

AIUCD 2016

Quinto convegno annuale, [Venezia](#), 6-9 settembre 2016

Tutorial per il rilevamento automatico di riusi testuali in documenti storici

Marco Büchler, Emily Franzini, Greta Franzini
Istituto di Informatica, Università di Gottinga

Gli organizzatori

[eTRAP \(Electronic Text Reuse Acquisition Project\)](#) è un nuovo gruppo di ricerca finanziato dal Ministero dell'Istruzione e della Ricerca tedesco (BMBF), con sede all'Istituto di Informatica dell'Università di Gottinga. Il progetto ha una durata di quattro anni, dal 1 marzo 2015 al 28 febbraio 2019. Il gruppo, con una forte natura interdisciplinare, si occupa del fenomeno linguistico e letterario del riuso testuale, con un interesse particolare per lingue storico-naturali; in particolare su come autori del passato copiavano, parafrasavano, traducevano e alludevano ad altri autori diffondendo così la conoscenza tramite lo strumento della scrittura.

Lo scopo è quindi quello di chiarire e definire in cosa consista il riuso testuale, il perché e come i testi vengano riutati e come questa pratica si sia evoluta nella storia.

Maggiori informazioni su eTRAP, il gruppo di ricerca e i relativi progetti sono disponibili su

<http://www.etrp.eu>

Il tutorial

Punto di partenza del tutorial saranno le attività di ricerca di eTRAP, molte delle quali fanno uso di [TRACER](#) (<http://www.etrp.eu/research/tracer/>), uno strumento di analisi testuale volto al rilevamento automatico di riusi testuali in vari corpora, siano essi in prosa o poesia. TRACER è stato concepito per l'utilizzo con lingue storico-naturali, quali il greco antico, il latino, l'arabo classico o il tedesco medievale, e fornisce all'utente un potente motore di ricerca per l'identificazione e la visualizzazione di diversi tipi di riuso, da citazioni a parafrasi. TRACER si basa su algoritmi di Natural Language Processing (NLP) ed è caratterizzato da un flusso di lavoro che permette di scomporre la non semplice identificazione del riuso testuale in compiti più piccoli e semplici. Grazie ad un file di configurazione TRACER l'utente può modificare filtri e parametri in base ai suoi interessi specifici.

I risultati della ricerca di TRACER possono poi essere visualizzati grazie a TRAViz, uno strumento integrato e scritto in JavaScript per la rappresentazione grafica di varianti testuali (<http://www.traviz.vizcovery.org/>). TRAViz permette una lettura sia close che distant per uno studio più intuitivo e immediato dei testi.

Durante il tutorial i partecipanti impareranno ad usare sia TRACER sia TRAViz. Nel primo pomeriggio di martedì 6 settembre verranno introdotti i comandi di base e si lavorerà con i testi in lingua inglese forniti da eTRAP per prendere dimestichezza con lo strumento. Il giorno successivo, in base ai risultati

raggiunti, i partecipanti potranno cominciare ad interrogare i propri testi e a usare TRACER in base alle loro esigenze. Naturalmente i testi utilizzati dovranno essere conformi al formato supportato.

Programma (Aula 2, Ca' Dolfin)

Martedì 6 settembre, 14:00-19:00

- Sessione 1 (1 ora) - Introduzione al rilevamento automatico di riusi testuali: sfide, problemi, e possibilità.
- Sessione 2 (3 ore) - Introduzione a TRACER e TRAViz: esercitazioni con dati (in inglese) forniti dagli organizzatori.

Buffer: 1 ora (pause, richieste d'assistenza, ecc.).

Mercoledì 7 settembre, 09:00-14:00

- Sessione 3 (3 ore) - Uso di TRACER con dati o corpora propri.
- Sessione 4 (1 ora) - Conclusione: presentazione dei risultati e passi successivi per analisi più sofisticate.

Le pause caffè verranno liberamente gestite dal gruppo in base alle esigenze giornaliere.

Requisiti & Aspetti tecnici

Si richiede che i partecipanti abbiano piena dimestichezza perlomeno con un sistema operativo. Si aspetta inoltre una conoscenza di base dei principi degli studi intertestuali e del riuso testuale.

Una lista completa di requisiti tecnici e istruzioni verrà inoltrata ai partecipanti prima della conferenza con un sufficiente anticipo. Tutti gli aspetti tecnici dovranno essere approntati prima del tutorial così da evitare ritardi il giorno stesso.

Sommario riassuntivo

Questo tutorial è rivolto a coloro che nutrono un interesse per lo studio dei riusi testuali tra due o più corpora (nella stessa lingua). Verrà insegnato ai partecipanti come utilizzare TRACER per il rilevamento semi-automatico e la visualizzazione di riusi testuali in vari corpora. Lo scopo è quindi quello di incoraggiare i partecipanti ad adattare e applicare le competenze acquisite nelle proprie ricerche, così da instaurare collaborazioni che potranno portare a eventuali pubblicazioni.

Per poter seguire adeguatamente ciascun partecipante, **il tutorial prevede 12 posti**. Non è previsto nessun costo e **si svolgerà in lingua inglese**. Per eventuali chiarimenti sarà possibile utilizzare la lingua italiana.

Questo tutorial è già stato organizzato con successo a Gottinga (luglio 2015), Londra (settembre 2015), Tartu (ottobre 2015) e Galway (febbraio 2016) e verrà riproposto a Cracovia (luglio 2016).

Richiesta di partecipazione

Per partecipare si prega di inviare un curriculum e una lettera di motivazione a **etrap-applications@gcdh.de** entro la mezzanotte del **31 luglio 2016**. Informazioni sull'accettazione verranno comunicate entro il 10 agosto 2016.