenuto ed è riportato solo nel **tag di apertura**



# Un file .xml ben formato

```
<?xml version="1.0" encoding="UTF-8"?>
<collezione>
  <oggetto n="1">
    <titolo>The Winters Tale.</titolo>
    <autore identificativo="https://viaf.org/viaf/96994048">Shakespeare, William, 15
  <accesso>
    <licenza>Distributed by the University of Oxford under a
      <URL indirizzo="https://creativecommons.org/licenses/by-sa/3.0/">C
        3.0 Unported License</URL>
    </licenza>
    <formati>Download: XML; HTML; ePub; mobi (Kindle); plain text</formati>
    <analisi>Analysis: Explore this text with Voyant Tools (this link takes you
      <URL>voyant-tools.org</URL> website - find out more here)
    </analisi>
  </accesso>
  <lingua>English</lingua>
  <parole-chiave>
    <soggetto vocabolario="LCSH" IDvalore="sj96006296">Plays</soggetto>
    <soggetto vocabolario="LCSH" IDvalore="sh85146167">England</soggetto>
    <soggetto vocabolario="LCSH" IDvalore="sh85123070">16th century</soggetto>
  </parole-chiave>
  <descrizione-fonte>
    Revised version of http://ota.ox.ac.uk/id/0119. The texts were originally prepared
    single volume concordances to Shakespeare (OUP, 1969f). They have since been reformatted
    proofread by staff of Oxford University Press' Shakespeare Department for use in the new
    the general editorship of Dr Stanley Wells: The complete works / William Shakespeare; ge
    Taylor ; editors Stanley Wells ... [et al.] ; with introductions by Stanley Wells. -- Ox
    Shakespeare). -- ISBN 0-19-812926-2.
    <fonte>The Winters Tale. Shakespeare, William, 1564-1616 Lee, Sidney, Sir, 1
    Clarendon Press Oxford: 1902</fonte>
  </descrizione-fonte>
  <identificativo>http://purl.ox.ac.uk/ota/5729</identificativo>
</oggetto>
</collezione>
```

The Winters Tale.

Please use the following text to cite this item or export to a predefined format: <span>EXPORT</span> <span>CMDR</span>	
Shakespeare, William, 1564-1616, The Winters Tale, Oxford Text Archive, <a href="http://hdl.handle.net/20.500.12024/5729">http://hdl.handle.net/20.500.12024/5729</a> .	
Share: <a href="#">f</a> <a href="#">t</a> <a href="#">g</a> <a href="#">+</a>	
dc.contributor	Oxford Text Archive
dc.contributor.author	Shakespeare, William, 1564-1616
dc.coverage.placeName	Oxford
dc.date.accessioned	2018-06-14
dc.date.accessioned	2019-07-04T10:36:48Z
dc.date.available	2019-07-04T10:36:48Z
dc.date.created	1623
dc.identifier	ota:5729
dc.identifier.citation	<a href="http://purl.ox.ac.uk/ota/5729">http://purl.ox.ac.uk/ota/5729</a>
dc.identifier.uri	<a href="http://hdl.handle.net/20.500.12024/5729">http://hdl.handle.net/20.500.12024/5729</a>
dc.description.abstract	"One thousand copies of this facsimile have been printed"--verso of half t.p. Facsim. reprint of ed. published. London : printed by Issac Iaggard and Ed [ward] Blount, 1623 with original t.p.: Mr. William Shakespeares comedies, histories, & tragedies Original colophon reads: Printed at the charges of W[illiam] Iaggard, Ed [ward] Blount, I [ohn] Smithweeke [i.e. Smethwicke], and W[illiam] Aspley, 1623 Contents: The tempest. The two gentlemen of Verona. The merry wives of Windsor. Measure, for measure. The comedie of errors. Much adoe about nothing. Loves labour's lost. A midsommer nights dreame. The merchant of Venice. As you like it. The taming of the shrew. All's well, that ends well. Twelfth night, or what you will. The winters tale. The life and death of King John. The life and death of King Richard the second. The first part of Henry the fourth. The second part of Henry the fourth. The life of Henry the fifth. The first part of Henry the sixth. The second part of Henry the sixth. The third part of Henry the sixth. The tragedy of Richard the third. The famous history of the life of King Henry the eighth. The tragedy of Troilus and Cressida. The tragedy of Coriolanus. The lamentable tragedy of Titus Andronicus. The tragedy of Romeo and Iuliet. The life of Tymon of Athens. The tragedy of Iulius Caesar. The tragedy of Macbeth. The tragedy of Hamlet, Prince of Denmark. The tragedy of King Lear. The tragedy of Othello, the moore of Venice. The tragedy of Anthonie, and Cleopatra. The tragedy of Cymbeline
dc.format.medium	Digital bitstream
dc.format.mimetype	text/xml
dc.language	English
dc.language.iso	eng
dc.publisher	University of Oxford
dc.relation.ispartof	Oxford Text Archive Core Collection
dc.relation.replaces	<a href="http://purl.ox.ac.uk/ota/0119">http://purl.ox.ac.uk/ota/0119</a>
dc.rights	Distributed by the University of Oxford under a Creative Commons Attribution-ShareAlike 3.0 Unported License
dc.rights.uri	<a href="http://creativecommons.org/licenses/by-sa/3.0/">http://creativecommons.org/licenses/by-sa/3.0/</a>
dc.rights.label	PLUB
dc.subject.lcsh	Plays -- England -- 16th century
dc.subject.lcsh	Plays -- England -- 17th century
dc.subject.lcsh	Comedies -- England -- 16th century
dc.subject.lcsh	Comedies -- England -- 17th century
dc.subject.lcsh	Tragedies -- England -- 16th century
dc.subject.lcsh	Tragedies -- England -- 17th century
dc.title	The Winters Tale.
dc.type	Text
has.files	yes
branding	Oxford Text Archive
files.size	1288927
files.count	5
otaterms.date.range	1600-1699

# **Dal paratesto al full-text. Il ruolo di TEI**



# Text Encoding Initiative (TEI)

Per agevolare la portabilità dei prodotti digitali in ambito umanistico e per consentire la condivisione dei formati di scambio è stato realizzato un progetto denominato TEI (*Text Encoding Initiative*).

La TEI è una DTD (ma anche uno schema) o modello di codifica che vuole contemplare tutta la serie dei fenomeni di interesse umanistico e trovare, per ciascuno di essi, un **vocabolario unico** al fine di arrivare ad una formalizzazione utile a normalizzare i criteri, le modalità e il lessico del *markup*, di fronte alla polisemia del linguaggio naturale.

Potremmo definire la TEI come uno Schema avente il compito di assegnare un vocabolario normalizzato all'XML, metalinguaggio che fornisce esclusivamente norme sintattiche.



# Origini della TEI

A partire dal 1987 le tre maggiori associazioni mondiali di studiosi di scienze umane attraverso metodologie informatiche, la *Association for Computers and the Humanities* (ACH) la *Association for Computational Linguistics* (ACL) e la *Association for Literary and Linguistic Computing* (ALLC) hanno avviato un progetto internazionale per sviluppare un **modello di codifica normalizzato** che mettesse ordine fra modelli e linguaggi di rappresentazione dell'informazione testuale in formato elettronico.

- ✓ Questo progetto è stato denominato *Text Encoding Initiative* (TEI), tutte le informazioni utili sono reperibili al sito ufficiale <http://www.tei-c.org/>, e ad esso hanno attivamente collaborato un gruppo di studiosi provenienti da differenti paesi





# Guidelines TEI

- ✓ Il lavoro delle commissioni, iniziato nel 1989, ha condotto alla realizzazione di una vasta e complessa DTD (*Document Type Definition*) per SGML, le cui caratteristiche sono state per la prima volta presentate nella pubblicazione, uscita nel 1994, col titolo *Guidelines for Electronic Text Encoding and Interchange (TEI P3)*.
- ✓ Nel 2002 è stata rilasciata la TEI P4, per venire incontro all'esigenza di prevedere, anche nelle *Guidelines*, l'adeguamento della rinnovata DTD alla sintassi XML.
- ✓ Nel 2008 è stata rilasciata la TEI P5, che ha comportato una sostanziale revisione delle *TEI Guidelines* (compresa la realizzazione di una versione della TEI basata su XML Schema che si affianca alla DTD).



# Caratteristiche della TEI

Al fine consentire un'universale applicabilità, si è optato per una divisione dello schema in molteplici **frammenti**, ognuno impiegabile:

a seconda della **tipologia** testuale da indagare (testo in prosa, testo poetico, testo drammatico, fonte manoscritta)

o

a seconda degli **scopi** di ricerca (codifica a fini editoriali, a fini di analisi linguistica, tematica, ecc.).



# TEI *Lite*

È stato elaborato anche un sottoinsieme della TEI, denominato *TEI Lite*, che dovrebbe consentire la diffusione del progetto di codifica, senza impegnare l'utenza ad apprendere lo schema nella sua interezza.

Si tratta infatti di una versione semplificata dell'intero schema di codifica definito dalla TEI, che permette di facilitare la realizzazione di testi in formato elettronico, senza richiedere lo studio dell'intero schema e consentendo quindi la creazione di documenti *tei-compliant*, cioè compatibili.

Trad. It: [http://www.tei-c.org/Vault/P4/Lite/teiu5\\_it.html](http://www.tei-c.org/Vault/P4/Lite/teiu5_it.html)



# Struttura TEI: testo e metadati

Scendendo nel dettaglio, diremo che ogni testo codificato conformemente alle specifiche della TEI è costituito da due parti:

- un *TEI header* (“testata”), contenente le informazioni editoriali concernenti, a diverso livello, il documento elettronico;
- un *TEI text* contenente la trascrizione codificata del testo in versione integrale.



# TEI *header*

Diremo che nella “testata” della TEI sono raccolte tutte le informazioni utili per:

- ✓ descrizione bibliografica del testo elettronico e del suo esemplare originale di riferimento,
- ✓ modalità e caratteristiche della codifica,
- ✓ note non bibliografiche e revisioni.

Queste notizie possono definirsi come i **metadati della risorsa**, cioè le meta-informazioni utili a descrivere il testo elettronico prodotto e il documento materiale utilizzato come esemplare per la trascrizione digitale



# TEI *header*. Macro struttura ed elementi obbligatori

```
<teiHeader>
  <fileDesc></fileDesc>
  <encodingDesc></encodingDesc>
  <profileDesc></profileDesc>
  <revisionDesc></revisionDesc>
</teiHeader>
```

```
<teiHeader>
  <fileDesc>
    <titleStmt>
      <title>The Wild Ass's Skin: an electronic edition</title>
    </titleStmt>
    <publicationStmt>
      <p>Published as an example for the header module of TBE.</p>
    </publicationStmt>
    <sourceDesc>
      <p>Honoré de Balzac (1906) . The Wild Ass's Skin.</p>
    </sourceDesc>
  </fileDesc>
</teiHeader>
```



# TEI header.

## Un fileDesc più ricco

```
1 <fileDesc>
2   <titleStmt>
3     <title type="main">Title</title>
4     <title type="version">An Electronic Edition</title>
5     <author>
6       <name>name</name>
7       <date>date</date>
8     </author>
9     <respStmt>
10      <resp>Creation of machine-readable text by</resp>
11      <name>name</name>
12    </respStmt>
13    <respStmt>
14      <resp>Header creation by</resp>
15      <name>name</name>
16    </respStmt>
17    <respStmt>
18      <resp>Encoded by</resp>
19      <name>name</name>
20    </respStmt>
21  </titleStmt>
22  <publicationStmt>
23    <publisher>
24      <name>name</name>
25    </publisher>
26    <pubPlace>
27      <address>
28        <addrLine>
29          <name>name</name>
30        </addrLine>
31      </address>
32    </pubPlace>
33    <date>date</date>
34    <availability>
35      <p>Copyright info</p>
36    </availability>
37  </publicationStmt>
38  <sourceDesc>
39    <bibl>bibl info</bibl>
40  </sourceDesc>
41 </fileDesc>
```





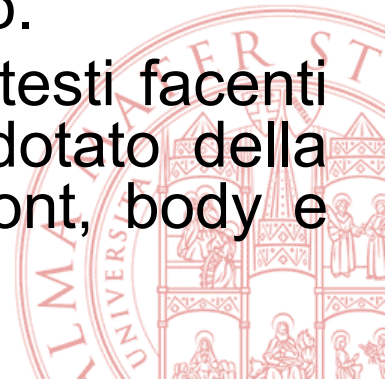
L'elemento *text* si divide, a sua volta, in quattro elementi:

<front> (opzionale),  
<body> (obbligatorio),  
<group> (obbligatorio)  
<back> (opzionale)

<Front> e <back> sono necessari qualora il documento riporti delle informazioni, a livello paratestuale, all'inizio o alla fine del testo vero e proprio (per esempio all'inizio si possono trovare un indice, un'intestazione, una prefazione, una dedica, ecc.; alla fine un glossario, un indice, un'appendice, una postfazione, ecc.), delle quali si necessita operare una codifica.

<body> è l'elemento che introduce il corpo del testo.

<group> è da impiegarsi nel caso di una serie di testi facenti parte della medesima pubblicazione, ognuno dotato della propria individualità (ognuno con un proprio front, body e back)



# TEI *text*

```
<text>
  <front>
    <div type="dedication">
      <p>In memory of Lisa Wheeman.</p>
    </div>
    <div type="contents">
      <head>Table of Contents</head>
      <list>
        <item>I. The Decision</item>
        <item>II. The Fuss</item>
        <item>III. The Celebration</item>
      </list>
    </div>
  </front>
  <body>
    <p>For the first time in twenty-five years, Dr Burt Diddledygook decided not to turn up to the annual meeting of the Royal Academy of Whoopledywhaa (RAW). It was a sunny day in late September 1960 bang on noontime and Dr Burt was looking forward to a stroll in the park instead. He hoped his fellow members of the RAW weren't even going to notice his absence.</p>
  </body>
  <back>
    <div type="colophon">
      <p>Typeset in Haselfoot 37 and Henry 8. Printed and bound by Whistleshout, South Africa.</p>
    </div>
  </back>
</text>
```



# Le suddivisioni interne del corpo del documento

All'interno dell'elemento <body> ci saranno ovviamente una serie di suddivisioni ulteriori, necessarie a definire la **struttura interna** del testo indagato, in modo direttamente proporzionale alla **tipologia** di testo, ovvero del genere (prosa, versi, testo teatrale, dizionario, parlato, ecc.).

Tali suddivisioni evidenzieranno la scansione logica del testo e la sua articolazione in **sezioni** di vario livello (<div>), in **paragrafi** (<p>) se necessario in **versi** (<l>), **titoli** (<head>), ecc.

Saranno necessari marcatori differenti a seconda del tipo di testo codificato e a seconda dei fenomeni che si intendono marcare.

Allora alla codifica della struttura si aggiungerà quella di alcuni elementi ritenuti necessari alla descrizione della fonte o agli scopi della ricerca: **nomi di persona, discorsi diretti, parole in lingua straniera, note, date, citazioni bibliografiche, liste, etc.**

**Elenco dei progetti che usano TEI:**

<http://www.tei-c.org/Activities/Projects/>





# TEI P5 Guidelines - TOC

<http://www.tei-c.org/Guidelines/P5/>

The screenshot shows a web browser window displaying the TEI P5 Guidelines Table of Contents. The browser's address bar shows the URL [www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html](http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html). The page has a blue header with the TEI logo and the text "< Text Encoding Initiative >". Below the header is a navigation menu with links: Home, Guidelines, Activities, Tools, Membership, Support, About, and News. A search bar is also present. The main content area is titled "P5: Guidelines for Electronic Text Encoding and Interchange" with the version "2.9.1. Last updated on 15th October 2015, revision 46ac023". Below this, there are language links: [English] [Deutsch] [Español] [Italiano] [Français] [日本語] [한국어] [中文]. The Table of Contents is organized into three columns: Front Matter, Text Body, and TEI sourcecode. The Front Matter section includes links to Title, Releases of the TEI Guidelines, Dedication, Preface and Acknowledgments, About These Guidelines, A Gentle Introduction to XML, and Languages and Character Sets. The Text Body section lists 23 numbered items, including The TEI Infrastructure, The TEI Header, Elements Available in All TEI Documents, Default Text Structure, Characters, Glyphs, and Writing Modes, Verse, Performance Texts, Transcriptions of Speech, Dictionaries, Manuscript Description, Representation of Primary Sources, Critical Apparatus, Names, Dates, People, and Places, Tables, Formulae, Graphics and Notated Music, Language Corpora, Linking, Segmentation, and Alignment, Simple Analytic Mechanisms, Feature Structures, Graphs, Networks, and Trees, Non-hierarchical Structures, Certainty, Precision, and Responsibility, Documentation Elements, and Using the TEI. The TEI sourcecode section includes links to the Github Repository and Bug Reports, Feature Requests, etc. At the bottom of the page, there are more language links: [English] [Deutsch] [Español] [Italiano] [Français] [日本語] [한국어] [中文]. The browser's taskbar at the bottom shows various application icons and the system clock indicating 10:52 on 16/02/2016.

File Modifica Visualizza Cronologia Segnalibri Strumenti Aiuto

Scientia rerum. La scienza ... x SCIENTIA RERUM x The TEI Guidelines x +

www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html

TEI < Text Encoding Initiative >

Home Guidelines Activities Tools Membership Support About News

P5 Guidelines — English Search

**P5: Guidelines for Electronic Text Encoding and Interchange**

Version 2.9.1. Last updated on 15th October 2015, revision 46ac023

[English] [Deutsch] [Español] [Italiano] [Français] [日本語] [한국어] [中文]

**Front Matter**

- Title
  - i. Releases of the TEI Guidelines
  - ii. Dedication
  - iii. Preface and Acknowledgments
- iv. About These Guidelines
- v. A Gentle Introduction to XML
- vi. Languages and Character Sets

**Back Matter**

- Appendix A Model Classes
- Appendix B Attribute Classes
- Appendix C Elements
- Appendix D Attributes
- Appendix E Datatypes and Other Macros
- Appendix F Bibliography
- Appendix G Prefatory Notes
- Appendix H Colophon

**Text Body**

- 1 The TEI Infrastructure
- 2 The TEI Header
- 3 Elements Available in All TEI Documents
- 4 Default Text Structure
- 5 Characters, Glyphs, and Writing Modes
- 6 Verse
- 7 Performance Texts
- 8 Transcriptions of Speech
- 9 Dictionaries
- 10 Manuscript Description
- 11 Representation of Primary Sources
- 12 Critical Apparatus
- 13 Names, Dates, People, and Places
- 14 Tables, Formulae, Graphics and Notated Music
- 15 Language Corpora
- 16 Linking, Segmentation, and Alignment
- 17 Simple Analytic Mechanisms
- 18 Feature Structures
- 19 Graphs, Networks, and Trees
- 20 Non-hierarchical Structures
- 21 Certainty, Precision, and Responsibility
- 22 Documentation Elements
- 23 Using the TEI

**TEI sourcecode**

- Github Repository
- Bug Reports, Feature Requests, etc.

[English] [Deutsch] [Español] [Italiano] [Français] [日本語] [한국어] [中文]

TEI Consortium Feedback

# TEI P5 Modules

module name	description
analysis	Simple analytic mechanisms
certainty	Certainty and uncertainty
core	Elements common to all TEI documents
corpus	Header extensions for corpus texts
declarefs	Feature system declarations
dictionaries	Dictionaries and other lexical resources
drama	Performance texts
figures	Tables, formulae, and figures
gaiji	Character and glyph documentation
header	The TEI Header
iso-fo	Feature structures
linking	Linking, segmentation and alignment
msdescription	Manuscript Description
namesdates	Names and dates
nets	Graphs, networks and trees
spoken	Transcribed Speech
tagdocs	Documentation of TEI modules
tei	Declarations for datatypes, classes, and macros available to all TEI modules
textcrit	Text criticism
textstructure	Default text structure
transcr	Transcription of primary sources
verse	Verse structures



# Un tutorial online: DARIAH TEACH- <https://teach.dariah.eu/course/>



DH teaching material  
open-source, high quality, multilingual teaching materials for the digital arts and humanities

About Courses Workshops Contact

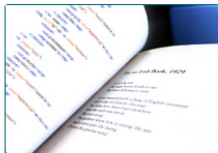
search



Guest (Log Out)

## Courses

Courses represent the equivalent level of student effort as a 5 or 10 ECTS module.



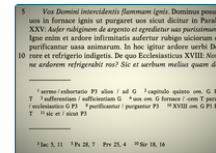
Text Encoding and the TEI - FR



Introduction to Digital Humanities



Text Encoding and the TEI



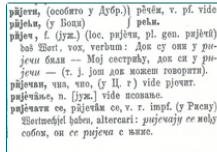
Digital Scholarly Editions: Manuscripts, Texts and TEI Encoding



Multimodal Literacies



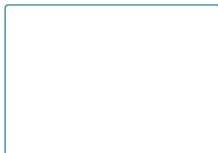
Sound Studies



Digitizing Dictionaries

## Workshops

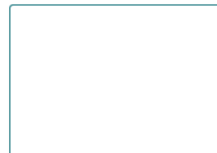
Workshops should take anywhere from several hours to two days to complete. They represent a short, focused introduction to a topic, method, or approach.



Spatial Image Analytics



An introduction to Conceptual Modelling



Spatial Image Analytics - GR



# Un ambiente di addestramento: TEI by example - <http://www.teibyexample.org/>

[Tutorials](#)[Examples](#)[Tests](#)[Exercises](#)[Tools](#)[About](#)

Send this on Google

## Welcome to the TEI By Example project!

TEI By Example offers a series of freely available online tutorials walking individuals through the different stages in marking up a document in TEI (Text Encoding Initiative). Besides a general introduction to text encoding, step-by-step **tutorial modules** provide example-based introductions to eight different aspects of electronic text markup for the humanities. Each tutorial module is accompanied with a dedicated **examples section**, illustrating actual TEI encoding practise with real-life examples. The theory of the tutorial modules can be tested in **interactive tests** and **exercises**. The tutorial materials are contextualised with a **tools section**, providing both an annotated overview of state-of-the-art XML encoding technology, and a **TBE validator application**, allowing you to test your TEI encoding as you type! For a full overview, see the section [below](#), or the drop-down menu above.

The TBE project team would be delighted to hear your feedback, comments, suggestions, but most of all: **examples**! If you happen to have any text samples that you want to share under a [Creative Commons Share Alike license](#), please do get in touch at [teibyexample@kantl.be](mailto:teibyexample@kantl.be).

### Tutorials

0. Introduction to Text Encoding and the TEI
1. Common Structure and Elements
2. The TEI header
3. Prose
4. Poetry
5. Drama
6. Primary Sources
7. Critical Editing
8. Customising TEI, ODD, Roma

### Examples

0. Introduction to Text Encoding and the TEI
1. Common Structure and Elements
2. The TEI header
3. Prose
4. Poetry
5. Drama
6. Primary Sources
7. Critical Editing
8. Customising TEI, ODD, Roma

### Tests

0. Introduction to Text Encoding and the TEI
1. Common Structure and Elements
2. The TEI header
3. Prose
4. Poetry
5. Drama
6. Primary Sources
7. Critical Editing
8. Customising TEI, ODD, Roma

### Exercises

0. Introduction to Text Encoding and the TEI
1. Common Structure and Elements
2. The TEI header
3. Prose
4. Poetry
5. Drama
6. Primary Sources
7. Critical Editing
8. Customising TEI, ODD, Roma

### Tools

1. TEI validator
2. XML toolkit



Except where otherwise noted content on this site is licensed under a [Creative Commons Attribution ShareAlike 3.0 License](#)

Contact: [teibyexample@kantl.be](mailto:teibyexample@kantl.be)

Page last modified on: 2010-07-09



**Aggiungere conoscenza ai dati.  
I vocabolari per il controllo di  
autorità**



# Liste di termini e forme controllate dei nomi

## Authority control.

- ✓ **VIAF:** <http://viaf.org/> (personal names, corporate names, geographic names, works, expressions, headings, titles).
- ✓ **Library of Congress authorities:** <http://authorities.loc.gov/> (subject, name, title, keywords). E in generale Linked Data Service, <http://id.loc.gov/>
- ✓ **Getty vocabularies:** <http://www.getty.edu/research/tools/vocabularies/>.
  - Art & Architecture Thesaurus (AAT)
  - Cultural Objects Name Authority (CONA)
  - Getty Thesaurus of Geographic Names (TGN)
  - Union List of Artist Names (ULAN)
- ✓ **Geonames.** <http://www.geonames.org/>
- ✓ **Sistemi di Classificazione**
  - CDD – Decimale Dewey (<https://www.oclc.org/en/dewey/features/summaries.html>)
  - UDC – Decimale Universale (<http://www.udcsummary.info/php/index.php>)
- ✓ **Rete semantica Wordnet** <http://wordnetweb.princeton.edu/perl/webwn>
- ✓ **Soggettazione BNCF** (Nuovo soggettario), <https://thes.bncf.firenze.sbn.it/ricerca.php>
- ✓ **Controllo di autorità generale Wikidata**  
[https://www.wikidata.org/wiki/Wikidata:Main\\_Page](https://www.wikidata.org/wiki/Wikidata:Main_Page)



# Da XML/TEI a HTML



# Trasformare per il Web

- ✓ Ci vogliono linguaggi ad hoc per convertire un file XML/TEI in un file HTML per la visualizzazione (ma anche in pdf, epub, etc.)
- ✓ Uno dei linguaggi più usati per fare questa operazione si chiama XSL(T) - esistono applicazioni (e.g. Saxon) incorporate negli editor XML (e.g. Oxygen) per fare questa trasformazione. Si usano poi i CSS per la gestione degli aspetti di layout, come in HTML
- ✓ Ma ci sono anche strumenti utili come Boilerplate, <http://teiboilerplate.org/> o Tapas project, <https://tapasproject.org/>

Ioannes Dantiscus' Texts and Correspondence Alfonso de VALDÉS  
Ioannes DANTISCUS

The screenshot displays the Tapas project web interface. On the left, the 'XML View' is shown, displaying the raw XML/TEI code. The code includes a header with the title 'Ioannes Dantiscus' Texts and Correspondence: Alfonso de VALDÉS I', the editor 'Anna Skolimowska et al.', and the date '2010-2014'. The main body of the XML contains a paragraph of Latin text, which is rendered on the right. The right side of the interface shows the 'Reader' view, which includes a 'Choose Stylesheet' dropdown set to 'TAPAS Generic'. Below this, there are controls for 'HIDE PAGE BREAKS' and 'VIEWS' (set to 'diplomatic'). The rendered HTML output shows the Latin text 'Salutem plurimam. Quasi quisquam me magis a caeremoniis abhorreat, ita te apud me purgare voluisti, quod insalutato hospite abieris, immo eo in genere (ut ingenue fatear) nihil mihi gratius facere potuisti, quare absint omnino a nobis hae vulgares caeremoniae, cum nihil minus professioni nostrae correspondeat, quam vulgum sequi. Quod tantum de me tibi polliceris'. At the bottom right, there is a 'view page image(s)' button and a small thumbnail image of a manuscript page.

Choose Stylesheet

XML View

☐ Toggle Soft Wrap  
☐ Toggle Invisibles

1 <?xml version="1.0"?>  
2 <TEI xmlns="http://www.tei-c.org/ns/1.0" xml:id="IDL0421">  
3 <!-->  
4 <!-->  
5 <!-->  
6 <!-->  
7 <!-->  
8 <!-->  
9 <!-->  
10 <!-->  
11 <!-->  
12 <!-->  
13 <!-->  
14 <!-->  
15 <!-->  
16 <!-->  
17 <!-->  
18 <!-->  
19 <!-->  
20 <!-->  
21 <!-->  
22 <!-->  
23 <!-->  
24 <!-->  
25 <!-->  
26 <!-->  
27 <!-->  
28 <!-->

Reader

Choose Stylesheet

TAPAS Generic

HIDE PAGE BREAKS ☐  
VIEWS diplomatic

Page BK 222, No. 50, p. 192

Salutem plurimam.

Quasi quisquam me magis a caeremoniis abhorreat, ita te apud me purgare voluisti, quod insalutato hospite abieris, immo eo in genere (ut ingenue fatear) nihil mihi gratius facere potuisti, quare absint omnino a nobis hae vulgares caeremoniae, cum nihil minus professioni nostrae correspondeat, quam vulgum sequi. Quod tantum de me tibi polliceris

Page BK 222, No. 50, p. 193



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

# Thank you!

Contacts:

DHDK, CHeDE & DH.arc Head

Francesca Tomasi (FICLIT)

[francesca.tomasi@unibo.it](mailto:francesca.tomasi@unibo.it)



Biblioteca "Ezio Raimondi" del Dipartimento di Filologia classica e Italianistica – FICLIT  
via Zamboni 32 – Bologna

Copyright ©University of Bologna