

Maratona di trascrizione

Il Cosmo di Cosimo: Cosimo Bartoli e l'informazione nel Rinascimento

23-25 marzo 2022

Sara Mansutti
University College Cork
Dottoranda in Digital Humanities

sara.mansutti11@gmail.com

Iniziativa organizzata da Marco Sartor (Università di Parma), patrocinata dal Venice Centre for Digital and Public Humanities e finanziata con i fondi per le attività studentesche dell'Università Ca' Foscari Venezia

ad agio a far tal deliberatione, et
et conquista openione sene va aguo
no far male officio essendo inverno
oppinione di religioni. a me
molte offerte fece V. C. et ha mostro
fatto quanto alle cose della sua Corona
fedr.
Lo Imfasciador.

ar tal deliberatione, et guardi a no si provacare un'
ta openione sene va a quella corte dove si prensa ch
ufficio essendo inverno tutto intrinsecamente volto
elligione. a me nondimeno stamattina in Palazzo
. E. et ha mostro con parole partirsi molto
cose della sua Corona, et offertos di farne in du

Cosimo Bartoli

- 1503-1572
- Preposto del Battistero di San Giovanni, Firenze
- Membro dell'Accademia Fiorentina
- Al servizio del cardinale Giovanni de' Medici
- Agente del Duca a Venezia tra il 1562 e il 1572
- Uomo di lettere
- Traduttore del *De Re Aedificatoria* di Leon Battista Alberti
- Amico di Giorgio Vasari





Progetto su Cosimo Bartoli

- Edizione digitale delle lettere e degli avvisi inviati da Cosimo Bartoli mentre era agente a Venezia (1562-1572), oggi all'Archivio di Stato di Firenze
- Comprendere come Bartoli otteneva le informazioni, dove e da chi
- Capire la differenza tra le notizie scritte negli avvisi e quelle scritte nelle lettere
- Rendere queste fonti disponibili online
- Sperimentare l'uso di vari strumenti digitali per creare l'edizione

Il riconoscimento automatico del testo manoscritto

OCR (Optical Character Recognition)

DRURY LANE.
By his MAJESTY's Company.
At the Theatre Royal in Drury Lane,
This Day before he Play a f.r urit
Ptcwt ca led O vt 3 I? ftp
The E L E C T I O N :
The principal Charaer n b,
Mr. VERNON,
Mr. CUDRY
And Mr. BANNISTER.
Mrs. LOVE,
And Mrs COLLETT.
After which will be performed a Comedy called
The M I S E R.
The Miser, Mr. YATES ;
Frederick, Mr. Brereton; Clermont, Mr. Farren
Decoy, M . W a d o n ; James, Mr.
Mr. Chap. is ; Upholderer, Mr wing
Sattin, Mr. N o i « ; Devrei.e.,
Ta) lor, Mr Ho croft.
ABO Ramjlif, Mr. PALMER.
Mariana hy Mrs. BR-RE O.V;
Harriet, M.s. S atp ; Mrs. wife].
W..eidle, Mi's Kirb ;
Aad Lappet, Mils POPE.
To which will be added a Pernaj.
Part.)

HTR (Handwritten Text Recognition)

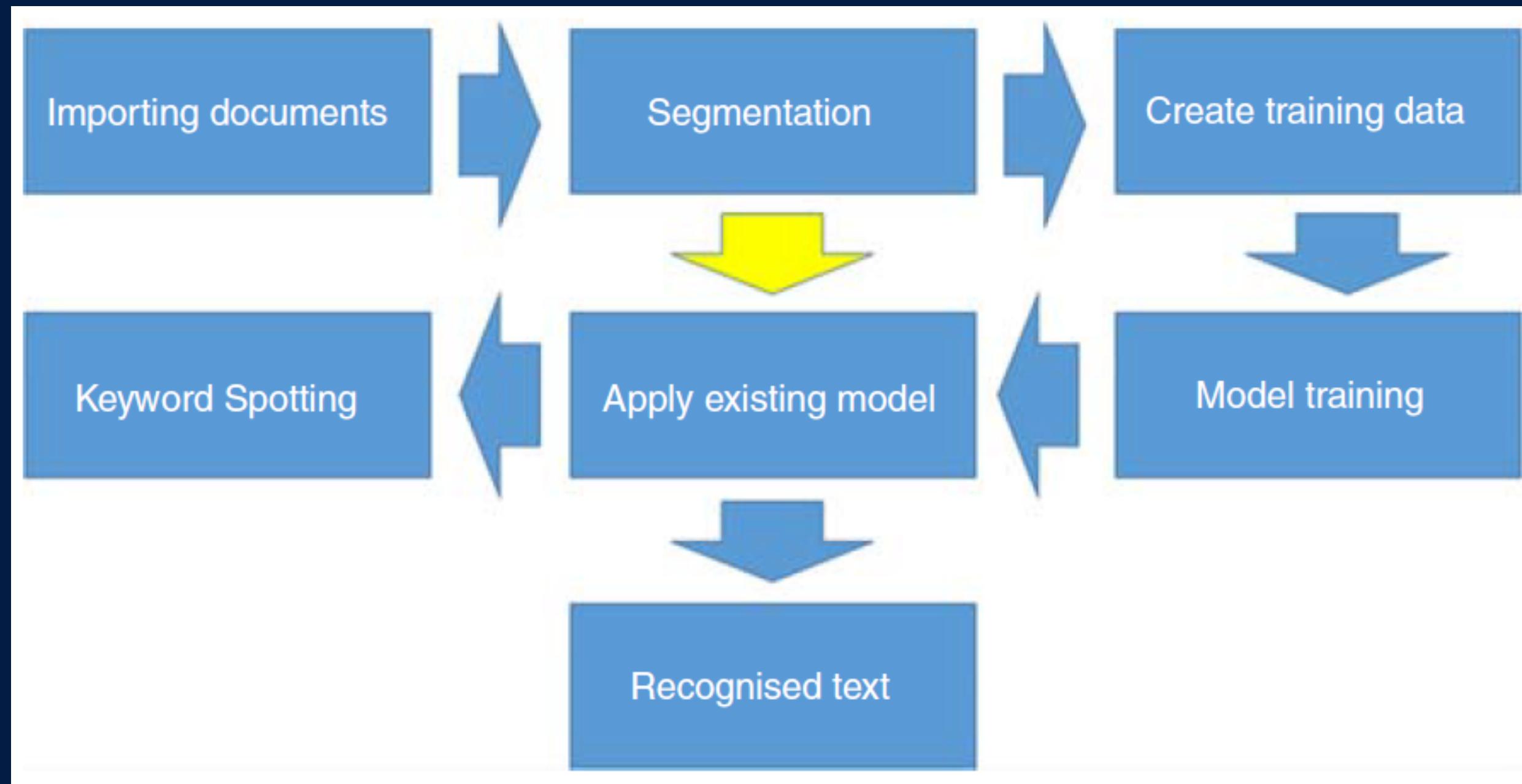
1-3 Che Maurilio certificato ch la massa delle genti del Conte di Mansfelt era di poco mo
1-4 mento, ne si partiva decosini di Brema lontano dallo stato suo nuovamente a cui
1-5 stato 200 miglia, senera tutto sicero ritornato in Corte il cui ritorno diede
1-6 gran contento a em. la q^ale con tutto q.o pensava di proveder' à tutto q.i ch posesse
1-7 fare il Mansfelt cò mandarvi Mons^re di Bura havendo ma xo q. ich saspetto

Transkribus

- piattaforma per la digitalizzazione, il riconoscimento del testo manoscritto basato sull'IA, la trascrizione e la ricerca di documenti storici
- <https://readcoop.eu/transkribus>



Workflow di Transkribus



from Mühlberger, G., et al., "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study", Journal of Documentation, Vol. 75 No. 5, pp. 954-976. <https://doi.org/10.1108/JD-07-2018-0114>

Transkribus e il riconoscimento automatico del testo manoscritto

Primo modello HTR (modello base): Italian Administrative Hands, 1550-1700

- Collaborazione tra Jake Dyble (University of Exeter e Università di Pisa), Antonio Iodice (University of Exeter e Università di Genova), Rachel Midura (Virginia Polytechnic Institute and State University)
- Modello pubblico in Transkribus
- Varie mani
- Training Set: 67361 parole
- CER (Character Error Rate) del Validation Set: 10.80%
- Descrizione del modello: <https://emdigit.org/tool/2020/07/21/italian-administrative-hands.html>

Transkribus e il riconoscimento automatico del testo manoscritto

9 versioni per ottenere un soddisfacente modello per i documenti di Bartoli:

- Selezione casuale di alcune pagine dei 9 volumi di Bartoli
- Applicazione del modello Italian Administrative Hands
- Trascrizioni corrette manualmente
- Addestramento di un nuovo modello (versioni 1 e 2)
- Aggiunta di più pagine
- Addestramento di un nuovo modello (versione 3)
- Aggiunta di più pagine
- Addestramento di un nuovo modello (versioni 4, 5 e 6)
- Correzione del testo e del layout di tutte le pagine
- Addestramento di un nuovo modello (versioni 7, 8 e 9)

Differenze tra le versioni:

- presenza del modello base;
- numero di iterazioni;
- numero di pagine trascritte su cui allenare il modello;
- Correzione del testo e del layout.

Transkribus e il riconoscimento automatico del testo manoscritto

Version	Base Model	Epochs	Nr. of words	CER on Validation SET
1	yes	200	17828	8.36%
2	no	200	17828	9.56%
3	yes	200	34009	6.96%
4	yes	200	52124	6.43%
5	no	200	52124	6.68%
6	yes	400	52124	6.45%
7	yes	200	51736	5.70%
8	yes	400	51736	5.81%
9	yes	800	51736	5.82%

Ch vadia adagio a far tal deliberatione, et guardi a no si provacare un
tanto Re. et con questa openione sene va a quella corte dove si prensa ch pos
sa anzich nō far male officio essendo invero tutto intrinsecamente volto
alla nuova opinione di Relligione. a me nōdimeno stamattina in Palazzo
ha fatto molte offerte p V. E. et ha mostro con parole partisimo molto
ben satisfatto quanto alle cose della sua Corona, et offertosi di farne in du

1-1 Ch vadia adagio a far tal deliberatione, et guardi a no si provacare un'

1-2 tanto Re. et con questa openione sene va a quella corte dove si prensa ch pos

1-3 sa anzich nō far male officio essendo invero tutto intrinsecamente volto

1-4 alla nuova opinione di Relligione. a me nōdimeno stamattina in Palazzo

1-5 ha fatto molte offerte p V. E. et ha mostro con parole partisimo molto

1-6 ben satisfatto quanto alle cose della sua Corona, et offertosi di farne in du

Transkribus e il riconoscimento automatico del testo manoscritto

Script "abbrevSolver", creato da Ismail Prada, per taggare e espandere le abbreviazioni più frequenti:
<https://github.com/raykyn/abbrevSolver>

Abbreviation	Expansion	Insecure?	Context
Cons̄ pli	Constantinopoli		
Consp̄ li	Constantinopoli		
Const̄ pli	Constantinopoli		
Cōnst̄ pli	Constantinopoli		
Const̄ poli	Constantinopoli		
Constā pli	Constantinopoli		
Constan.li	Constantinopoli		
Constan̄ li	Constantinopoli		
Constanli	Constantinopoli		
Constantī pli	Constantinopoli		
Cost̄ pli	Costantinopoli		
Cōst̄ pli	Constantinopoli		
Costan.li	Costantinopoli		
Cōstanp̄ li	Constantinopoli		
Cōstātī pli	Constantinopoli		
Cōstp̄ li	Constantinopoli		



Il cosmo di Cosimo

Cosimo Bartoli e l'informazione nel Rinascimento

<https://cosimobartoli.saramansutti.com/>

Crowdsourcing

Il crowdsourcing è una tipologia di **attività online partecipativa** in cui un individuo, un'istituzione, un'organizzazione no profit, una compagnia propone a un **gruppo di individui** eterogeneo, tramite un annuncio aperto e flessibile, la **realizzazione libera e volontaria** di un compito.

La realizzazione di tale compito, di complessità e modularità variabile, in cui i partecipanti dovrebbero contribuire apportando lavoro, soldi, conoscenza e/o esperienza, **prevede sempre un beneficio per entrambe le parti**.

Il partecipante riceverà, in cambio della sua partecipazione, il soddisfacimento di una necessità, che può essere economica, di riconoscimento sociale, di autostima, di sviluppo di capacità individuali, mentre l'organizzatore, dal canto suo, otterrà e utilizzerà a proprio beneficio il contributo offerto dai partecipanti, la cui forma dipende dall'attività realizzata.

Estellés-Arolas, Enrique, and Fernando González-Ladrón-de-Guevara. “Towards an Integrated Crowdsourcing Definition.” *Journal of Information Science*, vol. 38, no. 2, Apr 2012, pp. 189–200

Compiti richiesti ai partecipanti

- Correggere le trascrizioni generate automaticamente
- Taggare le abbreviazioni e aggiungere la forma espansa
- Nelle lettere: taggare con il tag “person” le persone che Cosimo Bartoli menziona di aver incontrato di persona o le persone con cui ha avuto uno scambio epistolare
- Negli avvisi: due elementi da taggare ed entrambi si trovano nell’ “intestazione” dell'avviso.
- Tag “place” per taggare il centro indicato nell'intestazione, cioè il luogo da cui provengono le notizie. Scrivi il nome del centro nella sua lingua originale nella casella “place”.
- Tag “date” per taggare la data dell'intestazione e scrivi la data nel formato aaaa-mm-gg nella casella “when”.

Norme di trascrizione

- TESTO
- ORTOGRAFIA
- FORMATTAZIONE
- ABBREVIAZIONI
- NUMERI
- INSERTI
- TESTO INCERTO

[https://cosimobartoli.saramansutti.com/it/come-correggere-e-taggere-i-
documenti/](https://cosimobartoli.saramansutti.com/it/come-correggere-e-taggere-i-documenti/)

Link da ricordare

- Google Sheet in cui segnare le pagine su cui si sta lavorando:
<https://docs.google.com/spreadsheets/d/1vjpyotlD-9NGk16OP-CEVH8CThMpeKoi5vytfue9w-0/edit#gid=827425675>
- Come correggere e taggare i documenti:
<https://cosimobartoli.saramansutti.com/it/come-correggere-e-taggare-i-documenti/>
- Lista delle abbreviazioni: <https://cosimobartoli.saramansutti.com/it/abbreviazioni/>
- Forum di discussione: <https://cosimobartoli.saramansutti.com/it/discussion-forum-2/>
- Email del progetto: cosimobartoli.project@gmail.com