

Linee guida per la redazione della Tesi magistrale in Informatica Umanistica

Approvata dal Consiglio del 21/1/2016

Introduzione

Questo documento contiene alcune linee guida, di natura non regolamentare, per la redazione delle Tesi che costituiscono la prova finale per la Laurea Magistrale in Informatica Umanistica. Le linee guida sono rivolte (1) agli studenti che stanno per intraprendere la redazione della loro tesi (tesisti); (2) ai loro relatori, correlatori, controrelatori; (3) alle Commissioni di Laurea.

Si tratta di consigli, non di prescrizioni. In quanto tali, le linee guida possono essere giudiziosamente interpretate per adattarle ai diversi casi particolari. Ai tesisti in particolare, si raccomanda di mantenere un rapporto costante e continuo con i loro relatori, e di non esitare a chiedere consigli e chiarimenti anche su come redigere la Tesi — oltre che, ovviamente, sul contenuto specifico del lavoro di tesi.

Tipologie di tesi

Una buona tesi in **Informatica Umanistica** conterrà elementi specifici della disciplina. Se avete prodotto una ottima tesi tipicamente letteraria, oppure una ottima tesi strettamente informatica, probabilmente la vostra non è una buona tesi di Informatica Umanistica. Esiste una specificità della nostra disciplina, che consiste nell'affrontare i problemi sfruttando insieme le conoscenze e le capacità acquisite durante gli studi di Informatica Umanistica. Non si tratta di una dicotomia, quindi non ci saranno elementi “solo” informatici o “solo” umanistici giustapposti: ma piuttosto, una fusione armonica di principi, concetti, metodi di entrambe le anime che ci caratterizzano. Preparatevi quindi a rispondere alla domanda “Perché la sua tesi si può considerare una tesi di Informatica Umanistica e non di Informatica o Lettere?”

Un lavoro di tesi può assumere forme molto variegata, ma tipicamente rientrerà in una o più delle seguenti tipologie. Capire che tipo di tesi si sta svolgendo è importante anche per prepararsi ad affrontare criteri di valutazione diversi. Non tutte le tesi vengono valutate secondo le stesse metriche, e ci possono essere differenze sostanziali fra i criteri di valutazione applicati in casi diversi.

Naturalmente è possibile che una tesi incorpori elementi di più tipologie: questo è normale, e se ne tiene conto in fase di valutazione.

Tesi teorica

Vengono elaborati dei principi teorici relativi a problemi tipici dell'informatica umanistica (spesso in aree quali analisi di testi letterari, filologia, tecniche di comunicazione, interazione uomo-macchina, ecc.); la tesi deve caratterizzarsi per rigore metodologico e considerazione delle basi teoriche già stabilite (ovvero lo stato dell'arte relativo al tema oggetto di tesi); deve presentare elementi innovativi sul piano teorico, e motivarli adeguatamente. È utile anche avere qualche forma di validazione, che può anche consistere in esperimenti di pensiero, dimostrazione di casi di applicazione della teoria, oppure nell'accettazione delle idee proposte da parte della comunità scientifica, per esempio in seguito all'accettazione di un contributo derivato dalla tesi in consessi con peer review (es. comunicazione a una conferenza)

Tesi sperimentale

In questo caso, l'oggetto primario della tesi è l'identificazione di una *ipotesi*, la progettazione di esperimenti tendenti a dimostrare se l'ipotesi è fattualmente vera o falsa, la loro esecuzione, e l'analisi dei risultati. Si tratta di un tipo di tesi di struttura molto diretta; gli elementi qualificanti sono l'interesse dell'ipotesi, il fatto che la sua verità non sia già conosciuta in letteratura, il rigore metodologico nella progettazione e nell'esecuzione degli esperimenti, la profondità e completezza dell'analisi dei risultati. Un criterio dirimente è che alla fine della tesi, si sappia qualcosa che non fosse già noto all'inizio. Naturalmente, la formulazione dell'ipotesi non deve necessariamente essere di tipo booleano: domande quali "che rapporto c'è fra la velocità di lettura di un testo e la ritenzione in memoria da parte del lettore, in relazione alla dimensione del font e alla spaziatura fra le righe?" sono perfettamente valide.

Spesso questo tipo di tesi richiedono qualche competenza di statistica, e vedono l'applicazione di strumenti informatici a problemi di tipo umanistico. Più raramente si vede il contrario, ovvero l'applicazione di metodi delle scienze umane a problemi più tipicamente informatici, ma proprio per questo tali studi sono ancora più interessanti. Un esempio potrebbe essere: "C'è correlazione fra il numero di lingue conosciute da una persona e la velocità con cui impara a programmare un calcolatore?"

Tesi progettuale

Oggetto di questo tipo di lavoro è la costruzione di un artefatto, solitamente di natura informatica, ma rilevante per problemi umanistici. L'artefatto può essere una singola applicazione, un sito, un sistema complesso, o una raccolta di dati. Rientrano in questa categoria lo sviluppo di portali o l'annotazione semantica di corpora testuali.

In una tesi progettuale il primo elemento di valutazione sarà la qualità dell'artefatto realizzato, e l'adeguatezza degli strumenti e delle tecniche usate per realizzarlo, sulla base dei criteri standard usati dalla comunità scientifica di riferimento. A differenza di un tirocinio, in una tesi progettuale l'autore si assume piena responsabilità per tutte le scelte tecnologiche, e dovete quindi essere preparati a rispondere a domande come "Ma perché ha scelto un database relazionale anziché serializzare gli oggetti in JSON?"

In una tesi progettuale riveste particolare importanza la documentazione del lavoro svolto, in termini di progettazione e implementazione. Come in tutte le tesi, occorre poi dimostrare la maturità professionale dell'autore; è dunque utile che le vostre scelte (o "punti di decisione") siano documentati, che esponiate anche quali alternative avete considerato e perché alla fine avete scelto quella che effettivamente è utilizzata nell'artefatto realizzato. Anche in questo caso, potete fare riferimento a criteri standard utilizzati nella letteratura scientifica.

La complessità, novità e "intelligenza" dell'artefatto realizzato è una caratteristica importante per queste tesi. Realizzare un sito web statico, per quanto possa essere utile ai suoi utenti, non dimostra grandi capacità. Tenete presente che una tesi magistrale non è una relazione di tirocinio: non si tratta di fare un "lavoro su commissione", con lo scopo di avere un cliente soddisfatto, ma di dimostrare ai relatori e alla commissione di laurea le vostre capacità e la vostra maturità professionale.

Tesi di rassegna o compilativa

Una tesi di rassegna fornirà una rassegna (il più possibile completa e sistematica) su qualche argomento rilevante per la disciplina. Dovrete dimostrare la vostra conoscenza approfondita su tutta la letteratura scientifica esistente su un determinato argomento, riferendo sulle modalità che avete seguito nella ricerca di riferimenti (es.: partendo da un articolo "seminale" e tracciando poi tutti quelli che lo citano, oppure leggendo gli indici di tutti i numeri degli ultimi tre anni di una particolare rivista del settore, ecc.), sui criteri di selezione (quali criteri avete seguito per decidere se includere o meno un lavoro nella vostra rassegna), e sul tipo di analisi effettuata (avete letto solo gli abstract o tutto l'articolo? Cosa cercavate nell'articolo?).

Il valore di una rassegna risiede in parte nell'opera di ricerca e selezione del materiale¹, e soprattutto nella struttura, organizzazione, classificazione e analisi di quanto avete trovato. Dovete in particolare dimostrare la vostra capacità di analisi e di sintesi, con l'obiettivo di gettare nuova luce, tramite l'opportuna organizzazione del materiale e la riflessione critica sullo stesso, sull'argomento prescelto.

Struttura della tesi

La struttura della vostra tesi, ovvero la sua organizzazione in capitoli e sezioni, e la sua struttura retorica/argomentativa, dicono molto sulle vostra capacità di organizzazione ed esposizione.

Tipicamente una tesi avrà i seguenti elementi (che possono variare a seconda della particolare tipologia, e molti dei quali sono opzionali):

- Introduzione
- Inquadramento dell'area/sottoarea
- Enunciazione del problema affrontato

¹ Possono fare eccezioni alcuni casi in cui la ricerca in sé è particolarmente difficoltosa: per esempio, rintracciare nell'archivio di un monastero di clausura una lettera autografa di un letterato minore del seicento contenente un precursore del *Candide* può gettare nuova luce su tutto lo sviluppo della letteratura successiva; in questo caso la ricerca (o meglio, il ritrovamento) è un contributo importante.

- Enunciazione delle tecniche/metodi/strumenti che si utilizzeranno
- Rassegna dello stato dell'arte
- Corpo della tesi: il vostro contributo
- Analisi e riflessione critica sul contributo
- Conclusioni
- Riferimenti bibliografici
- Appendici

Gli elementi elencati sopra sono esplicitamente vaghi, e non corrispondono di necessità a titoli di capitoli. Riportiamo di seguito alcuni esempi di adattamento.

In una **tesi progettuale** in cui si costruisce un artefatto informatico, l'*enunciazione degli strumenti* potrà includere un elenco di linguaggi o librerie presi in considerazione, insieme a una loro breve descrizione (magari accompagnata da riferimenti bibliografici o più probabilmente a siti online), e alle considerazioni che vi hanno indotto a scegliere un particolare linguaggio piuttosto che un altro. Il *contributo* potrà includere un elenco di requisiti, una progettazione architeturale del sistema, e qualche dettaglio sull'implementazione; possono comparire algoritmi o parti di codice se di particolare interesse e non di uso comune. La *riflessione critica sul contributo* potrà consistere in una validazione del software realizzato, che può essere ottenuta verificando la correttezza, o misurandone le prestazioni, o svolgendo dei test con degli utenti-tipo e chiedendo loro di compilare un questionario di soddisfazione, eccetera. I *riferimenti bibliografici* potranno constare di riferimenti a pubblicazioni scientifiche sui particolari algoritmi e metodi usati nel software e a metodi di validazione del prodotto, ma tipicamente conterranno anche una serie di riferimenti a siti di distribuzione di librerie, componenti, o raccolte di dati usati negli esperimenti.

D'altra parte, in una **tesi di rassegna**, la *rassegna dello stato dell'arte* coincide con il *contributo*, e quindi non c'è un capitolo separato; il *contributo* contiene la rassegna, che sarà però assai più dettagliata che nel caso precedente; in questo caso la selezione e l'organizzazione del materiale è il vostro contributo originale. L'*analisi critica* dovrà discutere le diverse posizioni scientifiche esistenti sul tema oggetto di tesi, secondo quanto emerso dalla letteratura, i rispettivi limiti e punti di forza, i problemi ancora aperti e le possibili prospettive di sviluppo, ecc. È anche utile considerare il rischio che le vostre conclusioni siano viziate da errori sistematici: per esempio, è possibile che ci siano pubblicazioni importanti che non avete identificato? La vostra selezione del materiale potrebbe essere stata limitata per problemi di accessibilità o di lingua (es.: avete davvero consultato la letteratura in Arabo sul rapporto fra social media e Crociate?). Se sì, in che modo questo può alterare le vostre conclusioni? Le *conclusioni* saranno spesso un capitolo "importante", perché in questa sezione dovrete trarre qualche lezione da tutte le letture fatte in fase di rassegna, che vorrete condividere con il lettore della tesi.

In una **tesi teorica** l'*inquadramento* dovrà essere più corposo, e situare il vostro contributo teorico all'interno di un quadro coerente di teorie già esistenti. Per esempio, se la vostra tesi si concentra sull'uso comunicativo delle emoticon sui social media, in questa sezione potreste parlare di semiotica, di speech acts, ma anche di ergonomia e di interfacce utente (può essere semplicemente più veloce inserire uno smiley che scrivere una lunga frase). Occorre che il vostro *contributo* sia dunque inserito in una posizione chiara all'interno del quadro teorico presentato nell'*inquadramento*. La *rassegna dello stato dell'arte* conterrà

dunque un breve esame delle teorie già proposte relative al vostro particolare argomento, con l'indicazione del perché le riteniate insoddisfacenti, e dunque del perché riteniate che la vostra proposta (da descrivere poi a fondo nella parte relativa al *contributo*) sia da preferire a quanto già esistente. Una teoria di tipo letterario raramente ammette confutazione oggettiva, quindi probabilmente non avrete da condurre esperimenti a supporto: anzi, dovrete affidarvi alla forza della vostra argomentazione. Al contrario, una teoria di tipo informatico si presta spesso a essere verificata (o falsificata) con degli esperimenti, che in questo caso potrebbero essere documentati in un capitolo separato di validazione (un caso di *analisi critica*).

In ogni tipo di tesi è fondamentale che ci sia un chiaro *enunciato del problema*. Spesso questa parte è molto breve, e può essere inclusa anche all'interno di un capitolo insieme ad altri elementi. È tuttavia importante che la vostra enunciazione sia ben marcata, in maniera tale che il lettore della tesi sappia, entro poche pagine dall'inizio, cosa intendete dimostrare o quale problema intendete risolvere. Una tesi che non abbia un chiaro *enunciato* si presta a innumerevoli critiche: se il problema è vago, sostenere che avete la soluzione "giusta" è difficile. Di converso, se il problema (letteralmente: la *tesi* che intendete difendere davanti la Commissione) è chiaro, è facile valutare se l'avete risolto bene o meno.

Una tesi senza *problema* rischia di diventare un guazzabuglio di idee, fatti e opinioni o un flusso di coscienza. Una tesi con un *problema* chiaro si presta solo a due tipi di obiezioni: che il vostro problema sia irrilevante o già risolto (e a questa obiezione potete rispondere tramite la *rassegna*: se altri se ne sono occupati il problema non è irrilevante; se qualcuno avesse già risolto il problema voi ne avreste trovato traccia nella letteratura scientifica), o che il vostro *contributo* non sia in realtà una soluzione al *problema* (e a questa obiezione potete rispondere mostrando nell'*analisi critica* che risolvete davvero il problema). Rimane fuori il caso in cui voi siate i primi a individuare un particolare problema: in questo caso, la *rassegna* sarà praticamente vuota (ma dovete comunque mostrare di aver fatto delle ricerche estese, anche se con risultati insoddisfacenti), e allora cade su di voi l'obbligo di discutere del perché il nuovo problema è rilevante. In questi casi, la sezione sull'*enunciato del problema* può diventare anche piuttosto lunga.

Lunghezza e stile della tesi

Non esistono regole rigide sulla lunghezza di una buona tesi. Tuttavia, una tesi di 30 pagine farà alzare molti sopraccigli, così come una di 1000 pagine. La lunghezza raccomandata è fra 100 e 300 pagine, considerando che oltre al vostro contributo propriamente detto, saranno presenti molti elementi "di contorno". Non sentitevi comunque obbligati a rientrare in questi limiti: semplicemente, preparatevi a rispondere a obiezioni come "la tal parte è assente o poco sviluppata" (in caso di tesi inopportuna breve) o "il candidato manca di capacità di sintesi" (in caso di tesi inutile lunga). Sta al candidato decidere quale sia il giusto compromesso fra capacità di analisi e capacità di sintesi, e decidere di conseguenza cosa è necessario e sufficiente dire per difendere efficacemente il suo contributo. La congruità di questa scelta costituisce parte della valutazione della tesi stessa. L'esperienza dimostra che il candidato è talvolta portato a credere che scrivere di più sia sempre meglio: agli occhi di molti commissari ciò è falso. Se avete materiali addizionali, che possono servire

da riferimento, questi possono essere inclusi in un'appendice da consultare solo in caso di necessità. Inutile diluire il sapore del vostro contributo in un brodo troppo allungato. Nella maggior parte dei casi, il codice sorgente o i dati grezzi di una raccolta possono essere resi disponibili separatamente (per esempio, su supporto ottico o come allegati alla tesi elettronica) e non necessitano di essere riprodotti, a stampa, come parte della tesi stessa. Dall'altro lato, un'eccessiva sintesi può andare a discapito della chiarezza dell'esposizione o della completezza del trattamento del problema o della descrizione dello stato dell'arte. Un semplice elenco di riferimenti bibliografici non è una rassegna, che richiede invece la loro presentazione e discussione critica.

Se sentite crescere in voi la tentazione di ricorrere a trucchi "tipografici" per gonfiare la dimensione della tesi, resistete. Una tesi scritta in corpo 14 con interlinea tripla e margini da latifondo manda un chiaro messaggio a chi la legge: "io non ho niente da dire".

Un laureando in Informatica Umanistica non dovrebbe avere necessità di particolari consigli riguardo allo stile di scrittura². Ricordate soltanto che una Tesi di Laurea è un documento formale a carattere scientifico: idealmente, ogni giudizio va sostenuto con una argomentazione; i fatti vanno presentati pianamente, e le opinioni personali (che certamente possono essere espresse) vanno distinte dai dati di fatto, come verificati dall'esperienza, e dal consenso della comunità di studiosi di riferimento, come comprovato dall'analisi della letteratura. Al di là delle differenze individuali, lo stile può oscillare fra la facondia più appassionata e il più asciutto resoconto fattuale, ma deve comunque mantenere un carattere adeguato a un documento formale di carattere scientifico. Vanno in particolare evitate le espressioni iperboliche, i resoconti troppo personali, lo stile "caro diario", gli errori sintattici e ortografici. La correttezza formale della tesi (ortografia e grammatica, stile e rigore scientifico della scrittura, ecc.) è un parametro fondamentale per la valutazione della sua qualità.

Non sempre il correttore ortografico è vostro amico. Non c'è nessun sostituto per una buona rilettura "a freddo" della tesi prima della consegna, quindi riservatevi qualche giorno per questa operazione. La maggior parte dei relatori, correlatori, controrelatori e commissari di laurea preferirebbe **non** dover leggere in una tesi frasi come *"A quel punto, mi resi conto che il database sarebbe venuto enormemente pesante, e piombai nella disperazione più nera per l'esilio della mia impresa."*

² Molti utili consigli al riguardo sono contenuti nelle analoghe linee guida per l'elaborato finale della Laurea (triennale) in Informatica Umanistica: M. Tavosanis, Regole di scrittura per elaborati di laurea triennale, Corso di studi in Informatica umanistica, Marzo 2009. Disponibile online: http://www.fileli.unipi.it/infouma/files/2014/11/Regole_elaborati_laurea_triennale2009.pdf