

**A Guide to the ALM Thesis
In Information Technology**

Third Edition

September 2006

Master of Liberal Arts Program

Harvard University

Extension School

Preface

Students seeking the degree of Master of Liberal Arts in Information Technology (ALM in IT) at the Harvard University Extension School, whose concentration is either Software Engineering or Mathematics and Computation, are required to complete a thesis project. The master's thesis can be both a demanding and a novel undertaking. This guide has been prepared to offer practical help and encouragement, and attempts to anticipate the problems candidates are likely to face. We hope it will enable you to avoid the most common pitfalls and make thesis writing the rewarding experience that it should be.

Most of the relevant procedures and guidelines for the ALM in IT thesis are described in the pages that follow. The guide includes ideas for identifying a suitable thesis problem, and guidelines for writing a thesis proposal, implementing the software project defined in the proposal, and formatting both the thesis proposal and the final thesis. All ALM in IT candidates are expected to be familiar with the contents of this guide before undertaking the thesis project. In particular, they should read it carefully before consulting with the research advisor for IT.

We would like to acknowledge the work of those who prepared the general guide for ALM candidates in the Extension School before Information Technology became a field of study for the ALM, as well as the students, faculty, and staff who have reviewed this guide.

Bill Robinson, Research Advisor

Henry Leitner, Program Director

September 2006

Table of Contents

Preface.....	iii
Table of Contents.....	iv
1 Introduction and Administrative Procedures	6
1.1 General Timeline for Completing a Thesis In Information Technology	6
1.1.1 Milestones for a Thesis Project.....	7
1.2 Role of the Research Advisor	8
1.3 Role of the Thesis Director	9
1.4 Administrative Procedures.....	10
1.4.1 The Thesis Course.....	10
1.4.2 Finding a Thesis Director.....	12
1.4.3 Submission of the Thesis	12
1.4.4 The Thesis Grade	12
1.4.5 Graduation.....	13
1.4.6 Illustration of Major Deadlines.....	13
1.5 Deadlines and Extensions	14
1.5.1 Five-year Enrollment Period.....	14
1.5.2 Thesis Course Deadlines.....	14
1.6 Resources	15
1.6.1 Computer Access	15
1.6.2 Library Privileges.....	15
1.6.3 Financial Resources	15
1.6.4 Document Templates	15
1.7 About the Rest of This Guide	16
2 The Thesis Process.....	17
2.1 Differences Between Course Work and the Thesis Project.....	17
2.2 Getting Started – Developing a List of Candidate Problems.....	18
2.2.1 Problem Statements	19
2.3 General Approaches to Academic Inquiry.....	20
2.3.1 Professional Organizations	21
2.3.2 Trade Journals.....	21
2.3.3 Online Searches	21
2.4 Recapitulation: The Basic Steps in Beginning a Thesis Project.....	22
3 The Thesis Proposal.....	24
3.1 At the Beginning.....	24
3.1.1 Introductory Reading	25
3.2 Identifying a Problem	25
3.3 How do You Choose a Thesis Topic?.....	26
3.3.1 Thesis Topics Related to Your Work.....	27
3.3.2 How High Should You Aim?.....	27
3.3.3 Human Subjects	28

3.4	Refining the Project Scope.....	28
3.5	Following Software Engineering Practice	29
3.6	Content of the Proposal.....	30
3.6.1	Tentative Thesis Title	30
3.6.2	Abstract.....	31
3.6.3	Thesis Project Description	31
3.6.4	Work Plan	32
3.6.5	Additional Sections.....	33
3.6.6	Glossary	33
3.6.7	Bibliography	33
3.7	Submitting Your Proposal.....	34
3.7.1	Revisions and Writing Style	34
4	The Thesis Project.....	36
4.1	Working on Your Project.....	36
4.2	Thesis Content and Organization.....	37
4.2.1	Organization.....	37
4.3	Mechanics of Writing	38
4.3.1	APA Publication Manual	39
4.3.2	Guidelines on Plagiarism	39
4.3.3	Being Clear about the Source	39
4.4	Preparation of the Manuscript.....	40
4.4.1	Front Matter	40
4.4.2	Text of the Thesis.....	45
4.4.3	End Matter	46
4.4.4	Formatting Source Code	46
4.5	Format Review and Format Approval	47
4.6	Binding the Thesis and Final Approval of the Bound Copy.....	48
Appendix 1 General Formatting Guidelines		50
Writing Style.....		50
Use and Format of Chapter Headings and Subheadings.....		51
Headings and Transitions.....		51
Formatting Headings.....		52
Examples of subsection headings		54
Use of Tables and Figures.....		54
Typing and Format.....		55
Margins and Pagination		55
Thesis Paper		55
Appendix 2 Guidelines for References and Citations.....		57
Bibliographies and References		57
Citations		58
Appendix 3 Common Problems In Grammar, Punctuation, And Usage		60
Terms Related to IT		60
Hyphenation.....		60
Trademarked Names		61
General Errors in English.....		61
Appendix 4 Bibliography.....		65

1 Introduction and Administrative Procedures

The Master of Liberal Arts in Information Technology (ALM in IT) is designed for individuals who already possess solid programming experience plus a solid understanding of discrete mathematics, and who wish to advance both their software engineering skills and their understanding of certain theoretical aspects of computer science, such as the mathematical analysis of algorithms.

The thesis project is the focal point of the master's program for students in the Software Engineering and Mathematics and Computation concentrations. This project provides an opportunity for significant work in an area of interest to the candidate, and demands a synthesis of the skills and knowledge acquired throughout the candidate's program of study. The final thesis document, along with working code, is the evidence that the project has been accomplished and communicated to others in a conventional format. In producing this document, the student is demonstrating mastery of general principles in Information Technology and the ability to apply these principles to a significant software problem. As a consequence, you should aim high with your thesis. At the very least, you should develop a significant piece of software and gain valuable insights about the use of an interesting technology or an approach that would be of interest to others. It is within your reach to write such a good thesis that you, with the collaboration of your thesis director, could produce a document of publishable quality.

It is important to begin to think about the project well before commencing the writing of a thesis proposal. In this early stage, you can eliminate unsuitable areas of research, articulate interesting questions and review relevant literature. With such advance preparation, the thesis will provide an even greater opportunity to achieve an integration of the work within the master's program. Prior theses are on reserve at the Extension School's Grossman Library (third floor, Sever Hall) and may be examined as examples of thesis results. We will have one or two completed theses for viewing on our website, <http://www.extension.harvard.edu/almit>.

1.1 General Timeline for Completing a Thesis In Information Technology

There are four major stages in the completion of a thesis in Information Technology.

1. Identifying a suitable thesis problem
2. Developing a thesis proposal
3. Implementing the thesis software project

4. Writing the thesis document

The thesis document, which is a detailed written statement of the results of the project, must include full citations, documentation, references, project code, and other scholarly devices as set forth in this guide, and it must be prepared in a suitable writing style. The code from the project and the descriptive text prepared by the student together constitute the master's thesis in Information Technology.

The length of the thesis depends on the nature of its topic. Since projects in Information Technology may take many forms, the nature of the project and views of the thesis director are the best guides to the length of a thesis, although the prose part is generally somewhere between 50 and 100 pages.

1.1.1 Milestones for a Thesis Project

In order to progress through the thesis proposal and thesis writing stages and graduate at a particular time, there are certain milestones you must meet. The dates for these milestones are listed in the table below for each of the three graduation dates, November, and March, and June.

Milestone	November Graduation	March Graduation	June Graduation
Final thesis proposal submitted	February 1	June 1	September 1
Thesis proposal approved	March 1	July 1	October 1
Registration for the thesis course	Within two weeks of notification of thesis proposal approval	Within two weeks of notification of thesis proposal approval	Within two weeks of notification of thesis proposal approval
Thesis Writer's Workshop (or meeting with thesis advisor)	Mid-July (dates TBD)	Mid- November (dates TBD)	Mid- February (dates TBD)
Project completed and thesis document submitted to thesis director and research advisor for final evaluation	September 1	Week of January 1	April 1
Final draft of thesis document approved	October 1	February 1	May 1
Project grade submitted	October 1	February 1	May 1
Bound copy of approved thesis document submitted	October 15	February 15	May 15

Table 1 Milestones for Completing a Thesis

N.B.: The thesis director is chosen only in the latter stages of your development of the thesis proposal, and it might take a month or more to identify the thesis director. Therefore compliance with the date for submitting a final draft of your proposal is mandatory.

Note also that it is possible for you to do your thesis project before completing all your other course work, in which case your thesis milestones will not be driven by your graduation date, and so will be different from the milestone dates in the table above. The second scenario in section 1.4.1 describes such a case.

This list of milestones does not specify when you should begin your preliminary work of identifying a thesis problem and writing a thesis proposal. You should allow at least two to three months for this part alone, although several students have spent more than a year developing an idea for a thesis project. All degree candidates must complete their thesis project and submit their thesis document within nine months of the date on which they register for the thesis course. Note that about six weeks before a candidate submits the thesis document, he/she should attend a thesis writer's workshop (or arrange a meeting with the research advisor if they are distance students or cannot attend the workshop for other reasons).

Suggested steps in identifying a suitable thesis problem are described in the chapter entitled *The Thesis Process* and the process of developing an acceptable thesis proposal is described in the chapter entitled *The Thesis Proposal*. The administrative procedures, including the milestones above, are described below in the subsection *Administrative Procedures*.

There are two key resources to help you get started and to work on your project. These are the research advisor and the thesis director. Their roles are described in the following two sections of this chapter.

1.2 Role of the Research Advisor

The research advisor assists ALM in IT candidates in four areas:

- Identification of the thesis topic
- Preparation and presentation of the thesis proposal, which is a prerequisite to writing the thesis
- Identification of a faculty member whose research and teaching interests coincide with proposals on specific topics, and who may be recruited to serve as thesis director
- Review of the final thesis document

The research advisor conducts a workshop at the beginning of the fall and spring

semesters. One of the main purposes is to outline the administrative details and strategies appropriate to the ALM in IT thesis. All ALM in IT candidates who expect to work on a thesis project during the subsequent academic year are urged to attend the meeting. In addition, the research advisor holds a workshop for thesis writers, open to all candidates whose thesis project is nearing completion.

The research advisor, in collaboration with the program director, will arrange for consultation with a member of the Faculty of Arts and Sciences who is eligible to serve as the thesis director for each student. As soon as a faculty member accepts the role of thesis director, he or she then becomes responsible for guiding the thesis. Students should be aware that, in *extremely rare instances*, the research advisor might not be able to find an eligible, available thesis director – despite the acceptability of the project. Factors such as sabbaticals, leaves of absence, changes in appointment, or too many concurrent responsibilities can sometimes affect a potential director's availability. In such cases, it may be necessary for the candidate to develop another thesis proposal on a different topic in order to obtain a thesis director.

1.3 Role of the Thesis Director

The thesis director should be an expert in your thesis topic area who is eligible to serve as an ALM in IT thesis director. The fact that you had a certain instructor in an Extension School course does not mean that this person is eligible to direct an ALM in IT thesis.

The thesis director may request modifications in the topic or other changes in the thesis problem or approach; he or she may be especially helpful in redefining the project, providing focus, and obtaining certain resources. But the responsibilities for acquiring basic technical skills and identifying an initial thesis problem lie primarily with the student.

The thesis director meets with the student at appropriate intervals during the implementation and documentation phases of the thesis project to assist with problems as they arise and to ensure adequate progress. These meetings are likely to be most frequent at the early stages of the project. In general, you should not expect your thesis director to teach you software you're not familiar with or to help you solve routine software problems.

Upon completion of the project, the student demonstrates the working code developed during the project, and submits a written thesis to the thesis director. Some thesis directors prefer to review this work in successive stages, chapter by chapter, while others prefer to wait until the entire manuscript has been completed. In either case, it is the thesis director's role to read the thesis, to request revisions as necessary, and to approve the final version, submitting to the Extension School a letter grade and brief narrative evaluation of the project.

1.4 Administrative Procedures

The thesis is the most demanding phase of the ALM in IT program and almost always takes longer than expected. Requiring close familiarity with the content and methods of Information Technology, it cannot be commenced until at least 24 units – the equivalent of six courses – have been earned. Most ALM in IT students have chosen to do their thesis project after all other course work has been completed.

1.4.1 The Thesis Course

There are three administrative steps in registering for the thesis course. First, the candidate consults with the research advisor, who assists the student in gaining familiarity with the relevant literature, formulating a suitable project and writing a thesis proposal. The thesis proposal usually must be revised several times before the research advisor approves it. Allow up to one month for each draft you submit.

Second, when the thesis proposal is nearing completion, the research advisor and the program director identify an appropriate person to supervise the project as thesis director. See Section 1.4.2 for more details on finding a thesis director. At this point, the prospective thesis director may request some modifications of the proposed project.

Third, after a candidate has had a satisfactory meeting with the thesis director, the ALM in IT office sends the candidate a proposal approval form. When candidate and the thesis director have signed it, the candidate submits it to the ALM in IT office for final approval by the Dean of Continuing Education. After final approval, the candidate receives notification by mail of this approval and a request for tuition payment for the thesis course. If a student fails to pay for the thesis course within two weeks of the date they receive notification of the Dean's final approval of their proposal, the research advisor will be notified and will not be allowed to give formal approval to the completed thesis until the tuition is paid – in this circumstance, the student is not allowed to graduate.

These arrangements are made independently of normal course registration periods. Thus the thesis need not begin at the start of a term. However, if your desire is to graduate at a particular time, you should be aware of the various milestones described in Table 1.

Thus, a side effect of your registration is that you should know which set of milestones you are working towards. There are four dates that, in theory, can be different:

- The target thesis submission date – This is the date by which you must submit a first draft of your thesis document to the thesis director and research advisor. In most cases, this is a date approximately two months before your target graduation date, as specified in Table 1.

- The target thesis completion date – This is the date by which you must submit your final draft of the thesis to the thesis director for grading purposes and to the research advisor for final format review. In most cases, this is a date approximately one month before your target graduation date, as specified in Table 1.
- The target graduation date – This is the month and year in which you expect to graduate. You specify this on the proposal approval form.
- The end-date for the thesis course – This is the date specified in the thesis acceptance letter from the Dean of the Division of Continuing Education, and is nominally nine months after the proposal approval form is sent to the student.

The target submission date should be earlier than the end-date for the thesis project. The target graduation date can be before or after the end-date for the thesis project. Here are two scenarios that illustrate the dates involved. The first scenario is more typical than the second.

Scenario 1. Ms. Z is completing her ninth course in the spring of 2004. She wants to work on her thesis proposal over the summer and start her thesis project in the fall. Her goal is to graduate in June 2005. She completes the proposal in early September, is assigned a thesis director and submits the proposal approval form on October 2, 2004. The end-date for her thesis course is July 2, 2004. Her target thesis submission date is two months before her target graduation date of June 2005. To meet the June 2005 target she must complete her project, her first thesis draft, and the final thesis draft according to the milestones in Table 1. Note that Ms. Z could have chosen March 2005 as her target graduation date, although she would have had less time to complete the work.

Scenario 2. Mr. X is completing his eighth course in the spring of 2004. He wants to start his thesis project in June of 2004 with the goal of graduating in June of 2005. He plans to take his last course in the spring of 2005, and wants to finish the thesis in January 2005 so the thesis work and the course do not overlap. Mr. X completes his proposal in April 2004, is assigned a thesis director, and submits the proposal approval form on July 1, 2004. He registers for the thesis course shortly thereafter. The end-date for his thesis course is set at April 1, 2005, his target thesis submission date is February 2005, and his target graduation date is June 2005. Since his thesis submission steps are not tied to a specific graduation date, he should meet the following milestones:

- a) Complete his project and his first thesis draft by February 1
- b) Complete his final thesis draft by March 1

He does not have the option of graduating in March 2005 because he will not have completed all his courses by then.

1.4.2 Finding a Thesis Director

When a candidate has prepared a reasonable draft of the proposal, the program director and the research advisor will work together to recruit a suitable thesis director. Under no circumstances should a candidate approach a prospective thesis director without the prior approval of the ALM in IT research advisor. Premature discussions of a proposed thesis project with a faculty member may inadvertently jeopardize the candidate's opportunity to work with that person later on. Locating a suitable director is a process that may go quickly, but most often it takes at least four to six weeks. You may begin work on the high-level design or implementation of your software project. It is frequently desirable to do so in order not to lose the momentum you have gained while researching and crafting the proposal. Students should not complete the entire thesis project before a director is found. Most faculty members expect to assist the student in mapping out the design of the overall work and might well want some change in direction to improve the quality of the work.

1.4.3 Submission of the Thesis

The first draft of the thesis is to be submitted to both the thesis director and the research advisor about two months before the target thesis completion date. This draft should be as complete as possible. There are almost always many comments and corrections from the thesis director and the research advisor to improve both the content and structure of the document. Be prepared to revise and resubmit a new version as quickly as possible. In some cases, several drafts have been required to bring the document up to an acceptable level.

The thesis director evaluates the final draft and forwards a grade and a narrative evaluation of the entire project process to the ALM in IT Program Office. The research advisor also does a format review. For this purpose, the research advisor checks the thesis against the specific requirements stated in this guide and notes any discrepancies. After the student has incorporated all of the requested changes and approvals by the thesis director and research advisor have been obtained, the final, master copy of the thesis is printed, bound, and delivered to the ALM in IT Program Office. The research advisor inspects the bound copy one last time before the thesis is placed on the shelves of Grossman Library. Only after the thesis passes this final inspection will a candidate be allowed to graduate.

1.4.4 The Thesis Grade

The thesis director will assign a grade to your thesis project based on the criteria for grades stated in the Extension School catalog. As part of your discussion of the thesis project with your thesis director, you should understand how the thesis director will apply the criteria to grading your final project and thesis document.

A thesis awarded a grade of B-minus or below by the thesis director cannot be accepted by the Extension School as fulfilling the ALM in IT thesis requirement. Ordinarily, the candidate will be permitted to register for the thesis again, preparing another or a modified proposal, submitting another tuition, and, in conjunction with the

research advisor, obtaining a thesis director again, who might or might not have served on the earlier thesis.

A thesis awarded a grade of INC (incomplete) at the end of the allotted period also is unacceptable for the ALM in IT requirement. In this case, the student may register again without preparing a new proposal and the previous thesis director may continue as director, but the student must pay tuition for reregistering.

1.4.5 Graduation

As summarized in Table 1, Extension School degrees are awarded in November, March, and June. The major graduation activities take place in June. Candidates who earn the degree in November or March may participate in the ceremonies held in the following June. In planning a date for graduation, the student should expect to complete the thesis document – including the revisions suggested by the research advisor and the thesis director, final preparation of the manuscript, and binding of the finished copy – according to the milestone date shown in Table 1. No student will be allowed to graduate until after the research advisor has approved the bound copy.

Graduation confirmation forms are mailed to the student by the ALM in IT Program Office. It is important to return this form to the ALM in IT Program Office so that your name and your thesis title appear correctly at your graduation. Candidates who fail to meet graduation requirements must submit another form for a later graduation date.

In exceptional circumstances, the candidate can make a petition for a short extension of time. However, such an extension will affect the date of graduation. See Section 1.5 for a discussion of policies on extensions of time for the thesis project.

It is the responsibility of each candidate to monitor his or her own progress towards completion of the specific ALM in IT degree requirements and to notify the ALM in IT Program Office of a realistic anticipated date of graduation. Students must be mindful of the fact that a desire to graduate by a certain date neither obviates the need to fulfill the various degree requirements nor imposes any obligation on the ALM in IT Program Office to bend its policies so that the desired graduation deadline can be reached.

1.4.6 Illustration of Major Deadlines

The following list is an illustration of the deadlines for a student who has targeted his/her thesis completion for graduation in June of a given year.

- 1) Submission of the final thesis proposal by September 1 of the previous year
- 2) Registration in the thesis course within two weeks of receiving a letter approving the thesis proposal from the Dean of the Division of Continuing Education
- 3) Completion of the software project and submission of a draft of the thesis

document to the thesis director and research advisor for review by April 1

- 4) Submission of a final thesis draft to the thesis director for grading, with a copy to the research advisor for format review, by May 1
- 5) Submission of the bound thesis to the ALM in IT office by May 15.

1.5 Deadlines and Extensions

This section describes the policies regarding obtaining an extension to various deadlines in the process.

1.5.1 Five-year Enrollment Period

When students enroll in the ALM in IT program, they have five years in which to complete the program, calculated from the end date of the third course they take to qualify for the program. Students who have not completed both their course work and their thesis project before the end of the five-year period can petition the ALM in IT Program Office to obtain an extension of time. Such an extension is not granted lightly, and a student should be able to show definite signs of progress toward completing the degree. One such sign of progress would be that a student has submitted a thesis proposal or is working on the thesis project. If a student has not completed a thesis proposal or has completed all other course work by the end of the five-year period they will have to pay a fee to continue their enrollment for one year.

1.5.2 Thesis Course Deadlines

If a student has registered in the thesis course and has targeted a particular graduation date, they might find that they will not be able to meet the thesis submission milestones identified in Table 1. Students in this situation should discuss this matter with their thesis director. With the thesis director's concurrence, the student can ask the research advisor to move the target thesis completion date. If the new target completion date is before the end-date of the thesis course, then the student can simply notify the research advisor of the change. If the new target completion date is beyond the end-date of the thesis course, then the student should send a more formal request for an extension of time for the thesis course. This is normally an extension of only three months. If a student is not able to complete the thesis course within that extended time, he/she might be required to re-register for the thesis course.

If a student is not able to complete the final thesis draft in time for the thesis director to provide a grade, the student will not be able to graduate on their target graduation date, and will have to follow the steps above to get an extension of time and move the graduation date. This applies whether the student has been unable to work enough on the thesis or the thesis director, after reviewing the initial thesis draft, has said that the student needs to do additional work on the project as well as make revisions to

the thesis document that will take more time than is remaining.

1.6 Resources

Students undertaking a thesis project often need resources, such as hardware and software, or other kinds of support such as access to a library. This section contains a description of the resources available.

1.6.1 Computer Access

ALM in IT candidates who are currently matriculated toward the degree will have access to university computer systems at 53 Church Street and at the Science Center, whether or not they are registered in courses. If students have special computer resource requirements to carry out their thesis project, they should make these known in their proposal. However, Harvard Extension School provides no funds or grants to students to offset the cost of their thesis projects.

1.6.2 Library Privileges

Two science libraries offering books and journals to anyone with a valid Harvard ID, which includes all Extension School degree candidates, are the McKay Science Library in Pierce Hall and the Cabot Science Library in the Science Center. The latter offers research services, which are available to Extension School students.

Electronic access to the University's Hollis system is available at the URL <http://lib.harvard.edu/libraries/index.html>. You can navigate from Hollis to the ACM Digital Library and several IEEE journals if you have obtained your Harvard ID and a University PIN.

1.6.3 Financial Resources

The Extension School program does not have funds to support the thesis project. Each candidate must be prepared to invest not only time but also money in this project. In preparing a budget, it is useful to include allocations for some or all of the following items: books, equipment, photocopying, postage, computer searches, thesis paper, and binding of the final copy of the thesis.

1.6.4 Document Templates

It is a good idea to train yourself, from the outset, to employ the format required for the proposal and, later, for the thesis – from correct spacing of margins and headings to the correct style of documenting references. Note that the guidelines for such things as citations and references given in Bibliographies and References apply to both the thesis proposal and the final thesis document. Such practice will facilitate the preparation of your manuscript later on, when time and patience may suddenly grow short. Properly

formatted versions of the outlines for the thesis proposal and the thesis document are available in Microsoft® Word© format. Look on the ALM in IT website.

1.7 About the Rest of This Guide

The rest of this guide is organized as follows:

Chapter 2 contains a description of the thesis process, especially the beginning stage of identification of suitable topics for a thesis in the ALM in IT program.

Chapter 3 contains a description of the content required in the thesis proposal.

Chapter 4 contains a description of the content required in the thesis document.

Appendix 1 has a set of general formatting guidelines that apply to both the thesis proposal and the thesis document.

Appendix 2 has a discussion of guidelines for formatting references and citations.

Appendix 3 has an enumeration of common problems with English grammar, punctuation, and usage.

Appendix 4 is a bibliography of useful books on writing, including books that are applicable to writing a thesis.

2 The Thesis Process

Developing and carrying out a thesis project is quite unlike taking a course. For that reason, and because many ALM in IT candidates have little previous independent project experience, this chapter presents a brief overview of the development of a thesis topic.

Thinking about a problem and writing about a problem are virtually inseparable. Hence, this chapter has some overlap with the first part of Chapter 3, which deals with preparing the thesis proposal. But the perspective here is broader and includes general considerations not mentioned in the following chapter.

2.1 Differences Between Course Work and the Thesis Project

The process of completing a thesis project is one that can be learned and understood and in which almost anyone can achieve competence. Most ALM in IT candidates have performed capably in courses. Being aware of the ways in which a thesis project differs from course work can markedly improve your success at (and enjoyment of) the thesis project.

The principal differences between course work and the thesis project involve logistics (definition of goals, ease of obtaining materials, ease of carrying out the work) and time use (gauging how long things will take). In a course, someone else provides the structure for you. The course topic, the course syllabus, assignments, text(s), and other materials are developed for you. With them come a timetable and an implied work schedule. Read so many pages and do so much work each week, prepare for and take so many exams, write programs – and success in the course is yours. Furthermore, in thirteen to fifteen weeks the course is over, regardless of how well you have done.

By contrast, for the thesis part of your graduate studies, you are the person with primary responsibility for producing all of the above.

Taking the steps of developing a proposal, implementing a major software system, and describing your work in a thesis does not follow an altogether orderly process. False leads, insufficient information, contradictory findings, changes of goals, and other frustrations are inescapable parts of the process.

Because of this lack of external structure, it is difficult to predict the amount of time needed to implement a project and to write the thesis. Thus, although nine months is the amount of time allotted for the actual carrying out of the thesis project (implementation, analysis, and manuscript preparation), no specified amount of time has

been set for the preparations leading up to that. You must begin your thesis planning well before you can register for the thesis course. During this initial period you will be doing preliminary investigation of interesting topics, finding and developing your thesis problem, doing bibliographic research, and reading the background materials necessary for writing a proposal. Students who want to do a project in certain fields, such as artificial intelligence, might have to spend time acquiring or brushing up on a working knowledge of a programming language such as LISP before they can progress. Waiting until you've entered the nine-month independent project phase might be too late to acquire the fundamental skills you need and then to complete the project as well. It has been the case that a potential thesis director turned down an opportunity to direct a thesis because the student did not have adequate preparation in the appropriate technology.

The nature of this pre-thesis research and planning makes it difficult to schedule rigorously. It can involve many activities (web searches, library searches, reading books or manuals, prototyping, consulting with your research advisor), and it does require some creativity. And, although you do want to make progress, you do not want to be rushed. *Your primary goal should be to complete a solid, thorough piece of work, not just to graduate at a certain time.* In the process, you will learn something about your subject (and also about yourself – particularly your intellectual and organizational abilities). In the best case, your project will have the side effect that you can enjoy the whole process. Indeed, if there is any part of your graduate career that should be seen as process rather than product, it is the thesis project. This is a learning experience in the fullest sense, and can be an exciting challenge rather than an overwhelming obstacle, if it is approached realistically.

Nine months is the amount of time considered appropriate for completing a thesis in Information Technology, assuming that you have a proposal that has been carefully thought out, screened, revised, and approved by the research advisor and the thesis director. During this time, you will be designing and implementing a significant software program or suite of programs that provide a solution to the problem(s) stated in your proposal, and then documenting your efforts. You are responsible for your schedule during that period – with help from your research advisor and thesis director. But the work itself is up to you. Your best insurance for completing the thesis on time is to have your proposal clearly thought through beforehand in every possible respect, from a detailed work schedule to producing the final draft.

Another aspect that differentiates a thesis project from course work is that you will be working much more on your own. You may, of course, talk to someone about your work – your thesis director, the research advisor, someone who sat next to you during a class, a colleague at work.

2.2 Getting Started – Developing a List of Candidate Problems

A thesis project requires independent thought on your part, rather than simple regurgitation or the "book report" style of work. As a result of your project, you will be

able to make an original contribution, however limited in scope, to the subject by bringing new insight or a fresh perspective to your topic. But you must

1. Formulate a clear problem statement
2. Implement software that provides a solution to the problem
3. Describe how your software supports your approach to solving the problem

All of these steps require critical thinking, in the positive sense of carefully evaluating, rather than accepting without questioning.

2.2.1 Problem Statements

All good thesis projects start with one or more problem statements. Information Technology, and the myriad places where Information Technology is applied, is full of rich problems. In spite of this richness, identifying a good problem for your thesis requires an active, thoughtful mindset, and above all, curiosity.

As you begin searching for a topic area and, within that area, a suitable thesis problem, you might feel uncertain about which problems will lead to an acceptable thesis. Certainly not all problem statements are created equal: some will in fact not grow into satisfying thesis topics. Your initial problem statements might simply be too trivial or based on a lack of familiarity with a particular discipline. However, the simplest test to differentiate between potentially productive and non-productive thesis ideas is this: "Has the topic been so worked over that there are no opportunities for fresh insights or fresh applications within that area?" Stated more succinctly, you could ask if the topic area is just too "old hat," so that there are few challenges in digging into it. For instance, if you want to work in distributed systems, just writing a simple two-tier application with a user interface application and a database would not be considered particularly challenging.

While ALM in IT policy states that you must write a significant program, doing a thesis is not just about hacking some code, even clever code. You should be prepared to give an explanation of what you're about to do, make adjustments as you proceed, and present a clear description of what you have done at the end.

Your thesis problem can fit into several categories

- Novel Problem, sound solution – e.g., defining an interface to a digital camera to automatically track and shoot photographs of a solar eclipse
- Technology Driven (new technology for old problems) – e.g., applying Microsoft's distributed computing infrastructure, .NET, to construct a significant distributed application
- Foundations of Computer Science – e.g., analysis of algorithms in a particular context

- Surveys/comparisons of technology (an expository thesis) – e.g., comparing .NET to J2EE, with significant coding examples to back up your claims
- Human Factors – e.g., defining two or more interfaces to a system and doing analysis of the relative effectiveness of each
- Multi-disciplinary – e.g., pursuing an interest in another field, such as chemistry or biology, where a significant computer application is required

2.3 General Approaches to Academic Inquiry

ALM in IT candidates are expected to be familiar with important publications in their topic area. And, while they may not be sufficiently expert to judge the reliability of a piece of research, they should at least be sufficiently familiar with standard works in the field to recognize an unconventional analysis and obtain assistance in evaluating it.

Every ALM in IT thesis project will include at least bibliographic work, since such work is considered to be the background preparation for formulation of a suitable thesis problem. This bibliographic work will involve both printed matter and electronic matter. In other words, this is the literature-based reading and thinking essential to the development and elaboration of a thesis problem. At the earliest stages, it can be a way of finding your thesis problem, and later it can be a way of developing a concrete proposal for the thesis problem.

All methodical investigation should include a careful analysis and critique of the published sources related to your topic. The research advisor may be helpful here, but one of the student's chief tasks is to acquire a thorough knowledge of the literature relevant to the thesis problem. This knowledge is mandatory in all areas of research, including Information Technology. Any responsible exploration of a thesis topic requires an understanding of others' investigations as they appear in the relevant literature.

A simple overview or description of the literature is not sufficient or appropriate as a thesis project. Usually, a review of the literature is only the first stage in an investigation. It is the preamble to a problem statement, an analysis or interpretation designed to add something to the special field of knowledge.

In using any literature, the candidate is reminded to pay particular attention to how recently it has been published and how reliable it is, and in some cases to the scholarly reputation of the author or publisher.

While browsing through bookstores and searching the Web might turn up some interesting ideas, there are two other key sources that you should consider – professional organizations and trade journals.

2.3.1 Professional Organizations

You should begin to familiarize yourself with the work of other professionals involved in computer science or information technology. There are two major organizations to which professional computer scientists and engineers belong. These are the Association of Computing Machinery (ACM) (see www.acm.org) and the International Electric and Electronic Engineers (IEEE) Computer Society (see www.ieee.org). Both publish numerous journals and magazines that have recent articles on the latest and greatest technology and on issues related to information technology. These include

- *Communications of the ACM*
- *ACM Computing Surveys*
- *IEEE Software*
- *IEEE Computer*

In addition, each of the ACM and the IEEE Computer Society has journals published by special interest groups devoted to specific areas of technology (e.g., languages, graphics, database management, etc).

Student memberships are available at a reduced rate. There are numerous conferences as well that publish conference proceedings.

2.3.2 Trade Journals

In addition to the journals published by professional societies, there are numerous trade journals and magazines that have information on the latest developments and plans within the realm of computer systems and technology. Examples of trade journals are

- IBM Systems Journal
- Microsoft Systems Journal
- ComputerWorld
- Business Integration Journal

In these magazines and others of their ilk, beware of 'marketecture'. As a successful ALM in IT graduate, you should be able to separate the hype from the reality. In particular, if you rely on Web searches for information on vendor products, you will find that many sources are not objective in discussing their own technology versus others – be wary of uncritically accepting such as "Product X is the most popular choice for <whatever>."

2.3.3 Online Searches

In developing a complete list of relevant materials, candidates should consider an online search of the literature through the University Library System. For Information

Technology students, there are two electronic libraries that should be considered. Both allow you to search a vast collection of articles and papers on all aspects of Computer Science and Information Technology. The first is the IEEE Electronic Library (see <http://www.ieee.org>). The second is the ACM Digital Library (see <http://www.acm.org>). Both of these are available for on-line access with your Harvard ID and PIN.

In addition, students should familiarize themselves with the Harvard Library system as soon as possible. It can be quite helpful to learn how to use the library's various catalogues, the main periodical/journal index, the HOLLIS computerized catalogue, as well as the various specialized dictionaries, encyclopedias, and abstracts available to researchers in IT and Computer Science. HOLLIS is an excellent, no-cost way to begin your review of books in the Harvard libraries and a more limited review of recent journal articles (choose AI from the menu). You can search sources by author, title, subject, or keyword. A free pamphlet describing all of the various libraries is available in the Widener reference room. The Cabot Science Library, located in the Science Center, is available to ALM in IT candidates.

The underlying methodological principle in all disciplines is essentially the same: to organize our thinking about the subject matter, whatever it may be, in such a way that we feel we have an active command of it, a recall of it that enables us to explain it to someone else. Implicit in the description just given is the assumption that we continually strive for better and deeper understanding of our total subject matter, and that our finding the definitive explanation is less crucial than the process of searching for it. Our methods allow us to join a community of scholars, for our work must be communicable to members of the community.

2.4 Recapitulation: The Basic Steps in Beginning a Thesis Project

To reiterate, the procedure for beginning a thesis project consists of the following steps:

1. Consider a variety of topics that interest you and about which you have some questions.
2. Do some preliminary reading that will enable you to select a single topic and a leading problem or two related to the topic upon which you will focus.
3. Engage in more extensive bibliographic research that will help you focus, refine, and eliminate uninteresting topics. If the latter, return to step 1; otherwise, continue.
4. Formulate an initial problem statement derived from the initial questions.
5. Write a preliminary proposal that outlines:
 - a. The thesis project

- b. Its background or theoretical framework
 - c. The software to be written as part of the project
 - d. Project plan
6. Write a final proposal, incorporating comments and suggestions by the research advisor (there may be several interim drafts as well).
 7. Obtain a thesis director through the research advisor and program director, and obtain proposal approval.
 8. Begin the nine-month thesis period.

Chapter 3 elaborates on the steps leading up to completing the proposal.

3 The Thesis Proposal

Presentation of an acceptable proposal is not the beginning of the work on the thesis. On the contrary, it marks the end of the second major stage of the thesis process. In particular, a thesis proposal that is clear, complete, and concise will demonstrate that you have prepared yourself well by discovering the key issues and technology needed to undertake the actual thesis work, and that you have sized the thesis problem appropriately for the time allotted. Both of these factors will be key in gaining approval for your thesis project.

Before a proposal can be submitted, the student must have identified an appropriate thesis problem and developed realistic ways of exploring it. This process generally takes several months, and it should begin, as suggested earlier, while the student is still taking courses for ALM in IT credit.

The student should realize that preparing the proposal is both a formal and an intellectual exercise. Success in obtaining a thesis director will be significantly influenced by the quality and topic of the proposal. Therefore, all aspects of the proposal should be prepared as carefully as possible.

This chapter elaborates on the process of making a selection of a thesis topic. It also describes the content expected to be in the proposal. A sample proposal is can be found on the Extension School website under the description of the ALM in IT program.

3.1 At the Beginning

In reply to a survey of recent ALM graduates, a respondent emphasized:

Major commitments, both professional and personal, should be resolved before beginning work. You must think your way through completing the entire project, particularly in the early stages.

This is excellent advice. A thesis cannot be successfully completed with less than full dedication to the task. Planning and completing a thesis, though ultimately among the most rewarding academic experiences, are also among the most difficult. You should not underestimate the time and energy, nor the sheer determination that it will require for you to reach your goal, nor the rewards in pride and personal satisfaction that await you when you do.

3.1.1 Introductory Reading

The most efficient way to explore a topic and to test one's interest is to read about it. In the initial stages, the reader skims or surveys a broad variety of materials with the purpose of gaining background knowledge in several fields of inquiry.

As a rule, much of this introductory reading may not be related directly to the final thesis project. Nonetheless, the reader is advised to make a few brief notes on each work read – library call number, author, title, publisher, date, URL, etc, and a brief synopsis of the most interesting points – so that the material can be found again if it is needed later. Annuals, surveys, general histories, reviews, and specialized bibliographies, such as the archives of *Computing Reviews* published by the (ACM), are useful sources for the exploration of many topics. Section 2.3 contains an initial list of references. Textbooks, research papers and notes from previous courses also will help you acquire sufficient background to begin identifying a suitable thesis problem.

At the beginning of this reading, you should establish a consistent method of documenting what you have discovered. For each item, record full bibliographical information, and make reference notes of some sort. Few things are more exasperating than the discovery that the documentation on a significant piece of research is missing. A methodical recording of the necessary information, though tedious, should save much time and discomfort at later stages in the project.

This is a good point in your work to begin keeping an engineering notebook, or project journal, in which you can record the bibliographic information described above. Later you can record ideas you have in formulating a thesis problem, decisions made in consultation with your thesis director, and detailed design decisions you make in the course of the project. All this data will be invaluable when you sit down to write your thesis.

3.2 Identifying a Problem

It is most important to begin with an area or a topic of genuine and lasting interest. An ALM candidate wrote:

Love your research There is no heavier cargo to bear than a waning interest. Read and re-read as much as you can before submitting your research proposal to ensure that you have a significant curiosity and commitment to the subject matter.

Another said:

Choose your problem as you would choose a mate. Be sure it is one you can live with day after day and that it has new and interesting facets to uncover as time goes by. Even the best thesis topic will try your

patience from time to time; you should therefore attempt to ensure that you and the topic are as compatible as possible from the very start.

Usually the thesis topic is established on the basis of prior reading and personal experience, but the interests of possible thesis directors should also be considered carefully. You should therefore already have given some thought to how closely your proposal matches the research interests and capabilities of the Harvard faculty.

3.3 How do You Choose a Thesis Topic?

In the ALM in IT program, your thesis should demonstrate mastery of Information Technology. You have taken courses that cover several areas of the field, and you have had to write many small programs that illustrate the principles you are learning. For your thesis, you need to go beyond the kind of programs that might be assigned in a typical course and do a significant programming project that has two key characteristics:

- a) It uses interesting, up-to-date technology
- b) It involves an application that requires a good design effort on your part in mapping the application requirements to the technology.

You can approach the selection of a topic from either of the perspectives a) or b). If you have an interest in a technology that you've partially explored in your courses, you might use that as the driving force behind the project and select an application that makes good use of the technology. Or, you might have interests in certain application areas and can choose the application first, and then select a technology or set of technologies that allow you to provide an interesting implementation of the application.

Like any large undertaking, this selection process is easiest to approach if broken down into smaller steps. Your first step should be simply to do some brainstorming and make a list of topics that might interest you. At this stage, the topics are somewhat general in nature, perhaps comprising whole areas of Computer Science, such as algorithms, compilers, or distributed computing, and some might be a little more narrowly scoped, such as "using XML for data interchange" or "component-based architectures". They may be related to programs or papers you have written for Harvard Extension courses you have taken, to your reading, or to issues raised in connection with your work. Let the list take shape over several days or weeks until you have a number of topics. Write a paragraph about each of them. Then go down the list and rank the topics in order of your preference.

Step two involves discussing your list with others. Take the topics that interest you the most and show them to fellow students and *especially* to the research advisor. Bearing their reactions in mind, you are now ready to select one topic for step three. Take a blank piece of paper and write briefly what you know and what you would like to find out about the topic. It will be helpful to keep the following questions in mind:

- What has been said about this subject already?
- Which aspects of this topic remain unexplored or unresolved?
- Do any questionable or erroneous assumptions characterize the previous writings on this topic?
- Is there a particular technology or an application that uses such technology you can identify within this topic that might shed new light on it?
- If you choose this topic, what software projects would be relevant and what is their complexity?

Writing in this way tests your interest and also suggests how much you need to discover in order to begin to prepare a proposal on the chosen topic. If, for example, you are interested in Internet security but discover that there are technologies or concepts that you do not know well, then you could have a lot of work ahead of you, and you may want to rethink the topic. If your short essay proves to you that a topic is not worth pursuing, move to the next on your list of interests and write briefly again, and so on, until you feel satisfied with the topic you have selected.

3.3.1 Thesis Topics Related to Your Work

In Chapter 2 we discussed general sources for project ideas. One that we haven't discussed is your current work environment. In many cases, your company might be willing to have you do useful work for them that would qualify as a *bona fide* thesis project. There are a couple of caveats to keep in mind in this case:

- State who else is involved, what their roles will be, and what contributions **you** will be making on your own. Situations where you are designing and implementing only a small piece of a major project are not sufficient for a thesis
- Be sure your project code and your written description can be placed in the public domain without violating the company's copyrights – it is best to get this in writing from your employer when you submit your proposal

3.3.2 How High Should You Aim?

The general description of the thesis requirement given in Chapter 1 states

“At the very least, you should do a significant piece of software development and gain valuable insights about the use of an interesting technology or your approach that would be of interest to others. It is also within your reach to write such a good thesis that you, with the collaboration of your thesis director, could produce a document of publishable quality.”

In order to honor work of the highest quality, the ALM in IT program has a thesis prize that has been awarded at commencement since 2003. Every year, research directors and the research advisor nominate outstanding theses for the prize. These are then evaluated based on the following four criteria:

1. Originality of the thesis topic
2. Quality of the thesis document (e.g., are the problem requirements well-stated, is the system design clearly articulated, is the writing appropriate for an audience in the field?)
3. Quality of the implementation (Elegance in design and clarity of code – is it a pleasure to read?)
4. External Factors
 - Publishability
 - Potential for long-term use

These are criteria you might consider as you ponder your choices for a thesis project.

3.3.3 Human Subjects

If your work involves human subjects (e.g., development of different styles of GUI that require tests with human subjects) you should follow, when they are applicable, the published guidelines of the Committee on the Use of Human Subjects. This was prepared by the Faculty of Arts and Sciences and presented to the Harvard Institutional Review Board. The current URL for guidelines on research is <http://www.fas.harvard.edu/~research>.

3.4 Refining the Project Scope

Once a general thesis topic has been chosen, the next step should be to find ways to limit it, give it clearer focus, and shape it into a tightly defined thesis project. If the problem is too broad, the project is likely to remain superficial. There will simply be too much material for you to investigate. Conversely, the topic may be so narrow that not enough material exists to do it justice.

All this does not mean, however, that the thesis cannot deal with important issues. It should. But the mark of a careful and competent researcher is to know how to choose the proper angle from which to approach an issue.

Consider, for example, the problem of Internet security. No thesis could possibly encompass all the ramifications of this topic. Instead, the investigator might examine one specific aspect of security (e.g., authentication, access controls, non-repudiation, encryption, ...), and even within that one aspect narrow the project either to specific algorithmic approaches or to a focused area of application (stock trading, financial services, catalog ordering systems).

Here are some examples of topics that are too broad or unfocused, and some

suggestions for narrowing them down:

Unfocused:	"Internet Security"
Focused:	"Algorithms for Detecting Computer Intrusion"
Unfocused:	"Computer Based Software Engineering"
Focused:	"Building Reusable Components with C# in the .NET® Environment"
Unfocused:	"The eXtensible Markup Language (XML)"
Focused:	"Using XML to Define Data Interchange for Stock Trading Applications"
Unfocused:	"Computer Languages"
Focused:	"Comparing the Performance of Java JIT Compilers"
Unfocused:	"Mobile Devices"
Focused:	"Using WAP/WML to Integrate a Mobile Phone into a Field Technician's Workflow"

Developing a specific thesis problem out of a more general topic is perhaps the most difficult and important phase of the entire thesis project.

A major resource during this time is the research advisor, who assists students in formulating their thesis proposals after the completion of their formal course work, but before they have begun working with a thesis director from the Harvard faculty. One of the chief functions of the research advisor in your field is to aid you in producing a successful proposal; another, equally important, is to help you find a thesis director once the proposal has been accepted. You are strongly encouraged to meet with the research advisor BEFORE you begin writing your proposal. Since the research advisor must approve your proposal before sending it on to your eventual thesis director, it is clearly in your interest to find out how he or she reacts to it before you have invested a great deal of time and energy in the project.

It should be added that the first version of a proposal, for even a well-defined thesis problem, is likely to be modified or amplified in one way or another over the course of the investigation. One never knows entirely what will be found until the research is complete. But a clear focus at the outset is essential for the project to succeed at all.

3.5 Following Software Engineering Practice

Your work should follow standard software engineering practice. This means you should

- Be clear about your requirements
- Develop reasonable estimates and a plan for the work to be done
- Apply good design techniques to the implementation
- Plan for testing and demonstrating your solution – it should go without saying that the code you write should work according to your specifications

3.6 Content of the Proposal

While the style and detailed content of the proposal will depend upon the specific thesis problem and method, there are several common elements that should be present, and a format that you should follow in your proposal document. These are embodied in a proposal template that can be found at the ALM in IT website.

Each proposal should have a title page that includes the *field of concentration*, the *candidate's name, address, telephone number*, and the *date* on which the proposal is being submitted. In terms of content, all proposals should include the following:

Tentative Thesis Title

Abstract

Thesis Project Description

Work Plan

Glossary

Bibliography

Each of these is described in the sections that follow.

3.6.1 Tentative Thesis Title

The first section, the tentative title, probably should be written last. A successful title will emerge only after it has been determined, often by trial and error, just what the investigator hopes to accomplish. The title should be specific and clear; you may want to accompany it with a subtitle. Ideally, the title should summarize the thesis problem with efficiency and style. Avoid titles that are pretentious, vague, or wordy. Expressions such as "An Investigation of" are redundant and should be omitted. Titles of just one or two words, on the other hand, are too brief to indicate the scope of the thesis problem. An overall rule is that the title should be explanatory but concise when standing by itself.

The title needs to be succinct, reflecting the appropriate scope of work you are

undertaking. See section 3.4 for a list of examples of titles for topics. Consider your title a working title, since changes in direction along the way will affect the results you finally achieve.

3.6.2 Abstract

Describe, in one or two sentences, what the central theme, or thesis, of this project is to be. Then, in only two or three paragraphs, give a *summary* of what you hope your project will show, what technical problems you will have solve, what technologies are relevant, and what your approach is. Feel free to give a name to your application or system. Having a name for the system to be developed will make it easier to write about. Two examples of application names given in ALM in IT theses are ‘JAWS’ (for Java Auction Web Service) and ‘SciencePeer’ (a collaboration tool for scientists using a peer-to-peer implementation). It’s much easier to write ‘SciencePeer provides the following six services ...’ than ‘The thing I’m building in this project provides the following six services ...’.

3.6.3 Thesis Project Description

In this section, elaborate on the summary given in the abstract above. This gives you a chance to describe how much background work you have already done in narrowing down your thoughts to a thesis topic of a reasonable scope. Include references to material you have used to help you define this proposal (use the Bibliography section to give a complete description of the references).

Begin by describing the context in which this project is being done. This includes relevant course work, research sources, your own background and readiness for the topic, other research that is like yours or upon which your project is based. Cite references to any external sources.

Next, describe how you will approach the project and related implementation work. For instance, you can give a description of background information you have already gathered, and other sources you expect to tap into. Then describe the implementation vehicle (for instance, an application you will build) that will either illustrate your thesis or form the core of your demonstration. By now, you should be able to describe preliminary requirements and design information.

This chapter should cover the following:

- Background of the project
- Comparison of your approach to other approaches
- A description of the application or software components that will be written as part of the project
- A brief description of your technology choices

3.6.4 Describing Your Software Components

Too often, a student will present a diagram showing a set of tiers, such as Presentation, Business Logic and Data Access. This is usually accompanied by text that explains the well-known purpose of each tier. While this is a useful way to organize one's implementation that is supported by many platforms, there are two problems with this diagram as "the architecture for my system". First, most people who use this don't include a single word on the diagram that refers to any specific feature of the application; that is, the diagram is application-independent. An analogy is that of a housing architect sending you a model for your new house that shows four tiers: a foundation, two floors with walls, and the roof. What you really need to see are diagrams that show how the house fits into its surroundings and the elements of the house and their relationships to one another – the style of the doors and windows, the location of rooms, stairs, hallways, and so on, so that you can imagine how you might live in the house. Second, although these multi-tier diagrams have been around and are available from many sources, no citation of the source used is present.

So you need to think carefully about the major functional components of the application you're building from a user perspective (e.g., a Customer Information Service, an Order Subsystem, a Business Rule Interpreter) and also indicate how these components depend on one another. Only then can you convey to a potential user what the system will do when it's built and how you are planning to design and stage the implementation.

3.6.5 Work Plan

In this chapter, describe what your overall work plan is. Be as specific as you can as to how the parts of your project will come together, so that you and your thesis director can make better decisions about changes as new information comes to light. A work plan is more than just a schedule containing a list of activities and dates. You should describe the assumptions you are making and the perceived technical risks, and then give the schedule itself.

Assumptions, Risks and Alternatives. Describe the development environment you require (language, OS, system) and other tools you expect to use. Also describe any assumptions you have made about what it will take to finish your work.

Describe the risks you now see as inherent in your work and alternatives you might have to take to ameliorate these risks (e.g., project scope and alternatives for scope reduction).

Preliminary Schedule. The schedule we are discussing here refers only to what happens after the proposal is accepted. Since the exact starting time for your project is unknown, your tentative plan should comprise a list of work activities based on the application components identified above, with a duration for each expressed in weeks. As part of your plan, allow six weeks for writing the first complete draft of your thesis document. Once you've estimated each of the activities in your plan, you can see whether or not the total duration fits within a nine-month period – if not, you need to consider modifying the scope. In other words, understand your project well enough so that you

can give a good account of how long each of the major components will take to complete and how long the entire project will take. It should go without saying that as a good engineer, you will include within your estimates adequate time for ensuring that your application works correctly.

3.6.6 Additional Sections

You may feel that the proposal requires additional sections in order to explain your project fully. You should feel free to add such supplementary sections.

3.6.7 Glossary

You should not assume that all readers are familiar with the technology or terminology referred to in your thesis proposal. This section should include definitions of major terms and an explanation of acronyms that are not commonly understood – for instance, it is not necessary to include entries for MySQL, SQL Server, or J2EE in your glossary. If you're using some new standard dialect of XML (XYZXML) you should include that in the glossary.

3.6.8 Bibliography

Your thesis project, like most efforts, is built on the work of others, whether it is the vendors whose platform you are using or those who define algorithms that you are using or extending. It is important that your thesis proposal indicates your level of knowledge about any prior art on which you are relying. This set of references is a “working bibliography,” in that it is likely a subset of the references you will have when your thesis is done. The working bibliography should be selective, in that it should not simply include all the materials that might conceivably be used in the finished research. Rather, it should demonstrate that you have actually read the sources you cite, know which further sources you will need to consult, and why. The bibliography, in other words, represents an interim tally of your progress.

Note that the working bibliography for the thesis proposal will likely include works that are not cited in the proposal, and it should include most of the materials that will be used in the finished thesis project. It should list, under separate heading, all works cited in the proposal (Works Cited); all works consulted in preparation of the proposal (Works Consulted); and all works that the student intends to consult in further research and writing (Works to Be Consulted). In your thesis document you will list only references that are cited in your thesis document.

You will not be able to satisfy the requirement for a working bibliography by simply appending a printout of titles generated by a computer search related to your topic; i.e., you should not haphazardly compile a lengthy but meaningless bibliography. The bibliography should represent actual work done, and it should reflect an organized approach to the thesis problem. Thus it will assist the research advisor in assessing the nature and direction of the project and will facilitate suggestions for additional reading.

The working bibliography should be prepared in the precise form required for the finished thesis. All entries should be complete. All ALM in IT students should use the APA style for citations and references, as described below.

3.7 Submitting Your Proposal

Submit each version of your proposal via email to the research advisor at robinson@fas.harvard.edu. Microsoft Word is the preferred format, but an Adobe PDF file is acceptable. There is a standard proposal template available on the Harvard Extension School website.

The attached proposal should be in one file whose name has the following format:
<last name><first initial>ALMinITProposal<yyyymmdd>.doc

For a large document, use a zip file. For instance, J.Q. Student, submitting a proposal on December 15 of 2004 would have a proposal file attached with the name StudentJALMinITProposal20041215.doc. The next revision, in say, January, would be something like StudentJALMinITProposal20050120.doc.

Once you submit your proposal, four weeks are normally required for its evaluation by your research advisor, who will be working on several other proposals simultaneously – as many as two dozen in peak periods. Occasionally the research advisor may be able to read the proposal more rapidly; but your tentative schedule should allow for the full four weeks. Students should not pressure the research advisor to rush their particular proposal along; the research advisor is often working with many candidates at a time, all of who have an equal claim on his/her attention. If the proposal needs revision, which is likely, you should allow for a second four-week period for the second reading after you have resubmitted the revised proposal, and so on.

3.7.1 Revisions and Writing Style

It usually happens that the first draft of the proposal is not acceptable. Typically, two or more attempts are necessary, each of which will require up to a few weeks for evaluation by the research advisor in your field, before a thesis director can be found and the proposal is approved.

Here are some of the reasons for which first proposals might be sent back for further work.

First, the thesis might not be properly focused, in that the advisor thinks the proposal covers too broad an area. In this case, the submitter will be asked to reduce the scope of the proposal.

Second, there might not be enough detailed description of the intended purpose of the software project. In this case, the submitter will be asked to do a little more work and define the software project, for instance, by providing a better rationale for the project, or

by including more information about the design of the software.

Third, some of the required content of the proposal might be missing, in which case the submitter will be asked to complete the proposal.

Fourth, the set of references cited in the text does not provide proper attribution to the sources of the material used.

Fifth, the proposal might lack the proper format or consistently demonstrate sub-standard usage of English. Be sure that all the sections required by the guide have been included in the proposal and that the spelling, grammar, and usage are accurate. Since the ALM in IT is both a graduate degree and the highest academic degree awarded in Extension, standards for English usage are correspondingly strict. Non-native English speakers may need to allot extra time for the preparation of the thesis proposal in order to ensure that their language use meets ALM in IT standards. If necessary, students can hire, at their own expense, an editor to aid with formatting and language issues.

The writing style in all instances must follow the guidelines established in standard writing handbooks and must be at the same level expected in the final thesis. Acceptance of the proposal will depend at least in part on the adequacy of the presentation. Correct grammar, spelling, format, citations, and a meticulously proofread text are essential. The pages of the proposal must be numbered.

Do not assume that because the thesis proposal is only a first step in the thesis process, form is not important. This assumption is a mistake. It is the overall nature of the proposal – i.e., its content and its form – that will determine whether a professor will consent to become your thesis director.

There are two appendices devoted to aspects of formatting and grammar that are relevant to the proposal. Appendix 2 contains guidelines for citations and references. Some of the most common mistakes in grammar, punctuation, and usage are listed in Appendix 3 Appendix 4 . The bibliography in Appendix 4 contains references to writing and style guides. Additional help is available through the Writing Center in the Extension School.

4 The Thesis Project

Once you have a thesis director who has signed the proposal approval form and you have submitted it, you have nine months to complete the thesis project and to write the thesis document. You, like most thesis candidates, will spend the bulk of the nine months allotted designing and implementing a significant software system or application, but you also need to allow sufficient time to describe what you have done in a thesis document.

4.1 Working on Your Project

We reiterate an important caveat for thesis work. The thesis project is not the same as taking a conventional class. A conventional class has a specified time limit and a specified set of tasks to accomplish. In the thesis project you, working with your thesis director, are responsible for the pace, the design decisions, and the development of the final results. There is no guarantee that you will finish the project on time and no guarantee that you will receive a passing grade.

The most important first steps are to review the thesis proposal with your director, make any changes that come out of that discussion, and then discuss the working relationship with your director, including scheduled meetings. It is also appropriate to discuss the thesis director's expectations relative to the quality of the finished work and the grading criteria the director will apply when you are finished. This will help to avoid surprises at the end of project.

It is also advisable that you keep a journal or engineering notebook in which you can record notes from your reading (building your bibliography), design decisions made, changes in direction because of things you have learned along the way, and key discussions with the research advisor or thesis director.

Note that the design documentation should not be thought about for the first time in writing the thesis document – rather, it should be a natural side effect of your implementation work in the software project. Keep your design up to date – it will save you time in the long run.

Also note that your code (including html files, xml files, sql files, etc) will be included in the thesis document, which will be placed in the Grossman Library. Keep your code organized and well commented from the start.

4.2 Thesis Content and Organization

Thesis candidates frequently ask how they should begin their thesis document.

One basic principle in writing is to write for a particular audience. In the case of the ALM in IT thesis, the primary audience comprises your thesis director, the research advisor, and your peers in the ALM in IT program.

A second principle for an expository document such as your thesis is to “tell people what you’re going to tell them, tell them, then tell them what you just told them.” At the highest level, the Introduction should include an overview of the rest of the document. The Summary and Conclusion chapter should summarize the salient points of your thesis, and describe what other things you might have done given more time. Each chapter should begin with a description of what it contains, and end with a brief summary and optionally a brief reference to the next chapter.

4.2.1 Organization

Throughout the process – from the moment you begin to consider a thesis problem to the time of the first draft of a thesis – some rough organizational scheme guides both the project and the writing. As the work continues, this scheme inevitably becomes more coherent, more logical, and more orderly. The code, interesting in isolation, becomes fully significant only when your description reveals something about the problems or ideas that prompted the project initially or the technical problems encountered during the implementation. Prepare a broad preliminary draft as quickly and efficiently as possible, even during the last stages of the implementation of the software project, with the understanding that successful writing inevitably requires multiple revisions. Just as software development often proceeds in iterative cycles of 'designing-coding-testing', a thesis manuscript will go through cycles of 'writing-reviewing-revising'.

At the end of the thesis software project, the candidate will be confronted with a wealth of information. The most difficult and often the most satisfying part of the process is about to begin or to begin anew: the organization of these materials into a coherent thesis document that documents not only the kinds of problems you solved but also the insights gained from your experience that might be of interest to others.

In organizing your document, you want to have a clear separation of concepts (e.g., requirements vs. implementation). Each chapter should have a distinct purpose in explaining your thesis project. Avoid repetition – many times students describe the same thing more than once, but use different terms or names for things. This creates confusion.

Many writers want to use almost all these materials in the first draft of the thesis. And many thesis directors agree that this approach is reasonable. Including as much as possible in the first draft allows the writer to organize all the material in some preliminary way. Further, these observations are less likely to be lost if they are included in a draft.

Beginning researchers should realize that they probably will not use all the

information they have accumulated. As draft succeeds draft, materials will be deleted, rearranged, or added – a normal part of any writing process. It is useful, however, to keep copies of earlier drafts, for they then can be compared and information in them retrieved easily. If you have a word processor, consider keeping drafts as separate documents or files so that you can subsequently "cut and paste" from one draft to another.

Many writers feel that the preparation of an outline is of great help. Often they prepare an outline before they begin to write, and they use it as a way of keeping track of their use of research data.

An outline can be used in another way. After a first draft, many writers make an outline based strictly on the thesis itself. This outline is often a sentence outline – a one-sentence summary of each paragraph. It helps the writer recognize a paragraph that is too long, too full of information, or out of place. If it is difficult or impossible to summarize a paragraph in one sentence, divide the material into two or more paragraphs. In addition, the sentence outline helps the writer gain a quick overview of the entire thesis. Material that appears in several places is easily identified and consolidated. Work to strengthen your argument and to sharpen your introduction and conclusion. The organization of a thesis, as revealed through the outline, is always subject to revision.

4.3 Mechanics of Writing

Once a preliminary draft has been completed, the writer can turn to the mechanical details of grammar, spelling, and punctuation, and to the more subtle questions of prose style. *The Elements of Style* (Strunk, White, and Angell, 2000) is useful here. This and other basic guides to writing style are listed in the references at the back of this guide. Students should carefully review the rules governing the use of basic marks of punctuation; the more specialized elements such as hyphens, italics, parentheses, brackets, and ellipses. In addition this guide specifies how your working source code should be organized as an appendix to your thesis document.

As with the thesis proposal, good English is essential in the final thesis. Errors in language, spelling, or usage quickly destroy a reader's confidence. The writer is urged to check these elements of style meticulously at all times. Theses have been rejected in the past in part because their authors clearly had not achieved an appropriate standard of grammatical and spelling accuracy. It is as much your responsibility to write carefully as to research accurately. Non-native English speakers should be especially careful in crafting their written work, and consider having their drafts reviewed by a native speaker if they feel the need for some assistance. Most word processors now have spell-checks and grammar-checks that can aid writers in achieving correctness in their work. If necessary, a student can hire, at their own expense, an editor to aid with formatting and language issues. Candidates are also cautioned to use nondiscriminatory, nonsexist language in their writing. (See the item "Nonsexist Language." Appendix 3)

Periodicals in the field can be efficient guides for content, organization, and style. The ALM in IT candidate is encouraged to consult these journals in the preparation of the

final thesis and to discuss these matters with the thesis director. The tone or style in a professional publication, however, which derives from certain assumptions about the reader, is not necessarily appropriate for the ALM in IT thesis. A student intending to submit the thesis for publication in a professional journal may, with the permission of the research advisor and thesis director, prepare the thesis in a particular journal style. In this case, on an extra page preceding all front matter including the title page, the following should be stated:

The style and format of this thesis follow the standards in <Title of Journal>, to which this work is being submitted for publication.

Note that even in this case, you will have an appendix with source listings for your project, and you should not leave out the other key sections required in the thesis.

4.3.1 APA Publication Manual

All thesis proposals and thesis documents should follow the guidelines of the APA Publication Manual, Fifth Edition. Copies of the guide are available in some bookstores and are on reserve in Grossman Library. The handbook can be ordered directly online at www.apastyle.org.

4.3.2 Guidelines on Plagiarism

ALM in IT candidates are expected to be thoroughly familiar with all University Extension regulations involving plagiarism, the proper use of sources, and the preparation of academic papers. These regulations are described in the current Extension catalogue under the section entitled "Academic Policy." On reserve in Grossman Library is a folder labeled "Plagiarism," which explains in detail what constitutes plagiarism and how it can be avoided. You can find an online copy of the Harvard University publication "Writing with Sources" at <http://www.fas.harvard.edu/~expos/sources/>.

4.3.3 Being Clear about the Source

What's important is that the reader can tell from the text itself whether a statement is a documented fact (with appropriate attribution of the source) or an opinion you developed during the course of the thesis project.

In theses that mention or cover algorithms, patterns, engineering processes, and so on, include a citation of a reference on the topic. Here is a simple example: If you say "The design of the X Module uses the Model-View-Controller (MVC) pattern," put in a citation to a book or article that explains MVC. In general, when a reference in a book or article exists, use that rather than a URL. If you have a long discussion that relies on external sources, use several citations that mention the chapter or page number for each reference.

If you mention features of a product that you used in your work, include a citation of a suitable reference to that product. These references can be to a book, user documentation, or an online tutorial.

In a related vein, avoid making unsubstantiated claims. A bold assertion such as “Most programmers would rather use an object-oriented language” will require either a credible reference or deletion. A modest “Based on my experience, it appears that most programmers would rather use an object-oriented language” could suffice.

4.4 Preparation of the Manuscript

The following comments are basic guidelines; in cases of ambiguity the final arbiter should be the thesis director in matters of content and the research advisor in matters of format.

4.4.1 Front Matter

The materials preceding the text, such as the preface, table of contents, title page, and frontispiece, are collectively known as the front matter. There are precise requirements for the format and sequence of the front matter in theses. With the exception of the title page and blank page, all headings for these pages must be formatted in a consistent manner, with each occurring 1 1/2 inches from the top of the page, followed by *quadruple spacing* before the first line of text.

The manuscript begins with a title page, followed by a blank page, and then the abstract page, on which appears a one-page summary of the whole thesis. These first three pages are unnumbered but counted in the pagination. Thus, in the following order and on separate pages, there appear:

1. Title page
2. Blank page
3. Abstract
4. Author's biographical sketch (optional)
5. Dedication (optional)
6. Acknowledgments (optional)
7. Table of Contents
8. List of Tables
9. List of Figures (including graphs/other displays)

Pages four through nine of the front matter are counted in the pagination and numbered consecutively with lower-case Roman numerals at the bottom of the page, beginning with "iv" for the optional biographical sketch or dedication.

If the student wishes to copyright the thesis, the statement "Copyright 20__ Author's Name"(but reproduced without quotation marks) should appear at the bottom of the blank page, item 2 in the list above. Or the word "Copyright" may be replaced by an uppercase "C" within a circle, (©) followed by the date and author's name (e.g., "Copyright 2003 John Q. Student", or "© 2003 John Q. Student")

Title page On the title page, which is entirely double-spaced, the title of the work appears two inches from the top, in appropriate uppercase and lowercase letters. The author's name, including initials, appears centered on the page, 2 1/2 inches below the title. A standard statement concerning the field of concentration appears two inches below the author's name. The name of the university and the author's date of graduation appear one inch below the field of concentration with the last line, showing the date, 2 1/2 inches from the bottom of the page. Nothing on the title page is underlined or reproduced in boldface. The date is always November, March, or June depending upon the time of graduation. The thesis template contains a title page.

An Efficient Web-Caching Algorithm

John Q. Student

A Thesis in the Field of Information Technology
for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

June 2003

Abstract The abstract is a one-page, double-spaced summary. It builds on the abstract of the thesis proposal, but focuses on what has actually been done in the project. Thus, it presents a succinct overview of the thesis including: (1) what the central theme, or thesis, of this project is, (2) what the project has shown, (3) what technical problems you had to solve, (4) what technologies were relevant, and (5) what your approach was. If the rest of the thesis were missing, the abstract could stand alone as a summary of the project, and would be comprehensible to a reader unfamiliar with the text.

Table of Contents The heading for the table of contents should be typed 1 1/2 inches from the top of the page, with quadruple spacing before the first item shown. The table of contents is entirely double-spaced. Each *chapter title* should appear just as it does in the text, using uppercase and lowercase with no underlining or boldface. Each *a-head* should be indented two *spaces* under its numbered chapter title, and each *b-head* should be indented two spaces under the preceding a-head, and so on. The number of the beginning page should be indicated in each instance, connected to the title by a continuous line of periods. Page numbers should be aligned.

A sample table of contents, based on a fictional thesis, is found on the following page. This illustrates format only. Note that the table of contents is itself a numbered page and is shown in the actual table of contents. Note also the indentions used to distinguish *Level 2 headings* and *Level 3 headings* from major chapter headings. Chapter numbers are represented by Arabic numerals, with the word "Chapter" preceding them. Two blank spaces separate the Arabic numeral and the chapter title.

Table of Contents

Dedication.....	iv
Acknowledgments.....	v
Table of Contents.....	vi
List of Tables	vii
List of Figures.....	viii
Chapter 1 Introduction.....	1
Chapter 2 Deficiencies in Previous Approaches.....	8
Chapter 3 High-Level Design	10
System Components.....	14
Object Model	17
Data Model.....	24
Chapter 4 The Caching Algorithm.....	28
Description.....	28
Examples.....	30
Performance Characteristics	35
Chapter 5 Summary and Conclusions.....	38
Lessons Learned.....	39
Future Work.....	40
References.....	41
Appendix A Source Code	42

List of Tables The list of tables should follow the format of the table of contents. For each table, show its number and title or other appropriate legend, and the page number on which the table occurs. If you have no tables in your document, omit this section.

List of Figures The list of figures should follow the format of the table of contents. For each figure, show its number and title or other appropriate legend, and the page number on which the figure occurs. Since you will almost certainly document the design of your system using graphical models, this section is not likely to be omitted.

4.4.2 Text of the Thesis

The text of the thesis follows the front matter. This can have as many chapters as you need. The only required chapters are the “Introduction”, which is always chapter 1, and the “Summary and Conclusions,” which is the last chapter before the list of references.

Pages in the main body of the thesis are numbered consecutively with Arabic numerals, following the formatting explained in the section "Margins and Pagination." Double spacing is expected throughout the manuscript. All new paragraphs must be indented one-half inch from the left margin, with no additional spacing between paragraphs.

Single-spacing is acceptable in those few instances where it may improve readability, as in long bibliographic entries, long quotations (more than four typed lines), lines of code used for illustrating technical points, and extensive figure captions. Introduce the section with a sentence, typically followed by a colon; then double-space before beginning the single-spaced insert. In order to distinguish them visually from ordinary new-paragraph indentions, long quotations, but not code, are indented ten spaces (or two tabs) from the left-hand margin. Code inserted into the text of a thesis should be the same font size as the surrounding text, but the font should be Courier New.

Headings, citations, the use of tables and figures, entries in the list of references, and all other mechanics must be presented in a consistent manner throughout the thesis. Unless the thesis is being prepared for publication in a journal that uses a particular notational style, students should follow the direction of this guide, or the *APA Publishing Manual*. See Appendix 1 for general formatting guidelines, and Appendix 2 for detailed information on formatting citations and references.

In most cases, theses in the ALM in IT program will not have recourse to footnotes or endnotes. Their use is rare in most Computer Science and Information Technology publications. If the thought is important enough to mention, either insert a citation to an item in the bibliography or place the text you would have in the note in the text body.

4.4.3 End Matter

Like the front matter, the materials at the close of the thesis have a special format and sequence. Known as end matter, their headings are formatted and paginated in a manner consistent with the other major headings throughout the thesis, beginning 1 1/2 inches from the top of the page, with quadruple spacing after the heading. They appear in the following order and begin on separate pages:

1. References
2. Appendices – (Optional)
3. Appendix – Source Code
4. Index (optional)

The list of references contains only those works specifically cited in the thesis. The list of references should be double-spaced between entries but single-spaced within entries.

Any appendix should be prepared and formatted as a chapter, including the heading, which comprises the word “Appendix”, followed by the appendix number, which is followed by the title. All theses in the ALM in IT will have at least one Appendix, which contains the source code from the software project.

4.4.4 Formatting Source Code

Source code should be in a single appendix. It should be ordered by language type (e.g., Java, C#, HTML, SQL, XML, etc). Each of these has a level 2 header. Files have level 3 headers and should be listed alphabetically by file name. Figure 1 below has a sample that shows what the file would look like. Note that use of this formatting allows all file names to appear in the table of contents. However, it is somewhat time consuming, so that you can defer adding the code appendix until near the end of the review cycle. Do not include generated files in your thesis, just files that you create or edit yourself.

Appendix 1 Application Code

This Appendix contains all source code, including HTML files, Java files, and database scripts (SQL files).

Java Files

The Java files are ordered alphabetically by class name. Each file is given its own level three heading so that it will appear in the Table of Contents.

AardvarkClass.java

```
/*  
 *  
 * File      :   AardvarkClass.java  
 *  
 * Author   :   Your Name  
 *  
 * Contents :   This is the driver for the Zoo application  
 *  
 */
```

Figure 1 Formatting Application Code

4.5 Format Review and Format Approval

The candidate submits one copy of the thesis to the thesis director, who evaluates the overall work. A second copy is submitted to the research advisor for evaluation of the format. The latter copy should be as nearly complete as possible, printed exactly as it will be when it is bound, so the research advisor can precisely evaluate the appearance of the thesis.

Format review involves more than making sure the margins are correct. It is, in fact, the final crucial stage in the editing process. The research advisor will carefully examine the thesis and will not give it final approval if there are errors in the format of the front matter or end matter, or if the text contains errors in spelling, punctuation, grammar, and usage. Such errors must be corrected and the corrections resubmitted to the research advisor before final approval will be given. **All theses inevitably contain some errors**, so students should not be alarmed if they are asked to make some corrections at this final stage. Most are quickly and easily rectified, especially if the student pays

careful attention to the instructions given in this chapter. **It is therefore wise to submit the final copy printed on ordinary paper, not the more expensive thesis paper that will be used for the bound copy. Students should submit a sample of the thesis paper they intend to use for the approved bound copy when they submit the thesis to the research advisor for final approval.** The research advisor will want to see **both** the original copy of the thesis and the corrected copy so that comparisons can be made quickly and efficiently. The corrected copy can be safely printed on thesis paper once the research advisor approves the thesis.

Students wishing to graduate in June should be prepared to work quickly and hard through the final weeks of March and the entire month of April, being mindful of the April 1st deadline to submit the thesis to the research advisor and the May 15 deadline to submit the bound copy to the ALM in IT office. Students targeting another graduation date should adhere to the timeline given in Table 1 Milestones for Completing a Thesis.

The research advisor will likely keep your thesis for up to two weeks (remember, s/he is probably reading about 20 during this period). *Students should not pressure the research advisor with unnecessary telephone calls during this period to inquire whether their thesis is ready.* The research advisor will contact you as soon as the format review has been completed. Once it is returned to you, work rapidly to make the necessary corrections and to resubmit the thesis to the research advisor for what will most likely be final approval this time. Then you must take your manuscript to the bindery without delay. At this stage, **no additional changes** should be made that have not been approved by the research advisor.

4.6 Binding the Thesis and Final Approval of the Bound Copy

The revised and approved final copy of the manuscript must be bound in a sewn Class A library binding with a traditional Harvard-red buckram cover. The author's name, the title of the thesis (abbreviated if necessary), and the year of graduation must be stamped in gold lettering on the spine of the binding. One copy of the bound thesis is then submitted to the Extension Office for placement on the Grossman Library shelves.

The cost of binding the thesis varies in accordance with the speed of service that the student requests. Two-week service costs as little as \$18 per copy, while one-day service can cost up to \$250 per copy. Current price lists are available in the ALM in IT office. There are a number of binderies in the Boston area. Two have extensive experience with theses for Harvard University:

Acme Bookbinding
Company
100 Cambridge Street
Charlestown, MA 02129
Telephone: 617-242-1100

Wells Bindery
54 Stearns Street
Waltham, MA 02154
Telephone: 781-893-3050

When the bound copy is submitted to the ALM in IT office, the research advisor will check to ensure that all of the format requirements have been met, including the appropriate binding. If the format is incorrect, the thesis will not be approved and the candidate will be required to make the necessary corrections, which may involve rebinding the thesis. Fortunately, this happens rarely. For the archival purposes of the University, the most important pieces that must be error free are the binding and the cover page. Be as thorough as possible in proofreading your final copy before binding.

Appendix 1 General Formatting Guidelines

Thesis proposals and thesis documents have nearly identical formatting guidelines, though the intent and the content of the two types of document are very different. Thus, the formatting guidelines given here are applicable to both the thesis proposal and the thesis document. Unless the thesis is being prepared for publication in a journal that uses a particular notational style, students should follow the direction of this guide, or the *APA Publishing Manual*. Appendix 2 contains a discussion of the format for citations and references.

Writing Style

Any successful thesis proposal or thesis document will appear in a consistent and recognizable style. Interested readers will be able to examine it efficiently and will be stimulated to learn more about the subject. Also, they will be able to find the resources, allowing them to think for themselves about these materials.

The document should be written primarily in the third person, active voice. Avoid using the passive voice, which leads to dull and soporific writing. For instance, rather than writing "The algorithm was run 100 times in the test program" write "The test program ran the algorithm 100 times." As a rule, use the first-person singular ("I") sparingly in the proposal or in the thesis.

Parenthetical expressions and successive one-sentence paragraphs should also be used rarely. Examine lengthy paragraphs carefully for coherence and cohesion, making sure that each paragraph conveys a separate, unambiguous point. Unless the author of a particular work is a highly significant figure, references to that work should emphasize the topic of inquiry, not the author's name or the title of the work. Do not simply refer to the "Smith thesis"; state what Smith's thesis is. In other words, do not indulge in scholarly name-dropping. Demonstrate that you have read and understood the work in question by referring to its substance rather than its form or authorship.

Also, in a thesis based upon careful investigation and solid technical work, the style should be matter-of-fact; in particular, you should avoid expressions that are emotional in tone, such as "Linux is the only operating system worth considering." Statements such as this make the thesis into a polemic or a diatribe, and the results become less convincing. Persuasion should be achieved with facts and a careful, reasoned interpretation, not through force of rhetoric. The presence of biased opinions or sarcasm can only detract from the seriousness of your presentation and prejudice the reader against your point of view.

Double spacing is expected throughout the manuscript. All new paragraphs must be indented one-half inch from the left margin, with no additional spacing between paragraphs.

Single-spacing is acceptable in those few instances where it may improve readability, as in long bibliographic entries, long quotations (more than four typed lines), and extensive figure captions. Introduce the quotation with a sentence, typically followed by a colon; then double-space before beginning the quotation. In order to distinguish them visually from ordinary new-paragraph indentions, long quotations are indented one inch from the left-hand margin.

Headings, references, bibliographic entries, and all other mechanics must be presented in a consistent manner throughout.

In most cases, theses in the ALM in IT program will not have recourse to footnotes or endnotes. Their use is rare in most Computer Science and Information Technology publications. If the thought is important enough to mention, either insert a citation to an item in the bibliography or place the text you would have otherwise placed in the note in the text body.

Use and Format of Chapter Headings and Subheadings

It is helpful to begin with an outline of the thesis – your ability to do so will indicate how well organized your material is. The outline leads naturally to division of the document into chapters, sections, and subsections, each of which has a particular heading style. In most cases, you will need at most three levels, and in some circumstances a fourth level. Level 1 is for major organizational units, such as chapters and appendices. Level 2 is for major subdivisions of the chapter or appendix, and level 3 is for subsections of the level 2 sections. Level 4, when it is needed, is for identifiable subsections within a third level section. Usually, three levels of heading is sufficient, and a fourth level should be used sparingly.

Avoid the extremes in the use of subsections. If you find that a chapter of more than two or three pages has no subsections, look at the flow of the text and ask what the structure is. On the other hand, if you find that you are using many fourth level or higher headings, rethink your organizational structure.

Headings and Transitions

In no instance should headings appear consecutively without interspersed text. After the chapter title, and before the first subsection, there should be some description of what the chapter is about and how it is organized. Recursively, after any heading, there should be some general introductory statements before any lower level heading appears. The issue here is successful transitions, which are a mark of clear and careful writing. Without such transitions, the presentation becomes an outline.

Formatting Headings

The ALM in IT thesis is divided into chapters. Titles for chapters are known as **Level 1 headings**. Generally, chapters are subdivided, and each subdivision is marked with its own heading (a **Level 2 heading**). Headings for subsections of these are called **Level 3 headings**; and headings for the subsections of these are called **Level 4 headings**.

The table below shows how you should format your headings. These formats are defined as styles in the Word thesis template.

Level	Name	Header Formatting
1	Chapter heading	<p>Chapters always begin a new page. The chapter title, along with the prefix "Chapter n", where n is the number of the chapter, should be centered and placed 1 1/2 inches down from the top of the page. If the chapter title is more than one line long, it should be double-spaced. After the chapter title, <i>quadruple spacing precedes the text</i>. The chapter title uses uppercase and lowercase letters, with no underlining or boldface, as in the following example</p> <p>Chapter 1 Introduction</p> <p>In a word processor, the Level 1 heading can be formatted as follows: Font: Arial, 16 point, regular Alignment: Centered Paragraph: 18 point before, 36 points after, page break before, double spaced</p>
1	Appendix	<p>This is like a chapter heading, except that the heading begins with 'Appendix n', where n is the number of the appendix.</p>
2	Level 2 headings	<p>The level 2 headings are centered using uppercase and lowercase, with <i>double spacing</i> after the heading. For example,</p> <p>Data Model</p> <p>In a word processor, the level 2 heading can be formatted as follows: Font: Arial, 12 point, regular Alignment: Center Paragraph: 12 point before, 24 point after, single-spaced</p>
3	Level 3 headings	<p>The level 3 headings are also uppercase and lowercase with <i>double spacing</i> below, and appear flush to the left margin.</p> <p>In a word processor, the Level 3 heading can be formatted as follows: Font: Arial, 12 point, regular Alignment: Left Paragraph: 12 point before, 24 point after, single-spaced</p>
4	Level 4 headings	<p>Level 4 headings are run-on heads, beginning a paragraph. They appear in lowercase, except for the first letter, and</p>

		<p>they are followed by a period. The Level 4 headings are underlined in order to separate them from the rest of the paragraph.</p> <p>In a word processor, the Level 4 heading can be formatted as follows:</p> <p>Font: Times New Roman, 12 point, regular, underlined</p> <p>Alignment: Left</p> <p>Paragraph: 0 point before, 0 point after, single spaced</p>
--	--	--

Table 2 – Heading Formats

Examples of subsection headings

The example below shows the use of the three subheading styles within a chapter. All subheadings appearing in the thesis must appear in the table of contents and must be formatted accordingly.

Reasoning and Problem Solving [Level 2]

We have previously considered the formation and use of concepts, noting that concepts are employed in thinking and that they are developed through thinking. We now consider the process of reasoning and problem solving.

Computer Simulation [Level 3]

The computer is an assembly of mechanical and electronic components that can engage in symbolic processes. Given some numbers, letters, or other symbols, the computer manipulates them in various ways and produces a result.

Use of computers [Level 4] In a broad sense, computer techniques and the programs associated with them are used for two general purposes. They are primarily designed for solving problems. But psychologists and other scientists also use computers for another purpose – to study the reasoning process.

Figure 2 -- Heading Examples

Use of Tables and Figures

Displays such as tables and figures can play a significant role in the presentation of information. Tables simply organize textual or numerical material into parallel columns, while figures are *graphical* or *pictorial*. The purpose of these displays is to amplify but not to repeat the text. The text should indicate the main points of the topic in question; further details are presented in the display. Tables and figures should be placed as soon as conveniently possible after their first mention in the text. If it is small, it may appear on the same page; if large, the next full page may be appropriate. However, if there are many tables or figures, they are sometimes presented in an appendix following the text of the thesis, if inserting them into the text would be too disruptive or distracting. If they appear on the same page with the text, *double space both before and after the table or figure* to clearly distinguish it from the text. All tables or figures interleaved with text should be counted in the pagination.

All table or figures should have a caption that includes a title (Figure or Table), a number, and a title (e.g., Figure 4 The object model for the Casino Gambling System). The caption should be centered beneath the table or figure. Tables should be numbered serially throughout the thesis with a list of tables shown in the front matter of the thesis. Figures should be treated similarly, but with numbering independent of the table numbers.

Each table or figure must be connected with a specific portion of the text in a systematic fashion, and labeled with sufficient clarity to be understood independently of the text description. In the APA style the reference commonly occurs in parentheses at the end of the sentence, like any other reference. Thus it is available to the interested reader, but it does not interfere with the flow of the text, as in the headings example above (Figure 2).

Photographs and other illustrations (e.g., maps, manuscript pages, musical scores, or autographs) included in the thesis should be scanned in when possible and presented as figures.

Typing and Format

The final manuscript should be prepared using a word processor, using double spacing, and printed to a laser printer using thesis-quality paper. The thesis should be printed on one side of the page only. The bound copy submitted to the Extension School may be either an original or a thesis-quality photocopy. The font must be 10- or 12-point. The manuscript must be free of mechanical errors and meticulously proofread – manual correction of mistakes to the page is not acceptable.

If your printer does not have an em dash (–), then use two hyphens (--) to simulate it. Use one space on either side of the em dash or hyphens ("this – that").

Margins and Pagination

The left-hand margin should be 1 1/2 inches. For the thesis document, this will accommodate the binding. The top, bottom, and right-hand margins should be one inch (except on pages with new chapter headings or other major headings, requiring a 1 1/2-inch top margin). The right margin should not be justified, unless your formatting program and printer have proportional spacing (most do). Otherwise, a ragged-right margin is preferred.

The first page of the main body of the text is always page one (after the title pages, table of contents, etc). Page numbers are centered at the bottom of each page, 1/2 inch from the edge. It is customary to omit the page number from a page containing a new chapter heading; however, the number may be centered at the bottom of the page (1/2 inch from the edge) if the student wishes to retain the page number.

Thesis Paper

The bound thesis must be printed or electrostatically photocopied onto a high-quality, long-lived, and durable paper, 8 1/2 x 11 inches in size and at least 16-pound

substance. All pages of the thesis, irrespective of printing procedure, must be a high-contrast, dark image on white paper. For typed or offset printed copies, the best available paper is Crane's Thesis Paper, a 100% cotton rag content, acid-neutral paper of long life expectancy and high durability. Other papers such as Permalife, Perma-Dur, Hollinger Acid-Free, and Xerox XXV Archival Bond also work well. *Erasable or corrasable paper is not acceptable for any part of the thesis.* Students should make sure that the photocopying agency they use to reproduce the thesis uses a paper that is at least 25% rag content, acid-neutral, and of permanent, durable quality. Paper lacking any of these attributes is not satisfactory thesis paper.

Appendix 2 Guidelines for References and Citations

All thesis proposals and thesis documents for the ALM in IT program will follow the American Psychological Association (APA) method for documenting sources. For all other issues concerning content and format of both the proposal and the thesis, students must follow the guidelines specified in this *Guide to the ALM in IT Thesis*. You should obtain the *most current edition* of these guides since each is subtly different from its predecessor; forms that were acceptable five years ago are not likely to be acceptable at the present time.

Bibliographies and References

Recall that the thesis proposal has a bibliography, which contains a list of references to material related to your work whether or not you have cited them in your proposal. By contrast, the thesis document will have a list of references only to those works that are cited in the thesis. The APA method requires that books, articles, and other materials appearing in the references be listed alphabetically according to the author's last name. Where there is more than one author of a work, it is the first author's last name that determines where the entry appears.

Sample Entries

Berners-Lee, T., Hendler, J., and Lassila, O. (2001). The Semantic Web, *Scientific American*, New York, 284, 5, 34-43.

Hammer, M., and Champy, J. (1993) *Reengineering the Corporation*, Harper Collins, New York.

Paithankar, A. (1996). *AINT: A Tool for Simulation of Shared-Memory Multiprocessors*, Master's Thesis, University of Colorado, Boulder.

Note that in the APA method, only the author's last name is spelled out in full; first and middle names are given as initials. The publication date follows in parentheses. In the title of an article or a book, only the first letter and proper nouns are capitalized. The title of the journal itself follows the usual rules of capitalization. The place of publication comes next. In the case of a journal article, the volume number, which is underlined, and the page numbers occur at the end of the entry.

To cite the Berners-Lee article in the text, one would have "According to Tim Berners-Lee <some quote or paraphrase> (Berners-Lee, Hendler, & Lassila, 2004)."

Care should be taken with references to articles or papers taken from the Web.

Often these have no date of publication and no author. Even in these cases, it is not enough just to plop in a URL. Here are some examples taken from the sample proposal that show how the reference should be listed and how a citation of that reference should be inserted in the text:

References:

Martin, R. (1994) OO Design Quality Metrics - An Analysis of Dependencies, Position Paper, Workshop on Pragmatic and Theoretical Directions in Object-Oriented Software Metrics, OOPSLA'94, <http://www.objectmentor.com/resources/articles/oodmetrc.pdf>, retrieved December 2005.

Schneider, J., & Mergenthaler, M. (2005) Code Analysis Plug-In (CAP), <http://cap.xore.de/>. retrieved December 2005.

The Eclipse Foundation (2004). Eclipse Platform Plug-In Developer Guide. <http://help.eclipse.org/help30/index.jsp>, retrieved December 2005.

Note the specificity of the URL reference to Martin's paper – using only <http://www.objectmentor.com> leaves the reader having to search for the particular article; imagine how difficult it would be for a user to retrace your steps if you had gone to an internal page of the Sun Microsystems website to find a tutorial on the latest JSR, but left the reader only a reference to <http://www.java.sun.com>.

References should be formatted with single-spacing within entries and double-spacing between entries.

Example Citations:

a) Martin introduced the concepts of instability and abstractness as ways to measure whether the module structure of an application is appropriate (Martin, 1994).

b) This application will be written as an Eclipse Plug-in. See (The Eclipse Foundation, 2005) for a tutorial on how to develop plug-ins.

c) The Code Analysis Plug-In (Schneider & Mergenthaler, 2005) is another example of a metrics tool written as a plug-in for Eclipse.

Citations

There are no endnotes or footnotes as such in the APA method; instead, one uses citations made in brief form directly in the text. This brevity constitutes one of the major advantages of the APA approach to documentation. After each citation, the author's last name, or the names of all the authors, and the date of the publication are included. Thus the citation for a work written in the year 2000 by an author named Green would be (Green, 2000). There are a few simple rules to follow that cover other cases.

When a citation refers to two or more works by different authors within the same parentheses, all with reference to the same topic or point in the text, then use semicolons to separate them (Able, 1980; Baker, 1982; Charlie, 1984).

When there are two or more authors of the same work, an ampersand (" & ") is used before the name of the last author, as in (Berners-Lee, Hendler, & Lassila, 2001).

After the first citation for three or more authors, in subsequent citations use only the name of the first author followed by "et al" and the year of publication; e.g., (Berners-Lee et al., 2001). For all citations, as demonstrated in the prior examples, the final punctuation appears outside the parentheses.

If the same author has published more than once in a single year, differentiate the separate works by appending a lowercase letter ("a," "b," and so forth) to the date. Thus if an author named Smith had published two articles in 2001, and both were to be cited within the text, we would have (Smith, 2001a) and (Smith, 2001b). These same designations would also be matched properly with the reference in the list of references. If both articles are cited within in the same parentheses, the author's name appears once and the year references are separated by commas, as in (Smith, 2001a, 2001b).

Answers to further questions about APA form can be found in the fifth edition of the *APA Publication Manual* (2002), on reserve in Grossman Library.

Here's an example of how a citation would appear in one of your documents. Suppose your thesis proposal is based on your readings about ontology and you want to relate this to the concept of the semantic web. You could insert an extended quotation with a proper citation as follows:

Tim Berners-Lee and others (Berners-Lee, Hendler, & Lassila, 2001) have given strong indications as to how they see the current World-Wide Web evolving:

To date, the Web has developed most rapidly as a medium of documents for people rather than of information that can be manipulated automatically. By augmenting Web pages with data targeted at computers and by adding documents solely for computers, we will transform the Web into the Semantic Web.

Appendix 3 Common Problems In Grammar, Punctuation, And Usage

This Appendix has some examples and guidelines for dealing with English usage. The first section covers some terms specific to IT. The second section deals with common errors in English. Be sure to run a spell and grammar checker on your thesis document before submitting it. Note that not all errors flagged by the best spelling checkers are correct and in some cases errors will not be caught.

Terms Related to IT

Since IT evolves somewhat rapidly and uses many novel terms that aren't found in current dictionaries, you have to be careful to use the novel terms correctly. This section addresses some IT-specific issues with hyphenation and with the use of trademarked names.

Hyphenation

In IT we use a lot of terms that are combinations of two or more words. Sometimes these require hyphens, sometimes they are left as separate words, and sometimes the words are concatenated. It often depends on whether the combined words are used as an adjective or not. Here are some examples. When in doubt, look things up in a recent dictionary or use a Web resource such as dictionary.com.

email, offline, online, website, webpage	<p><i>From dictionary.com</i></p> <p>Usage Note: The transition from <i>World Wide Web site</i> to <i>Web site</i> to <i>website</i> seems to have progressed as rapidly as the technology itself. The development of <i>website</i> as a single uncapitalized[sic] word mirrors the development of other technological expressions which have tended to evolve into unhyphenated forms as they become more familiar. Thus <i>email</i> has recently been gaining ground over the forms <i>E-mail</i> and <i>e-mail</i>, especially in texts that are more technologically oriented. Similarly, there has been an increasing preference for closed forms like <i>homepage</i>, <i>online</i>, and <i>printout</i>.</p> <p>[Note: 'uncapitalized' is not a word according to both the Microsoft Word spell checker and dictionary.com itself]</p>
--	---

run-time vs. runtime	When used as an adjective, use a hyphen; otherwise, use 'runtime'. "Polymorphism requires run-time binding rather than compile-time binding." "The runtime of the algorithm in these circumstances is of order n^2 ."
real-time vs. real time	When used as an adjective, use a hyphen, otherwise separate the words "The real-time system..." or "The data is analyzed in real time". Note that 'realtime' is not a word.

Trademarked Names

Avoid creative spelling of trademarked names. In particular, pay attention to the correct spelling of official names for companies, their products, and product components. Don't let 'Java' drift down to 'java', 'C#' become 'c#', and so on. Check the documentation of the product or company you're referring to for their correct spelling and stick to it.

General Errors in English

The following is a list of common errors in usage.

Affect and Effect. These two words differ by only one letter, but have completely different meanings. As a verb, to affect something means to influence it ("The cold weather affected his health"), or else to pretend something ("He affected indifference to the cold weather"). To effect something, on the other hand, means to accomplish it ("The cold weather effected his speedy departure"). As a noun, "affect" is used almost exclusively as a psychological term meaning strong emotion, whereas "effect," a far more common noun, means a result ("His affect was an effect of the cold weather").

Capitalization. Besides capitalizing all proper nouns, one should also capitalize abstractions used to represent a proper noun. For example, when speaking about "the Catholic Church," one can simply use "the Church." One should maintain a distinction between "the church on the corner" (a physical object) and the abstract concept of a "Church."

A Century Used Adjectivally. When used as an adjective, the form is "nineteenth-century" and "twentieth-century," as in "nineteenth-century poetry" or "twentieth-century architecture."

Changes or Additions in Quotations. Any change in, or addition to, a quotation should be represented within brackets: "Any change in, or addition to, a [direct] quotation should be represented within brackets."

Commas in a Series. Commas should appear after all the items in a series, including the penultimate item: "Run, jump, and throw."

Consistency in Abbreviations. Be consistent with abbreviations: either US or U.S.; either USSR or U.S.S.R. Abbreviations occurring within direct quotations are left

as they appear in the quotation.

Digitizing Numbers. The numbers "ninety-nine" and below are written out. The numbers 100 and above are digitized. For instance, "If I've told you once, I've told you fifty times not to write 99."

Ellipses. If you omit something from a direct quotation, you must indicate the omission with ellipsis points, i.e., three spaced periods: "If you omit something from a direct quotation, . . . indicate the omission with ellipsis. . . ." If the ellipsis occurs at the end of a sentence (as in the example above), four spaced periods must be used (one of which is a sentence period).

Fewer and Less. "Fewer" applies to units that you can count, "less" to a quantity that you can measure. If what you are counting is made up of individual things such as milk bottles or people, use "fewer" ("Fewer people drink milk in France than in America"). If what you are counting does not have individual units but is a collective noun such as "milk" or "butter" or "money," then use "less" ("People drink less milk in France than in America"). The phrase "less people" is incorrect, since you can count people. The phrase "fewer money," for obvious reasons, is also incorrect.

Identify What "This" Is. Avoid beginning a sentence with an unspecified "this": "This led to unspeakable atrocities." Be specific about what "this" might be: "This issue came before Congress."

Identifying the Author of a Quotation in Your Text. It is important that you identify in your text the author of any quotation you use. Quotes should never be loosely and randomly displayed in the text, without proper attribution.

Identifying Centuries. Write out the names of the centuries: "nineteenth century," "twentieth century," "twenty-first century," *not* "1800s", "1900s," or "2000s"

Imply and Infer. "To imply" means to suggest or to hint at something ("She implied that she would be late"). "To infer" means to draw a conclusion or to make a deduction based on an implication ("He inferred from what she said that she would be late"). Put another way, "to imply" means to send a veiled message, while "to infer" means to interpret one.

Incomplete Sentences. Complete sentences always contain a subject and a verb. Rarely, and principally for emphasis, an accomplished writer may use a sentence fragment. But for a professional writing style in a thesis, sentences should be complete.

Insolvable and Insoluble. "Insolvable" means "incapable of being solved," as in the sentence "This puzzle is insolvable." "Insoluble," on the other hand, means "incapable of being dissolved," as in "Lead is insoluble in water." You could, of course, say "This puzzle is insoluble," but that would simply mean that it does not dissolve in liquid.

Introductory Adverbial Clauses. The rule for introductory adverbial clauses is to

set them off from what follows in the sentence with a comma: "When Irwin was ready to eat, his cat jumped onto the table." Notice what would happen to the meaning of this sentence if the comma were omitted. If, however, the clause is short and no possible ambiguity exists, a comma can be omitted after an introductory adverbial clause: "In 1919 the poet toured America."

Its and It's. One of the most common mistakes in ALM in IT theses is the confusion between "its" and "it's." "Its," with no apostrophe, means "of it" ("The cat ate its food"). "It's," with an apostrophe, is a contraction meaning "it is" ("It's about time that the cat ate its food"). One reason for this mistake is that possessives formed from proper nouns, such as "John's" or "Mary's," end in apostrophe "s." But possessive pronouns do not. Please remember, whenever you write "it's," you are really writing "it is." When you mean to say "of it," use "its."

Long Quotations. Single space and indent quotations of more than four typed lines. Indent one inch from the left margin, and omit the quotation marks. A colon generally introduces a quotation displayed thus.

Methodology. This word means the study of method or methods ("Professor Jones has a strong interest in methodology in nutrition"). It should not be used in place of the word "method" ("Professor Smith used an incorrect method"). Using "methodology" when you mean "method" is not only inaccurate, but also sounds pretentious. And most of the time you will almost certainly mean "method," since it is a far more common noun.

Modification of Superlatives. Superlatives such as "unique" (meaning "the only one") should not be modified. Something cannot be "almost unique" or "very unique"-it is either unique or not. Qualifying superlatives is equivalent to using "most" with adjectives ending in "-est"; it would be an error to write "most biggest." By the same token, you should not write "most unique." Other superlatives that should never be modified include "endless," "final," "foremost," "highest," and (despite the example offered by the preamble of the U.S. Constitution) "perfect."

Nonsexist Language. Despite the practice of grammarians since the eighteenth century, it is no longer considered appropriate to use the masculine pronouns "he," "him," and "his" as generic pronouns because of their obvious exclusion of women. Instead, one should use "he or she" or "him or her" or "his and her": "Each student should thoroughly research his or her topic." If this form seems awkward, it can be avoided by using a plural noun and a plural pronoun: "Students should thoroughly research their topic." For further assistance, consult the MLA publication *Language, Gender, and Professional Writing: Theoretical Approaches and Guidelines for Nonsexist Usage*, by Francine Wattman Frank and Paula A. Treichler.

Run-On Sentences. Avoid run-on sentences, which are two sentences thrust together without any intervening punctuation or without the use of a coordinate conjunction between them: "I finished my thesis I had it bound."

Significantly. This adverb has a technical definition in the sciences. It does not

mean "highly" or "very," but rather a statistically significant association between two facts or occurrences (one whose chance of random occurrence is .05 or less). So a sentence such as "The measured performance of data retrieval using the algorithm described in this thesis is significantly better than the previously used algorithm" implies that you can demonstrate a statistical difference in the performance of the two algorithms. If this is not your intention, be sure to use the word in such a way as to differentiate it from a statistical context.

That and Which. The relative pronoun "that" introduces restrictive clauses containing information necessary to the correct meaning of the main clause: "An office manager for a corporation that had government contracts asked her supervisor if she could reprimand her co-workers for smoking." The relative pronoun "which" introduces non-restrictive clauses containing extra but unnecessary information about the noun that precedes it: "The Saudis buy spare parts, which cannot be produced locally, from Great Britain." A comma precedes the use of "which," while no comma precedes "that."

Their, There, and They're. These three words are frequently confused. "Their" is possessive ("Their house is white"). "There" is demonstrative ("Their white house is over there"). "They're" is a contraction for "they are" ("They're in their white house over there").

Word Hyphenation. Students are often careless about where to place hyphens when dividing words at the end of a line. Words should be divided only at the end of a syllable, as defined in a standard English dictionary (such as *Webster's* or *American Heritage*). Thus the word "syl-la-ble" has only two places where it can be divided correctly; to divide it as "syll-a-ble" would be incorrect and therefore unacceptable. Please consult a standard dictionary if you are in doubt about where to divide a word. Some words whose second syllable contains only a silent "e" cannot be divided at all. Examples include: "tamed," "played," and "fanned." Fortunately, most modern word processors provide built-in capabilities for hyphenating words on your behalf. You need to be careful that words specific to Computer Science and Information Technology are not hyphenated incorrectly, as described in Section "Terms Related to IT" above.

Appendix 4 Bibliography

Helpful discussions of research, writing, style, and organization can be found in the works listed below. All are on reserve in Grossman Library.

Barzun, J. (1975) *Simple and Direct, A Rhetoric for Writers*. New York, Harper and Row.

Chicago University Press. (1993) *Manual of Style*. 14th ed. Chicago, U of Chicago Press.

Harvey, G. (1996) *Writing with Sources, A Guide for Harvard Students*. Rev. ed. Cambridge, President and Fellows of Harvard University.

Howard, V. A., and Barton, J. H. (1986) *Thinking on Paper*. New York, William Morrow.

Light, R., and Pillemer, D. (1984) *Summing Up, The Science of Reviewing Research*. Cambridge, Harvard Univ. Press.

Mann, T. (1987) *A Guide to Library Research Methods*. Oxford, Oxford Univ. Press.

American Psychological Association (2001) *Publication Manual of the American Psychological Association*. 5th ed. Washington, D.C.

Strunk, W., Jr., White, E. B, Angell, R.. (2000) *The Elements of Style*. 4th ed. Pearson Higher Education, Upper Saddle River, NJ .

Turabian, K.L. (1982) *A Manual for Writers of Term Papers, Theses, and Dissertations*. 4th ed. Chicago, U of Chicago Press.

Writing Handbook (1997), <http://karn.ohiolink.edu/~sg-ysu/writehb.html>, retrieved in 2006.