



ALMA MATER STUDIORUM | DIPARTIMENTO
UNIVERSITÀ DI BOLOGNA | DI STORIA CULTURE CIVILTÀ

Dottorato in Scienze Storiche e Archeologiche

9, 10 e 13 maggio 2024, ore 10.00-13.00
Piazza San Giovanni in Monte 2

Aula Morandi

Transkribus: formazione e applicazione di *Handwritten Text Recognition* software in fonti storiche dal Medioevo all'Età contemporanea



Edward Loss (UniBo/MemoBo)

Accesso libero previa l'iscrizione.

Informazioni: edward.loss2@unibo.it

Programma

09/05/2024

- Cos'è la *Handwritten Text Recognition*
- Cos'è *Transkribus* e quello che si può fare con il software
- L'interfaccia di *Transkribus*
- Creare una collezione – *upload* dei documenti
- Gli elementi del Layout (*Text Regions, Baselines e Polygons*)
- Cos'è un "Modello"
- Cos'è la *Ground Truth (GT)*, il *Training* e il *Validation Set*
- Come allenare un *Baselines Model*
- Come allenare un *Fields Model*
- Come allenare un *Table Model*
- Prima sessione pratica

10/05/2024

- Riconoscimento del testo (*Text Recognition*)
- I "Modelli Pubblici"
- Come allenare un *Text Recognition Model*
- Come valutare un *Text Recognition Model*
- La *Learning Curve* e gli *Epochs*
- *Tags* Testuali e Strutturali
- Seconda sessione pratica

13/05/2024

- Opzioni avanzate di allenamento dei *Text Recognition Models*
- Opzioni avanzate di *upload* dei documenti
- Strumenti di ricerca (*Smart Search*)
- Esportazione
- Opzioni avanzate di esportazione
- Terza sessione pratica