

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

Angelo Mario Del Grosso

angelo.delgrossos@ilc.cnr.it

CNR-ILC

Istituto di Linguistica Computazionale “A. Zampolli”,
21st February 2022

Piano della presentazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Di cosa mi occupo

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Filologia Digitale e Computazionale

Attività di ricerca scientifica e tecnologica volta allo sviluppo di sistemi di linguistica e filologia digitale e computazionale finalizzati alla produzione, rappresentazione, analisi, fruizione e interrogazione di testi di tradizione medievale, a stampa e di autori moderni e contemporanei.

Presentazione del Corso

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia



Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Introduzione al Corso di Codifica di Testi

Obiettivi, competenze e conoscenze

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Obiettivo

Illustrare i principi di modellazione e le prassi di codifica del testo per una adeguata rappresentazione ed elaborazione digitale di risorse testuali.

Rationale

Fornire gli strumenti e le conoscenze necessarie per progettare e realizzare criticamente una codifica digitale di testi complessi, in particolare testi letterari e di interesse storico-culturale, usando le linee guida della Text Encoding Initiative (TEI).

Argomenti trattati

Obiettivi, competenze e conoscenze

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Cosa ci aspettiamo alla fine del corso

- valutare il metodo di codifica più appropriato allo scenario d'interesse
- creare uno schema di codifica TEI
- usare gli strumenti più idonei per la codifica di una risorsa testuale
- elaborare e visualizzare il testo codificato tramite fogli di stile e browser web

Principali Argomenti

Obiettivi, competenze e conoscenze

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Codifica dei caratteri e di testi
- I linguaggi di markup e introduzione a XML
- Creazione di schemi di codifica
- Le norme TEI (Text Encoding Initiative)
- Alcuni specifici Moduli TEI
- Definizione di schemi di codifica personalizzati
- Introduzione ai fogli di stile XSLT
- Elaborazione documenti XML-TEI (XSLT2.0, DOM)
- Esempi, esercitazioni e seminari

Calendario e Orario del Corso

Lunedì e Venerdì dalle 14.15 alle 15.45

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

TERZO ANNO

Ora	Lunedì	Martedì	Mercoledì	Giovedì	Venerdì
8:30					
10:15					
12:00	Letteratura italiana contemporanea <i>Fib PS4</i>	Letteratura italiana contemporanea <i>Fib PS4</i>	Filosofia della scienza <i>Fib PS4</i>		Filosofia della scienza <i>Fib PS4</i>
14:15	Codifica di testi <i>Fib M-Lab</i>	Sviluppo dei Servizi Web <i>Fib M-Lab</i>	Sviluppo dei Servizi Web <i>Fib M-Lab</i>		Codifica di testi <i>Fib M-Lab</i>
16:00		Filosofia della scienza <i>Fib PS1</i>	Progettazione grafica <i>Fib M-Lab</i>		
17:45			Progettazione grafica <i>Fib M-Lab</i>		

Bozza piano del Corso: Calendario 2020/2021

Lunedì e Venerdì dalle 14.15 alle 15.45

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Introduzione al corso e strumenti software utilizzati
- Elementi teorici relativi alla rappresentazione digitale del testo
- Introduzione ai Linguaggi di Markup ed elementi di XML
- Introduzione alla definizione degli Schemi XML (DTD, XSD)
- Elementi di Codifica XML-TEI

Bozza piano del Corso: Calendario 2020/2021

Lunedì e Venerdì dalle 14.15 alle 15.45

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Moduli Editoriali TEI
- Introduzione ai Fogli di Stile
- Elaborazione documenti XML-TEI (XSLT e DOM)
- Approfondimenti Codifica TEI e XSLT
- Esercitazione e discussione sui progetti d'esame

Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Materiale Didattico e strumenti di lavoro

Studio e approfondimento

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Studio

Sono disponibili numerose risorse, anche gratuite e online, per studiare ed approfondire i principali temi trattati dal corso

Pratica e strumenti di lavoro

Utilizzeremo strumenti open source, liberi da licenze a pagamento, per implementare le risorse oggetto di studio

Materiale Didattico

Studio e approfondimento

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Supporto allo studio XML-TEI

- Slide delle lezioni (Moodle, GitHub)
- Libri Burnard, Pierazzo e Ciotti
- Linee guida TEI (online)
- Riviste scientifiche (jTEI, UD, Balisage)

- Libri tecnologie XML
- Risorse online e specifiche ufficiali (W3C)
- Esempi ed esercitazioni

Slide delle lezioni

Moodle, GitHub

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022
A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows the Moodle LMS interface for the 'Codifica di Testi (IU, LINTRA, 2020/21)' course. The top navigation bar includes links for 'Miei corsi', 'Accesso', 'Assistenza', 'Attivazione corsi', and language selection ('Italiano (it)'). A user profile for 'ANGELO MARIO DEL GROSSO' is visible on the right.

The main content area displays course information: 'Codifica di Testi (IU, LINTRA, 2020/21)', 'Home / I miei corsi / DISCIPLINE UMANISTICHE - Corsi ordinari / AA 2020 - 2021 / Il semestre / 299IL_20', and a 'Attiva modifica' button.

The left sidebar contains a 'Navigazione' menu with the following structure:

- Home
 - Dashboard
 - E-LEARNING AREA UMANISTICA
 - I miei corsi
 - DISCIPLINE UMANISTICHE
 - Corsi ordinari
 - AA 2020 - 2021
 - Il semestre
 - 299IL_20
 - Partecipanti
 - Badge
 - Competenze
 - Valutazioni
 - Introduzione
 - Argomento 1

<https://elearning.humnet.unipi.it/>

Slide delle lezioni

Moodle, GitHub

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

The screenshot shows a GitHub repository interface. At the top, there's a header with a dropdown for 'master', the repository name 'corsoCodifica / CTaa20-21 /', and buttons for 'Go to file', 'Add file', and three dots. Below the header is a list of files and folders:

File/Folder	Description	Last Commit
..		2 hours ago
others	add saxon and xerces tools	2 hours ago
progetto-esame	init repo CdT aa20-21	7 hours ago
slides	init repo CdT aa20-21	7 hours ago
src	add shell code	2 hours ago
README.md	add saxon and xerces tools	2 hours ago

Below the file list is a section for 'README.md' containing the text:

Corso Codifica di Testi aa 2020 / 2021

<https://github.com/angelodel80/corsoCodifica>

Materiale di studio

Linee Guida della TEI - Online

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows the homepage of the Text Encoding Initiative (TEI) at <https://tei-c.org/release/doc/tei-p5-doc/en/html/index.html>. The header features the TEI logo and navigation links for Home, Guidelines, Activities, Tools, Membership, Support, About, and News. A search bar is present. The main content area is titled "P5: Guidelines for Electronic Text Encoding and Interchange" and includes a note about the version being updated. Below this, there are language links for English, Deutsch, Español, Italiano, Français, 日本語, 한국어, and 中文. The page is divided into sections: "Front Matter" (with links to Title, Releases of the TEI Guidelines, Preface and Acknowledgments, About These Guidelines, A Gentle Introduction to XML, and Languages and Character Sets), "Text Body" (with links to various chapters like 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 Related Music, Language Coverage, Linking, Segmentation, and Alignment, Simple Analytic Mechanisms, and Feature Structures), and "TEI sourcecode" (with links to Getting and Using the TEI Sources, TEI GitHub Repository, and Bug Reports, Feature Requests, etc.).

Materiale di studio

Open Edition Book - Lou Burnard

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

The screenshot shows the OpenEdition Press website. At the top left is the logo 'OpenEdition press'. To its right is the title 'OpenEdition Press' in a large, bold, pink font. On the far right, there is a navigation bar with links: 'HOME', 'ENCYCLOPÉDIE NUMÉRIQUE', 'READ/WRITE BOOK', 'SCIENCES PO | MÉDIALAB', and 'BRÉSIL / FRANCE | BRASIL / FRANÇA'. Below this is a large image of the book cover for 'WHAT IS THE TEXT ENCODING INITIATIVE?'. The cover has a colorful, abstract background with the title in white text. To the left of the book image is a search bar with the placeholder 'CERCA NEL LIBRO' and an 'OK' button. Below the search bar are buttons for 'INDICE DEI CONTENUTI', 'CITA', 'CONDIVIDI', and 'CITATO DA'. To the right of the book image, the title 'WHAT IS THE TEXT ENCODING INITIATIVE?' is displayed in large, bold, black capital letters. Below it is a short description: 'How to add intelligent markup to digital resources' by 'Lou Burnard'. The text 'Encyclopédie numérique' is also present. A section titled '...' follows. The main text begins with: 'The Text Encoding Initiative (TEI) Guidelines have long been regarded as the de facto standard for the preparation of digital textual resources in the scholarly research community. For the beginner, they offer a daunting range of possibilities, reflecting the huge range of potential applications for text encoding, from traditional scholarly editions, to language corpora, historical lexicons, digital archives and beyond.' Below this is a link 'Drawing on many examples of TEI-encoded text from a variety of res...'. To the right of this link is a button '→ Leggi il seguito'. At the bottom of the page, there is a row of text details: 'Editore : OpenEdition Press', 'Collana : Encyclopédie numérique | 3', 'Luogo di pubblicazione : Marsella', 'Anno di pubblicazione : 2014', 'Pubblicato su OpenEdition Books : 11 avril 2014', 'EAN (edizione cartacea) : 9782821834590', 'EAN digitale : 9782821834606', 'DOI : 10.4000/books.oep.426', and 'Numero di pagine : 114 p.'. To the right of these details is a vertical sidebar with icons for 'LEGGI' (a computer monitor), 'ACCESSO APERTO' (a person icon), 'LETTURA CONFORTEVOLI' (a person reading a book icon), 'EPUB' (an e-reader icon), and 'PDF DEL LIBRO' (a document icon). Below each icon is a small red circle with a white dot, likely indicating an update or notification.

<https://books.openedition.org/oep/426>



Materiale di studio

Libro Digital Scholarling Editing - E. Pierazzo

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows a digital scholarly editing book page on the HAL archive. The title is "Digital Scholarly Editing: Theories, Models and Methods" by Elena Pierazzo. The page includes a detailed abstract, keywords (Digital Humanities, Philology, Digital Editing, Scholarly editing), document type (Pré-publication, Document de travail), domain (Sciences de l'Homme et Société / Littératures), and a complete list of metadata. The right side of the page features a sidebar with sections for FICHIER, LICENCE, IDENTIFIANTS, COLLECTIONS, and CITATION, each containing specific details about the publication.

hal-01182162, version 1 Pré-publication, Document de travail

Digital Scholarly Editing: Theories, Models and Methods

Elena Pierazzo Détails

GERCI - Groupe d'étude et de recherche sur la culture italienne

Abstract: This book provides an up-to-date, coherent and comprehensive treatment of digital scholarly editing, organized according to the typical timeline and workflow of the preparation of an edition: from the choice of the object to edit, the editorial work, post-production and publication, the use of the published edition, to long-term issues and the ultimate significance of the published work. The author also examines from a theoretical and methodological point of view the issues and problems that emerge during these stages with the application of computational techniques and methods. Building on previous publications on the topic, the book discusses the most significant developments in digital textual scholarship, claiming that the alterations in traditional editorial practices necessitated by the use of computers impose radical changes in the way we think and manage texts, documents, editions and the public. It is of interest not only to scholarly editors, but to all involved in publishing and readership in a digital environment in the humanities.

Keywords: Digital Humanities | Philology | Digital Editing | Scholarly editing

Type of document: Pré-publication, Document de travail

Domaine: Sciences de l'Homme et Société / Littératures
Sciences de l'Homme et Société / Héritage culturel et muséologie
Sciences de l'Homme et Société / Sciences de l'information et de la communication

Liste complète des métadonnées Voir

FICHIER
DSE-revised-29sept2014.pdf
Fichiers produits par l'(les) auteur(s)

LICENCE
Copyright (Tous droits réservés)

IDENTIFIANTS
HAL Id : hal-01182162, version 1

COLLECTIONS
GERCI | UGA

CITATION
Elena Pierazzo. Digital Scholarly Editing: Theories, Models and Methods. 2014. (hal-01182162)

<https://hal.archives-ouvertes.fr/hal-01182162>

Materiale di studio

Rivista jTEI

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows the homepage of the OpenEdition Journals platform. At the top, there is a logo for "OpenEdition Journals" and links for "HOME", "CATALOGUE OF 554 JOURNALS", and "OPENEDITION SEARCH". On the right, there are language and sharing options. Below the header, the "Home" link is visible. The main content area features the "JOURNAL OF THE TEXT ENCODING INITIATIVE" logo and title. To the left, there is a sidebar with a search bar, an "INDEX" section containing "Authors" and "Keywords" links, and a "FULL TEXT ISSUES" section listing "Issue 11 | July 2019 - June 2020 Selected Papers from the 2016 TEI Conference" and "Issue 12 | July 2019 - May 2020 Selected Papers from the 2019 TEI Conference". The main text area describes the journal as the official journal of the Text Encoding Initiative Consortium, publishing proceedings of annual conferences and thematic issues. It also mentions the latest issue (Issue 11) and selected papers from the 2016 TEI Conference.

<https://journals.openedition.org/jtei/>

Materiale di studio

Ulteriori risorse

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Ciotti F. (2007). Il testo e l'automa. Saggi di teoria e critica computazionale dei testi letterari. Aracne.
- Driscoll, M. J., and Pierazzo, E. (Eds.). (2016). Digital Scholarly Editing: Theories and Practices. Open Book.
- Lenci, A., Montemagni S., Pirrelli V. (2016). Testo e Computer. Elementi Di Linguistica Computazionale. Aulamagna. Carocci.
- Mancinelli, T., Elena Pierazzo. (2020). Che cos'è un'edizione scientifica digitale. Roma: Carocci.
- Orlando, T. (2010). Informatica testuale: teoria e prassi. Laterza.

Materiale di studio

Tecnologia di rappresentazione e di elaborazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

<https://www.w3.org/standards/xml/>

XML TECHNOLOGY



On this page → technology topics • news • upcoming events and talks

XML Technologies including XML, XML Namespaces, XML Schema, XSLT, Efficient XML Interchange (EXI), and other related standards.

XML Essentials

XML is shouldered by a set of essential technologies such as the infoset and namespaces. They address issues when using XML in specific applications contexts.

Efficient Interchange

XML standards are omnipresent in enterprise computing, and are part of the foundation of the Web. Because the standards are highly interoperable and affordable, people have wanted to use them in a wide variety of applications. However, in some settings (on devices with low memory or low bandwidth, or where performance is critical) experience has shown that a more efficient form of XML is required.

Schema

Formal descriptions of vocabularies create flexibility in authoring environments and quality control chains. W3C's XML Schema, SML, and data binding technologies provide the tools for quality control of XML data.

Security

Manipulating data with XML requires sometimes integrity, authentication and privacy. XML signature, encryption, and xkms can help create a secure environment for XML.

Transformation

Very frequently one wants to transform XML content into other formats (including other XML formats). XSLT and XPath are very powerful tools for creating different representations of XML content.

Query

XQuery (supported by XPath) is a query language for XML to extract data, similar to the role of SQL for databases, or SPARQL for the Semantic Web.

Components

The XML ecosystem is using additional tools to create a richer environment for using and manipulating XML documents. These components include style sheets, xlink xml:id, xinclude, xpointer, xforms, xml fragments, and events.

Processing

A processing model defines what operations should be performed in what order on an XML document.

Internationalization

W3C has worked with the community on the Internationalization of XML, for instance for specifying the language of XML content.

Materiale di studio

Tecnologia di rappresentazione e di elaborazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

XML.com

Home Articles Authors News Job Board About Search

News

Articles

Foundations

Popular tags: articles

XFORMS XSLT VOCABULARY

UBL XPROC NISO RELAX-NG

XSD CONFERENCE XQUERY

Popular tags: news

XML XSLT DITA DOCBOOK

XQUERY XHTML DOCK

XMLMIND HTML EPUB

Announcing the XML slack community

Markup UK 2021 Conference Dates Announced

Markup UK is pleased to announce that the next edition of the conference will take ...

18 December 2020 [Read](#)

Release of XMLmind Word To XML v1.8.3

Release of XMLmind Word To XML v1.8.3 (<http://www.xmlmind.com/w2x/changes.html>)

Enhancements:

- Upgraded XMLmind Web Help ...

2 December 2020 [Read](#)

3 Useful Things DITA XML Can Do for Your Content

Tim Ludwig
You know what DITA is, but what can you actually do with it?

29 December 2020 [Read](#)

XProc 3.0 - Strategies for merging documents

Erik Siegel
Erik Siegel continues his series on XProc 3.0 with an analysis of potential strategies for merging documents.

16 November 2020 [Read](#)

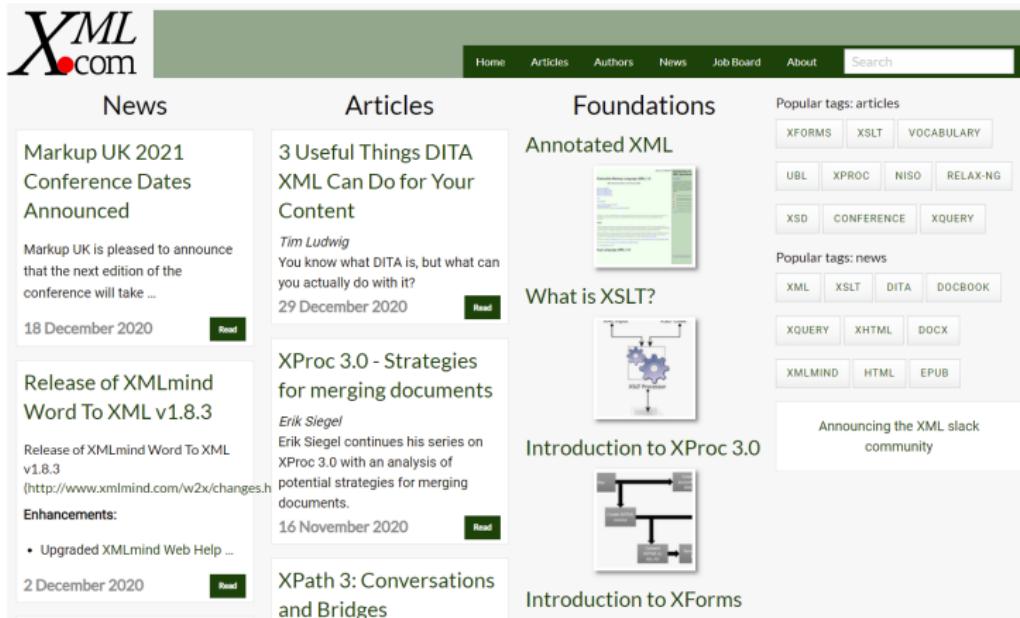
XPath 3: Conversations and Bridges

Annotated XML

What is XSLT?

Introduction to XProc 3.0

Introduction to XForms



<https://www.xml.com/>

Materiale di studio

Tecnologia di rappresentazione e di elaborazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows a web browser displaying the Saxonica documentation. The URL in the address bar is <https://www.saxonica.com/documentation/>. The page title is "Reference: XSLT Elements". The left sidebar contains a navigation tree with categories like "Using XML Schema (XSD)", "The Gizmo Utility", "Saxon on .NET", "Extensibility Mechanisms", "Saxon Extensions", "Sample Saxon Applications", "Saxon SQL Extension", and "Reference: XSLT Elements". Under "Reference: XSLT Elements", there is a list of XSLT elements: xsl:accept, xsl:accumulator, xsl:accumulator-rule, xsl:analyze-string, xsl:apply-imports, xsl:apply-templates, xsl:assert, xsl:attribute, xsl:attribute-set, xsl:break, xsl:call-template, xsl:catch, xsl:character-map, xsl:choose, xsl:comment, xsl:context-item, xsl:copy, xsl:copy-of, xsl:decimal-format, xsl:document, and xsl:element. The main content area starts with a heading "Reference: XSLT Elements" and a paragraph about the list of standard XSLT elements. It then discusses "Conformance" and "XSLT Patterns". Below these, there is a section titled "Standard attributes" which lists attributes like default-collation, default-mode, default-validation, exclude-result-prefixes, expand-text, extension-element-prefixes, use-when, version, and xpath-default-namespace. The "Section index" at the bottom lists various XSLT elements grouped into columns.

- [xsl:accept](#)
- [xsl:accumulator](#)
- [xsl:accumulator-rule](#)
- [xsl:analyze-string](#)
- [xsl:apply-imports](#)
- [xsl:apply-templates](#)
- [xsl:assert](#)
- [xsl:attribute](#)
- [xsl:attribute-set](#)
- [xsl:break](#)
- [xsl:call-template](#)
- [xsl:catch](#)
- [xsl:character-map](#)
- [xsl:choose](#)
- [xsl:comment](#)
- [xsl:context-item](#)
- [xsl:copy](#)
- [xsl:copy-of](#)
- [xsl:decimal-format](#)
- [xsl:document](#)
- [xsl:element](#)
- [xsl:function](#)
- [xsl:global-context-item](#)
- [xsl:if](#)
- [xsl:import](#)
- [xsl:import-schema](#)
- [xsl:include](#)
- [xsl:iterate](#)
- [xsl:output](#)
- [xsl:output-character](#)
- [xsl:override](#)
- [xsl:package](#)
- [xsl:param](#)
- [xsl:perform-sort](#)
- [xsl:preserve-space](#)

<https://www.saxonica.com/documentation/>

Materiale di studio

Tecnologia di rappresentazione e di elaborazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

25 / 71



▶ Technologies

▶ References & Guides

▶ Feedback

Search MDN



Sign in

XSLT: Extensible Stylesheet Language Transformations

Web technology for developers ▶ XSLT: Extensible Stylesheet Language Transformations

English (US)

Change language

Table of contents

Documentation

Documentation

XSLT/XPath Reference: XSLT elements, EXSLT functions, XPath functions, XPath axes

Extensible Stylesheet Language Transformations (XSLT) is an XML-based language used, in conjunction with specialized processing software, for the transformation of XML documents.

Related Topics

▶ XSLT

▶ EXSLT

▶ XPath

▶ Guides

▶ XSLT Elements

Although the process is referred to as "transformation," the original document is not changed; rather, a new XML document is created based on the content of an existing document. Then, the new document may be serialized (output) by the processor in standard XML syntax or in another format, such as HTML or plain text.

XSLT is most often used to convert data between different XML schemas or to convert XML data into web pages or PDF documents.

<https://developer.mozilla.org/en-US/docs/Web/XSLT>



Materiale di studio

Tecnologia di rappresentazione e di elaborazione

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

XML
3RD EDITION



COMPREHENSIVE

Carey :: Vodník

New Perspectives on XML, Comprehensive, 3rd Edition.
Carey/Vodnik. Cengage Learning.

Materiale di studio

Ulteriori risorse

- Williams, I. (2009). *Beginning XSLT and XPath: Transforming XML Documents and Data*. Wiley.
- Kay, M. (2011). *XSLT 2.0 and XPath 2.0 Programmer's Reference*. Wiley.
- Kelly, D. J. (2015). *XSLT Jumpstarter: Level the Learning Curve and Put Your XML to Work*. Pragmatic Bookshelf.
- XML Guild Team. (2006). *Advanced XML Applications from the Experts at The XML Guild*. Course Technology Press.
- Fawcett, J., Ayers D., Quin L. R. E. (2012). *Beginning XML 5th Edition*. Wrox.

Strumenti di lavoro

requisiti di base

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Buona conoscenza dell'inglese
- Nozioni di HTML / CSS / JS
- Computer personale a disposizione per svolgere gli esercizi assegnati
- Uso di editor di testo
- Uso di editor per programmatori o editor XML
- Eseguire comandi shell da terminale per lavorare sui documenti XML

Strumenti di lavoro

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

Overview

SETUP

GET STARTED

Intro Videos

Tips and Tricks

User Interface

Themes

Settings

Key Bindings

Display Language

USER GUIDE

LANGUAGES

NODE.JS /

JAVASCRIPT

PYTHON

JAVA

AZURE

EXTENSION

AUTHORING

EXTENSIBILITY

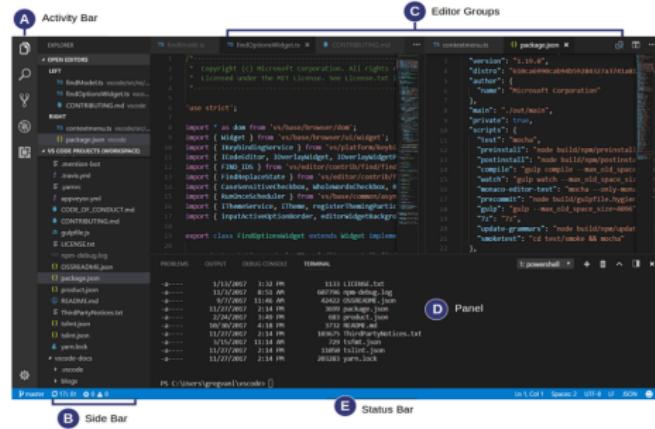
REFERENCE

OTHER

User Interface



At its heart, Visual Studio Code is a code editor. Like many other code editors, VS Code adopts a common user interface and layout of an explorer on the left, showing all of the files and folders you have access to, and an editor on the right, showing the content of the files you have opened.



IN THIS ARTICLE

Basic Layout

Side by Side Editing

Minimap - outline view

Explorer

Open Editors

Views

Command Palette

Configuring the Editor

Tabs

Preview mode

Editor Groups

Grid editor layout

Working without Tabs

Window Management

Next Steps

Common Questions

Tweet

Subscribe

Ask questions

Follow @code

Request features

Report issues

Watch videos

Basic Layout

<https://code.visualstudio.com/>

Strumenti di lavoro

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

30 / 71

The screenshot shows the Editix XML Editor 2015 interface. The main window displays an XSLT file named "personal.xsl". The code editor pane contains the following XSLT code:

```
<xsl:stylesheet version="1.0">
  <xsl:output method="html"/>

  <xsl:template match="person">
    <h1>
      <xsl:value-of select="@id"/>
    </h1>
  </xsl:template>

</xsl:stylesheet>
```

The XML Data Source pane shows a schema named "personal-schema.xml" with a single element "personnel". The XSLT Result Preview pane displays two results: "one.worker" and "two.worker". The bottom status bar shows the time as 10:12.

<http://www.editix.com>

Editor XML

Cosa è consigliato usare

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- un buon editor open source: XML Copy Editor (<http://xmlcopy-editor.sourceforge.net/>)
- per Mac: XMLSpear, Textmate, Eclipse, IntelliJIDEA
- un ottimo editor: Oxygen (<http://www.oxygenxml.com/>)
- multi piattaforma, ma non gratuito (in prova gratuita per un mese)
- altri editor: funzioni fondamentali sono la validazione, l'autocompletamento, l'esecuzione di fogli di stile

Strumenti di sviluppo

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows the official Java website at <https://www.java.com/en/download/manual.jsp>. The page features a large red header with the Java logo (a steaming coffee cup) and the word "Java". Below the header is a search bar. The main content area has a white background with a large, bold, sans-serif font displaying the text "JAVA + YOU, DOWNLOAD TODAY!". To the right of the text is a photograph of a smiling woman with curly hair lying on a white surface, looking at a silver laptop. Below the main title is a red button labeled "Java Download". Underneath the download button are three small links: "What is Java?", "Need Help?", and "Uninstall". At the bottom of the page, there's a horizontal row of six smaller links with icons: "About Java" (with "Go Java" icon), "Alice" (with "Alice" icon), "Greenfoot" (with "Greenfoot" icon), "JavaOne" (with "JavaOne" icon), "Oracle Academy" (with "Oracle Academy" icon), and "Java Magazine" (with "Java Magazine" icon).

<https://www.java.com/en/download/manual.jsp>

Strumenti di sviluppo

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

The screenshot shows the Apache Xerces Project website. At the top is the Apache Xerces logo, featuring a stylized butterfly above the word "Xerces" in blue, with "apache" in smaller letters to the right. Below the logo is a navigation bar with links for "Home", "Wiki", and "Mailing List Archives". The "Home" link is currently selected. A breadcrumb trail indicates the page is at apache.org > [XML Federation](http://xml.apache.org) > xerces.apache.org. The main title is "The Apache Xerces™ Project - xerces.apache.org". The left sidebar contains a navigation menu with sections like "Xerces" (Overview, Charter, News, Downloads, Get Involved, Who We Are, License, Legal Stuff, Mailing Lists, Code Repositories, Reference, Security, Sponsorship, Thanks), "Local Projects", "Hibernated Projects", and "Related Projects". Below the sidebar is a "Build with Apache Forrest" button. The main content area has a section titled "Welcome to the Apache Xerces™ Project" which states: "The Apache Xerces™ Project is responsible for software licensed to the [Apache Software Foundation](#) intended for the creation and maintenance of: XML parsers, related software components". There is also a section titled "Over a Decade at the Apache Software Foundation" and a banner for "10 YEARS OF The ASF Celebrating a Decade of Open Source Leadership" with the URL <http://www.apache.org/>. At the bottom of the page is a footer with various links and icons.

<http://xerces.apache.org/index.html>

Strumenti di sviluppo

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

34 / 71

SAXONICA
XSLT AND XQUERY PROCESSING

Our Products Technology Download Buy Documentation Support About us Site map

SAXONICA > Download > Java

Download | Java | .NET | C/C++, PHP & Python | JavaScript | Evaluation License | Downloads Info

Java Platform Downloads

To use Saxon-PE or Saxon-EE you need a license key. If you don't have a current license you can request a 30-day evaluation license.

Saxon-HE is open source software and can be downloaded from SourceForge or Maven.

Use the buttons below to download Saxon-PE and Saxon-EE directly. The Saxon-HE buttons redirect to SourceForge.

Java	Saxon-EE	Saxon-PE	Saxon-HE
	10.3 (19.2Mb)	10.3 (17.0Mb)	10 (SourceForge)
	9.9.1.8 (19.8Mb)	9.9.1.8 (17.6Mb)	9.9 (SourceForge)

The latest release for Java is Saxon 10.3, released on 28 October 2020. **Saxon 10 is now considered the most stable and reliable release.** For more information see [Latest Releases](#).

The latest maintenance release of 9.9 is Saxon 9.9.1.8, released on 22 October 2020.

On the download info page:

- Release Notes
- Installing the Software
- Earlier releases

See also:

- Download .NET Saxon products
- Download C/C++, PHP and Python Saxon/C products
- Download saxon-resources

Strumenti di sviluppo

esercitazioni, esempi, progetto

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

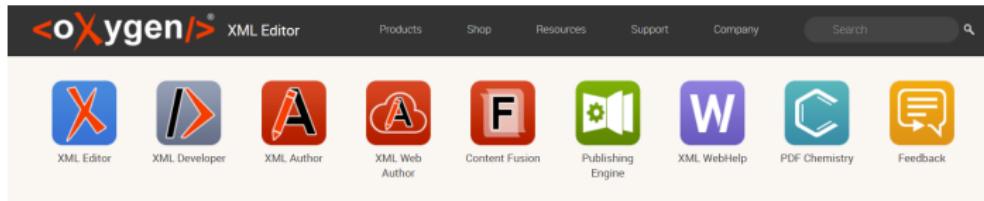
Ecosistema
XML

Bibliografia

A.M. Del Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

35 / 71



Oxygen XML Editor 23

The Premier All-In-One XML Editing Suite

Oxygen XML Editor is the best XML editor available and provides a comprehensive suite of XML authoring and development tools. It is designed to accommodate a large number of users, ranging from beginners to XML experts. It is available on multiple platforms, all major operating systems, and as a standalone application or an Eclipse plug-in. You can use Oxygen XML Editor in conjunction with all XML-based technologies and it includes a large variety of powerful tools for creating, editing, and publishing XML documents.

SEE WHAT'S NEW

Read more

Download

Buy (from \$99)



<https://www.oxygenxml.com/>

Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Programma d'esame

Cosa bisogna fare per superare l'esame

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Studiare i lucidi delle lezioni
- Padroneggiare gli esercizi svolti durante il corso
- Studiare i testi indicati dal docente
- Studiare i capitoli delle Guidelines TEI che riguardano il corso e il progetto
- Realizzare il progetto di codifica concordato con il docente

Modalità d'esame

In cosa consiste l'esame

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- Invio del progetto (obbligatoriamente tramite github)
- Colloquio
 - Discussione del progetto
 - Verifica delle conoscenze di base XML, XSD, XSL
 - Verifica delle basi teoriche
 - Conoscenza di TEI P5 (moduli principali, parti spiegate a lezione, moduli particolari utilizzati nel progetto)

Tirocini, Tesi

Pronti a far parte della comunità?

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

A.M. Del Grosso

Daniele Gambarara: SGL
Beyond the Course. A Knowledge Site for Rethinking Saussure's General Linguistics

home about corpora indexes views

Hi, your actual role is SCHOLAR ▾

SIDE BAR:

[view the source in ext software](#)

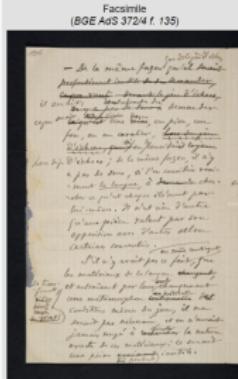
ENGLISH TRANSLATION ▾

Interpretative Edition of the Source

Just in chess it would be absurd to ask what a queen, a pawn, a bishop or a knight would be. They are outside the game of chess, so there is no sense. If language is really what is being considered, in seeking what each element is in itself. It is nothing other than a piece whose value depends on its opposition to others within certain conventions.

LEXICAL INFO ▾

- LANGUE
- INTERNE
- PANCHRONIQUE
- VALEUR



ERC-2020-ADV Grant

© 2020 Beyond the Course. A Knowledge Site for Rethinking Saussure's General Linguistics

SGL project

[Open Sandbox](#)

Explore Saussure Thoughts

Dive Into Échec from BGE AdS 372/4, f. 135

Diplomatic

(BGE AdS 372/4 f. 135)

Interpretative

(Science of the language, p. 231)

– De la même façon que dans le jeu d'échecs il serait absurde de demander / ce que serait une dame, un pion / un fou, ou un cavalier, si on le considérait / hors du jeu d'échecs; de la même façon, il n'y a pas de sens, si l'on considère vrai / ment la langue, à chercher ce qu'est chaque élément par / lui-même. Il n'est rien d'autre / qu'une pièce valant par son / opposition avec d'autres selon / certaines conventions.

[show similar texts](#)

- 1 – De la même façon <que dans le jeu d'échecs> qu'il serait profondément absurde de demander / ce que serait une dame, un pion / un fou, ou un cavalier, si on le considérait / hors du jeu d'échecs; de la même façon, il n'y a pas de sens, si l'on considère vrai / ment la langue, à chercher ce qu'est chaque élément par / lui-même. Il n'est rien d'autre / qu'une pièce valant par son / opposition avec d'autres selon / certaines conventions.
- 2 profondément absurde de demander;
- 3 ce que vaut devant le jeu d'échecs;
- 4 <il serait> n'y a pas de sens à <est absurde de> demander
- 5 <ce que serait> ce qu'est <dans le jeu d'échecs> une reine <dame>, un pion, un fou, ou un cavalier, hors du jeu
- 6 d'échecs quand <so> on <les> considérait le jeu
- 7 <hors du jeu> D'échecs; de la même façon, il n'y a pas de sens, si l'on considère vrai / ment la langue, à demander cher=
- 10 cher ce qu'est chaque élément par / lui-même. Il n'est rien d'autre
- 11 cher ce qu'est chaque élément par
- 12 lui-même. Il n'est rien d'autre
- 13 qu'une pièce valant par son
- 14 opposition avec d'autres selon
- 15 certaines conventions.

Tirocini, Tesi, EVT

Pronti a far parte della comunità?

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

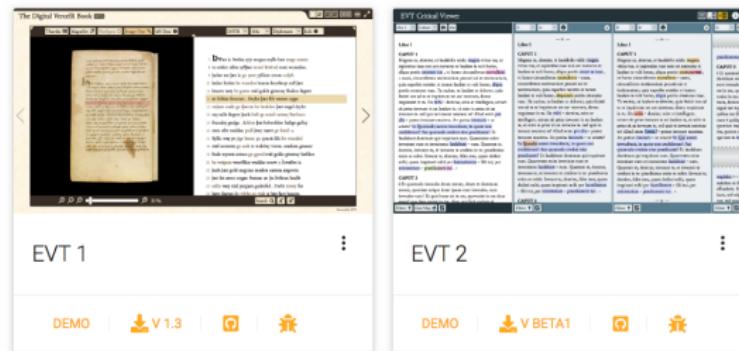
Codifica dei
Testi

Ecosistema
XML

Bibliografia

Edition Visualization Technology

A light-weight, open source tool specifically designed to create digital editions from XML-encoded texts, freeing the scholar from the burden of web programming and enabling the final user to browse, explore and study digital editions by means of a user-friendly interface.



EVT tool <http://evt.labcd.unipi.it/>

Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Perché è importante la codifica dei testi

Motivazioni pratiche

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Dove troviamo i testi

Nella nostra cultura tradizionale la quasi totalità dei testi è “registrata”, “trasmessa” e quindi “conservata” mediante supporti e materiali fisici di varia natura e forma (iscrizioni su pietra, manoscritti di pergamena, papiri, carta, libri a stampa, incunabula, cinquecentine, etc).

Perché è importante la codifica dei testi

Dove troviamo i testi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

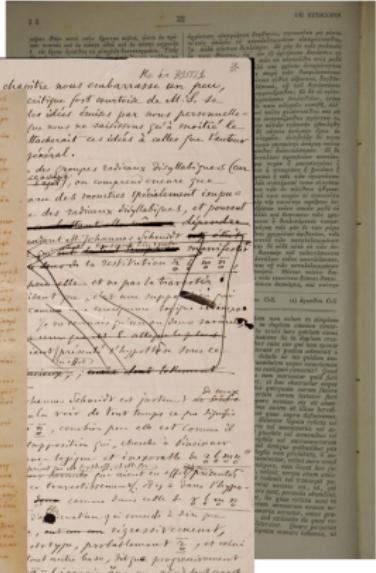
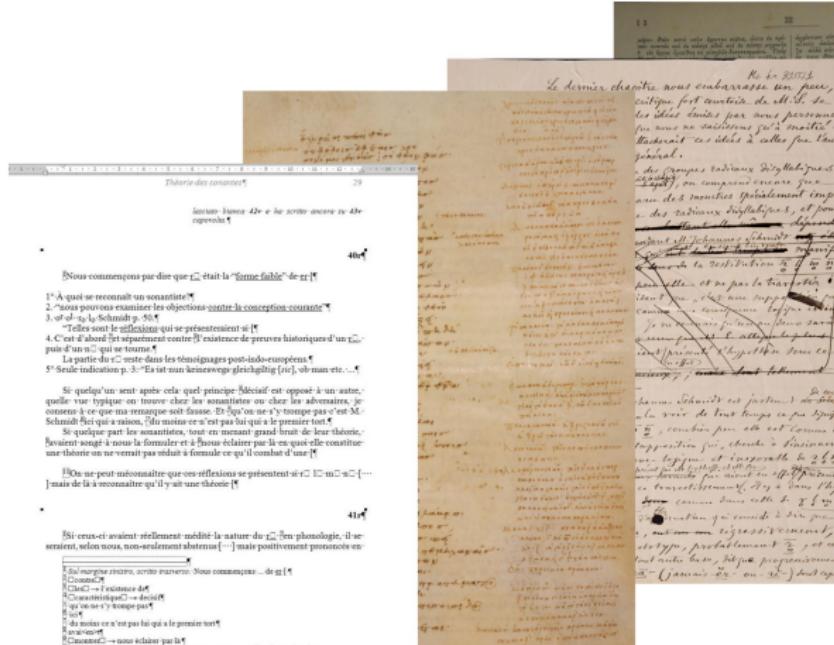
Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia



Perché è importante la codifica dei testi

Dove troviamo i testi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

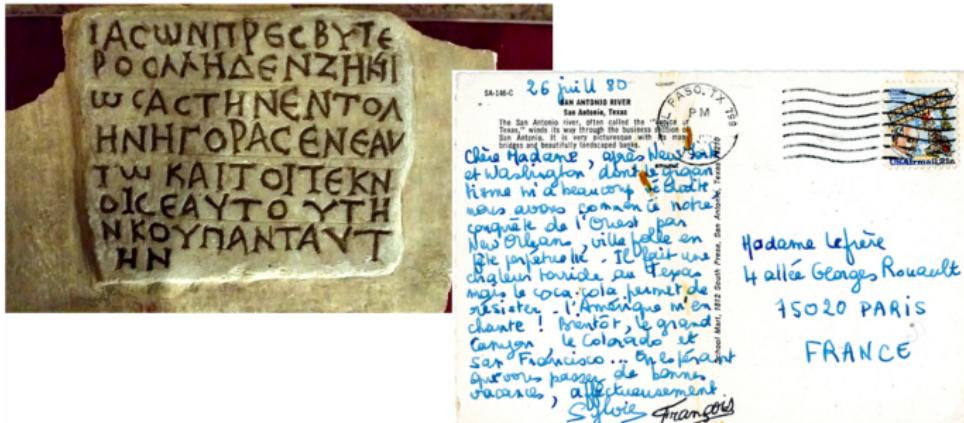
Ecosistema
XML

Bibliografia

A.M. Del Grosso

Presentazione del Corso Codifica di Testi a.a. 2021-2022

44 / 71



Perché è importante la codifica dei testi

Dove troviamo i testi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

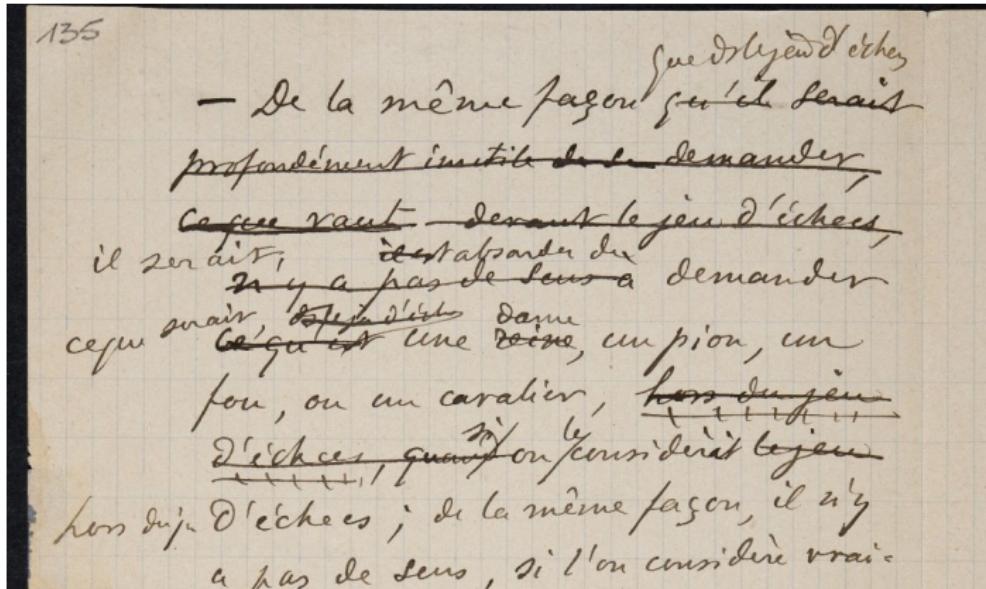
Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia



Perché è importante la codifica dei testi

Motivazioni pratiche

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Perché codificare i testi

Per rendere disponibile questo patrimonio attraverso i sistemi per la gestione dell'informazione digitali e computazionali è necessario effettuare una trasposizione/transcodifica* dei testi dal loro supporto originario verso il nuovo supporto elettronico (*Machine Readable Form*).

* *procedimento di conversione dei dati codificati secondo un sistema verso un sistema diverso*

Perché è importante la codifica dei testi

Motivazioni teoriche

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Scienze umane vs Scienze informatiche

Il rapporto tra sapere umanistico e informatica non è solo una questione meramente strumentale:

l'informatica non è solo un utensile dalle notevoli capacità.

Salto di paradigma sia da un punto di vista teorico sia
metodologico.

Perché è importante la codifica dei testi

Motivazioni teoriche

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

La codifica come metodologia

L'attività di codifica diviene metodo di analisi della conoscenza nell'ambito delle discipline che si occupano del testo.

La codifica come analisi teorica

Il linguaggio di codifica adottato può essere considerato come un linguaggio teorico:

*Esplicitare e formalizzare le ipotesi interpretative su un certo
oggetto di studio*

Perché è importante la codifica dei testi

In sintesi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Rappresentare il testo

Il focus del corso sarà incentrato sulla rappresentazione digitale del testo.

Esistono dibattiti e controversie

Per ottenere una rappresentazione digitale del testo ci sono diversi formati, formalismi e prassi:
la nostra scelta ricade sulle norme suggerite dal consorzio TEI.
Molte questioni non sono risolte altre sono controverse, sia dal punto di vista teorico-metodologico, sia pratico-tecnologico.

Rappresentazione digitale dei testi

basso e alto livello di codifica

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Codificare un testo

La codifica dei caratteri evidentemente non esaurisce i problemi per una opportuna rappresentazione delle caratteristiche interne ed esterne di un testo.

Codificare un testo

Difatti la codifica del testo è una questione molto più complessa di una semplice riproduzione meccanica di un dato.

Rappresentazione digitale dei testi

basso e alto livello di codifica

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Rappresentare un testo

La rappresentazione digitale di un testo è una operazione intrinsecamente assai difficile perché coinvolge una pletora di aspetti, a varie dimensioni, a varie granularità e a vari livelli di astrazione sia teorici, sia metodologici, sia tecnologici e sia pratici.

Rappresentazione digitale dei testi

basso e alto livello di codifica

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Rappresentare un testo

**Prima di poter fare qualsiasi ipotesi su come compiere
una codifica di un testo e su come rappresentarlo
digitalmente, bisogna stabilire cosa si intende per testo.**

Rappresentazione digitale dei testi

Modello dati di un testo

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Un testo non ha una struttura rigida, predefinita:

- Non è rappresentabile solo come un insieme di record di un archivio elettronico.
- Non è rappresentabile solo come un insieme di tabelle di una banca dati.
- Non è rappresentabile solo come un albero o un insieme di sotto-alberi (struttura gerarchica)
- Non è rappresentabile solo come un grafo o come un insieme di sotto grafi (struttura reticolare)

Molteplici modelli per diverse esigenze

Strutture dato e testo

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

La rappresentazione di un testo

- modello lineare: sequenza di dati non strutturati
- modello a record: enumerazione delle proprietà
- modello tabulare: insieme di dati omogenei
- modello ad albero: gerarchie di dati e insiemi di dati
- modello grafo: rete di strutture informative interconnesse tra loro

Elementi di Codifica del testo

Tabella Formalismi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Formalismi

	Data	Text	Hierarchies	Presentation	Validation	References	Annotations	Overlapping
CSV	Light Green	Light Red	Light Red	Light Red	Light Red	Dark Green	Light Green	Light Red
JSON	Dark Green	Light Green	Dark Green	Light Red	Light Red	Light Red	Light Green	Light Red
RDF	Dark Green	Light Red	Light Red	Light Red	Dark Green	Dark Green	Dark Green	Light Red
Markdown	Light Red	Dark Green	Light Red	Light Green	Light Red	Light Green	Light Red	Light Red
HTML	Light Red	Dark Green	Dark Green	Dark Green	Light Red	Light Green	Light Green	Light Red
HTML+RDFa	Dark Green	Dark Green	Dark Green	Dark Green	Light Green	Dark Green	Dark Green	Light Red
XML	Dark Green	Dark Green	Dark Green	Dark Green	Dark Green	Dark Green	Dark Green	Light Red
Overlapping formats				Light Red	Light Red	Dark Green	Dark Green	Dark Green

courtesy of *Fabio Vitali*

Elementi di Codifica del testo

Formalismi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Formati di rappresentazione

Un formato è un insieme di regole e convenzioni formali per rappresentare un insieme di dati, nel nostro caso un testo.

Importanza dei formati

Seppur isomorfi la scelta dei formati condiziona molto l'efficienza delle operazioni e l'efficacia delle dichiarazioni.

Elementi di Codifica del testo

Esempio di codifica del testo utilizzando CSV

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Why it's worth giving up .
The human_body is remarkably resilient and will start to repair the damage caused by smoking almost as soon as you start giving_up .
The benefits to your health can be seen in as little as 20 minutes from the time you put_down your last cigarette .
The whole process remains only marginally_safe on average for women_and_children .
Both men_and_women want the process to be safer .
Both men_and_women are involved in developing and exploiting the sciences and technologies to make_this_happen .
The gender_dynamics of public_perceptions on DNA_paternity_testing are likely to become more important as the tests become more widespread .
Some perceptions of paternity_testing will obviously be dependent on the inherently_different reproductive_roles of men_and_women .

Assume VB assume ''
that IN that
a DT a
natural JJ natural
cyclic JJ cyclic
phenomenon NN phenomenon
has VBN have
been VBN be
measured VBN measure
,

but CC but
the DT the
data NN data
is VBN be
corrupted VBN corrupt
by IN by
errors NNS error
SENT .

```
<$ s_id="1">
<w surface="Cathy" lemma="Cathy" pos="N" syn="su:2" />
<w surface="Zag" lemma="zile" pos="V" syn="ROOT:0" />
<w surface="hen" lemma="hen" pos="Pron" syn="obj1:2" />
<w surface="wild" lemma="wild" pos="Adj" syn="mod:5" />
<w surface="zwaaien" lemma="zwaai" pos="N" syn="vc:2" />
<w surface="." lemma="." pos="Punc" syn="punct:5" />
</s>
```

```
{
  "@context": "http://www.w3.org/ns/anno.jsonld",
  "id": "http://574heritago.com/annotations/2/",
  "type": "Annotation",
  "created": "2017-03-12T12:03:45Z",
  "body": {
    "type": "TextualBody",
    "purpose": "tagging",
    "format": "text/plain",
    "language": "en",
    "value": "Where is the origin of this statue?"
  },
  "creator": {
    "id": "http://heritago.com/users/1",
    "type": "Person",
    "email_shai": "5bad08927902ff9307b621c54716dcc5083e339"
  },
  "selector": {
    "type": "TextPositionSelector",
    "start": 15,
    "end": 25
  }
}
```

<<http://example1.com/citingwork>>
cito:cites <<http://example2.com/citedwork>>;
cito:inTextCitationFrequency [
 a cito:IntTextCitationCount ;
 cito:inTextCountValue "10"^^xsd:integer ;

Elementi di Codifica del testo

Formalismi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Formati come formalismi

Data l'importanza metodologica accanto ad un formato di rappresentazione del dato testuale, abbiamo la necessità di adottare un vero e proprio formalismo. Si parla cioè di linguaggi di codifica basati su un insieme di istruzioni rigorose definite da grammatiche formali.

Elementi di Codifica del testo

Formalismi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Formati e formalismi di codifica

Quindi ogni pezzo di informazione aggiunta ad un testo grezzo attraverso l'inserimento di dati metatestuali (markup, annotazione, codifica), constituisce il risultato di una analisi e di una interpretazione che è stata condotta (da un umano o da una macchina) al fine di esplicitare e rappresentare nel modo più accurato e completo possibile le informazioni da veicolare attraverso il formato digitale prescelto (anche in modo incrementale).

Progress status

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

1 Presentazione

2 Introduzione Corso Codifica di Testi

3 Materiale Didattico e Strumenti di lavoro

4 Progetto Esame

5 Codifica dei Testi

6 Ecosistema XML

Codifica ad alto livello

Sistemi di marcatura

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Metodi e tecniche per la codifica di testi

La riflessione sui metodi e le pratiche migliori per la codifica elettronica dei testi è stato uno dei temi fondamentali della ricerca e della sperimentazione nel dominio dell'Informatica Umanistica per molti anni.

Markup language e XML

soluzione corrente per la codifica dei testi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

XML per la descrizione e la codifica

Ad oggi la soluzione considerata ottimale per una corretta rappresentazione del testo è l'adozione dei markup language descrittivi basati su XML.

TEI-XML

Standard de facto per la codifica dei testi è considerato lo schema XML messo a punto dalla Text Encoding Initiative (TEI-XML).

Lo schema TEI-XML

Estratto di documento TEI-XML

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

```
1  <?xml version="1.0" encoding="utf-8" standalone="yes"?>
2  <!DOCTYPE div SYSTEM "divTEIsnippet.dtd">
3  <divTEIsnippet>
4      <div type="destination">
5          <ab>
6              <stamp type="postmark">
7                  <placeName>El Paso</placeName>
8                  - TX 799 -
9                  <date notBefore="1980-07-26">
10                     <unclear>PM JUL</unclear>
11                 </date>
12             </stamp>
13             <stamp type="postage">
14                 Profil masculin, avec un avion et un radar au second
15                 plan:
16                 <mentioned>US Airmail 21 c.</mentioned>
17             </stamp>
18         </ab>
19         <ab>
20             <address>
21                 <addrLine>
22                     Madame
23                     <name>Lefrère</name>
24                 </addrLine>
25                 <addrLine>4, allée George Rouault</addrLine>
26                 <addrLine>75020 Paris</addrLine>
27                 <addrLine>France</addrLine>
28             </address>
29         </ab>
30     </div>
31 </divTEIsnippet>
```



Perché TEI

La Text Encoding Initiative (TEI) è un autorevole progetto internazionale, a cui afferiscono varie organizzazioni e università, il cui scopo è fornire agli studiosi di informatica umanistica uno strumento il più espressivo e flessibile possibile per rappresentare qualsiasi aspetto di interesse relativo alla risorsa testuale da rappresentare digitalmente.

Impiego di XML

Benefici

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Perché XML

- separazione dei dati dall'applicativo di authoring/editing
- separazione della rappresentazione dei dati dalla presentazione dei dati
- possibilità di trasformare i dati in qualsiasi altro formato compatibile
- leggibilità dei documenti XML da parte di esseri umani

Impiego di XML

Benefici

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Perché XML

- standard w3c testuale, aperto, personalizzabile e liberamente utilizzabile
- semplicità di condivisione e scambio dati (interoperabilità e portabilità)
- adatto per codificare dati semistrutturati oltre che dati strutturati
- validazione del documento attraverso grammatiche formali

Ecosistema XML

Le tecnologie XML per la definizione ed elaborazione di documenti XML

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

- XSD: XML Schema Definition Language
- XPath: XML Path Language
- XSL: eXtensible Stylesheet Language
- XSL-T: XSL – Transformations
- XSL-FO: XSL – Formatting Objects
- XQuery: XML Query Language for XML Databases
- XInclude: XML inclusion Language
- DTD: Document Type Definition Language
- RelaxNG: Regular Expression Language for XML (New Generation)

Linguaggio di marcatura XML

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Perché XML

Adottando la tecnologia e il linguaggio XML abbiamo la possibilità di creare linguaggi di marcatura personalizzati e specifici per ogni esigenza e dominio.

XML come linguaggio per la codifica di testi

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia

Vantaggi

Attraverso XML è possibile strutturare i dati, gestire in modo nativo strutture gerarchiche, elaborare e presentare i dati con strumenti XML nativi, validare i tipi di strutture e i tipi di dati, gestire riferimenti incrociati tramite opportuni meccanismi di dereferenziazione, aggiungere e gestire annotazioni a vari livelli di granularità.

References

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione
Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia



Lenci, A., Montemagni S., and Pirrelli V. (2016). *Testo e Computer. Elementi Di Linguistica Computazionale*. Aulamagna. Carocci.



Pierazzo, E. (2015). *Digital Scholarly Editing : Theories, Models and Methods*. Farnham Surrey: Ashgate.



Orlandi, T. (2010). *Informatica testuale: teoria e prassi*. Laterza.



Driscoll, M. J., and Pierazzo, E. (Eds.). (2016). *Digital Scholarly Editing: Theories and Practices* (Vol. 4). Open Book Publishers.



Ciotti F., e Crupi G, a c. di. (2012). *Dall'Informatica umanistica alle culture digitali*. ROMA : Casa Editrice Università La Sapienza. open access: <http://www.editricesapienza.it/node/7688>



Williams, I. (2009). *Beginning XSLT and XPath: Transforming XML Documents and Data*. Wiley.



Kay, M. (2011). *XSLT 2.0 and XPath 2.0 Programmer's Reference*. Wiley.

References

Presentazione
del Corso
Codifica di
Testi
a.a.
2021-2022

A.M. Del
Grosso

Presentazione

Introduzione
Corso Codifica
di Testi

Materiale
Didattico e
Strumenti di
lavoro

Progetto
Esame

Codifica dei
Testi

Ecosistema
XML

Bibliografia



XML Standards Reference, MSDN.

[https://msdn.microsoft.com/en-us/library/ms256177\(v=vs.110\).aspx](https://msdn.microsoft.com/en-us/library/ms256177(v=vs.110).aspx)



IBM XML Tutorial,

<https://www.ibm.com/developerworks/xml/tutorials/xmllintro/xmllintro.html>



W3School Tutorial <https://www.w3schools.com/xml/default.asp>