

Descrizione del corso

Docenti	Insegnamento	Abstract	Data ora	Aula	Riferimenti bibliografici
<p>Docenti del progetto H2IOSC: Monica Monachini, Francesca Frontini, Federico Boschetti, Angelo Mario Del Grosso, Anas Fahad Khan, Giulia Pedonese</p>	<p>L'infrastruttura CLARIN per la gestione FAIR dei dati linguistici e testuali</p>	<p>Questo corso è destinato ad un pubblico di linguisti ma anche filologi e a chiunque debba approfondire la gestione dei dati linguistici e testuali. Al termine del corso gli studenti padroneggeranno nozioni quali quelle di scienza aperta, buona gestione dei dati della ricerca, piano gestione dei dati, e sapranno come rendere i loro corpora e risorse FAIR (findable, accessible, interoperable, e reusable); inoltre padroneggeranno i principali strumenti dell'infrastruttura CLARIN per le risorse linguistiche. Il corso articola in moduli distinti, pensati per fornire ai partecipanti una solida base teorica, strumenti pratici avanzati e applicazioni specifiche per la gestione dei dati. I 4 moduli di 4 ore ciascuno si svolgeranno i giovedì pomeriggio di marzo 2025, per un totale di 16 ore. La partecipazione al corso sarà da remoto. Il corso è sviluppato nel contesto del progetto H2IOSC che mira a sviluppare i nodi nazionali di quattro infrastrutture di ricerca nel settore delle scienze umane e del patrimonio culturale.</p>	<p>6, 13, 20, 27 marzo 2025 Ore 14:00 - 18:00</p>	<p>Online</p>	<p>Monachini, Monica, and Francesca Frontini. 2016. 'CLARIN, l'infrastruttura Europea Delle Risorse Linguistiche per Le Scienze Umane e Sociali e Il Suo Network Italiano CLARIN-IT'. IJCoL - Italian Journal of Computational Linguistics, Special Issue on NLP and Digital Humanities, 2 (2): 11–30. https://doi.org/10.4000/ijcol.387.</p> <p>Jong, Franciska de, Dieter Van Uytvanck, Francesca Frontini, Antal van den Bosch, Darja Fišer, and Andreas Witt. 2022. 'Language Matters. The European Research Infrastructure CLARIN, Today and Tomorrow'. In CLARIN. The Infrastructure for Language Resources, by Darja Fišer and Andreas Witt, 1:31–58. Digital Linguistics. Berlin, Boston: De Gruyter. https://doi.org/10.1515/9783110767377-002.</p>

Programma dettagliato dei moduli:

Primo Modulo: Fondamenti di FAIR e Open Science - Ricerca e Deposito

6 MARZO 2025, Giovedì 14:30 – 18:30

Giulia Pedonese, Federico Boschetti

- **Contenuti:**
 - Concetti di FAIR (Findable, Accessible, Interoperable, Reusable)
 - Open Science (OS)
 - Tecniche di ricerca e deposito dei dati
 - Come “farsi aiutare” e accedere all’expertise della rete di centri CLARIN – l’esempio del Knowledge Centre for Digital and Public Textual Scholarship (DiPText-KC)
 - Esercitazioni pratiche
- **Durata:** 4 ore (2 ore di lezione frontale, 2 ore di esercizi)

Secondo Modulo: Strumenti CLARIN e NLP - Applicazioni Pratiche

13 MARZO 2025, Giovedì 14:30 – 18:30

Giulia Pedonese, Angelo Mario Del Grosso

- **Contenuti:**
 - Utilizzo di strumenti e servizi CLARIN
 - Servizi di ricerca
 - Virtual Language Observatory (VLO)
 - Servizi di analisi testiale
 - Switchboard
 - WebLicht
 - Voyant Tools
 - Esercitazioni pratiche
- **Durata:** 4 ore (2 ore di lezione frontale, 2 ore di esercizi)

Terzo Modulo: Interrogazione e utilizzo di corpora annotati – l'esempio ParlaMint

20 MARZO 2025, Giovedì 14:30 – 18:30

Monica Monachini, Giulia Pedonese

- **Contenuti:**
 - CLARIN-IT come esempio di nodo CLARIN, e i suoi servizi
 - Esempi di corpora di riferimento dai vari nodi CLARIN, incluso CLARIN-IT
 - La Federated Content Search
 - Approfondimento sull’interrogazione di corpora scritti con l'esempio del corpus ParlaMint
- **Durata:** 4 ore (2 ore di lezione frontale, 2 ore di esercizi)

Quarto Modulo: Interrogazione di risorse lessicali e ontologiche – l'esempio di LiLa

27 MARZO 2025, Giovedì 14:30 – 18:30

Giulia Pedonese, Francesca Frontini (o Fahad Khan)

- **Contenuti:**

- Breve introduzione alle risorse lessicali digitali e loro formati, con particolare riferimento ai Linguistic Linked Open Data e al paradigma del Web Semantico
- Le risorse terminologiche e i lessicali di CLARIN
- LiLa: Una risorsa del consorzio CLARIN-IT
 - Esercitazione
- **Durata:** 4 ore (2 ore di lezione frontale, 2 ore di esercizi)