

# Memoriarchivio

---

**L'informatica al servizio della Memoria**

Progetto dell'Università di Pisa (Informatica Umanistica) 2021

Coordinato dalla Prof.ssa Marina Riccucci



# INTRODUZIONE

---

L'applicazione web «memoriarchivio» consente di operare in tre macro aree:

- 1) Gestione e consultazione del database delle testimonianze non letterarie
- 2) Interfaccia per il caricamento e l'elaborazione dei testi in codifica XML/TEI
- 3) Analisi e confronto lessicale fra testi

# Database - 1

Inserimento, modifica, cancellazione, consultazione dei dati relativi alle:

1. Anagrafiche
2. Fonti
3. Testimonianze

## ARCHIVIO DELLA MEMORIA: Testimonianze

^ Anagrafica ^ Fonti ^ Curatori ^ Testimonianze ^ Navigazione ^ Elaborazione +

Nuova

Testimonianze

◀ ◀◀ 1 2 3 4 ▶▶ ▶

id	Titolo ▾	Nome ▾	Cognome ▾	Fonte ▾	Data ▾	Altro ▾
50	L'Inferno	Dante	Alighieri	Divina Commedia		
3	Il bel viaggio	Antonio	Falanga	L'ultimo treno. Racconti del viaggio verso il lager, Roma, Donzelli Editore		Poesia in dialetto milanese
4	Ricordo di Bruno Piazza	Bruno	Piazza	Perché gli altri dimenticano. Un italiano ad Auschwitz, Milano (Feltrinelli)		Ricordo di Bruno Piazza del momento precedente alla



# Database - 2

Ricerca trasversale di informazione nelle tabelle del database mediante parole chiave, in tre modalità:

1. AND
2. OR
3. Stringa Intera

## ARCHIVIO DELLA MEMORIA: Anagrafica

^ Anagrafica ^ Fonti ^ Curatori ^ Testimonianze ^ Navigazione ^ Elaborazione ↗

Cerca in tutto il sito:  OR ▼ Cerca

Anagrafiche inserite

Cerca...

◀ ◀ 1 2 3 4 ▶ ▶

id	Nome	Cognome
32	Rosa	Cantoni
35	Alberto	Mieli
4	Carlo	Greppi
37	Vittorio	Cremisi
38	Germana	Del Mare

Testimonianze

Cerca...

◀ ◀ 1 2 3 4 ▶ ▶

id	Titolo
1	Figini Ines
2	Relazione di una reduce da un campo di concentramento
5	Un'infanzia perduta
6	Oggi vi racconterò l'inferno
7	Virginia Gattengo Cipolato

Fonti inserite

◀ ◀ ▶ ▶

id	
2	Archivio Vicissitudin
19	Hanna prigioniera a
4	Perché gli altri dime Auschwitz, Milano (I
7	Se chiudo gli occhi i Stampa
23	Alberto Mieli soprav viato l'arrivo della co

# Codifica XML / TEI - 1

Caricamento nel database del file XML / TEI, a fianco del corrispettivo testo piano della testimonianza, da cui è possibile, tramite le funzionalità di Saxon-JS, inviare il file al software open-source EVT, per un'ulteriore elaborazione, visualizzazione e infine salvataggio del risultato nel database

File .xml inseriti				
File name	Corpo	ID del file destinazione TXT		
main.xml	<?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE TEI SYSTEM "myTEI.dtd"> <TEI xmlns="http://www.tei-c.org/ns/1.0"> <teiHeader> <fileDesc> <title>Edizione digitale della testimonianza di	48		



# Codifica XML / TEI - 2

## Edizione digitale della testimonianza di Romana Feld

<b>transcription by :</b>	Giulia Causarano
<b>Ente di Appartenenza :</b>	Fondazione Centro di Documentazione Ebraica Contemporanea
<b>Edition:</b>	Edizione diplomatico-interpretativa per la tesi di Laurea Triennale in Informatica Umanistica, Università di Pisa 2020
<b>encoding by :</b>	Giulia Causarano
<b>Publisher:</b>	Unipi Pisa 2020
<b>Repository:</b>	CDEC Digital Library
<b>Collection:</b>	Archivio - Fondo Massimo Adolfo Vitale Serie: "Ricerca sulla Deportazione" (CRDE) Sottoserie: "Vicissitudini dei singoli"
<b>Identifier:</b>	Busta 3 Fascicolo 101 000167/34
<b>Support:</b>	Fascicolo cartaceo
<b>Measure:</b>	10 pagine
<b>Condition:</b>	Buono
<b>Note:</b>	Nella parte superiore destra della prima pagina è presente il numero romano 15. Possibile numero di inventario aggiunto dal catalogatore. Nella parte superiore destra della prima pagina è presente il nome della testimone in lettere maiuscole
<b>Licenza:</b>	Contenuti soggetti alla licenza Creative Commons - Attribution 4.0 International (CC BY 4.0)


# Analisi lessicale -1

---

- Selezionare due gruppi di testi da analizzare singolarmente o a confronto
- Per ciascuno gruppo di testi:
  - Creazione di un dizionario di parole usate, con la loro frequenza
  - Ricerca di parole singole
- Confronto fra gruppi testi:
  - Per ogni parola nel testo di riferimento ricerca di corrispondenze nel testo da analizzare
- Possibilità di fuzzy matching e analisi mediante il software Voyant



# Analisi lessicale -2

Testi di riferimento			
Titolo	Autore	Fonte	
L'Inferno	Dante Alighieri		

## Crea dizionario

[Visualizza testo / Apri Voyant](#)

Testi da confrontare		
Titolo	Autore	For
Il bel viaggio	Antonio Falanga	

## Crea dizionario

[Visualizza testo / Apri Voyant](#)

Raggruppamento:

1

Numeri:

 Frequenza%: 

Raffina:

☒ Fuzzy match:

## Confronta i testi

## Ordina

Cerca parole o radici:

Fuzzy match:



Cerca

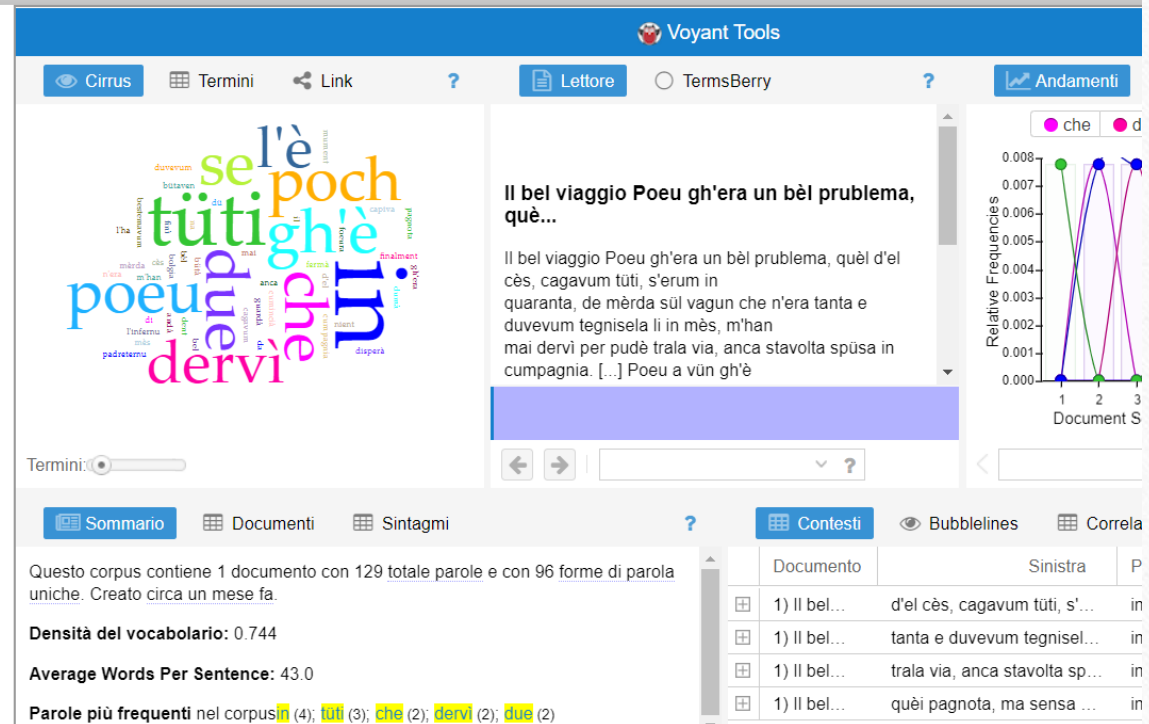
R-> a: [708]  
C-> a: [5]  
R-> anca: [3]  
C-> anca: [1]  
R-> bel: [8]  
C-> bel: [1]  
R-> bolgia: [18]  
C-> bolgia: [1]  
R-> che: [1192]



# Analisi lessicale -3

## Testo di riferimento

Poeu gh'era un bèl problema, quèl  
m tütì, s'erum in quaranta, de  
n che n'era tanta e duvevum  
ès, m'han mai dervi per pudè trala  
ta spüsa in cumpagnia. [...] Poeu a  
ura i sentiment, l'ha cumincià a  
perà e a poch a poch el mal l'è  
tütì e se capiva pù nient, la pareva  
nferu e bestemavum tütì el  
alment el trenu el s'è fermà, han  
per un mument, gh'è vegnù dū  
n bütà dent na quèi pagnota, ma  
i due i bütaven, el pan l'è andà a  
a, due...se po no di.





AIUCD 2021

# Testimoniare il lager: l'informatica al servizio della memoria

G. Causarano, M. Riccucci, F. Valecchi,

A. Del Grosso

Università degli Studi di Pisa

ILC-CNR - CLARIN-IT

## PROGETTO "VOCI DALL'INFERNO" (Coordinamento Marina Riccucci)

Obiettivi: - Creazione di un database che raccolga le testimonianze non letterarie dei sopravvissuti ai lager nazisti. (*Memoriarchivio* - Frida Valecchi)  
- Censimento della presenza del lessico dantesco all'interno di quelle testimonianze



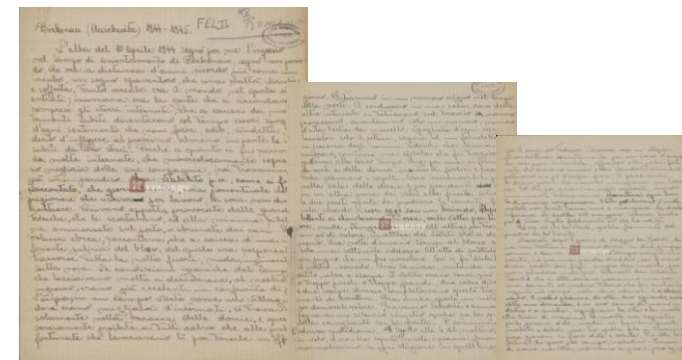
## CODIFICA DEL TESTO (Coordinamento Angelo Mario Del Grosso)

Codifica (XML-TEI) del corpus eterogeneo delle testimonianze e integrazione nell'ambiente *Memoriarchivio* di un componente client-side basato sull'innovativa libreria JavaScript, Saxon-js 2.0



## UN CASO: LA TESTIMONIANZA DI ROMANA FELD (Giulia Causarano)

Individuazione, codifica ed elaborazione della testimonianza autografa e inedita.





## **Progetto di ricerca “Voci dall’inferno: le parole per dirlo”**

Assunto di base: I sopravvissuti all’Olocausto trovano nel Dante della *Commedia* un vocabolario da cui attingere immagini e concetti per riferire l’indicibile orrore del Lager.

Obiettivi del progetto:

1. La costituzione di un *corpus* digitale delle testimonianze non letterarie dei sopravvissuti ai Lager nazisti (resoconti orali, in forma di intervista, e resoconti scritti, in forma di diario, memoriale, epistolario, solo di rado in tangenza con la letterarietà).
2. La verifica della presenza, all’interno di quelle testimonianze, di lessico dantesco



A questo scopo è stata creata l’applicazione web *Memoriarchivio*, strutturata in modo da consentire:

- La raccolta e l’archiviazione delle testimonianze italiane non letterarie
- La consultazione da parte dell’utente delle testimonianze, delle fonti da cui sono state tratte, delle anagrafiche dei testimoni e dei curatori delle fonti.
- L’analisi delle risorse testuali raccolte e la loro interrogazione in merito alla presenza di lessico dantesco. È disponibile la funzionalità di conteggio e ricerca delle parole nei testi (in formato *plain text*) nonché la funzionalità di confronto fra i testi, già presenti nel sito o caricati direttamente dall’utente.



# AIUCD 2021

ARCHIVIO DELLA MEMORIA: Testimonianze

^ Anagrafica

^ Fonti

^ Curatori

^ Testimonianze

^ Elaborazione

⌂

Nuova

Testimonianze

1 2 3 4

id	Titolo ↕	Nome ↕	Cognome ↕	Fonte ↕	Data ↕	Altro ↕	Elaborata		Testo	
							sì	no		
3	Il bel viaggio	Antonio	Falanga	L'ultimo treno. Racconti del viaggio verso il lager, Roma, Donzelli Editore		Poesia in dialetto milanese	Falso			
4	Ricordo di Bruno Piazza	Bruno	Piazza	Perché gli altri dimenticano. Un italiano ad Auschwitz, Milano (Feltrinelli)		Ricordo di Bruno Piazza del momento precedente alla	Falso			
2	Relazione di una reduce da un campo di concentramento	Lina	Verona Valabrega	Archivio Vicissitudini dei singoli	30-11-1945		Falso			
5	Un'infanzia perduta	Liliana	Segre	Voci dalla Shoa. Testimonianze per non dimenticare, Scandicci(Firenze), La Nuova Italia Editrice			Falso			
6	Oggi vi racconterò l'inferno	Nedo	Fiano	Voci dalla Shoa. Testimonianze per non dimenticare, Scandicci(Firenze), La Nuova Italia Editrice			Falso			
1	Figini Ines	Ines	Figini	Amministrazione Comunale di Nova Milanese e della città di Bolzano, Lager e deportazione	21-7-2004	Intervista del 21/07/2004 ad Ines Figini	Falso			
7	Virginia Gattengo Cipolato	Virginia	Gattengo Cipolato	Meditate che questo è stato. Testimonianze di reduci dai campi di sterminio, Firenze, Giuntina	1-12-1993	intervista Venezia, dicembre 1993	Falso			
15	Sultana Razon intervista 2014	Sultana	Razon	Sultana Razon intervista 2014	10-4-2014		Falso			
10	Auschwitz, un mondo rovesciato	Marta	Ascoli	RACCONTI DAL LAGER Testimonianze dei sopravvissuti ai campi di concentramento tedeschi			Falso			
11	La mia vita ad Auschwitz	Samuele	Modiano	Sami Modiano: la mia vita ad Auschwitz	27-1-2015		Falso			

Schermata principale del software *memoriarchivio* dalla cui barra del menu si evincono le principali funzioni



ARCHIVIO DELLA MEMORIA: Elaborazione

^Elaborazione

Testi di riferimento

Titolo	Autore	Fonte	
TESTIMONIANZA.txt		caricamento da file	-

Crea dizionario

Testi da confrontare

Titolo	Autore	Fonte	
INFERNO.txt		caricamento da file	-

Crea dizionario

Raggruppamento: 1 Numeri: ☐ Frequenza%: ☐ Raffina: ☒ Fuzzy match: ☐ Confronta i testi

Ordina Cerca parole o radici:  Fuzzy match: ☐ Cerca Distribuzione

R-> a: [31]  
C-> a: [688]  
R-> acqua: [10]  
C-> acqua: [24]  
R-> ad: [1]  
C-> ad: [41]  
R-> ai: [2]  
C-> ai: [3]  
R-> al: [14]  
C-> al: [191]  
R-> albero: [1]  
C-> albero: [2]  
R-> alla: [6]  
C-> alla: [1]  
R-> alle: [6]  
C-> alle: [1]  
R-> allora: [1]  
C-> allora: [8]  
R-> altra: [6]  
C-> altra: [60]  
R-> altre: [6]

Sono visualizzate le parole del testo da confrontare (indicate a schermo con la lettera "C") che ricorrono nel testo di riferimento (le cui parole sono indicate a schermo con la lettera "R")

Schermata del software memoriarchivio nella sezione di elaborazione: un esempio di confronto fra testi



AIUCD 2021

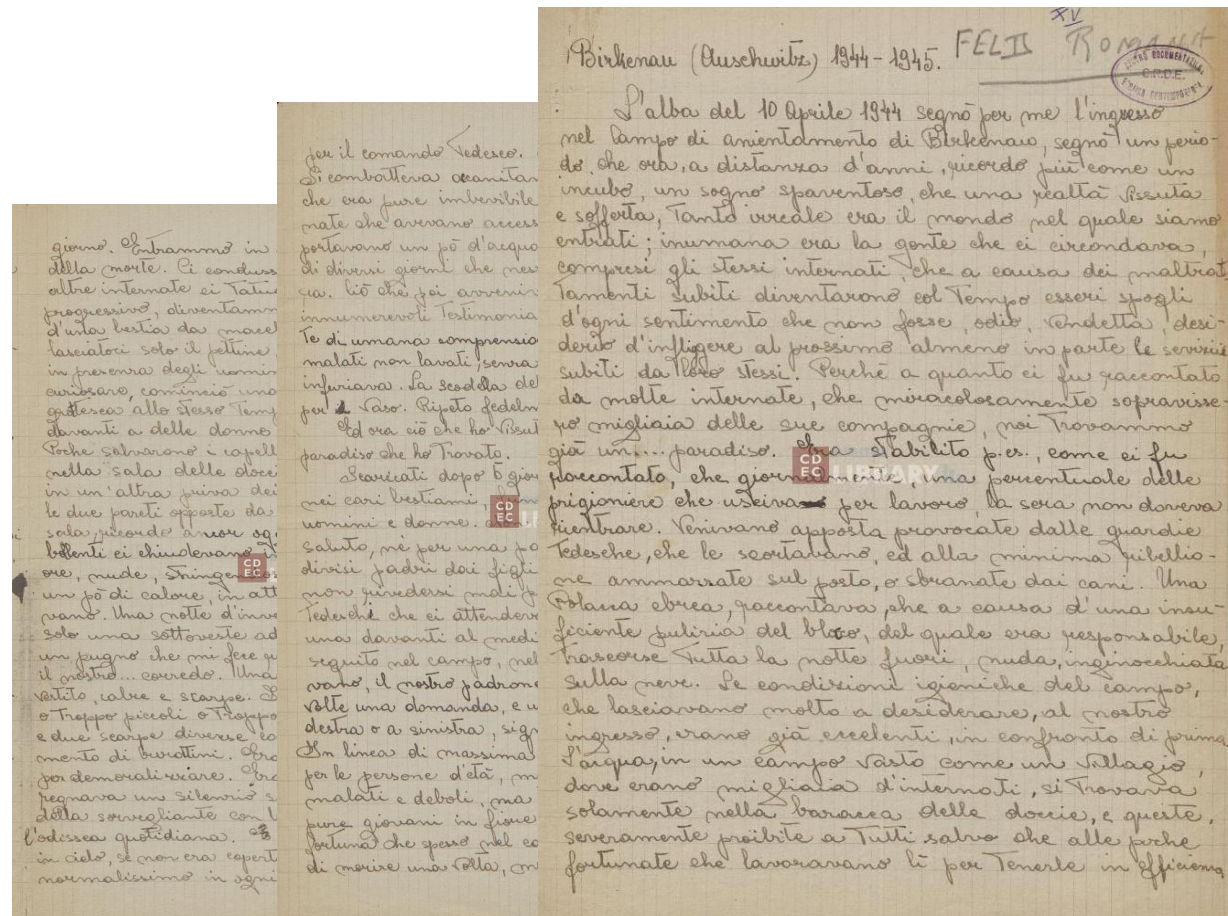
## Selezione del materiale:

La ricerca del documento da codificare e da eleggere, pertanto, a oggetto di indagine è stata condotta sul materiale attualmente disponibile sul portale web CDEC *Digital Library*, sviluppato appositamente per consentire l'accesso e la consultazione dei documenti custoditi nell'archivio della **Fondazione CDEC** (Centro di Documentazione Ebraica Contemporanea), ente che ricopre un ruolo fondamentale nel panorama storico-culturale ebraico nazionale e internazionale.



## Oggetto di studio e indagine:

La **testimonianza manoscritta inedita** di Romana Feld, donna di origini polacche deportata nel campo di concentramento e sterminio di Birkenau nel 1944.





## La codifica della testimonianza

Le pagine autografe che compongono il testo (dieci fogli) sono state codificate seguendo le linee guida del progetto TEI (versione P5), queste dotate di grande flessibilità ed espressività, fattori che consentono di personalizzare il proprio schema di codifica secondo le specifiche esigenze. Il documento così codificato è stato in seguito incluso all'interno del database *Memoriarchivio*.

### Schema di codifica:

Per la rappresentazione digitale della testimonianza è stato definito un modello di codifica XML-TEI costituito dai moduli TEI di base (*tei*, *header*, *core*, *textstructure*), che definiscono elementi utili in qualsiasi tipo di documento, e dai moduli opzionali necessari per:

- Descrivere il manoscritto (*msdescription*)
- Trascrivere la fonte primaria (*transcription*)
- Codificare i nomi e le date (*namesdates*)
- Rappresentare analisi di particolari strutture del testo (*linking*)
- Associare interpretazioni semantiche o sintattiche agli elementi del testo (*analysis*)

### Resa grafica del documento:

In seguito ad un'attenta lettura del documento digitalizzato si è scelto di realizzare **un'edizione diplomatico-interpretativa** della testimonianza, ovvero di rappresentare il testo tenendo fede all'originale (rispettando impaginazione e punteggiatura) e apportando solo gli interventi editoriali necessari a:

- Sciogliere le abbreviazioni
- Normalizzare le lettere maiuscole
- Correggere gli errori più evidenti
- Marcare aggiunte e cancellature autoriali



AIUCD 2021

## EVT (Edition Visualization Technology)

Terminato il processo di codifica della testimonianza si è fatto uso del software *open-source* EVT, strumento che permette di navigare, esplorare e visualizzare edizioni digitali create a partire da testi codificati secondo lo standard XML-TEI, rendendo graficamente, all'interno di un browser, i fenomeni testuali precedentemente esplicitati attraverso la codifica.

Mediante l'uso di tecnologie web e di fogli di stile (presenti di default in EVT) tale strumento software permette di:

- Visualizzare solo la trascrizione diplomatica del testo, o solo quella interpretativa.
- Confrontare in parallelo i due livelli di edizione e distinguerli tramite la diversa resa grafica dei marcatori XML-TEI utilizzati.
- Visualizzare contemporaneamente una pagina della trascrizione e la corrispondente immagine del manoscritto, con la possibilità, in aggiunta, di muovere il cursore su una specifica area dell'immagine digitale ed evidenziare la relativa riga di testo o viceversa.

Collegamento "testo-immagine" reso a video grazie al software EVT





AIUCD 2021

## Sviluppo di un componente *client-side*

Per consentire all'applicazione *Memoriarchivio* di ospitare e di analizzare non solo testi in formato *plain text* ma anche testi codificati, in formato XML, è stato necessario estendere le funzioni originarie dell'applicativo.

### Modifiche lato server:

- Integrazione all'interno della piattaforma di un pulsante per la selezione e il caricamento dei documenti in formato XML
- e di una tabella che conserva questi documenti e ne permette la gestione

### Modifiche lato client:

Sviluppo e integrazione nell'ambiente

*Memoriarchivio* di un componente che permette

1. Recuperare dal server la testimonianza in formato XML
2. Elaborare l'XML ed estrarre sia il contenuto testuale che i metadati relativi alla risorsa
3. Mostrare all'utente i dati estratti in una pagina web
4. Restituire all'interno del database la testimonianza in formato TXT

Testo.txt:\* + Testo.txt

Testo.xml:\* + Testo.xml

Altro:

Testo.txt

<?xml version="1.0" encoding=

Annulla Salva

Schermata del software *memoriarchivio* relativa all'inserimento di una testimonianza, con relativo file TEI xml, se presente.

| Testimonianze |                              |       |          |       |      |       |           |
|---------------|------------------------------|-------|----------|-------|------|-------|-----------|
| id            | Titolo                       | Nome  | Cognome  | Fonte | Data | Altro | Elaborata |
|               |                              |       |          |       |      |       | si no     |
| 48            | Testimonianza di Romana Feld | Frida | Valecchi |       |      |       | Falso     |

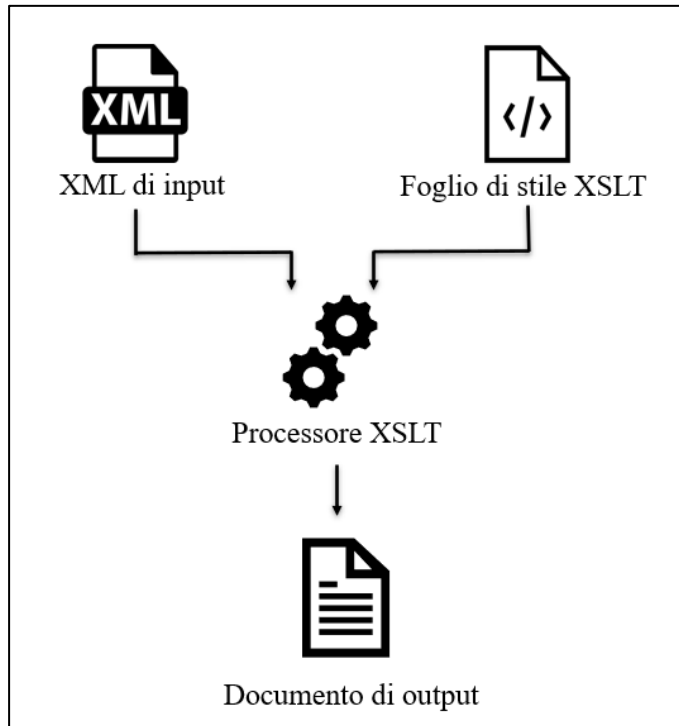
| File .xml inseriti       |                              |           |  |                              |
|--------------------------|------------------------------|-----------|--|------------------------------|
| ID del file sorgente XML | Testimonianza relativa       | File name | Corpo  | ID del file destinazione TXT |
| 20                       | Testimonianza di Romana Feld | main.xml  | <?xml version="1.0" encoding="UTF-8"?> <!DOCTYPE TEI SYSTEM "myTEI.dtd"> <TEI xmlns="http://www.tei-c.org/ns/1.0"> <teiHeader> | 48                           |

Schermata del software *memoriarchivio* relativa alla gestione dei file TEI

## Tecnologie utilizzate

Il funzionamento del componente *client-side* poggia su due tecnologie in particolare:

- La tecnologia **RESTful**, utilizzata per stabilire un canale di trasferimento dati tra il client e il server basato sul protocollo http.
- La tecnologia **XSLT**, utilizzata per elaborare l'XML e in particolare per trasformare un documento XML di input in un altro documento di diverso formato.



Trasformazione XSLT

### Saxon-JS 2.0

Per garantire il funzionamento sul lato client del componente sviluppato e permettere dunque che le trasformazioni XSLT fossero eseguite tramite codice javascript (lato client), è stato utilizzato l'innovativo **processore XSLT Saxon-JS 2.0**, scritto in puro javascript e in grado di funzionare sia sul Browser, lato client, sia su Node.js (la piattaforma javascript che ne permette l'esecuzione anche lato server, permettendo al nuovo processore di funzionare in modo completamente indipendente dal prodotto Java, l'unico disponibile per le versioni precedenti).

Principale aggiornamento della versione 2.0, rispetto a quanto consentito dalle versioni precedenti dello stesso prodotto, è l'inclusione all'interno del pacchetto 2.0 di un nuovissimo **compilatore in javascript**, per la conversione dei fogli di stile XSLT nel formato di esportazione riconosciuto da Saxon-JS, il formato SEF (*Stylesheet Export File*), indispensabile per la loro esecuzione.





## AIUCD 2021

Per invocare una trasformazione XSLT ed eseguirla dal codice javascript, la libreria Saxon-JS rende disponibile la funzione `SaxonJS.transform(options[, execution])`, costituita da: un primo parametro (options), che fornisce al processore le opzioni di trasformazione, quali il foglio di stile SEF da compilare e il documento di input da trasformare, e un secondo parametro (execution), che stabilisce la modalità sincrona o asincrona della trasformazione.

```
SaxonJS.transform({  
  stylesheetLocation: "stylehtml.sef.json",  
  sourceNode: xmlDoc,  
  destination: "serialized",  
}, "async")
```

↓

Trasformazione asincrona

Foglio di stile SEF  
Input della trasformazione  
Modalità di consegna del risultato

```
<!DOCTYPE html>  
<html>  
  <head>  
    <meta charset="UTF-8" />  
    <title>SAXON-JS</title>  
    <script type="text/javascript"  
      src="Saxon-JS-2.0/SaxonJS2.js"></script>  
    <script type="text/javascript"  
      src="App.js"></script>  
    <link rel="stylesheet" href="style.css">  
  </head>  
  <body>  
    <div id="Info"></div>  
    <div id="Testo"></div>  
  </body>  
</html>  
  
<xsl:template name="main" match="/">  
  <html>  
    <head></head>  
    <body>  
      <xsl:result-document href="#Info">  
        <xsl:apply-templates select=          "//tei:fileDesc"/>  
      <xsl:apply-templates select=        "//tei:profileDesc" />  
    </xsl:result-document>  
    <xsl:result-document href="#Testo">  
      <xsl:apply-templates select=        "//tei:text"/>  
    </xsl:result-document>  
  </body>  
</html> </xsl:template>
```

Esempio di utilizzo della libreria Saxon-JS. Sulla sinistra il documento HTML e sulla destra il documento XSLT.

## Indagini sul testo

Durante l'ultima fase del lavoro è stata condotta un'analisi critico-interpretativa della testimonianza, volta a determinare e a quantificare l'apporto del lessico dantesco.

### Prima indagine sul testo:

- Realizzata senza il supporto di strumenti informatici
- Ha permesso di rilevare la presenza nel testo di:
  - Un insieme di allusioni a situazioni e immagini proprie dell'*Inferno* dantesco, ovvero episodi raccontati dalla testimone che rimandano implicitamente a circostanze descritte da Dante, per quanto non richiamate con lessico propriamente dantesco.
  - Una citazione esplicita alla *commedia* che compare invece nelle ultime pagine del testo, quando la testimone racconta di aver visto con i propri occhi trasporti di ragazze giovanissime condotte alle camere a gas e di aver associato quell'immagine ai cortei delle anime dannate nell'*Inferno* dantesco, che in quel momento diventavano realtà.

### Seconda indagine sul testo:

- Realizzata tramite lo strumento di confronto tra i testi (disponibile sulla piattaforma *Memoriarchivio*), che ha permesso di realizzare un confronto lessicale tra la testimonianza di Romana Feld e la cantica dell'*Inferno*, restituendo le parole comuni ai due testi.
- Tra queste, le parole utilizzate dalla testimone con un'accezione prettamente dantesca sono:
  - Il termine "bestia", che ricorre nel testo per indicare la disumanizzazione delle detenute.
  - Il termine "tormento", usato per indicare il perpetuo dolore che affligge le prigioniere nel campo.
  - La parola "fortuna", che la testimone impiega con il significato di "sorte".
  - La parola "martirio", che come nell'*Inferno* dantesco ha l'accezione di "patimento", "sofferenza".
  - L'espressione "inumana gente".