

# Le Digital Humanities tra accademia e industria

Simone Abbiati

*Università di Trento · 17 Aprile 2026*

OGGI

# Struttura della mattinata

---

- |           |   |        |
|-----------|---|--------|
| <b>01</b> | <b>Chi sono</b>                             | 15 min |
| <b>02</b> | <b>Cosa sono le Digital Humanities</b>      | 20 min |
| <b>03</b> | <b>Lo spazio per gli umanisti nel tech</b>  | 30 min |
| <b>04</b> | <b>AI tools per l'accademia — live demo</b> | 40 min |

01

# Chi sono

*Il percorso da umanista a builder di AI*

“

*Un linguista che ha finito  
per costruire prodotti AI.*

— Simone Abbiati

IL PERCORSO

# Da linguista a builder.

Laurea  
Linguistica

PhD Ling.  
Computazionale

Startup &  
Ricerca

Think  
Tanks

Squirro AG



*Le radici*

*La svolta*

*Internazionale*

*Policy + AI*

*Oggi*

---

*Il percorso umanistico non è stato una deviazione — è stata la fondazione.*

# Cosa sono le Digital Humanities?

*Definizione, storia, metodi*

## DEFINIZIONE

# Non una disciplina. Un campo di tensione.

---

L'indagine umanistica incontra i metodi computazionali.

Non solo strumenti: un cambio epistemologico.

Scala, riproducibilità, distant reading.

## LO SPETTRO

- Digitalizzazione patrimonio
- Edizioni digitali critiche
- Analisi computazionale
- Cultural analytics
- AI applicata alle Humanities

STORIA

# Da Padre Busa ai Large Language Models.

---

**1949**

**Index Thomisticus**

Padre Busa & IBM. Nasce l'Humanities computing.

**1990–2000s**

**Svolta terminologica**

TEI, edizioni digitali, corpus linguistics. DH come campo.

**2005**

**Cultural Analytics**

Franco Moretti: la letteratura come dato su grande scala.

**2020s**

**AI Humanities**

LLMs come punto di inflessione. Il campo diventa urgente.

# I cinque pilastri del metodo DH.

---

**01. Analisi del testo & NLP**

Elaborazione del linguaggio naturale, sentiment, topic modelling

**02. Edizioni digitali & corpus**

TEI, filologia computazionale, archivi critici

**03. Network analysis**

Grafi sociali e storici, relazioni tra entità

**04. Computer vision & CH**

Riconoscimento immagini, OCR su patrimonio storico

**05. Visualizzazione dei dati**

La visualizzazione come argomento interpretativo

# Cambridge Digital Humanities.

*cdh.cam.ac.uk*

---

Fondato nel 2017, il Cambridge Digital Humanities (CDH) e' un centro interdisciplinare incardinato nelle Faculty of Arts and Humanities. Ha attivato MPhil (2022) e PhD (2024) in Digital Humanities. Collabora con la University Library e le collezioni di Cambridge.

— **(Anti)Colonial Archives Working Group**

Decolonizzazione degli archivi digitali e giustizia epistemica

— **Social Data School**

Formazione annuale su ML applicato al giornalismo investigativo e ricerca sociale

— **Future Nostalgia**

Salvaguardia della conoscenza su floppy disk: storia materiale del digitale

— **DH + RSE Summer School**

Ricerca Software Engineering applicata alle Humanities, aperta a PhD e postdoc

# Stanford CESTA & Literary Lab.

*cesta.stanford.edu / litlab.stanford.edu*

---

CESTA (Center for Spatial and Textual Analysis, 2012) e il Literary Lab sono i due centri DH piu' influenti di Stanford. CESTA integra spatial history, text analysis e data visualization; il Literary Lab produce ricerca computazionale sulla letteratura.

— **Grimke Archive AI Project**

AI per la trascrizione di archivi storici afroamericani (Howard Univ. + Stanford)

— **Chinese Railroad Workers**

Vite dei migranti cinesi che costruirono la ferrovia transcontinentale USA

— **Ottoman Fiscal Codex**

Reti finanziarie e politiche dell'Impero Ottomano (1750-1850) analizzate digitalmente

— **Literary Lab Pamphlets**

Analisi computazionale della letteratura: genere, stile, distant reading a larga scala

# ETH Zurich: DH distribuito.

*library.ethz.ch / turing.ethz.ch / ai.ethz.ch*

---

ETH non ha un unico centro DH: il lavoro e' distribuito tra ETH Library (Digital Scholarship Services), il Turing Centre, il gruppo di Prof. Menna El-Assady (visual analytics + DH) e il gruppo di Elliott Ash (NLP + scienze sociali).

— **Digital Scholarship Services / Collections as Data**

Digitalizzazione e analisi computazionale delle collezioni ETH Library

— **Turing Centre**

Centro transdisciplinare su Turing: AI, logica, digital humanities e societa'

— **El-Assady Lab — Visual Analytics**

Human-AI collaboration e visualizzazione per esplorare dati complessi nelle humanities

— **Elliott Ash — Law & NLP**

Analisi computazionale di testi legali e politici con tecniche NLP avanzate

# Lo spazio per gli umanisti nel tech.

*Necessari. Non solo benvenuti.*

# Dove il tecnico da solo fallisce.

---

## Dove i modelli mancano

- Significato e contesto
- Ragionamento storico e culturale
- Profondità interpretativa
- Etica applicata ai dati
- Sfumature linguistiche

## Dalla mia esperienza

- Il linguista che ha cambiato il prodotto
- La domanda di profili 'bilingui'
- Modelli che falliscono sulla sfumatura culturale
- Chi scrive i prompt migliori? Chi sa leggere.

ACCADEMIA

# I ruoli che esistono — e crescono.

---

## **DH Researcher / Postdoc**

Edizioni digitali, corpora, analisi computazionale di testi

## **Direttore DH Lab**

Infrastrutture, grant, team interdisciplinari

## **Linguista Computazionale**

Sintassi, semantica, corpus work, risorse linguistiche

## **Data Curator / Archivista**

TEI/IIIF, metadati, preservazione a lungo termine

## **Docente di DH**

Metodi e critical data literacy nei dipartimenti



*Finanziamenti, tenure e legittimità interna restano sfide reali.*

INDUSTRIA

# I ruoli che non cercano umanisti nell'annuncio.

---

## **NLP / Language Engineer**

Pipeline di testo, fine-tuning di modelli linguistici

## **AI Trainer / RLHF Specialist**

Addestrare i modelli — richiede basi linguistiche e culturali

## **Content & Knowledge Strategist**

Ontologie, tassonomie, struttura delle informazioni per AI

## **UX Writer / Conversation Designer**

Come parlano i sistemi AI: voce, tono, persona

## **AI Ethics & Trust Analyst**

Audit dei bias, fairness review, interlocuzione con policy

## **Prompt Engineer / AI Product Spec.**

Ruolo ponte tra modello e bisogno dell'utente

## **DH Consultant**

Musei, archivi, editori, istituzioni culturali

## ZOOM – NLP & LANGUAGE ENGINEER

---

### Cosa fai

- Costruisci pipeline NLP
- Fine-tuning su domini specifici
- Valuti qualità degli output
- Lavori su NER, disambiguation, information extraction

### Perché serve un umanista

- Il linguaggio non è solo statistiche
- Polisemia, ironia, variazione diatopica
- Valutare un testo richiede competenza filologica
- Multilinguismo: vantaggio competitivo enorme

# Il ruolo più strategico — e più trascurato.

---

## **Audit dei bias**

I dati di addestramento riflettono strutture di potere storiche. Chi li legge criticamente?

## **Fairness review**

Valutare equità tra gruppi diversi richiede teoria sociale, non solo matematica.

## **Policy e regolazione**

L'AI Act europeo, le linee guida UNESCO: tradurre le leggi in pratiche concrete.

## **Comunicazione pubblica**

Spiegare l'AI a cittadini e istituzioni: la formazione umanistica è la più adatta.

*“Come parla un sistema AI  
è una scelta retorica,  
non tecnica.”*

---

— Voce, tono e persona dei sistemi AI

— Guide di stile multilingue e multiculturali

— Fallback, errori, microtesti

— Persona design: identità per sistemi non umani

# Il ponte diretto tra accademia e mondo esterno.

---

**Musei & Archivi**

Digitalizzazione, AI per la catalogazione, accessibilità

**Case editrici**

Edizioni critiche digitali, data-driven publishing

**Istituzioni culturali**

Exhibit digitali, portali, corpus per il pubblico

**Enti pubblici & PA**

Patrimonio immateriale, policy culturale, AI governance

**Startup & Tech**

Domain expertise su contenuti storici, letterari, linguistici

## CONFRONTO

# Accademia vs Industria.

	Accademia	Industria
Ritmo	Lento, riflessivo	Veloce, iterativo
Impatto	Conoscenza e formazione	Prodotti, scala, utenti reali
Stabilità	Tenure (quando arriva)	Variabile, stipendi più alti
Libertà	Ampia autonomia di ricerca	Vincolata agli obiettivi
Rischio	Finanziamenti, precariato	Mercato, obsolescenza tecnica
Rete	Comunità scientifica internaz.	Startup / corporate

# Ora vediamo cosa può fare l'AI per voi.

*Una panoramica dei principali strumenti usati in accademia oggi.*

# Cosa è un Large Language Model.

---

Un modello statistico addestrato su miliardi di testi per predire quale parola viene dopo un'altra.

*Non 'capisce' il linguaggio nel senso umano. Impara pattern su scala industriale.*

## **Token**

L'unità minima di testo ( $\approx$  parola o frammento)

## **Parametri**

I 'pesi' appresi durante il training (GPT-4:  $\sim$ 1.7T)

## **Contesto**

Quanti token il modello 'ricorda' in una conversazione

## **Allucinazione**

Quando il modello genera testo plausibile ma falso

# Come viene addestrato.

---

## **01.** Pre-training

Il modello legge enormi corpus di testo (web, libri, codice) e impara a predire il token successivo. Mesi di compute su migliaia di GPU.

## **02.** Fine-tuning supervisionato (SFT)

Esempi scritti da umani mostrano al modello il comportamento desiderato: come rispondere a domande, seguire istruzioni, strutturare testi.

## **03.** RLHF — Reinforcement Learning from Human Feedback

Valutatori umani confrontano risposte e le classificano. Un modello di ricompensa impara le preferenze umane e guida ulteriore ottimizzazione.

# Come si usa nella pratica.

---

## API

*Chiamata via HTTP a un modello hosted (es. Claude, GPT-4, Gemini)*

Nessuna infrastruttura propria. Si paga per token. Il modello è aggiornato dal provider. Usato da quasi tutti i tool accademici che vedremo.

## Fine-tuning

*Ri-addestramento su dati propri*

Si parte da un modello base e si specializza su un dominio (es. testi medievali, clinical notes). Richiede dati annotati e compute.

## RAG — Retrieval Augmented Generation

*Recupero + generazione in tempo reale*

Prima recupera documenti rilevanti da un database (es. i tuoi paper), poi il modello genera la risposta basandosi su quei documenti. Riduce drasticamente le allucinazioni.

# Cosa è un Agent Framework.

---

*Un LLM da solo risponde a un messaggio. Un agente percepisce, pianifica, agisce e itera.*

- **Modello (brain):** LLM che ragiona e decide le azioni da compiere
- **Strumenti (tools):** Funzioni che il modello può chiamare: cerca sul web, leggi un file, esegui codice
- **Memoria:** Contesto a breve termine (conversazione) e a lungo termine (vector database)
- **Loop di esecuzione:** Il modello agisce → osserva il risultato → ri-pianifica → agisce di nuovo

# Cinque aree. Venti strumenti.

---

- 01. Scoperta & Literature Review**  
Trovare paper rilevanti, costruire mappe della letteratura
- 02. Scrittura & Revisione**  
Assistenza alla scrittura accademica, grammatica, tesi
- 03. Comprensione & Sintesi**  
Spiegare paper complessi, generare riassunti e note
- 04. Grant & Finanziamenti**  
Trovare opportunità di finanziamento, redigere proposte
- 05. Visualizzazione & Comunicazione**  
Infografiche, poster scientifici, dati visivi

# Trovare i paper giusti. Capire il campo.

---

## Consensus

*consensus.app*

Motore AI su 200M+ paper; sintetizza il consenso scientifico su un topic

## Elicit

*elicit.com*

Cerca, estrae dati e genera tabelle da 125M+ paper; ottimo per systematic review

## Semantic Scholar

*semanticscholar.org*

Database gratuito con AI reader, TLDRs e raccomandazioni personalizzate

## Research Rabbit

*researchrabbit.ai*

Mappa visiva di reti di citazioni; scopre connessioni tra paper e autori

## Litmaps

*litmaps.com*

Mappe di citazioni su 270M+ paper; monitora nuovi paper in automatico

## Connected Papers

*connectedpapers.com*

Grafico relazioni tra paper: asse x=data, y=citazioni; identifica paper fondativi

## Sourcely

*sourcely.net*

Suggerisce e inserisce citazioni rilevanti partendo dal testo già scritto

## SciSpace

*scispace.com*

Piattaforma all-in-one: ricerca, review sistematiche, drafting, raccomandazione journal

# Scrivere meglio, più velocemente, con rigore.

---

## Jenni AI

*jenni.ai*

Workspace per ricercatori: autocomplete, gestione citazioni, chat con PDF, integra Zotero

## Logically AI

*logically.ai*

Verifica di affermazioni e fact-checking; utile per la solidità argomentativa dei testi

# Capire paper complessi. Organizzare la conoscenza.

---

## Explain Paper

[explainpaper.com](https://explainpaper.com)

Carica un PDF: spiega termini, metodologie e passi complessi in linguaggio semplice

## NotebookLM

[notebooklm.google](https://notebooklm.google)

Google: carica documenti, genera sintesi, podcast audio, mappe concettuali e Q&A

## Scholarcy

[scholarcy.com](https://scholarcy.com)

Genera flashcard riassuntive di paper; estrae hypothesis, metodologie, risultati

## Moonlight AI

[moonlight.ai](https://moonlight.ai)

Workspace AI per organizzare note, paper e appunti di ricerca in modo unificato

# Trovare fondi. Scrivere proposte convincenti.

---

## Seamless AI

*seamless.ai*

AI per trovare opportunità di finanziamento e supportare la redazione di proposte di grant

---

## Come l'AI può aiutare con i grant

**Scoperta opportunità:** Ricerca automatica di bandi rilevanti per profilo e keyword di ricerca

**Prima bozza:** Generazione di sezioni (Aims, Significance, Innovation) da affinare

**Analisi grant finanziati:** Confronto con proposte simili già finanziate per calibrare il tono

**Nota bene:** I principali enti (UKRI, NIH, ERC) scoraggiano l'uso non dichiarato di AI generativa

# Comunicare la ricerca. Render the data visible.

---

## Eraser.io — DiagramGPT

*eraser.io/diagramgpt*


Genera diagrammi, flowchart e architetture visive da testo in linguaggio naturale; utile per visualizzare processi di ricerca e strutture dati

---

## Perché la comunicazione è parte della ricerca

- Diagrammi & flowchart: Eraser.io DiagramGPT, Miro AI
- Presentazioni: Gamma.app, SlidesAI, Beautiful.ai
- Mappe concettuali e mindmap: NotebookLM, Miro AI
- Video e reel da ricerca: Synthesia, HeyGen per divulgazione

*“Quale problema nel tuo campo  
potresti notare solo tu  
come umanista —  
ma che ha bisogno di  
una macchina per risolverlo?”*

 *È la tua proposta di valore.*

IL CAMPO

# Dove si colloca il campo DH oggi.

---

## Scienze Umane

Filologia

Linguistica

Storia

Letteratura

Filosofia

Archeologia

## Computer Science

NLP

Machine Learning

Data Engineering

Computer Vision

HCI

Database

## Design & Comunicazione

UX/UI

Visualizzazione

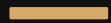
Storytelling

Editoria digitale

Architettura info

Accessibilità

# Grazie.



**Simone Abbiati**

*[linkedin.com/in/simoneabbiati](https://www.linkedin.com/in/simoneabbiati)*

→ Dialogo con Paolo Bonafede