

DigitalMemoryStone

Archivio Storico delle Persecuzioni Razziali a Modena

STUDENTI PRESENTATORI

Crispino Salvatore e Brugnano Federico

DOCENTI E TUTOR

Prof.ssa Piccinini Silvia e Prof.ssa Russo Giuseppina

SVILUPPATO PER

Archivio di Modena

ISTITUTO SCOLASTICO

I.T.I.S Enrico Fermi Modena



Introduzione e Contesto Storico

L'Emergenza della Memoria

Durante il periodo delle leggi razziali e delle persecuzioni, la città e la provincia di Modena furono testimoni di drammatici transiti, arresti e deportazioni di cittadini ebrei. La nostra base di partenza è un volume digitale in formato PDF pre-esistente e fornito dal committente.

La sfida consiste nel traghettare questa preziosa documentazione analogica in un ecosistema digitale interattivo, rendendola accessibile, strutturata e preservata nel tempo.



Obiettivi Principali del Progetto

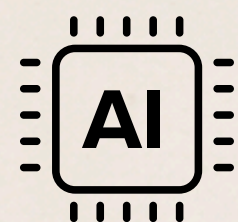
Conservazione Digitale

Preservare le informazioni storiche estratte dal PDF per assicurarne la stabilità documentale all'interno del sistema informatico d'archivio.

Estrazione Automatica

Abbandonare l'inserimento manuale in favore di una lettura semantica e logica automatizzata del testo basata su intelligenza artificiale locale.

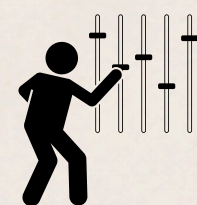
Architettura Tecnologia Utilizzata



Modello AI locale

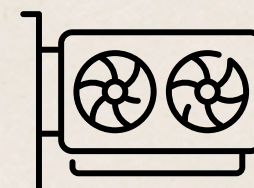
Qwen 3.5 — 0.5B parametri

Utilizzato in locale per analizzare il testo del PDF
senza inviare dati a servizi esterni.



Fine-tuning LoRA

- Specializzazione del modello
- Adattamento del modello al linguaggio storico del documento: nomi, date, luoghi, provenienze, esiti e fonti.



Infrastruttura hardware

GPU AMD Usata per accelerare l'elaborazione e
la fase di addestramento/fine-tuning.



Database e interfaccia web

HTML, CSS, JavaScript, Firebase ed API PHP.

Archiviazione dei dati strutturati e creazione
dell'interfaccia web per
consultazione, filtri e gestione delle schede.

Flusso di Funzionamento del Sistema

- ✓ **Lettura Base Documentale**
Il file PDF digitale pre-esistente e contenente il testo d'archivio viene caricato direttamente all'interno dell'ambiente locale del sistema.
- ✓ **Estrazione dei Campi Chiave Attraverso l'adattamento LoRA**
il sistema riconosce ed estrae in modo strutturato nomi, cognomi, date di nascita, luoghi di detenzione e spostamenti.
- ✓ **Interfaccia di Consultazione**
I dati strutturati e archiviati vengono messi a disposizione in tempo reale tramite un'interfaccia web responsive per ricerche rapide e mirate.

- ✓ **Elaborazione Semantica LLM**
Il modello locale addestrato analizza le righe di testo interpretandole logica per logica, superando la semplice e rigida lettura dei caratteri.
- ✓ **Normalizzazione e Importazione DB**
Le entità biografiche isolate vengono formattate (JSON/SQL) ed inserite in modo coerente nel [Database Utilizzato] relazionale.
- ✓ **Esportazione Reportistica**
La piattaforma web implementata consente la generazione automatica di schede riassuntive delle biografie per facilitare lo studio locale offline.

Diagramma di Flusso delle Informazioni

1

Sorgente Dati

Volume d'archivio digitale pre-esistente fornito in formato PDF

2

Elaborazione AI

Inference in locale del modello
[Qwen 3.5 LoRA] su hardware GPU

3

Estrazione

Parsing semantico delle informazioni
e popolamento database

4

Visualizzazione

Interfaccia Web di frontend per la
ricerca filtrata e l'accesso pubblico

Problemi Riscontrati e Soluzioni

✗ L'inadeguatezza dell' OCR Tradizionale

I primi test condotti con diversi metodi OCR standard hanno evidenziato forti limiti di lettura e interpretazione. I dati non venivano estratti adeguatamente a causa di layout d'archivio complessi e testi storici particolari.

✓ La Soluzione Adottata

Abbiamo abbandonato l'approccio OCR tradizionale, decidendo di addestrare in locale il modello LLM Qwen 3.5 (0.5B) tramite fine-tuning LoRA, ottenendo così una comprensione semantica accurata.

✗ Frammentarietà dei Dati

Manca di informazioni complete nei testi d'origine per alcuni soggetti (es. date di nascita parziali, destinazioni ignote o trascritte con toponimi obsoleti non più esistenti).

✓ La Soluzione Adottata

Progettazione di un database flessibile basato su campi opzionali controllati e indicizzazione flessibile delle località con note storiche integrative redatte in collaborazione con i committenti.

Risultati Conseguiti e Impatto



L'estrazione intelligente ha consentito di mappare in modo estremamente rapido e accurato centinaia di record biografici individuali partendo dal PDF originario.

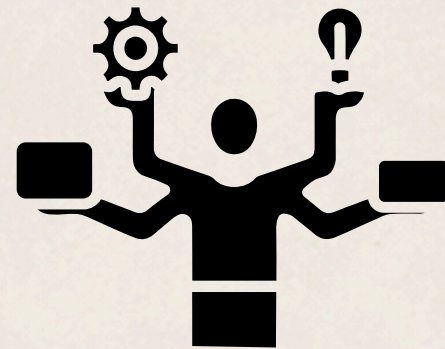
L'archivio è ora un patrimonio dinamico, interrogabile per nome, provenienza, data di arresto o specifico evento di persecuzione nel territorio modenese.

2500 min

TEMPO DI INSERIMENTO RISPARMIATO*

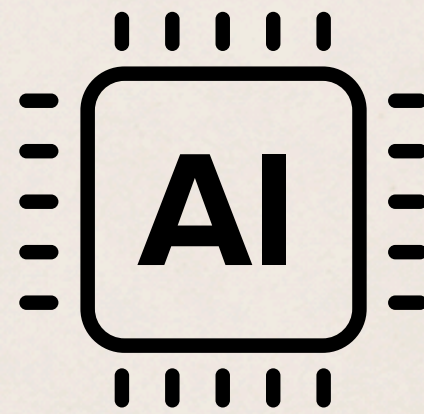
*CONSIDERANDO CIRCA 90 SECONDI PER L'
INDIVIDUAZIONE E LA TRASCRIZIONE DI OGNI EBREO

Competenze Sviluppate e Conclusioni



Soft Skills

- Lavoro in team
- Gestione delle richieste esterne
- Etica professionale della conservazione della memoria.



Hard Skills

- Fine-tuning LoRA di modelli LLM locali (Qwen 3.5 0.5B)
- Elaborazione dati offline
- Gestione di database relazionali SQL.

Domande e Risposte

Vi ringraziamo per l'attenzione dedicata

**“Se comprendere è impossibile,
conoscere è necessario.”**

~ Primo Levi

.....
Salvatore Crispino

.....
Federico Brugnano