

Cum se scrie o teză de doctorat în Informatică

Dan Cristea

Care este scopul unei teze de doctorat?

Să vă învețe să descrieți activitatea de cercetare depusă de dumneavoastră. Să vă învețe să căutați prin literatura de specialitate sau pe internet alte abordări pe aceeași temă. Să vă învețe să prezentați o problemă de la scop la soluție. Să vă învețe să vă exprimați în scris într-o manieră riguroasă. Un document scris este unul care rămâne și poate fi oricând consultat și criticat. De aceea reflectați de 10 ori și scrieți o singură dată.

Nu în ultimul rând, teza reprezintă și un certificat pentru școala doctorală care v-a format, precum că ați ajuns la maturitatea necesară pentru a începe o carieră științifică de cercetător. Este un document cu valoare de atestare, care se păstrează și care poate să probeze, în anii ce vin, că diploma de doctor care vi s-a eliberat este reală și nu măsluită.

În sfârșit, atunci când problema pe care o abordați în teză este atât de complicată încât nu aveți cum s-o epuizați, teza voastră poate fi folosită de alți studenți, în anii următori, ca punct de plecare în tentativa lor de a continua ceea ce ați început d-voastră. Și desigur, în astfel de cazuri, veți fi și citat.

Cînd încep să scriu la teză și în ce ritm lucrez la ea?

Începeți să lucrați la teză cu mult înainte de prezentare. Cel mai bine e ca în perioada în care citiți ce au făcut alții și încercați să sistematizați, în care acumulați ori triați date, sau în care lucrați la program, să faceți note. Acestea le puteți strânge într-un număr de dosare (ori directoare) egal cu numărul capitolelor lucrării, pe fiecare dosar înscriind titlul capitolului. În felul acesta, când vă hotărâți să treceți la operația de redactare a tezei, în locul unei operații de compunere, de creație, anevoioasă și supusă riscurilor, veți avea de făcut mai mult una de triere, de organizare și interpretare, de reformulare a materialelor adunate. Tot adunând note, concepția unui capitol se va forma de la sine. La un moment dat vă va fi astfel ușor să detaliați conținutul capitolului.

Făcând aceste note din timp, cu gândul la lucrare, din timp, vă ușurați scrierea pentru că totul acum vă este proaspăt în minte. Mai mult, în această fază se pot semnala greutăți, inconsistențe, cazuri ce nu pot fi luate încă în considerare din cauza unor limitări teoretice ale abordării pe care o dezvoltați, frânturi de demonstrație, chiar dacă acestea vi se par atunci evidente. Mai târziu, s-ar putea să rămâneți cu amintirea că sunt evidente dar să nu mai înțelegeți de loc de ce sunt ele atât de evidente¹. De asemenea puteți nota, într-un limbaj simplificat, algoritmi ce vi s-au părut interesanți în momentul în care i-ați construit.

¹ Mulțumesc colegei Anca Ignat pentru această observație.

Vă propun un experiment. Faceți un efort și scrieți, acum, mâine ori poimâine, o bucată dintr-un capitol, poate chiar un capitol întreg. Unul despre care aveți o imagine clară despre cum trebuie să arate, ce trebuie să conțină. Puneți-l apoi deoparte și nu vă mai atingeți de el o lună. Citiți-l apoi după o lună. Pariez că unele exprimări vă vor părea străine, ca și când n-ar fi fost scrise de dumneavoastră, unele străine în bine, altele în rău, și cu siguranță veți simți nevoia să formulați altfel unele pasaje, ca să nu spun nimic de conținut. De aceea un proces de scriere trebuie să fie unul de durată. O carte, ori o teză scrisă pe nerăsuflăte este de cele mai multe ori una slabă. Desigur, nu neg că puteți avea o inspirație divină care să vă facă să lucrați demențial și să dați gata într-un timp record lucrarea. Dar n-o publicați a doua zi. Lăsați-o la dospit un timp. Ocupați-vă în timpul acesta de altceva, eventual luați-vă o vacanță. Când veți reveni veți fi surprins să vedeți cât de mult veți dori să modificați la textul inițial. Cu siguranță că mintea noastră are o inerție care face ca anumite lucruri pe care le-am avut recent în minte să persiste un timp, implicarea lor în lanțuri de raționament să fie înțeleasă de la sine, fără o explicitare a lor. Recuperarea acestor scăpări poate fi făcută fie dând manuscrisul unui coleg, fie lăsându-l deoparte un timp. Alți ochi înseamnă o altă înțelegere...

Lucrul la o teză este un proces care seamănă uneori cu cel al scrierii unui roman. Sunt unii autori care reușesc să adauge câteva pagini noi pe zi romanului pe care îl au pe masa de lucru, dar adesea scrierea nu urmează o cale incrementală. **Scrierea unei teze este un proces neliniar și uneori chiar iterativ.** Cineva a spus că o carte, și deci și o teză, se scrie începând cu capitolul unu. Apoi se refacă acest capitol și se scrie capitolul 2. După ce capitolul 2 a fost scris, e posibil ca viziunea asupra capitolului 1 să se fi modificat oarecum. Ca urmare se refacă capitolul 1 și în consecință și capitolul 2. Apoi se scrie capitolul 3. Odată acesta scris, se refac capitolele 1, apoi 2, apoi 3, se scrie 4 ș.a.m.d.

Ce trebuie să conțină o lucrare?

O regulă adesea bună de urmat este regula IMRAD²:

- I - Introducere (*Introduction*)
- M - Metode (*Methods*)
- R - Rezultate (*Results*)
- A (*and*)
- D - Discuții (*Discussion*)

Recomandabil este și un cuprins de genul:

- Un cuvânt înainte, în care să formulați recunoștința față de persoanele pe care le considerați că v-au stat alături în realizarea cercetării doctorale, precum și orice alte gânduri adresate comunității științifice;
- Introducere, cu scop, motivație, context științific, anunțarea aportului propriu și o prezentare succintă a conținutului tezei;
- Starea artei, cu o conturare cât mai exactă a nivelului atins de cercetare în nișa în care se plasează cercetarea dv.;

² Robert A.Day: How to write and publish a scientific paper?, Cambridge University Press

- Metode, cu definiții și o descriere aprofundată a modelelor teoretice dezvoltate;
- Artefacte, cu o descriere a datelor acumulate pentru validarea experimentală a modelor, a corpusurilor adnotate create, a seturilor de date de antrenare etc.
- Experimente, cu indicarea realizărilor experimentale, a implementărilor menite să probeze modelele propuse;
- Evaluări, rezultate, discuții, cu prezentarea metricilor de evaluare utilizate, anunțarea rezultatelor obținute și compararea rezultatelor proprii cu vârfurile din *state-of-the-art*, care să dovedească succesul demersului științific;
- Concluzii, cu o reluare rezumativă a conținutului tezei, o subliniere a contribuțiilor originale și o enunțare a posibilelor dezvoltări de urmat în viitor;
- Lista bibliografică și publicații proprii în legătură cu subiectul tezei; dacă vreți să convingeți comisia asupra profilului dv, științific larg, respirației cu bătaie lungă, ați mai putea include aici, în secțiuni separate, și alte publicații care nu au o legătură directă cu teza, o listă a proiectelor în care ați participat și care au legătură cu subiectul tezei, precum și anexe.

Cum se scrie o introducere?

Introducerea este o parte foarte importantă a unei lucrări științifice și cu siguranța **una dintre părțile cele mai importante dintr-o lucrare de doctorat**. Pentru că ea trebuie să se constituie într-un rezumat al întregii lucrări. Sunt cititori care nu trec niciodată de introducere, restul fiind doar o răsfoială. În ideea că teza va fi publicată, puteți câștiga un cititor sau îl puteți alunga după modul în care este redactată introducerea. Aici trebuie să spuneți toate lucrurile importante. A păstra, în vederea obținerii unor efecte de surpriză, rezultatele spectaculoase pentru partea finală a lucrării este o greșeală. S-ar putea ca cititorul pe care ați vrut să-l cuceriți să nu ajungă niciodată la capitoul respectiv...

Cum se scrie capitolul despre *state-of-the-art*?

În acest capitol, care în unele cazuri poate fi completat cu informații și în următoarele, faceți o **trecere în revistă detaliată a câmpului științific ori tehnic pe care îl investigați** în lucrare. Pentru că o lucrare de doctorat urmărește și deprinderea candidatului la titlul de doctor cu "absorbirea" literaturii științifice, verifică capacitatea lui de a înțelege corect un domeniu, de a-i sesiza cât mai exact limitele la momentul actului cercetării.

În general, acest capitol trebuie să fie semnificativ ca întindere și aici nu trebuie să vă mărginiți să reproduceți în secvență o serie de lucrări, teorii, rezultate, ci trebuie să faceți o sinteză critică, semnalând evoluția domeniului ori partea care lipsește într-o teorie pentru a fi completată de alta. Adesea evoluția științei este la fel de palpitantă ca un roman polițist. Acest capitol, care este adesea răsfoit în diagonală de un cititor grăbit, poate deveni spectaculos dacă e scris mergând pe această idee. În acest caz veți adăuga încă o contribuție

originală la cele pe care teza trebuie obligator să le conțină (de exemplu, găsirea unor exemple noi și explicarea lor comparativă prin mai multe filtre interpretative).

Sunt excluse **tezele complet dedicate unei activități de sinteză**. Puteți scrie o astfel de lucrare, care e foarte utilă multor cercetători, dar nu ca o teză de doctorat.

În ce tip de teză se încadrează lucrarea mea?

Sunt **teze cu caracter pronunțat aplicativ** (în care primează o aplicație practică) și **teze teoretice** (în care primordială este o investigație de natură teoretică). **Ambele tipuri presupun aporturi originale.**

În funcție de tipul de teză, ponderea secțiunilor poate fi diferită. În ambele cazuri însă rezultate sunt foarte importante și teza trebuie să le pună în evidență, inclusiv în comparație cu ce s-a elaborat până acum la nivel mondial.

Ce scriu în capitolul de concluzii?

În orice caz trebuie revăzut pe scurt, în ordinea prezentării în lucrare, contribuțiile dv. originale. Apoi lucrarea trebuie plasată în contextul altor realizări și făcută o comparare: de ce este mai bună, ce aduce nou, ce completează la alte teorii, aplicații.

De obicei acest capitol se încheie cu o prefigurare a cercetărilor ulterioare celei începută de dv.

Ce mai conține o teză de doctorat?

Referințe bibliografice, numai acelea ce sunt referite în text, în ordinea alfabetică a autorilor, sau în ordinea referirilor în text. Luați un anumit format (*style*) și aplicați-l cu consecvență la toată lista. Nu ignorați numerele volumelor, numerele de pagină, coduri DOI sau ISBN/ISSN.

O secțiune de mulțumiri în care puteți sublinia aportul altor persoane la realizarea lucrării, de obicei plasată la început, într-un **Cuvânt înainte**.

Nu includeți în teză coduri sursă ale programelor realizate de dumneavoastră. Dar un gest frumos, care, în plus, vă poate propulsa în comunitatea științifică, este să le depuneți pe GitHub sau într-un portal al școlii doctorale, cu indicarea unei lucrări a dv. ce ar trebui citată în caz de folosire a codului. În teză însă, descrierea algoritmilor trebuie făcută succint, într-un limbaj pseudocod, fără detalii de implementare sau idiosincrazii ale limbajului sursă (cu excepția cazului în care propuneți un limbaj nou și în care e nevoie să insistați pe astfel de detalii). Părerea mea este că prea multe detalii supra tehnologiilor folosite în implementări, nu stau bine într-o teză de doctorat (ele pot fi date din abundență într-o teză de licență sau disertație). Tehnologiile se schimbă rapid, dar teza dv. ar trebui să-și păstreze prospețimea și peste mai mulți ani, când probabil tehnologiile folosite vor fi depășite. Puteți plasa însă astfel de detalii într-o anexă, ca să facilitați reproducerea experimentelor dv. de către persoanele

interesate. Dacă insistați să le includeți în teză, tot în anexă stau, mai bine decât în corpul ei, capturi ecran ale interfețelor realizate.

Cercetarea doctorală poate produce, de asemenea, multe alte artefacte științifice în afară de lucrări publicate: seturi de date statistice, date de antrenare a rețelelor neurale, corpusuri lingvistice adnotate, modele de limbă, aplicații ori servicii web etc. Acestea pot fi de un real ajutor altor cercetători, care, slujindu-se de ele, ar putea continua efortul vostru de cercetare. Dacă nu există restricții de drepturi de autor (ca, de exemplu, atunci când folosiți o resursă originală pentru care ați primit drept de utilizare în cercetare, dar nu și de expunere publică a ei, sau ca atunci când aveți un angajament față de o entitate care vă sponsorizează studiile ori cu care doriți să începeți o activitate lucrativă) atunci plasarea în domeniul public a acestor date reprezintă un gest de generozitate față de comunitatea științifică. Gândiți-vă doar la cât de mult a însemnat pentru dv. să găsiți în internet seturi de date de antrenare gata generate atunci când ați început cercetarea.

Care sunt fazele unei descoperiri?

Când ajungeți să întrezăriți cam pe **unde s-a ajuns cu cercetarea** într-un domeniu (sau „anvelopa domeniului”) înseamnă că ați depășit primul prag. Acest lucru trebuie să se producă în faza **Programului de Pregătire Universitară Avansată** (PPUA).

În mod normal, dacă sunteți pasionat de domeniu, după această fază urmează una în care **încep să vă vină idei**. Teoretic, înseamnă că ați intrat (cu un picior) în faza **Programului de Cercetare Științifică** (PCS). De cele mai multe ori ideile sunt inspirate de lecturi. Lucrurile se întâmplă cam așa. Citiți un articol cu un scop în minte. Pe măsură ce avansați cu lectura găsiți o slăbiciune a modelului prezentat de autor în directă legătură cu problema care vă frământă. Restul lecturii este de obicei una precipitată în care vreți să vedeți dacă autorul atinge această chestiune, sau ea scapă modelului. Dacă ajungeți la sfârșit și ea nu a apărut înseamnă că aveți deja o "teorie" proprie.

Următoarea fază e una de **consolidare a ideii**. Acum veți căuta cu înfrigurare exemple, cazuri, situații care să vă probeze teoria. O lectură o atrage pe următoarea. Pe măsură ce, citind, constatați că soluția dv. e validă deveniți tot mai îndrăzneț și mai... stăpân pe situație în același timp.

Acum **știți că aveți ceva de spus**. Aceasta este faza cea mai periculoasă. Este periculoasă pentru că, fără a avea o cultură științifică completă (sau măcar suficientă), sunteți foarte aproape de a combate alți autori, unii chiar consacrați. În forul dv. interior sunteți oarecum circumspect, recunoscând ciudățenia faptului că un autor atât de faimos ca Dr. Samuel W. Goodresearcher să facă o omisiune atât de mare (care în plus să nu fi fost recunoscută de nimeni până acum). Este faza în care faceți prima comunicare. Ca urmare vă duceți la conducător sau la un îndrumător și-i explicați ce ați descoperit. De obicei cam **aici se termină aventura** dv. pentru că îndrumătorul vă poate dărâma dintr-o frază întregul eșafodaj, punând în evidență un aspect care v-a scăpat până

atunci. În schimb, i-ați trezit interesul. Discuțiile sunt un prilej să vă afișați superioritatea tehnică de curând câștigată.

Din această experiență însă ați câștigat ceva. Ați învățat că un "clasic" este greu de dărâmat, că nu orice zboară se mănâncă și în orice caz ați înțeles mai bine teoria în cauză. Eșecul suferit nu trebuie însă să vă descurajeze. El e normal pentru orice tip înzestrat cu imaginație. Este absolut necesar și pozitiv să vă consolați cu gândul că așa ceva nu i se poate întâmpla niciodată cuiva lipsit de imaginație creatoare. Este primul **semnal că sunteți dotat pentru cercetare**. Un cercetător în embrion ca dv. cu greu poate urmări până la capăt un articol, și asta nu pentru că nu-l înțelege ci pentru că multe din ideile din articol îi sugerează noi posibilități de investigare, ori contra-exemple. Ca urmare este aproape întotdeauna mai dispus să-și urmeze propriile imbolduri sugerate de lectură decât să urmărească argumentația autorului până la capăt.

Dar faptul că sunteți pasionat pentru cercetare nu e suficient pentru a deveni cercetător. Partea cea mai grea abia începe, pentru că, de la Edison citire, **o descoperire e 1% inspirație și 99% transpirație**. Ca urmare, abordați alți autori, și munca abia începe...

Cât de mult trebuie să citesc despre ce au făcut alții?

Începeți cu ceea ce vă indică conducătorul și continuați să căutați mai departe singuri. **Lecturile se înlanțuie dacă urmăriți referințele bibliografice**. O imagine cât mai clară a limitelor unui domeniu este un lucru deosebit de important.

Desigur, cel mai frecvent sentiment la începutul unei cariere științifice este că nu veți ajunge niciodată să știți tot ceea ce s-a scris ori făcut în domeniul în care tocmai ați intrat. Acest sentiment este deosebit de frustrant și face viața amară multor cercetători tineri și cinstiți.

Îmi aduc aminte de o dificultate cu care m-am confruntat în primii ani din cariera de cercetător: influența pe care credeam că o am din partea altor autori. **Mi se părea că citind** ce au făcut alții, metodele lor mi se vor impregna în creier de manieră să-mi **toropească imaginația**, să mă aplatizeze, să nu mă lase să-mi folosesc propriile mele idei.

Nimic mai greșit decât acest sentiment. Imaginația ori o aveți ori n-o aveți. Probabil că ea se poate și cultiva, dar eu nu știu cum. Dacă o aveți, nimeni nu v-o poate dilua. Dimpotrivă cu cât citiți mai mult cu atât mintea dv. se va putea lega de mai multe idei, plecând de la care să vă aburcați mai sus. **Sfatul meu este să citiți cât mai mult**. Lucrul acesta vă dă și **competență și siguranța** că aprecierile dv. în domeniul ce vă interesează sunt cele corecte. Dar o competență adevărată este foarte greu de atins. Cei mai mulți tipi de oarece valoare se mărginesc să știe un pic mai mult decât ceilalți, moment în care încep să manifeste aroganță. Din această clipă ei sunt, din punct de vedere științific,... terminați.

Cât de sincer trebuie să fiu în teză?

O teză de doctorat în Informatică concentrează un efort foarte mare de cercetare. Orice cercetare poate să ducă la un succes sau la un eșec. Un eșec este apreciat în funcție de parametrii inițiali întrevăzuți. El trebuie recunoscut mai degrabă decât trecut cu vederea. **Ceea ce acum apare un eșec**, acel caz ce nu se supune teoriei pe care încercați să o formulați, excepția ce nu se supune regulii, **s-ar putea dovedi treapta inițială a unei noi ramuri a teoriei**, pe care dv. sau altcineva să îndrăznească ulterior o nouă construcție teoretică. Nu trebuie trecută cu vederea și posibilitatea ca un eșec să contureze o limită a teoriei pe care o experimentați, lucru foarte important.

Citați întotdeauna ceea ce au făcut alții. Fiți atenți cum introduceți în lucrare contribuția altora și cum vă referiți la a dv. Locul prezentării altor teorii, mai întâi pentru precizarea nivelului la care se situează domeniul pe care-l abordați și apoi pentru compararea a ceea ce faceți dv. cu ceea ce s-a mai făcut, este de obicei în introducere, în capitolul dedicat *state-of-the-art* și în concluzii. Nu este interzis să vorbiți despre ce au făcut alții în mijlocul lucrării, dar acolo riscul de a produce confuzii este mai mare. Începe să devină dificilă departajarea între contribuția dv. și cea străină. Ori de câte ori vă referiți la munca dv. spuneți acest lucru clar, folosind expresii de genul "*we have proved*", "*the method proposed by us*" etc.. Multă lume preferă pluralul când se referă la munca proprie, dar asta nu e o regulă. Este la fel de periculos ca cititorul să creadă că vă atribuiți un rezultat care nu este al dv. și ca el să creadă că un rezultat care este al dv. aparține de fapt altcuiva. În orice caz citați orice teoremă, rezultat, algoritm, arhitectură de sistem etc. pe care o preluați din altă parte. Lucrul cel mai grav ce vi se poate întâmpla la aprecierea tezei este să fiți învinuit că "ați tras cu ochiul" în cutare sau cutare carte ori lucrare și v-ați atribuit rezultate străine, un fapt aproape la fel de îngrozitor ca un plagiat...

În ce (fel de) limbă trebuie scrisă lucrarea?

Există o cutumă nescrisă ca tezele de doctorat în Informatică din România să se redacteze în limba engleză. Acest lucru nu reprezintă manifestarea vreunui orgoliu sau a unui sentiment de infatuare puerilă, ci înseamnă pur și simplu o ieșire în lume a candidatului și viitorului doctor, certifică putința de a afișa public, pe mapamond, rezultatele cercetării sale și deschide poarta către obținerea de citări, atât de dorite în configurarea unei cariere academice.

Când scrieți, în primul și în primul rând, **fiți atenți la gramatică**. Atenție la greșeli de ortografie și la exprimări. Maniera în care scrieți este oglinda cea mai elocventă a culturii dv.

Când redactați rezumatul în limba română, căutați cele mai adecvate traduceri pentru termenii tehnici ori sintagmele domeniului. Nu utilizați un termen sau o sintagmă preluată din engleză decât dacă vi se pare că respectiva traducere ar putea induce o confuzie în cititor, deci traduceți atât cât se poate, dar fără exagerări.

Cum trebuie să arate redactată o teză?

Fiți oricât de meticuloși cu forma lucrării dv. Grijă poate merge aici până la pedanterie. Cea mai bună formă a unei lucrări este dată de o prelucrare electronică profesionistă (o formă mai bună decât cea obținută sub Word, desigur, o oferă limbajul Latex).

Nu vă jenați să faceți apel la un corector ortografic. Dacă nu-l aveți deja în calculator, instalați-vă unul. Vă va semnala cu promptitudine multe greșeli făcute din neatenție.

O virgulă , un punct ,un semn de întrebare sau exclamare îl lipiți de cuvântul de dinainte ,iar spațiul îl lăsați după el , nu invers ori cu spații și de o parte și de alta . Fraza precedentă arată rău din acest punct de vedere. Și mai puteți avea și surpriza unui punct care a sărit, ca prim caracter, în partea stângă a unui rând nou.

Desenele și diagramele trebuie și ele realizate cu instrumente grafice adecvate. Dacă nu dispuneți de programe profesionale precum Visio, ori CorelDraw, panoplile de instrumente grafice ale Word-ului, ori Excel-ului, precum și cele ce pot fi realizate cu Google Drawing sau în PowerPoint, sunt perfect satisfăcătoare. Dacă trebuie pus un accent special pe o realizare practică cu interfață utilizator nu exagerați totuși cu capturile de ecran. Prea multe încep să dea de bănuț că ați încercat să ajungeți mai repede la un anumit număr de pagini pe care vi l-ați propus.

Cum se prezintă public o teză de doctorat?

Timpul pe care îl aveți la dispoziție este de 30-45 minute. În acest timp va trebui să concentrați toată activitatea dv. de câțiva ani! Ca urmare trebuie să fiți sintetic, urmând în general în expunere conținutul tezei. Punctați ceea ce considerați esențial și dați o atenție deosebită părților originale din lucrare ori muncii ce a însoțit elaborarea lucrării. Pentru prezentare puteți utiliza în anumite momente și tabla, ceea ce, în ziua de azi când toată lumea face apel la PowerPoint, poate crea un divertisment binevenit. Dacă vreți neapărat să expuneți și o aplicație practică, intrați direct în inima ei, fără a plictisi cu meniuri de logare, fiți foarte scurt și la obiect, pentru că o demonstrație vă poate aduce mai multă animozitate decât simpatie în public. În acest caz, organizați-vă demonstrația dinainte, pregătind exemple potrivite și spectaculoase, care să pună în evidență părțile tari ale programului pe care l-ați realizat și forța demersului teoretic care stă la baza aplicației. Ceea ce nu merge (încă) în program nu trebuie ascuns ci recunoscut cu dezinvoltură. Nimeni nu o să vă condamne dacă aplicația dv. nu e „la cheie”.

Succes!