Abilità informatiche

A.A. 2023/2024

06 - Digital Humanities

Sebastian Barzaghi

<u>sebastian.barzaghi2@unibo.it</u> <u>https://orcid.org/0000-0002-0799-1527</u>

Riassunto della lezione precedente (05c)

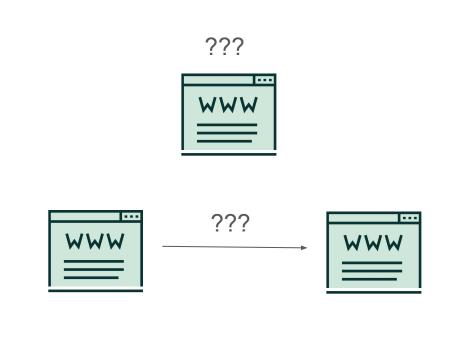
I contenuti sul Web sono (solo) per noi

Contenuti leggibili e comprensibili dagli esseri umani

Ma le macchine?

HTML \rightarrow come rappresentare (non cosa)

Alcuni tag sono semantici (es. <title>) ma il loro contenuto non è strutturato né standardizzato

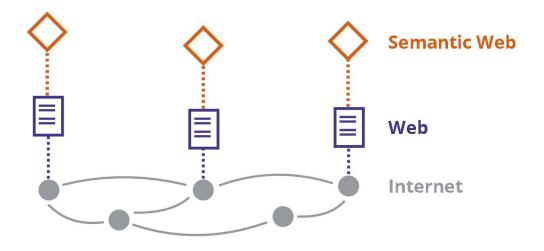


Il Semantic Web è uno strato ulteriore



Proposto da Tim Berners Lee nel 2001

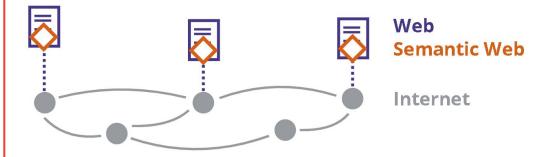
Ragionare sui dati disponibili sul Web in maniera automatica, estendendo il Web con informazioni semantiche



Il Semantic Web è integrato nel Web



- aggiungere informazioni
- aggiungere struttura
- permettere collegamenti semantici tra silos informativi
- permettere inferenze logiche (automatiche) sui dati

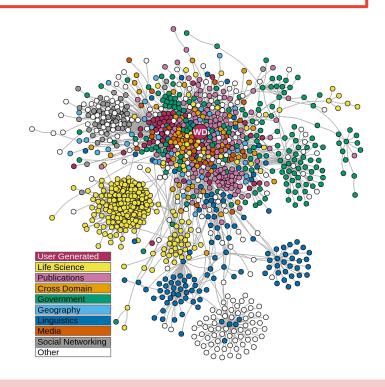




La base sono i Linked (Open) Data

Dati (semi-)strutturati in grafi, interpretabili dalle macchine (pubblicati in formato aperto)

- 1. Molteplici dataset con licenza aperta
- Stessi formati standard di riferimento ai e modellazione dei dati
- → interrogazioni incrociate su dataset interoperabili

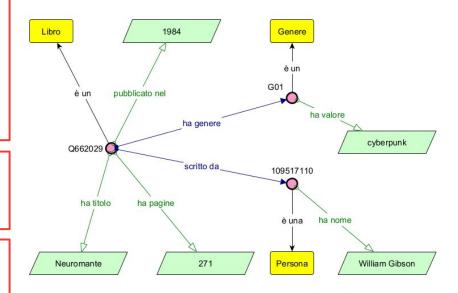


Gli identificatori possono esistere già...

In entrambi gli esempi, abbiamo utilizzato degli authority file (o risorse simili) per assegnare identificativi unici a William Gibson e a Neuromante

109517110 (William Gibson) è preso da <u>VIAF</u>

Q662029 (Neuromante) è preso da WikiData



... ma non sempre!

5-star rating system:

- disponibile sul Web con una licenza aperta (es. Creative Commons)
- 2. 1 + formato strutturato e leggibile dalle macchine (es. Excel)
- 3. 2 + formato non proprietario (es. CSV)
- 4. 3 + usa standard aperti per identificare (es. RDF)
- 5. 4 + link a dati esterni per fornire ulteriore contesto

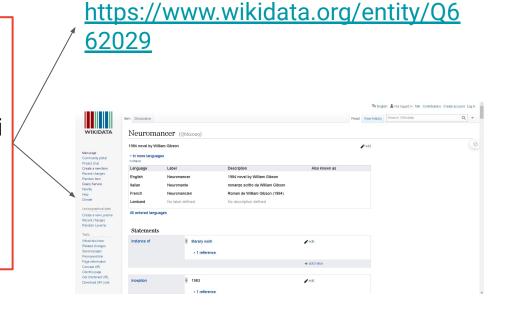




Quattro principi di pubblicazione LOD



- 1. usare **URI** come nomi per le cose
- 2. usare **HTTP** per permettere ai computer di cercare questi nomi
- 3. fornire **informazioni utili** al momento della ricerca
- 4. includere **link** ad altre cose

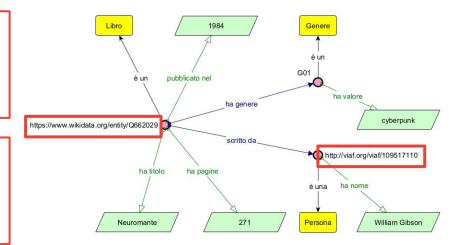


Usare URI per dare nomi alle cose: esempio

•••••

109517110 (William Gibson) e Q662029 (Neuromante) sono necessari ma non sufficienti

- http://viaf.org/viaf/109517110
- https://www.wikidata.org/entity/ /Q662029



Gli URI disambiguano su tutto il Web (perché i loro domini sono unici, perché registrati in maniera univoca nel DNS!)

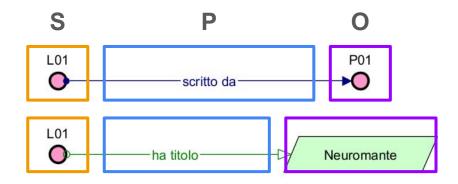
L'unità minima nei LOD: la tripla RDF

Costrutto astratto minimo di modellazione dei LOD

 $\begin{array}{l} \textbf{Soggetto} \rightarrow \textbf{il dominio del} \\ \textbf{predicato} \end{array}$

Predicato → una caratteristica del soggetto

Oggetto → il codominio del predicato



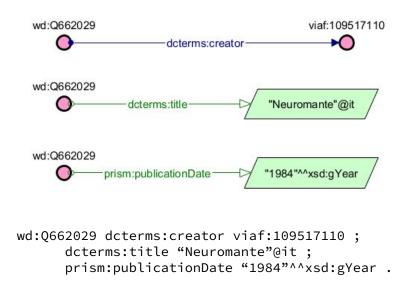
Resource Description Framework

Modello di dati standard che descrive i dati tramite triple SPO

• • • • • • • • • • • •

Dice cosa fare, ma non:

- come scrivere triple → per questo ci sono le serializzazioni (sintassi concrete di RDF)
- come esprimere entità e proprietà → per questo ci sono vocabolari, ontologie, ecc.



Il Semantic Web ha tanti dati pronti al riuso

Livello strutturale: vocabolari (**modelli semantici** di dati) che forniscono classi e proprietà da riusare

Livello contenutistico: dataset che forniscono identificativi e dati delle entità

- Dublin Core
- FOAF
- SKOS
- ...
- OpenCitations Meta
- OpenCitations Index
 - ...
- WikiData
- ..

Perché non usare i DB relazionali, allora?



Assunzione del mondo aperto: in un sistema logico, l'assenza di un fatto non lo rende falso

I database relazionali usano strutture rigide, poiché aspirano a dati completi (quindi mondo chiuso)

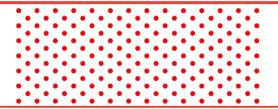
In LOD, nessuna fonte ha tutti i dati

Tassonomie, tesauri, ontologie strutturano i LOD con un'apertura alla continua integrazione ed estensione

Dati, modelli pensati per essere

- espressivi
- riutilizzati
- potenzialmente estesi

Tesauro



Vocabolario controllato nel quale sono presenti anche relazioni di varia natura, es. gerarchiche, associative (sinonimia, iperonimia, iponimia, olonimia, meronimia), ecc.



URI(s)

- http://id.loc.gov/authorities/subjects/sh2012000080 📮
- http://id.loc.gov/authorities/sh2012000080#concept

Variants

- Cyberprep fiction
- Cyberpunk novels
- Cyberpunk science fiction
- Cyberpunk stories
- Post-cyberpunk fiction
- Postcyberpunk fiction

Broader Terms

- Science fiction

Closely Matching Concepts from Other Schemes

Cyberpunk 2

kyberpunk 🗗

Littérature cyberpunk d

Sources

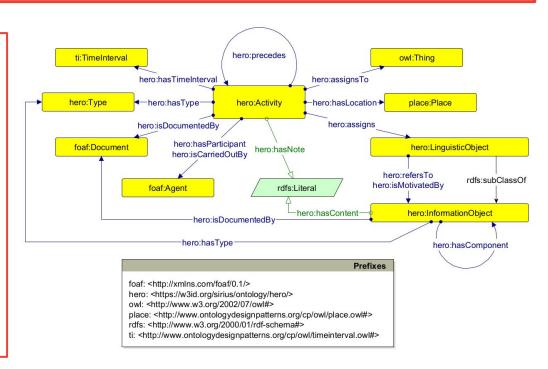
- found: Work cat.: 2009045559: Beyond cyberpunk: new critical perspectives, 2010:p. xi (Literary cyberpunk) p. xiii (cyberpunk, a subgenre [of science fiction]) p. 3 (cyberpunk SF) p. 96 (cyberpunk fiction) p. 195 (cyberpunk stories)

Ontologia



Modello di dati che descrive una particolare area di conoscenza definendo una terminologia comune per:

- entità
- proprietà (relazioni e attributi)
- vincoli logici e regole di inferenza





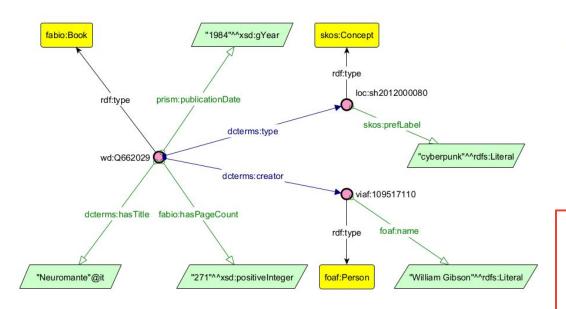
Riutilizziamo delle ontologie



- Libro → http://purl.org/spar/fabio/Book
- Genere → http://www.w3.org/2004/02/skos/core#Concept
- Persona → http://xmlns.com/foaf/0.1/Person
- è un/una → http://www.w3.org/1999/02/22-rdf-syntax-ns#type
- pubblicato nel → http://prismstandard.org/namespaces/basic/2.0/publicationDate
- ha genere → http://purl.org/dc/terms/type
- scritto da → http://purl.org/dc/terms/creator
- ha pagine → http://purl.org/spar/fabio/hasPageCount



Riutilizziamo delle ontologie



Prefixes http://purl.org/dc/terms/ dcterms: http://purl.org/spar/fabio/ fabio: http://xmlns.com/foaf/0.1/ foaf: http://id.loc.gov/authorities/subjects/ loc: http://prismstandard.org/namespaces/basic/2.0/ prism: http://www.w3.org/1999/02/22-rdf-syntax-ns# rdf: rdfs: http://www.w3.org/2000/01/rdf-schema# http://www.w3.org/2004/02/skos/core# skos: viaf: http://viaf.org/viaf/ https://www.wikidata.org/entity/ wd: http://www.w3.org/2001/XMLSchema# xsd:

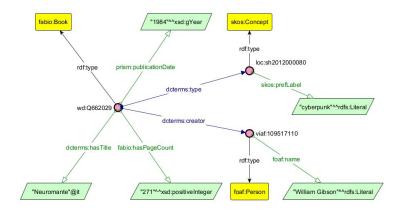
- dati (semi-)strutturati
- machine-readable
- semantici
- interoperabili
- riusabili

Riutilizziamo delle ontologie

```
wd:Q662029 rdf:type fabio:Book ;
    dcterms:creator viaf:109517110 ;
    dcterms:title "Neuromante"@it ;
    prism:publicationDate "1984"^^xsd:gYear ;
    fabio:hasPageCount "271"^^xsd:positiveInteger ;
    dcterms:type loc:sh2012000080 ;
    dcterms:creator viaf:109517110 .

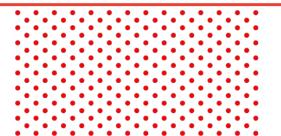
loc:sh2012000080 a skos:Concept ;
    skos:prefLabel "cyberpunk"^^rdfs:Literal .

viaf:109517110 a foaf:Person ;
    foaf:name "William Gibson"^^rdfs:Literal .
```

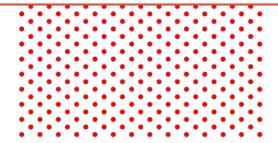


- dati (semi-)strutturati
- machine-readable
- semantici
- interoperabili
- riusabili

Quiz 5



https://forms.gle/zkQ Gysh12V5wG6wU9





6.1 Digital Humanities

Contesto Il progetto digitale Esempi

La tecnologia dell'informazione è strutturale



La tecnologia dell'informazione è sistemica, integrata nelle nostre strutture sociali, specialmente nella ricerca scientifica

•:•:•:•:•:•:•:•:•:•:•:•:•

Viene utilizzata per raccogliere, presentare, analizzare e trarre conclusioni dai dati

Inoltre, favorisce una crescente collaborazione tra professionisti specializzati in campi molto diversi



Questo vale anche per le Humanities



Disponibilità di risorse digitali e strumenti software per supportare:

- domande di ricerca esistenti
- metodologie esistenti
- nuove domande di ricerca
- nuove metodologie



Digital Humanities?



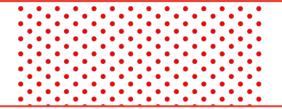
Un campo in rapida crescita che esplora il potenziale dei metodi e delle tecniche digitali nei progetti di ricerca umanistica

Coniato nei primi anni 2000, il termine descrive una molteplicità di realtà accademiche che operano all'interfaccia tra tecnologia e discipline umanistiche

Ancora difficile da definire



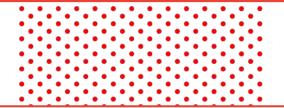
Una doppia definizione



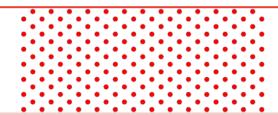
[...] a nexus of fields within which scholars **use computing technologies to investigate the kinds of questions that are traditional to the humanities**, or [...]
ask traditional kinds of humanities-oriented questions about computing technologies.



Una doppia definizione



[...] a nexus of fields within which scholars use computing technologies to investigate the kinds of questions that are traditional to the humanities, or [...] ask traditional kinds of humanities-oriented questions about computing technologies.



Educare a costruire

• • • • • • • • • • • • • • • • •

- digital: implementazione, creazione, programmazione = pratica
- humanities: interpretazione, speculazione, approcci comparativi = teoria

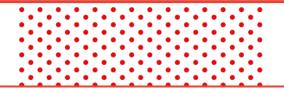
•••••

Compito di incoraggiare (cioè 'educare') nella 'costruzione' (processo intellettuale, sociale, ideologico) di qualcosa che genera conoscenza all'interno del dibattito accademico, pubblico e privato.

- Abilità pratiche
- Approcci interpretativi
- Valorizzazione, diffusione e utilizzo dei risultati della ricerca

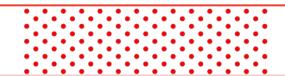


Progetti digitali

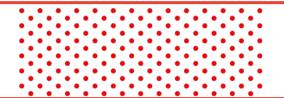


Un progetto che utilizza metodi e tecnologie digitali come parte integrante del processo di ricerca, della diffusione dei risultati della ricerca e dell'interazione con un vasto pubblico di utenti

Es. edizioni digitali, basi di dati bibliografiche, applicazioni di analisi dei dati, applicazioni GIS, ecc.



Progetti digitali



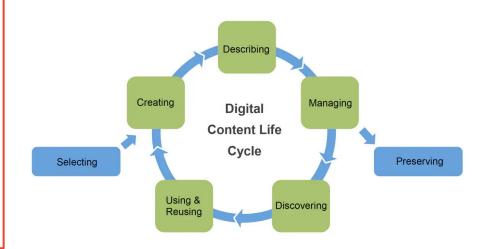
- Uso di metodi di analisi basati sui dati
- Gestione delle risorse digitali (ad esempio, un insieme di documenti o dati salvati in uno o più file in un formato specifico)
- Organizzazione strutturata, informativa e interpretabile delle risorse
- Pubblicazione digitale delle risorse

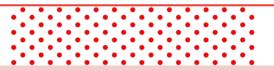


Ciclo di vita di un progetto digitale



Creare uno strumento scientificamente valido, affidabile e riutilizzabile, capace di affrontare domande di ricerca, metodologie e temi pertinenti alle discipline umanistiche (testo, interpretazione, incertezza, esperienza umana, ecc.)



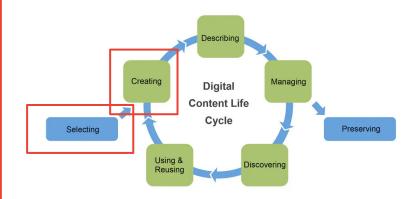


Produzione



- Selezione delle risorse
- Acquisizione digitale e/o trascrizione delle risorse
- Produzione/riuso di dataset

Esempio: raccolta e trascrizione di una serie di documenti per la digitalizzazione per un'edizione digitale

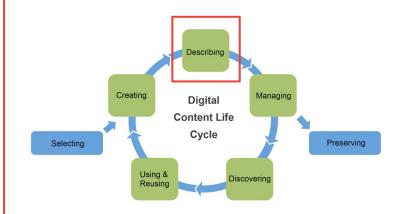




Descrizione

- Analisi del contenuto, contesto e utenti
- Selezione e/o progettazione di schemi di metadati, vocabolari controllati e altri standard

Esempio: definire lo schema di markup per la struttura del documento e le ontologie per i contenuti e le relazioni



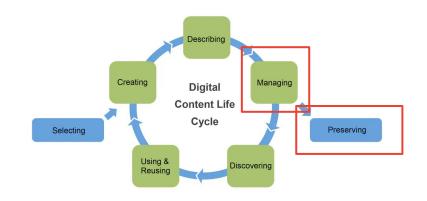


Gestione



- Riflessione sul servizio di archiviazione
- Controllo qualità
- Garantire la loro accessibilità a lungo termine

Esempio: scegliere il server per ospitare l'edizione digitale, pubblicare il dump del dataset su Zenodo

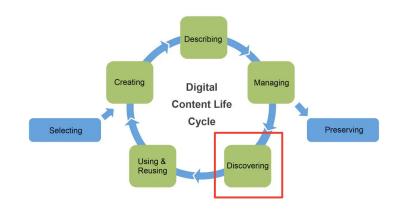




Scoperta



- Elaborazione delle risorse
- Analisi e interrogazione delle risorse
- Ragionamento a livello di infrastruttura per ospitare i risultati
 Esempio: conversione semi-automatica dei documenti in formati adatti alla pubblicazione semantica tramite un'applicazione e degli script

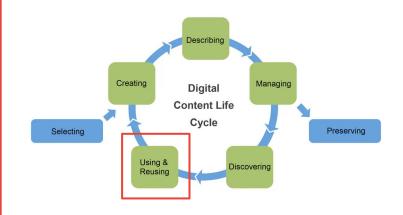




Uso e riuso

- Progettazione e sviluppo di interfacce
- Progettazione e sviluppo di servizi di query e download
- Sviluppo di politiche di condivisione dei dati e delle applicazioni

Esempio: implementazione di meccanismi di ricerca, visualizzazioni dei dati, download, ecc.



6.2 Tutto il resto

Abilità informatiche Progetti Ruoli

Abbiamo parlato di...

pensiero computazionale

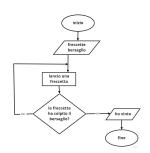
gestione dei dati

conoscenza di Internet e Web

.

digital literacy

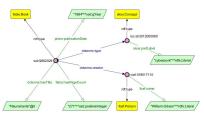
modellazione dei dati











Programmare



Scrivere algoritmi in un linguaggio comprensibile da un computer per svolgere determinate operazioni



- Analisi di dati
- Automatizzazione dei compiti ripetitivi
- Realizzazione di progetti
- ..





- Markup and web languages such as HTML, CSS and Javascript
- Asylicilionous requests and Ajax
- Specialized web editing software
- Image editing
- Accessibility
- · Cross-browser issues
- Search engine optimisation



Back End

- Programming and scripting such as Python, Ruby and/or Perl
- · Server architecture
- Database administration
- Scalability
- Security
- · Data transformation
- Backup

Progettare

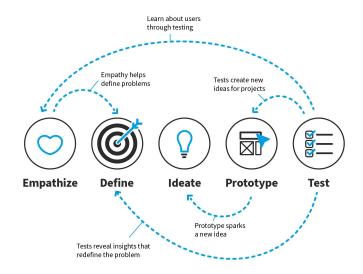


Ideare, pianificare e sviluppare una soluzione strutturata per un problema

- Confronto con gli utenti
- Definizione degli obiettivi
- Creatività
- Collaborazione
- ..

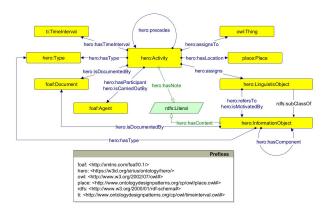
Design Thinking: a Non-Linear Process



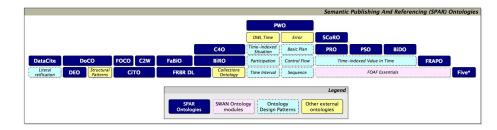


Interaction Design Foundation interaction-design.org

Modelli di dati

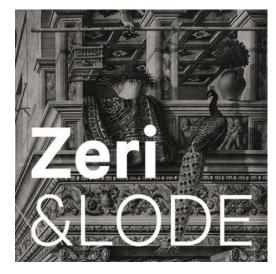


Heritage Risk Assessment Ontology (HeRO)

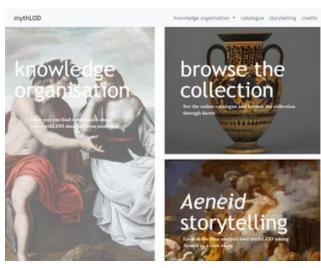


Semantic Publishing and Referencing Ontologies (SPAR)

Collezioni di dati



ZERI & LODE

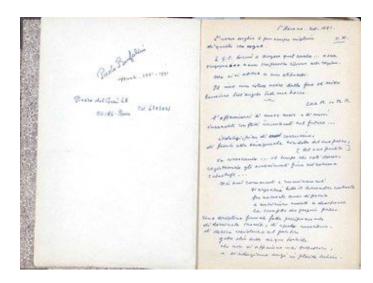


mythLOD

Edizioni digitali



Edizione nazionale delle opere di Aldo Moro



Semantic Digital Edition of Paolo Bufalini's notebook

Altro



DHDKey!

musoW



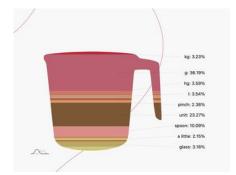




ODI e BACODI: uno studio su Il Castello dei Destini incrociati di Italo Calvino

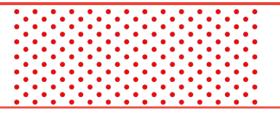
60% of participating a decir contribute des dentre accounté surrius de fiate Calonia BACKET A visa have if commence models in a fraction of the commence of the com ternal del special China sella tris dimensiona inclicata a caratta. In particulars, reports progents a progence of studions have managina to the curse on barracks in in ter-







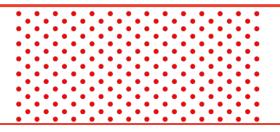
Sviluppatore/rice



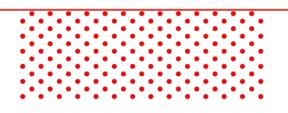
Crea applicazioni e strumenti software che supportano ricerche umanistiche, sviluppa piattaforme per la gestione dei dati e collabora con ricercatori per implementare soluzioni tecnologiche

- Siti web
- Applicazioni
- Script per automazioni
- Script di elaborazione dati
- Servizi Web
- Modelli di Machine Learning
- Sistemi di gestione dei contenuti
- ...

Specialista dei (meta)dati

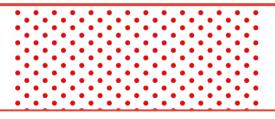


Gestisce, pubblica e contestualizza dati, sviluppa modelli di (meta)dati, sviluppa software per analizzare, produrre e riconfigurare dati



- Cataloghi di dati
- Modelli di dati
- Ontologie
- Strumenti di gestione dei dati
- Visualizzazioni
- Linee guida
- ...

UX Designer

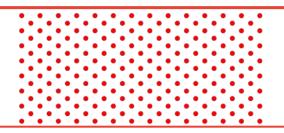


Progetta interfacce utente accessibili e intuitive per strumenti digitali, conduce test di usabilità, raccoglie feedback dagli utenti, crea wireframes e prototipi



- Manuali di stile
- Interviste
- Schizzi
- Wireframe
- Prototipi
- Modelli dei contenuti
- ...

Data steward

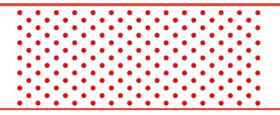


Gestisce i dati in un'ottica di governance: qualità, sviluppo di standard e politiche, documentazione, formazione e supporto

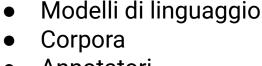


- Linee guida
- Modelli di (meta)dati
- Processi e metodologie FAIR
- Politiche di sicurezza
- Formazione
- Controlli
- ..

Linguista computazionale

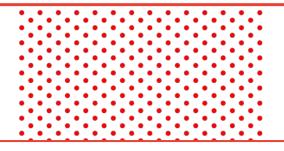


Implementa algoritmi e modelli per l'elaborazione automatica del linguaggio naturale, annota corpora testuali, implementa applicazioni per l'analisi del linguaggio

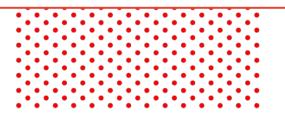


- Annotatori
- Sentiment analysis
- Estrazione delle informazioni
- Disambiguazione semantica
- Agenti conversazionali
- ..

Modellatore/rice 3D



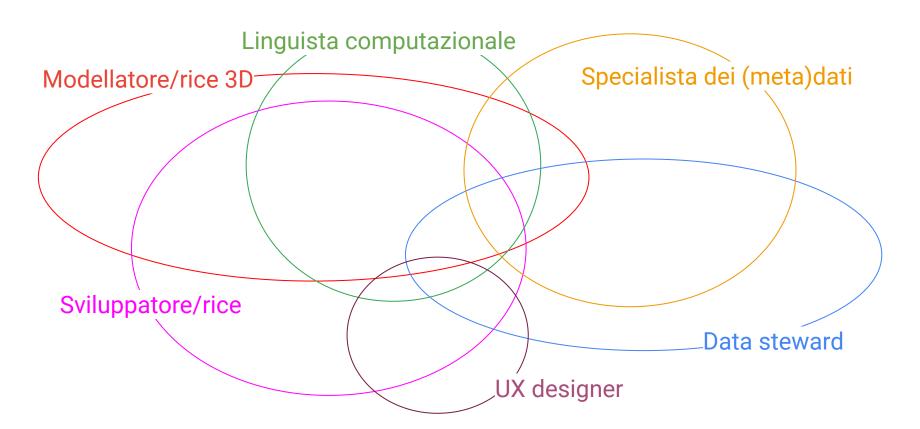
Crea modelli 3D, applica texture, prepara il rigging per l'animazione, sviluppa applicazioni VR e AR



- Modellazione
- Animazione
- Sviluppo VR e AR
- Digitalizzazione
- Acquisizione
- Ottimizzazione

• ..

Un po' quello che vi pare



Abilità informatiche

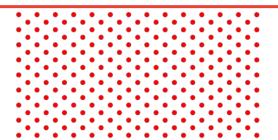
A.A. 2023/2024

06 - Fine

Sebastian Barzaghi

sebastian.barzaghi2@unibo.it
https://orcid.org/0000-0002-0799-1527

Valutazione del corso



https://forms.gle/nq8 Dmvk6hhrHLY876

