Information Organization Project
Fall 2015

General Description .................................................................................................................. 2
  Introduction .......................................................................................................................... 2
  Goal and objectives .............................................................................................................. 2
  Overview ............................................................................................................................... 3
  Writing and research ........................................................................................................... 3

Project Drafts ........................................................................................................................... 5
  Draft development process .................................................................................................. 5
  Topic approval ...................................................................................................................... 5
  Draft contents and deadlines ............................................................................................... 5
  Draft grading ........................................................................................................................ 6
  Draft submissions ................................................................................................................ 6
  Graded draft returns .......................................................................................................... 6
  Report Format ...................................................................................................................... 6

Project Report ........................................................................................................................ 8
  1. Project description .......................................................................................................... 8
    1.1. Collection and information objects ......................................................................... 8
    1.2. Users’ demographics and knowledge ..................................................................... 9
    1.3. Users’ problems and questions ............................................................................... 9
  2. Representation of information objects ............................................................................ 10
    2.1. Entity level ................................................................................................................ 10
    2.2. Metadata elements and semantics ......................................................................... 11
    2.3. Record structure and specifications ....................................................................... 11
    2.4. Record content and input rules .............................................................................. 12
  3. Access and authority control ............................................................................................ 13
  4. Representation of information content ............................................................................. 14
    4.1. Subject access .......................................................................................................... 14
    4.2. Thesaurus structure .................................................................................................. 14
    4.3. Classification scheme ............................................................................................... 15
  5. Name authority control .................................................................................................... 17
  6. System evaluation and development ................................................................................. 17
    6.1. Performance test ....................................................................................................... 17
    6.2. Change and Development ....................................................................................... 19
  7. Project summary ................................................................................................................ 20

Project Appendixes ................................................................................................................. 21
  Appendix A. Metadata elements and semantics ................................................................. 21
  Appendix B. Record structure and specifications .............................................................. 22
  Appendix C. Record content and input rules .................................................................... 23
  Appendix D. Sample thesaurus .......................................................................................... 24
  Appendix E. Classification scheme ..................................................................................... 25
  Appendix F. Name authority file ......................................................................................... 27
  Appendix G. Sample records .............................................................................................. 28

(Click section heading to jump inside.)

All SLIS 5200/4200 course materials are copyrighted and may not be copied, revised, or distributed in any form or venue, beyond their use by students for purposes of fulfilling course requirements, without prior permission of the authors or the University of North Texas.
General Description

Introduction

This assignment provides the opportunity to synthesize the knowledge gained in this course. It brings together all aspects of organizing a particular collection of information objects for a specific set of users, including hands-on experience with appropriate database software. The deliverable for this assignment is a written report in the form of a technical manual that describes the organization system and its implementation in a specific database management system (DBMS).

The project is developed in four stages during the semester, each aligned with class lectures, online modules, and readings. Students turn in drafts of the project report periodically for instructor review and comment, and then revise each draft before moving on to the next.

IMPORTANT: You must read closely all directions for this project. Following these directions will help you to develop a complete and accurate project and reduce unnecessary revisions.

Goal and objectives

The goal of the project is to engage you in planning and problem solving as you develop and implement a suitable organization system for a collection of information objects.

The objectives are for you to:

- Gain realistic experience in organizing a collection of information objects
- Demonstrate understanding of the concepts and principles of information organization
- Demonstrate ability to synthesize and apply these concepts and principles in the context of organizing a real collection for a particular set of users
- Demonstrate basic skills in implementing an information retrieval system to provide access to information objects in the collection
- Demonstrate basic skills in writing succinct explanations and technical instructions

The course exposes you to a range of challenges and decisions for organizing information. The goal is not to produce the perfect organization system, but rather to have you encounter and wrestle with key problems when designing and implementing a useful information organization system.

This assignment assumes you have information technology knowledge and skills. These include a basic knowledge of databases and concepts of files, records, fields, and indexes. You will learn and use Inmagic, a DBMS program, for implementing the database component of the project. A tutorial is available to introduce the features of Inmagic. You are expected to learn Inmagic to fulfill the requirements of the assignment. An Inmagic Frequently Asked Questions (FAQ) is also available on the course Learn site.

IMPORTANT: These instructions represent the MINIMUM level of effort necessary to secure an AVERAGE grade, and should be taken as a rough outline of your paper. Exemplary (grade of A) work requires effort, initiative, depth of thought, mastery of the material, and an ability to explain concepts and all their inter-relations in a manner well beyond the outline given in these guidelines.
Overview

The emphasis of this assignment is on the planning and intellectual effort that is necessary to organize a collection of information objects. During this project, you will:

- Identify a collection to organize for a specific set of users
- Build an organization system for the collection
- Implement the organization system in an information retrieval system
- Write a project report in the form of a technical manual for your systems
- Submit report drafts for review
- Revise report drafts according to instructor feedback

You begin by identifying a collection of information objects and a group of users who will access the collection. Assume this is a large collection, with a size (real or hypothetical) of at least 1,000 items. Choose 10 (real) objects to serve as a representative sample, or subset, of the full collection. Refer to the sample objects as you complete the steps in this assignment, but develop the organization system as if the full-size collection exists. You are expected to gear the design of your organization system toward meeting the needs of the users you have identified.

In a series of drafts throughout the semester, you will write a report describing the information organization system as you develop it. The system will include an information retrieval system (database) to provide access to data about the collection, along with input rules, thesaurus, classification scheme, and name authority file to control content and consistency of data in the system. All of these will give you direct experience with creating fundamental mechanisms for information organization. The final report consists of a tightly written narrative and set of associated appendixes. You must complete and pass each draft before you are allowed to proceed to the next. Except for Draft 1, no resubmissions will be allowed.

Several days prior to submitting each draft to your instructor, you will participate in a peer review exchange with a classmate. The student reviewer's comments, based on a simple checklist, will give you a chance to improve the draft before submitting it to your instructor. After your instructor approves your topic in Draft 1, you will post a brief description in Learn for everyone to read. Finally, after the formal review of each draft, you will receive the instructor's comments to guide your revision for the subsequent draft. These activities provide a maximum amount of feedback from both instructor and classmates to help you get the best grades possible.

Writing and research

The purpose of this project is to give you conceptual and hands-on experience with major processes and tools of information organization. The report is a vehicle for you to demonstrate your understanding of the concepts related to information organization. Throughout the report, it is vital that you use your own words to explain, define, and demonstrate your mastery of the concepts. Do not take wording from course materials: doing so fails to demonstrate your understanding of the concepts and constitutes plagiarism, which is illegal and unethical. Occasionally you will be directed to format part of the report exactly as in the instructions. This is format only; except for those instances, do not copy the assignment. When in doubt, ask your instructor.

The project report is in the form of a technical manual describing your organization system as if it is complete and operational. Part of your grade depends on the readability, clarity, and conciseness of this manual. With the information you provide, another person should be able to continue the organization process by following your description, specifications, rules, and examples. The reader must be able to understand not only the technical details, but also the rationale and decisions behind the organization system. All of this requires the basic skills necessary for technical writing. Technical writing style is clear, succinct, and impersonal, without creative embellishments. The assignment provides specific instructions for technical writing, while Tips for Coursework (available in the Learn course site) provides general guidelines.

The assignment instructions call for a strict technical manual format, consisting of specific sections in sequence. By following the prescribed format, you will not be distracted by the task of organizing your report in addition to organizing your collection. For some sections with appendixes, the instructions say to create the appendix first, before writing the narrative. This is typical of technical writing; it is necessary because the narrative summarizes the details shown in an appendix. Be prepared to skip back and forth between sets of instructions.
For each section of the report, the instructions provide an overview, list of tasks, and description of the content required in the narrative for the section. Before beginning a report draft, read the assignment for all sections of the draft. Attention to detail is as critical for composing the report as it is for organizing information: be sure to address every part of the instructions.

Guidance for completing the project is available in course materials; external research is not required. Many materials available in Learn, including lecture slides, content modules, tutorials, and help documents, along with the readings, relate directly to the project. It is important that you read course materials systematically as you develop your report drafts. Communications in class and in Learn provide further support throughout the semester. It is also important that you check the Learn course site daily for new messages and participate in online chats and discussions. Together, the assignment, readings, and communications will help ensure your success with this project.

IMPORTANT: These instructions represent the MINIMUM level of effort necessary to secure an AVERAGE grade. Exemplary (grade of A) work requires effort, initiative, depth of thought and explanation well beyond these guidelines. If your paper shows no depth of thought beyond “checklisting” these minimum requirements, you will not receive an A. A good rule of thumb to remember is “if you put in minimal work, you will get a minimal grade.”
Project Drafts

Draft development process

During the semester you will write four drafts of your project report. Each draft except Draft 1 will have a preliminary and a final version (7 submissions total). The preliminary draft (for Drafts 2, 3, and 4) goes to another student for peer review and the final draft goes to the instructor for grading. After Draft 1, each subsequent draft consists of:

- A revised version of previous sections based on instructor’s comments and your insights
- All sections and appendixes due for the current draft

Topic approval

The instructor must approve your choice of topic (collection and information objects) to ensure that it is appropriate for the assignment. This approval generally takes place at institute or in the first couple days of class in face-to-face sections.

Draft contents and deadlines

Each draft requires specific sections of the project report to be submitted. The contents of each preliminary and final draft are detailed below. You may fill in the due date column from your Semester Schedule in Learn.

IMPORTANT: No extensions will be granted for this project.
IMPORTANT: Preliminary drafts that are substantially incomplete will not be accepted and you will forfeit the peer review process for that draft.

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Content</th>
<th>Version</th>
<th>Due date (fill in yourself)</th>
<th>Evaluation</th>
<th>% of grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Draft 1</td>
<td>Section 1</td>
<td>Final</td>
<td>Instructor approval of topic</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td></td>
<td><em>(Post description in Learn Discussions)</em></td>
<td></td>
<td></td>
<td>1*</td>
<td></td>
</tr>
<tr>
<td>Draft 2</td>
<td>Revised section 1 plus section 2 and its appendixes</td>
<td>Prelim.</td>
<td>Peer review feedback</td>
<td>3*</td>
<td></td>
</tr>
<tr>
<td>Draft 3</td>
<td>Revised sections 1 and 2 and their appendixes plus sections 3 and 4 and their appendixes</td>
<td>Prelim.</td>
<td>Peer review feedback</td>
<td>3*</td>
<td></td>
</tr>
<tr>
<td>Draft 4</td>
<td>Revised sections 1-4 and their appendixes, plus sections 5-6 and their appendixes (Preliminary Draft), plus section 7 and all previous material (Final Draft)</td>
<td>Prelim.</td>
<td>Peer review feedback</td>
<td>3*</td>
<td></td>
</tr>
</tbody>
</table>

*Score for submitting a preliminary draft and participating in the peer review process, not for quality of IOP content. See Peer Review Process in Learn.
Draft grading

The purpose of submitting several drafts is to constantly improve the quality and consistency of your work. The IOP is a tightly structured technical manual with interrelated parts. Statements you make in section 1 affect later sections, and decisions you make later may require revision of earlier sections. It is vital that you address the instructor’s comments and corrections on each draft. Each draft is graded as a whole, which means that points will be deducted for uncorrected errors from a previous draft. If you have not attempted to correct major omissions or problems from the previous draft, the instructor may refuse to grade the current draft until you resubmit it with corrections. Resubmitting the draft will delay grading and possibly lead to a late penalty.

Draft submissions

Submit each draft as a single file in MS Word (.doc) attached to a Learn mail message (see Mail Attachments Guide in Learn). Write a descriptive Subject line for the message, such as “Preliminary Draft 1” or “Final Draft 2” (your name is unnecessary here). Send the mail message with attached file to 5200 Central.

We will note the time your submission was sent and check the format.

We will email you within 24 hours if:

- The file will not open or is not readable.
- Your name is missing in file name or at top of page 1.
- Draft number is wrong in file name or at top of page 1.

IMPORTANT: Check your email the day after you submit a draft to see if there are problems. Failure to have a draft accepted on initial submission may result in a late review of the draft by the instructor. This may delay your progress on the subsequent draft.

Graded draft returns

Draft 1 is returned online or in person, depending on your course section and instructor. Drafts 2 and 3 are graded and returned by email as soon as possible (generally about a week after the due date) to allow you to incorporate changes in the subsequent draft. Draft 4 is emailed shortly after final course grades are submitted to UNT.

Failure to complete the IOP project may result in being assigned a final class grade of “F”. You must complete and pass each section of the IOP in order to proceed to the next. All assignments and due dates are mandatory. Failure to submit a draft will result in a grade of “0” for that assignment, and you will not be allowed to proceed. Failure to correct drafts as per instructor’s comments and suggestions will result in further point deductions on the next draft.

Report format

The following file name and internal format conventions are required in order to help us manage and grade submissions. These instructions assume you understand basic word processing functions.

File name: Each project draft must be a single MS Word (.doc) file. After Final Draft 1, you will keep adding to the same document but rename it for each draft. Begin the file name with your last name, using this style:

    LastnameDraft#Prelim.doc  LastnameDraft#Final.doc

Example: MoenDraft2Prelim.doc  MoenDraft2Final.doc

Report format: The IOP report is strictly formatted with prescribed sections and appendixes. This format is realistic, saves you the trouble of writing your own section headings, and helps us find sections for grading. You set up the report format only once, because each draft simply expands the previous draft. You must download and use the IOP Template (available on the Learn course site) which follows the format guidelines below. Failure to use the template will result in failure of the assignment.
Margins: One-inch margins on all sides and half-inch margin for header.

Type: Body text (narrative) in Arial 10 point or Times Roman 12 point. All body text left-justified only. Headings boldfaced and one or two point sizes larger.

Page 1: Do not create a separate title page (wastes paper). At top of page 1, type four lines (do not insert a header) and right-justify the lines. Use the style below. Complete the information shown in square brackets (remove the brackets).

For Spring 2015 Web Institute and Thursday night sections, the section identifiers are:

TXWI_A, TXWI_B, TXWI_C, LUBBOCK, DENTON* and ELPASO

*Thursday night section only

Example:

William Moen
Fall 2011
SLIS 5200 TXWI_B
Preliminary Draft 2

Title: Center your project title (name of collection) above the first report heading, like this:

Denton County Historical Society:
Information Organization System

Header: When using the IOP Template, click View, Header and Footer, and then replace bracketed placeholder labels with your last name and section identifier as shown below. This is a true running header that will automatically print on all except page 1: do not type it on each page. If you create a new header, format it in the same font as body text, same point size or smaller, and right-justified.

Example:

[Lastname] / [section identifier] / p. [#]

Section headings: Headings should be worded and ordered exactly as in IOP Template and this assignment. In Drafts 1-3, omit headings for parts not yet developed; you can paste these in from the template later.

Body text: Type narrative text single-spaced within paragraphs and double-spaced between paragraphs. Do not indent first line of paragraph. Write succinctly; provide only the information requested. Spell-check the draft. Consult Tips for Coursework for guidance on grammar, punctuation, etc.

Section order and length: Place narrative sections together at beginning of draft and all appendixes together at end. Length of each narrative section may vary from about 1/3 to several pages. Appendixes may be several pages long. Length of completed report (Final Draft 4) is usually about 25 to 40 pages.

Failure to follow the formatting listed above will result in points being deducted from the final grade of the respective draft.

Note: A section identifier is not the same thing as your enrollment section. In other words, “Section” means TxWI_A, not “SLIS 5200.002”
Project Report

1. Project description

Every aspect of information organization presents problems that the organization system attempts to solve. Problems stem primarily from challenges of describing and representing objects in a way that responds to users’ requests. Section 1 sets the stage for making design decisions throughout the project.

1.1. Collection and information objects

Here you identify a collection of information objects to work with. Some objects work better than others for this assignment. You may discuss ideas for information objects to use with the instructor, who will assess whether the objects are appropriate for the assignment. Avoid choosing highly personal collections of information objects because these may be difficult for you to envision in a public setting. Real collections at your place of work may be acceptable. Hypothetical collections (i.e., realistic but not necessarily existing) are also acceptable.

Tasks: Choose a collection of information objects to organize. These objects must have these two attributes: proper names (personal and/or corporate entities as creators, publishers, etc.) and identifiable subjects (information or intellectual content, topics, aboutness). If your collection contains objects in multiple formats (e.g., books, videos, etc.), limit it to one format for this project to reduce your time and effort.

Obtain a sample of 10 actual objects from your collection. These objects must be real and readily available: you will be handling and examining them as you develop your organization system. The sample should be representative (not random), that is, it should include a carefully chosen variety of objects (by subject, format, etc.) that represent a cross-section of the collection.

Determine the setting or location of the collection, which must be a public place such as a library, museum, or store. The setting can be real or hypothetical. If the collection is a subset of a larger real collection (as in a library), treat it as independent for the duration of this project; that is, do not try to integrate it with the rest of the existing collection. Think about the purpose that your particular collection serves for the institution.

Determine the size of the full collection. The size can be real or hypothetical, but for this project assume that the collection contains at least 1,000 objects. Consider how the collection typically grows, such as through purchase or donation.

Provide a descriptive name for the collection. This is the basis for the title of your report at the top of page 1.

Narrative: Start with an introductory sentence that states the name of the collection and its public setting (i.e., where it is housed; its institutional context). State the purpose of the collection and why it exists.

Expand with a description of the collection, its scope, and the types of information objects it contains. Indicate the objects’ physical aspects (e.g., formats) and intellectual aspects (e.g., topics covered). State the size of the collection (minimum 1,000, real or hypothetical). Explain how it will grow and change in the future.

The description should be general here because you will write detailed descriptions of the objects later in the report. Focus on a description of the collection and why it exists. Do not list the titles or names of individual objects.

IMPORTANT: Technical writing style for this report requires the following (with a few exceptions noted elsewhere). Use present-tense verbs (“is”) as if your information organization system is complete and operational, not future- or conditional-tense verbs (“will be,” “would be”). Use third-person pronouns (“it,” “they”) that do not personalize the text; do not use first- or second-person pronouns (“I,” “we,” “you”). Write in a gender neutral style (no “he”, “she”, “her”) Write clearly and succinctly, without creative flourishes or emotional adjectives and adverbs. See Tips for Coursework for additional advice on grammar and spelling. Be sure to avoid common errors such as noun-pronoun disagreement (e.g., not “user . . . they,” but rather “users . . . they” or “user . . . he/she”). Use of words such as I, me, our, will, would be, etc., will result in loss of points on any assignment.
1.2. Users’ demographics and knowledge

It is essential that you identify and describe your users at the outset because the information organization system you develop must respond to their needs and behaviors. The more familiar you are with these users in real life, the more concrete and accurate your description will be. Certain factors may affect their ability to request information and use the organization system. A clear understanding of the users now will guide your thinking for the remainder of the project. In section 6 you present a plan for evaluating how well the resulting system serves these users and carry out a test of your system with a user who is a member of your defined user group. Your group must have several hundred potential users.

Note: For this project, it is important to understand that “users” perform two different types of tasks. First are the end users of the collection, people who have information needs and who search the database and the collection. You are designing the information organization system for these users. Second are the technical users of the organization system, people who create and maintain the database and the records. In a traditional library or information center, end users are the patrons who browse the collection, use the catalog, and approach the information or reference desk. Technical users are staff personnel including database experts and indexers or catalogers. For various sections, you will adopt the viewpoint of one or the other type of user.

In this section, you identify and describe the end users of the collection. Later, in Appendixes C and F, you provide input rules for the technical user on how to create records for the database.

Tasks: Identify the types of people who use the collection. They may fall into one relatively homogeneous group, or into primary and secondary groups. Limit groups to two, for simplicity’s sake. If you have two user groups, describe each group separately. The demographic and knowledge differences (see below) between groups should be significant (e.g., adults vs. children). At least one group must search the database directly.

Describe the demographics of the group. Estimate the percentage of each gender, age range, education level, and any other demographic characteristics that are relevant to their use of the collection and thus relevant to the design of the organization system. Consider characteristics that may particularly influence their information needs, abilities, and searching behavior (e.g., education, occupation, language, ethnicity, culture).

Determine whether the users’ level of each type of knowledge (general, domain, system, information-seeking) is high, medium, or low (i.e., whether users are experts, novices, or in between). Address each type of knowledge separately. Explain how users’ types and levels of knowledge affect their ability to request information and use the system.

Narrative: Start with an introductory sentence that generally describes the user group. Describe the demographics of the group and identify specific demographic characteristics you feel are relevant to user behavior and thus relevant to the design of the organization system. Explain why certain characteristics are important.

Define the four types of knowledge and then discuss the levels of each type of knowledge of the user group.

Explain how users’ demographics and their levels of knowledge serve as a basis for the design of your system.

You should have one user group.

1.3. Users’ problems and questions

This section describes why and how users request and interact with the collection through the organization system. You will repeat the user questions later in a system performance test in section 6.1. The information objects’ attributes will inform your decisions about metadata elements in section 2.2.

Tasks: Consider the kinds of situations that motivate users to seek information from the collection: their information problems or needs and how those are translated into requests to the system. Think about their purposes for using the collection (e.g., research for school assignments).

Brainstorm a list of typical users’ information problems and requests, or the kinds of questions they ask. To be a good test for the system, do not use questions that are answered with yes/no or a number. It should be possible to respond to each question by retrieving at least one of your 10 sample objects. For example, phrase a question for books by specifying the author’s name (e.g., “book by Hemingway,” not “book by certain author”).

Hint: To get ideas, have someone who might be a potential user ask you some questions.
Identify the attributes of the information objects that these user questions suggest. Attributes are general characteristics or properties (e.g., title, creator, size, format) of information objects. They are not specific data values. For example, in a request for a book "by Hemingway," the attribute is author or creator and the data value is "Hemingway." Attributes that are important to users, as indicated by their questions, suggest characteristics or properties of the objects that should be described in the system. Edit the list of questions to three or four that demonstrate the most variety in object attributes.

Consider also that user requests imply expectations for information system retrieval performance, such as levels of recall and precision. For each question, consider whether users desire information retrieval to be high, moderate, or low in precision and recall. For example, a request for "all books by Hemingway" suggests the attributes format and author. The user wants all books available, so he/she desires high recall. The user doesn't care if a few nonrelevant books are retrieved, so he/she desires moderate precision. On the other hand, a request for "A Farewell to Arms by Hemingway" suggests the attributes title and author. The user wants a specific title, so he/she desires low recall. The user wants only that title, so he/she desires high precision.

**Narrative:** Begin with a brief paragraph generally describing users’ information needs or information need situations, that is, what they are looking for and why they are looking for it.

List four user questions or requests, the information object attributes they suggest, and the desired precision and recall, using the format below. Compose questions as full sentences. Copy the format below for each question (number questions appropriately).

**User question 1:**
**Object attributes:**
**Desired precision:**
**Desired recall:**

End with a brief paragraph summarizing important attributes of the objects that should be represented in the records. Be sure to list attributes that are suggested by the questions. This can be done in as little as one sentence. If you think of other important attributes that are not suggested by the questions, mention these separately.

**Additional task to prepare Draft 1**

Use Draft 1 Checklist (available in Learn) on your own draft before submitting the draft. Note that there is no peer review for draft one. (face-to-face sections may incorporate peer review for this draft – check with your instructor)

### 2. Representation of information objects

In section 2, you begin the process of determining the metadata necessary for adequately representing the information objects to serve users’ needs. You also move into the role of database designer. The first step is to determine the entity level and determine the metadata elements to create a resource description of the object. Next you structure the actual database and explain its technical specifications, and then you write rules for data input.

For Draft 2, this section focuses on the description of the physical information container. Section 4 focuses on metadata elements necessary to represent information content (i.e., describe intellectual content, subjects, topics). When doing Draft 3, you will revise as appropriate the sections below to reflect a focus on representing both the container and the content.

**2.1. Entity level**

**Tasks:** Determine the entity level or unit of analysis represented by a single metadata record (e.g., book series, individual book, or book chapter). For simplicity, choose only one entity level.
Narrative: Explain the concept of entity level or unit of analysis and the purpose it serves. State the entity level for representing the objects in your collection. Explain why this entity level is appropriate for representing objects for your users. Your writing should reflect your understanding of the connection between entity level and metadata record.

2.2. Metadata elements and semantics

Representations of information objects are created using a set of metadata elements. These elements support a variety of user tasks, from finding database records to physically obtaining information objects. The metadata element set contains all the elements that represent the objects. Elements tend to be more specific than attributes or general characteristics of the objects you discussed in section 1.3. Attributes that are important to your users, as indicated by their questions, should suggest metadata elements for your metadata element set.

Tasks: Translate attributes of the information objects into specifically named metadata elements for describing the objects. For Draft 2, focus on physical description, or elements that represent the information container, but mention all your elements in some fashion. In Draft 3, you will revise to include in depth discussion on the elements that represent the information content. Avoid trying to describe every detail of the objects. If all objects share the same value of an attribute (e.g., all items are CDs, the attribute value for the attribute Format would be CD for all objects), that attribute has no discriminatory power for information retrieval and therefore should not be used as an element.

In addition to considering users’ questions in your choice of elements, consider how the elements assist users in these four tasks: finding information that corresponds to the search criteria, identifying the object described as being appropriate to the search criteria, selecting a specific object that meets the user’s needs, and obtaining or accessing the actual object for use. Note that the last three user tasks rely on the information contained in the metadata record. You may have some elements that do not support specific tasks, but these tasks provide a basis for choosing the most important elements. Consider how value can be added to your system by one or more of your elements, particularly elements that are not usually included in standard bibliographic description.

In section 4 you will focus on subject description, or representation of the information content of the objects. For now, simply add two placeholder elements called Subject and Classification. When you get to section 4, you may rename or add elements for representing intellectual content.

Create Appendix A: Metadata elements and semantics, which names and defines the elements in the metadata element set.

Write narrative.

Narrative: State the number of metadata elements that represent the information objects. Explain in general why the elements are appropriate for the users and objects. Define the four user tasks and discuss which elements or types of elements support each task. Do not list all the elements; summarize. Mention any less-traditional elements that add value to the system. Spell and capitalize element names exactly the same as they are listed in Appendix A.

Refer to Appendix A in an explanatory fashion, as in "Appendix A contains a complete list of metadata elements and semantics" or "(See Appendix A: Metadata elements and semantics.)." Use similar wording to refer to appendixes in subsequent sections.

When you are finished, it should be clear how many elements you have and what they are, how whichever ones support whichever of the four user tasks, and what purpose the elements that do not support any tasks serve.

2.3. Record structure and specifications

The metadata elements in Appendix A provide the conceptual foundation for structuring the main database file and metadata records that represent the information objects. In the main database file, each metadata element translates into one or more database record fields. Selected record fields provide access points for searching the database. The record structure and related technical specifications determine the functionality of an information retrieval system. For this report section, you create Appendix B to show the record structure of the main database file.
**Tasks:** Decide whether each element in Appendix A translates into one or more than one field (e.g., Creator element may become Author and Illustrator fields). Write names for the fields that are the same or parallel to names of the metadata elements.

Create **Appendix B: Record structure and specifications**, which shows all the fields in the database record and their database specifications. Note that two additional fields, RecordID and RecordDate, appear in Appendix B: these fields are not derived from metadata elements; they represent the record, not the information object. You must create these two fields along with all those that are derived from your elements. They should be the first two created. Determine the specifications for each field and state these in the table in Appendix B. In order to complete Appendix B, you actually create the record structure and set the specifications in Inmagic, and then print out the structure from Inmagic to include in Appendix B.

**Narrative:** State the total number of fields in the database record. State how metadata elements in Appendix A are structured into fields in the database record as shown in Appendix B. If each element maps one-to-one to a single field, simply say so. If not, state which elements translate into which multiple fields. (Hint: If the translation is complex, show it in a table with one column for elements and a second column for fields.) End this paragraph by explaining the purpose of the two additional fields, RecordID and RecordDate. **Spell and capitalize** field names exactly the same as they are listed in Appendix B.

Explain the four kinds of technical specifications in order of their appearance in the table in Appendix B: field type, indexing, entry validation, and content validation. Write a brief separate paragraph for each kind of specification (not for each field). Even if you don’t use some of the specifications in your database, **explain each type of specification and the options available for that specification. Make sure to indicate which specifications are used in your system.** Begin each paragraph by defining the specification and its options. Summarize your decisions (do not list every field). Expand on any unusual or difficult decisions.

Your paragraph on indexing should discuss a number of concepts and decisions. Generally, explain what indexes are (in the context of inverted indexes and how Inmagic creates the indexes) and the relationship between indexing and access points. Specifically, state why certain fields in your database are or are not indexed for the purpose of creating access points. Justify your choice of word and/or term indexing for various kinds of fields.

Generally, this section should demonstrate your understanding of the purpose and functional results of the database specifications listed in Appendix B. You need to provide sufficient information for the reader to make sense of Appendix B. Your instructor should be able to recreate your database exactly as you did it without consulting Appendix B.

Refer to Appendix B.

**When you are through with this section, your instructor should be able to reconstruct your database just as you did, by reading the narrative without referring to Appendix B. Further, the instructor should have a good understanding of your decision process for each specification you put on every field.**

**This is one of the most critical parts of the IOP. If this is less than a page, you are probably not addressing all that is expected.**

**IMPORTANT:** Throughout the report, use terms accurately (e.g., field vs. field name vs. field type). If you are referring to the database, talk in terms of fields. If you are referring to the metadata scheme, talk in terms of elements.

### 2.4. Record content and input rules

To provide consistency in the creation of records, you need to develop rules for the content of database record fields and how data are to be entered into the records. Record content refers to the data contained in the records; input rules address directions to put the data into records. The input rules are explicit instructions for the person who enters data values in individual records: the indexer or cataloger. These rules are used only by the technical user to create the records, not by the end user to search the records.
Tasks: Create Appendix C: Record content and input rules. The rules cover every field, including the automatic entry fields RecordID and RecordDate. Each rule identifies the chief source of information, or where data are obtained, for the field. Note that some rules will be revised or added later in the project.

Write narrative.

Narrative: Explain and discuss the purpose of input rules and what these rules address.

Define the concept of chief source of information and explain how it assists in data quality and consistency. State the most common chief source(s) of information in your system: the location of data for the majority of fields. Discuss any rule with unusual or problematic aspects related to your information objects (e.g., locating information that is not in/on the object) and the rationale for the rule. Do not reiterate each rule in the narrative.

When you are finished with this section, a cataloger should be able to catalog any book likely to be added to your collection, without having to come to you for questions on how to do it.

Refer to Appendix C. Also refer to Appendix G, sample records showing application of the rules (see below).

---

3. Access and authority control

Access points are often under some form of authority control (also called access control or terminology control). Authority control is a mechanism for bringing consistency to data values in an information organization system. Data entered in fields that are under authority control must come from a file or list of authorized (or controlled) terms. In your system, terms related to subjects and to names of people and corporations are under authority control. You can establish authority control in three forms for purposes of this assignment:

- content validation lists in the main database file (section 2.3; optional),
- thesaurus (external to the main database file) for subject terms (section 4.2, required)
- name authority file (external to the main database file) for names in the records (section 5, required)

In this section, you explain authority control in general and state which fields are under which type of control.

Tasks: Determine which fields (both physical description and subject description) are under some form of authority control. Consider the following:

- Fields with simple, predictable terms. These are usually physical description fields such as Format with terms such as "book" and "video." Decide whether any such field should be under control of a content validation list within the database (you may need to reconsider and revise your decisions about database specifications in section 2.3 and Appendix B). Content validation is optional. If you do specify it in Appendix B, you are not required to create actual content validation list(s) in the database. (Do not specify content validation for the thesaurus field or for name fields (next two points) because they have separate term lists.)

- The field with the greatest number of potential terms and the most semantically (conceptually) complex terms, especially how the terms are related to one another. Usually this is a subject field. Choose one field only for vocabulary control using a thesaurus.
• Fields with proper names. These may be personal names (people) or corporate names (companies, organizations). Usually **all name fields** are controlled by a single **name authority file**. The name authority file also controls the form of names used in subject fields.

Write narrative.

**Narrative:** Discuss the purpose of authority control and its importance in your system. Explain how it works. Explain why it is beneficial to have specific access points under authority control from the perspectives of the end user searching the system and the technical user creating the records. State the kinds of authority control in your system.

Discuss the fields under control of a content validation list (optional), a thesaurus, and a name authority file. State explicitly which fields are under which type of control mechanism.

**Hint:** If you have trouble completing this section, come back to it after completing section 4.

---

**4. Representation of information content**

Given the basic resource description for the information container developed in section 2, you now need to determine the metadata elements necessary for representing information content (or intellectual content, subjects, topics). Section 4 focuses on problems of describing subjects, including use of controlled vocabulary in section 4.2, and subject-based classification in section 4.3.

**4.1. Subject access**

**Tasks:** Determine how to provide subject representation, or how to represent the information content of the objects. The subject representations will be the basis for providing subject access in your system. Consider the kinds of subjects (e.g., topics, themes, time period, geographic area) of the information objects. Note that, although fields such as title and table of contents can provide clues to aboutness, these fields are considered physical description of the information container, not subject description of the information content.

Decide how many subject fields you need. You may translate your earlier placeholder element Subject into more than one field (e.g., Topics and Time Period) and/or you may rename the metadata element and database field. You may have some subject fields controlled by an InMagic content validation list, or controlled by a thesaurus, or fields that contain natural language terms (e.g., abstracts, summaries, etc.).

Your earlier placeholder element Classification should translate into just one field that contains a classification code (or call number) from your classification scheme. The classification code should be based in part on information content.

**Narrative:** Define subject representation and subject access. Explain the importance of subject access for your users. Describe how your organization system provides subject access by listing all fields in your records that contain subject-related data or information. Explain that classification is partially based on subject, identify the subject-based facet(s) in your classification scheme, and name the field that contains the classification code. (You may need to return to this after you complete section 4.3).

**4.2. Thesaurus structure**

This section addresses subject authority control (also called vocabulary control or terminology control) using a thesaurus. A thesaurus is a list of controlled vocabulary terms that provides data values (terms) for a single field under subject authority control. It serves both technical users (indexers, catalogers) as a source of terms to enter in the record and end users as a source of search terms.

**Tasks:** Review the *Thesaurus Tutorial* in the Learn course site. **Do NOT use InMagic’s online thesaurus function.**

Choose one subject field to contain controlled vocabulary terms from the thesaurus. If you have more than one subject field, this should be the field that is likely to have the greatest number of terms and the terms with the most complex meanings and relationships.
Review, discuss and demonstrate the three semantic (conceptual) relationships in the thesaurus, and understand how mandatory reciprocals are used to indicate these three relationships.

Determine the domain and scope of the thesaurus.

Make decisions concerning specificity and exhaustivity. Consider how each decision may affect information retrieval performance based on measures of precision and recall. (Be aware that the meaning of "indexing" here is different from its meaning in database specifications in section 2.3.)

Create Appendix D: Sample thesaurus. Do NOT use InMagic's online thesaurus function.

Write narrative.

Narrative: Explain the purpose of subject authority control, how it is implemented in your system, and why it is important for both end users and technical users of your system. State the name of the field that uses the thesaurus to provide authority control. If you have more than one subject field, explain why this field is the best choice for thesaurus-based vocabulary control.

Define the thesaurus as a kind of controlled vocabulary. Explain the purpose of its syndetic structure. Define and describe the three kinds of semantic relationships and how each is displayed. Explain mandatory reciprocals and how they are used.

Describe the domain and scope of the thesaurus.

Define specificity. State the level of specificity in the thesaurus (high, moderate, low) and explain why it is appropriate for the users and/or information objects. Discuss the probable effect of this level of specificity on precision and recall measures of information retrieval performance.

Define exhaustivity. State the level of exhaustivity for indexing, that is, whether the indexer should tend more toward depth indexing or summarization. Explain why this level is appropriate for the users and/or information objects. Discuss the probable effect of this level of exhaustivity on precision and recall measures of information retrieval performance.

Refer to Appendix D: Sample thesaurus.

Note: The instructor understands that your thesaurus is only a sample and that it is not comprehensive.

4.3. Classification scheme

Classification is a process of categorizing objects according to one or more attributes or characteristics. Formal classification systems such as Dewey Decimal and Library of Congress are called schemes. Classification codes are derived from schemes and assigned to objects to group items that are similar in one or more ways together. The primary function of bibliographic classification is to bring items together that contain similar intellectual content or subject matter. In the library world, bibliographic classification systems are also used as the basis for physical location. Classification schemes are used by technical users who create the codes and by end users who want to understand the organization of materials.

Tasks: Review Faceted Classification Tutorial and/or Hierarchical Classification Tutorial in the Learn course site. Determine your approach to classification: faceted (recommended) or hierarchical.

Choose three or four attributes of the objects (e.g., subject, creator, literary form or genre, media format, date) to be used in classification. Consider attributes suggested by users' questions and how these relate to users' expectations for physical arrangement of objects (e.g., whether to arrange objects first by subject or by format). For this project, at least one facet must relate to information content or subjects. Your first facet should not be Author or any other facet that merely alphabetizes the collection.

Develop a notation code (you may not use a pre-existing code such as Dewey or LC) to identify and group the objects by class. In order to physically organize the objects, make this a unique identifier (call number) by adding to the notation code a unique number to identify the individual object.

Note: Digital information objects may be treated somewhat differently; ask the instructor if you are uncertain.
Create Appendix E: Classification scheme.

Write narrative.

**Narrative:** Define classification and its purposes in general.

Describe the role of classification in your system with regard to providing intellectual access and physical access if appropriate. Define and describe the difference between faceted and hierarchical approaches to classification; state your approach and explain your choice.

State the primary facet and explain why you chose it with regard to providing intellectual access (subject-based). List the other facets in order. Explain why you chose these facets, including their effectiveness as a system for intellectual and physical organization of the objects (if applicable). Your primary facet should be derived from a field that uses a controlled vocabulary.

If you are adding a unique identifier to the classification code for physical arrangement, explain why that is necessary and the source of the unique identifier.

In a separate paragraph, illustrate your classification system by providing a complete example of one classification code:

- Briefly describe one of your 10 objects
- Show the classification code for that object
- Explain what each part of the classification code represents.

This can be exactly the same example as in Appendix E (see instructions there). Do not repeat classification rules here; refer to Appendix E.

### Additional tasks to prepare Draft 3

- Revise Appendixes A, B, and C to include any new subject elements and fields respectively. Note that you may retain one element called Subject in Appendix A and translate this into multiple fields in Appendixes B and C. Check that element/field numbers, names, and order are consistent throughout.

- Review Appendix B table to ensure it reflects decisions made in sections 3 & 4.
  - Beyond new/revised subject fields, check subject field specifications. Do not specify content validation for thethesaurus field, which has a separate term list, or for the classification field.
  - Make sure all cells are filled in: enter a dash in any cell for which you made no specification.

- Edit database and print out a new textbase structure report to insert in Appendix B. Suggestions:
  - If elements in Appendix A and fields in Appendix B are not parallel in name or order, it's easier to edit Appendix A than Appendix B.
  - If Appendix B field names are not exactly the same in table and in textbase printout, it is easier to edit the table than Inmagic.
  - Inmagic may let you edit a field name, but it will not let you change field order if you have records with data in them. The easiest solution, since you have only three records so far, is to delete the records. Once you have altered the record structure, recreate the records. The original record numbers will not be reused: if you incorporate RecordID number in classification, edit the classification codes.
  - Inmagic will let you add fields at the end of the record. The final order may not be esthetically pleasing, but don't worry about it. Or, if it bothers you too much, create a new textbase from scratch. In this case, you will also have to recreate all the records and revise all appendixes and narrative to match.

- In Appendix C, write/revise rules for subject field(s), especially fields under control of the thesaurus and content validation lists, and classification field (see Input Rules Tutorial). Check all other rules for consistency with any changes in fields, specifications, and other appendixes referenced.

- Revise narrative text in sections 2.2 through 4.1, if necessary, to reflect any changes in elements and fields, element and field names and numbers, and specifications.

**IMPORTANT:** Revise Appendix G: Sample records by creating 10 records to demonstrate your full set of
5. Name authority control

Name authority control is applied to fields that contain personal names (people) and corporate names (companies, organizations). A name authority file provides authorized forms of names for technical users to enter in records and, in some systems, for end users to identify as search terms. The name authority file you develop in this section builds on decisions you made in section 3.

**Tasks:** Identify all fields that contain personal or corporate names. Develop a name authority file to control data in all of these fields. This is a second database file, separate from the main database file described in Appendix B. Create *Appendix F: Name authority file* to describe this file.

Write narrative.

**Narrative:** Explain in general what name authority control is, how it works for technical users and end users, and why it is important. List the specific fields that are under name authority control in your system.

Refer to Appendix F.

**Additional tasks after completing this section**

- Review the *Appendix B* table to ensure that fields under name authority control do not have content validation lists specified, which would be redundant because they have a separate term file.
- Revise (or write) rules in *Appendix C* for each field under name authority control. See *Input Rules Tutorial*.
- Revise section 3 narrative, if necessary, to list fields under name authority control.
- **IMPORTANT:** Revise *Appendix G: Sample records* by completing the 10 records to demonstrate your final set of record fields, specifications, and input rules.

6. System evaluation and development

There are so many variables in the information world that no system can be expected to remain static over time or to respond perfectly to all requests at any given time. Given the dynamic characteristics of your collection and users described in section 1, section 6 focuses on the need for ongoing evaluation and development of both your overall information organization system and your IR system.

**NOTE:** Although your system responds to changes in the collection, this section addresses evaluation and development of the information organization and information retrieval systems, not changes to the collection, collection development policies or procedures (materials selection/deselection, etc.).

6.1. Performance test

This section addresses an analysis of database performance with respect to the user questions in section 1.3 and various information organization decisions throughout the project.

**Tasks:** Ask a prospective user—someone who shares the general characteristics of the user group described in section 1.3—to search the IR system while you observe. Before the searcher begins, decide whether to give him/her any instructions, including about the thesaurus and name authority file (i.e., would a typical user understand these search aids and use them?). Make a note about any instructions you give, so you can report
them later. If you do not give the thesaurus to the user, provide a valid reason, such as “the users are too young to read the thesaurus”. “They would not use it anyway” is not a valid reason.

Ask this person to conduct a search for each of the user questions in section 1.3. You may also ask the searcher to talk aloud what he/she is doing during the search, but do not bias the results by helping. The searcher may formulate multiple queries for a given question. During the search, keep track of fields searched, search terms, retrieval results, any problems encountered, and any comments by the searcher.

Estimate probable precision and recall for the search (high, moderate, low), based on the search results and what you now know about IR system functions. (Note that the database contains too few records for actual calculations of precision and recall to be meaningful. Simply make as realistic a projection as possible.)

Analyze and explain the search results. Consider factors related to representation of various attributes suggested by the question. For example, are the attributes suggested by the questions represented in the system? Is it difficult for an end user to formulate a particular kind of search query? Is retrieval affected by subject-related indexing decisions such as specificity of thesaurus terms? Are some searches better conducted using natural language rather than controlled vocabulary? Consider also the retrieval matching functions of the database. Is retrieval affected by technical decisions such as specifying term vs. word indexing?

Narrative: Begin with a brief paragraph that describes the research methodology. Explain how you conducted the test, including basic demographics of the searcher and the extent of instructions provided (e.g., was the searcher told about the thesaurus and name authority file?).

Report and analyze the test results. Copy all four subheadings for the first user question in section 1.3 exactly as they appear (do not edit them). Then add four subheadings for reporting the search test, beginning with the query formulation, as shown below. State the query formulation as field(s) searched and term(s) entered (e.g., Author "Poe" AND Form "poetry"). Write an analysis of the results of this one query. For a second query, repeat the last two subheadings, beginning with query formulation, and write a separate analysis. In this case, add a number to each query (e.g., Query formulation 1, Query formulation 2).

User question 1:
Object attributes:
Desired precision:
Desired recall:
Probable precision:
Probable recall:
Query formulation (n):
Analysis of results:
Repeat this procedure for each user question individually. The analyses should explain why the results were or were not as expected, minimally with regard to precision and recall.

End with a separate paragraph summarizing strengths and/or weaknesses observed in system performance. The narrative should demonstrate your basic understanding of the effects on system performance of relevant factors such as user abilities, vocabulary control, query formulation, precision and recall, and database matching functions. (Save any speculations on how the findings can be used for the next section.) Throughout this section, remember that no system is perfect; that poor retrieval results may actually contribute the most to the learning experience. This paragraph should not be a recant of what your user did during the test.

Note: This narrative is based on an actual test, so past-tense verbs are appropriate.

6.2. Change and development

This section concerns ways to improve and develop the organization system and its implementation in the IR system. It takes into account implications of evaluations discussed in the previous section, along with potential responses of the system to large-scale changes in its information environment.

IMPORTANT: Although your system does respond to changes in the collection, this section addresses only changes in the information organization system and IR system. It does not address everyday database maintenance (e.g., record additions/deletions) or collection development policies and procedures (e.g., materials selection/deselection).
Tasks: Consider the real or potential results the user test. Speculate on how specific findings from this evaluation can be incorporated into system improvements and development.

Review previous sections of the project. Think about how the information environment—collection, setting, and users—may change over time. Suggest how certain aspects of the system can evolve or expand to respond to these changes. For example, if the collection grows primarily through donations (section 1.1), could the donations cause large and unpredictable shifts in domain or scope that require the organization system to adapt? Feel free to idealize additional or advanced features that might add value to the system or overcome weaknesses. Your suggestions may involve:

- Metadata elements
- Access points
- Database fields and their specifications
- Record content and input rules
- Authority control
- Indexing language (specificity, technical/professional jargon)
- Depth of indexing
- Classification
- Any other aspects of the organization system and/or IR system (except screen display; you have worked only with the designer view of Inmagic, but an attractive user view can be created)

Narrative: Discuss and recommend ways in which the information organization system and IR system can be improved or developed. Organize this section any way that seems appropriate, but it must include suggestions that explicitly refer to:

- The results of the test in section 6.1
- Anticipated changes in the information environment

Be as specific as possible in suggesting changes in any of the areas listed above (or others), and in justifying your suggestions.

Note: This narrative is particularly geared toward your organization system as a whole and thus should cover only areas you feel can or should be improved. It should demonstrate your understanding of the implications of pertinent decisions made throughout the project. Future-tense verbs are appropriate here.

Additional tasks to prepare Preliminary Draft 4

- Review instructor’s marks and comments on Draft 3 to see that you addressed or considered all of them. If you have questions, discuss them with your instructor prior to submitting the new draft.
- Use Draft 4 Checklist on your own preliminary draft before submitting the draft for peer review.
- If Draft 4 is particularly long, you may add a table of contents under the project title (optional).
- IMPORTANT: Complete section 7 only for Final Draft 4. This goes to the instructor, not to the peer reviewer.

7. Project summary

Congratulations! You have created an original and comprehensive information organization system including a working database. Now to debrief . . .

Section 7 is an opportunity for you to describe your experiences with the project, both positive and negative. This section is a critical component of the report because it helps the instructor understand your experience this semester and improve the assignment for future semesters. Your comments will not affect your grade; however, this section does count toward your score for completeness of the final report draft.
Use this section to describe both your opinion of the outcome of the project and any problems you had in completing it. Because this may be the first time you have attempted the systematic organization of a collection of information objects, you should expect to be confused or frustrated at times. Here you can identify problems you found especially difficult to deal with and discuss how you might construct your system differently based on what you now understand about the concepts and practices of information organization.

Tasks: Review your experience with and thoughts about the project, such as:

- Your reasons for choosing to organize this specific collection
- Why or how your system is different or better than an existing or traditional system
- Any major problems you had with representing objects in this collection
- How you might construct your system differently based on what you now know
- Problems you had with certain parts of the assignment
- Problems overcome and skills learned
- Whether there is a chance that you might actually implement this system for a real collection

The above items are just suggestions to get you started. You may choose to discuss other concerns and ideas.

Narrative: This section is open-ended; discuss any or all of the points above (or others). Suggested length is up to 2 pages.

Note: Because this is your personal assessment, you should write in the first person.
Appendix A. Metadata elements and semantics

**Tasks (Draft 2):** Copy the table format below.

Name and list the metadata elements that represent attributes of the information container. Element names should be descriptive, simple, and distinctive. Use regular words, not abbreviations. Avoid meaningless generic names such as Type, Category, or Description, which can apply to any element.

Write semantics (definitions). Definitions should be broad and not circular (i.e., don't define a word with itself). Avoid listing data values unless necessary to clarify meaning. An example of semantics for an element named Creator is "Entity responsible for original content, such as author or illustrator."

Add two placeholder elements at the end: Subject and Classification. You will develop these later; you may write semantics for them now or wait until Draft 3.

**Tasks (Draft 3):** Complete the table with semantics for Subject and Classification (you may also rename these or add other subject-related elements).

**Appendix:** Insert this table under the Appendix A heading and fill it in with your own elements, adding as many rows as necessary. Subject and Classification are placeholders to be completed later.

<table>
<thead>
<tr>
<th>No.</th>
<th>Element name</th>
<th>Semantics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>Subject</td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>Classification</td>
<td></td>
</tr>
</tbody>
</table>
Appendix B. Record structure and specifications

Here you begin developing the main database for the IR system. Appendix B consists of two parts:

- A table that illustrates the plan for the record structure and specifications for each field
- A printout from Inmagic showing actual record structure and field specifications

**Tasks (Draft 2):** Copy the headings and table format below.

Begin with rows for two fields, RecordID and RecordDate; these fields related to the record itself and are automatic fields with limited specifications. RecordID provides a unique identification number for the database record and RecordDate contains the date the record was created or modified.

Add a row to the table for each field based on the elements in Appendix A. Decide which elements, if any, translate into more than one field (e.g., Creator element may become Author and Illustrator fields). Include fields to match the placeholder elements Subject and Classification.

Name the fields. Field names may be shortened or abbreviated, but should be descriptive and recognizable, preferably the same as or similar to the names of elements in Appendix A. **Fields must follow the same order as the elements in Appendix A.** Note that field numbers in Appendix B differ from element numbers in Appendix A due to the addition of RecordID and RecordDate fields and any multiple fields you structured from a single element. (Revise Appendix A and sections 2.2 and 2.3 if necessary for consistency.) **“In the same order” does not mean “numbered the same”.**

Use the columns to indicate specifications for each field. For guidance, consult the Inmagic walkthrough and Inmagic FAQ in the Learn course site and InMagic Help (located in Readings and Assignments). When deciding whether to index fields for user access, bear in mind that, although some fields may not be searched alone, they may be searched in Boolean queries with other fields. For the autonumber and autodate fields, dashes are already filled in where options are not available. You may delay decisions on subject fields, the classification field, and content validation until Draft 3: simply leave these cells blank.

Open Inmagic and create a new textbase (an Inmagic term) with your fields and specifications. Follow your table exactly, including order of fields and capitalization and spelling of field names. Print the textbase structure (an Inmagic term) showing fields and specifications. Insert it in Draft 2. Delete all information in the textbase printout below the last field in the list (saves paper), unless it shows some additional specifications you made.

**Tasks (Draft 3):** Complete the table, filling in every cell with a specification or a dash. Note that content validation is not required, and that if you do specify it, you are not required to create actual content validation list(s) in Inmagic. Open Inmagic, revise textbase as needed, and print the revised textbase structure to insert in Draft 3.

**Appendix:** Insert subheadings and table below under Appendix B heading and fill it in with your own fields, adding as many rows as necessary. Subject and Classification are placeholders to be completed later.

**IMPORTANT:** The specifications you list in the table must match the specifications in the printout of the Inmagic textbase structure. The textbase structure demonstrates the implementation of the specifications in every detail, including exact spelling of field names.

1. **Record structure specifications** [Insert from IOP Template or this assignment. Add rows as needed.]

<table>
<thead>
<tr>
<th>No.</th>
<th>Field name</th>
<th>Field type</th>
<th>Indexing</th>
<th>Entry validation</th>
<th>Content validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RecordID</td>
<td>Autonumber</td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>RecordDate</td>
<td>Autodate</td>
<td></td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>Subject</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>n</td>
<td>Classification</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. **Textbase structure** [Print from Inmagic and insert here; View Inmagic walkthrough for instructions.]
Appendix C. Record content and input rules

Tasks (Draft 2): Input rules follow a strict five-part format. They must be written for every field in the record structure, in the same order as in Appendix B. For guidance, consult the Input Rules Tutorial. To summarize:

Field #: The number of the field

Field name: Spellings of field names exactly as in Appendix B.

Semantics: Copy from Appendix A, or revise where an element has been structured into more than one field. Write semantics for RecordID and RecordDate.

Chief source of information: Primary location of the data needed for the field, usually in/on the object itself. For RecordID and RecordDate, simply state that input is automatic.

Input rules: Prescription for how to enter data, including spelling, capitalization, punctuation, etc. For RecordID and RecordDate, again state that input is automatic.

Example: One or two data values for the field drawn from sample objects.

Remember that these rules are for the technical user who creates the records, not the end user who searches them. Hint: Test rules by having someone else try to follow them.

Note: Rules are purely practical: write them as succinctly as possible. Use simple, active verbs in an imperative sentence structure (e.g., "Use sentence-style capitalization.")

Tasks (Draft 3): Write/revise rules for subject field(s), especially fields under control of the thesaurus and content validation lists, and classification field.

Tasks (Draft 4): Write/revise rules for name fields for which the name authority file serves as source of data.

Appendix: Copy the five-part format below for every field in the record structure in the exact order of Appendix B, inserting appropriate field numbers.

Field#:
Field name:
Semantics:
Chief source of information:
Input rules:
Example:

Use the format just above in your IOP.

The idea here is to provide enough rules to allow your cataloger to enter data without asking any questions. Simple rules such as "Enter author’s name as seen on source" are not well thought out and will result in point reductions.
Appendix D. Sample thesaurus

This appendix contains a sample thesaurus of controlled vocabulary terms to be entered in one designated subject-related field in the record. The thesaurus is created only in MS Word, not in Inmagic. Do NOT use InMagic’s online thesaurus function.

Tasks: Work through the Thesaurus Tutorial in the Learn course site. Examine your 10 sample information objects to begin building a list of subject terms for the thesaurus. You must develop at least 15 subject terms that you will use in your records. As you develop and edit the list, bear in mind the domain and level of specificity you determined in section 4.2. Do not try to be comprehensive; this is only a sample with sufficient terms to represent each of your 10 sample objects at the level of indexing you described in section 4.2.

Determine semantic (conceptual) relationships among terms. The thesaurus must contain several examples each of a hierarchical, equivalent, and associative relationship. All the terms in the field on which the thesaurus is exerted must be in the thesaurus. All the authorized terms in the thesaurus must appear in at least one record.

Develop the syndetic structure of the thesaurus. Sort and display terms alphabetically and use appropriate conventions (BT, NT, etc.) to show several sets of mandatory reciprocals for each kind of relationship. Not all terms in your sample thesaurus will be related to other terms in the sample. Do not force a relationship on every term; instead, let terms stand alone where appropriate.

Provide a key or legend to explain the abbreviations BT, NT, etc.

Note... Every authorized term must appear in at least one Appendix G record. No unauthorized terms may appear in Appendix G records. There may be no terms in the thesaurus field in Appendix G that do not appear as authorized terms in Appendix D.

Appendix: Sample thesaurus. Length is usually 1 to 3 pages. One or two columns are acceptable. Must contain at least 15 authorized terms. Do NOT use InMagic’s online thesaurus function.
Appendix E. Classification scheme

This appendix presents the structure of the classification scheme and instructions for use. It contains:

- A table or other illustration of the scheme
- Rules for notation
- Rule for unique number
- Example of the full classification code

**Tasks:** Review the classification tutorials in the Learn course site. Determine your approach to classification: faceted (recommended) or hierarchical. Identify three or four facets, as well as their order, to use in classifying the objects. At least one facet must be subject-oriented. In classification, subjects tend to be less specific than in thesauri and subject headings lists, and may even be some other kind of subject, such as genre. Note that in a faceted classification approach, a facet can be optional, or used only for some objects.

Create classes (or foci) for each facet. If a class is open-ended (e.g., year), do not list every class: the rule will explain how to enter the data. In choosing both facets and classes, strive for mutual exclusivity so there is no overlap in meaning between facets or between classes.

Write rules for assigning classification notation to an object. These rules follow a three-part format for every facet in order of synthesis. Here is a brief summary.

**Facet name:** Descriptive, simple.

**Chief source of information:** Primary location in/on the object where the information can be found, or more likely the field from which the value for the facet is drawn.

**Notation rules:** Instructions for synthesis, including how to choose class, how to capitalize and spell class codes, and how to punctuate after the code. If the facet is optional, state when it is used.

Write a separate rule, under a separate heading, for adding a unique number to the notation to form a unique identifier (call number) for an object. An easy solution is to add the RecordID number to the end of notation, but many other solutions are possible. A classification system that represents digital information objects and/or that is not related to physical arrangement of objects may require a different solution. If in doubt, ask your instructor.

Provide an example in which you:

- Briefly describe 1 of your 10 objects.
- Show entire classification code or unique identifier (call number) for the object.
- Explain what each part of the classification code represents.

You can use this same example in the section 4.3 narrative.

**Hint:** Test rules by having someone else try to follow them.
Appendix: Copy the four-part format below. Under "1. Scheme," the classification scheme can be presented as a table (columns for facets, rows for classes/foci) and/or a hierarchy, depending on the classification approach. Under "2. Notation rules," repeat the three subheadings for each facet.

1. Scheme [table and/or text description of facets and classes]
2. Notation rules [for each facet, repeat three-part format below]
   - Facet name:
   - Chief source of information:
   - Notation rules:
3. Rule for unique number
4. Example
Appendix F. Name authority file

The name authority file provides authorized forms of personal and corporate names that are entered in designated fields of records in the main database file. When creating records in the main database, the technical user finds a name in/on an information object and then consults the name authority file to determine the authorized form of that name to input. The name authority file itself is a **second** database file in which each record represents one authorized name.

For this appendix, you create a new Inmagic textbase file for name authority records. The appendix has four parts that mirror closely what you did for Appendixes B, C, and G for the main database:

- Specifications table for name authority file
- Textbase structure (Inmagic printout)
- Input rules
- Sample records

**Tasks:** Copy the headings and table below. All fields are predetermined. Note that no content validation specifications are required but choose appropriate entry validation specifications. **Fill in every cell** of the table with a specification or a dash. Create a new Inmagic textbase. Structure the records with the five fields in the table. Print the Inmagic textbase structure and insert it below the specifications table.

Write input rules for each field, following a format similar to that in Appendix C, but with only four parts: **Field name, Semantics, Input Rules, Example.** See guidelines in the online module *Name authority records and rules*. Here’s a brief summary:

- Spell all field names exactly as in the table below.
- Handle RecordID and RecordDate the same way as in Appendix C.
- The rule for the AuthorizedName field should include instructions for establishing the authorized form of a name, along with style limitations (e.g., whether to invert names as in "Picasso, Pablo").
- For the VariantNames field, invent variant names if none exist in/on the object.
- The SourcesUsed field cites source(s) used in establishing that particular authorized name. Citations usually consist of title and year.

Create **only 5 records**, each for a different name drawn from your 10 sample information objects. These should include records for at least one personal and one corporate name (if your objects contain both). When entering data in the sample records, follow your own rules. Print records and insert in this appendix.

**Appendix:** Copy the four-part format below with the subheadings and table shown.

1. **Record structure specifications** [Insert from IOP Template or this assignment. Add rows as needed.]

<table>
<thead>
<tr>
<th>No.</th>
<th>Field name</th>
<th>Field type</th>
<th>Indexing</th>
<th>Entry validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RecordID</td>
<td>Autonumber</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>RecordDate</td>
<td>Autodate</td>
<td></td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>AuthorizedName</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>VariantNames</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>SourcesUsed</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. **Textbase structure** [Print from Inmagic and insert here; View Inmagic walkthrough for instructions.]

3. **Record content and input rules** [Number fields 1-5]

   **Field#:**
   **Field Name:**
   **Semantics:**
   **Input Rules:**
   **Example:**

4. **Sample records** [Print from Inmagic and insert here.] (View Inmagic walkthrough for what goes here)
Appendix G. Sample records

This appendix contains a printout of records from your Inmagic database. The records should accurately reflect all specifications and rules. You submit a version of this appendix with every draft except the first.

Tasks (Draft 2): Create 3 database records to demonstrate the system so far. Enter data to represent 3 objects in your sample collection, following your own rules. Print records (see Inmagic walkthrough) and insert in draft. Check records as follows:

- Field order and field name spellings must match Appendix B.
- Formats of RecordID and RecordDate data (e.g., date, time) must match examples in Appendix C.
- Subject and classification fields may be left empty. Note that empty fields do not show in the record display.
- You may improve the appearance of records by boldfacing field names, etc. (optional).

Tasks (Draft 3): Create 10 database records to represent all 10 objects in your sample collection, following your additional rules and any other revisions. Print records and insert in draft. Check records as for Draft 2, plus:

- All terms in the field controlled by the thesaurus must come from the thesaurus (Appendix D) and must match the thesaurus in form, spelling, and capitalization.
- For each field with a content validation list, one or more terms from the Appendix C example must appear in at least one record.
- In the classification field, the classification code in one record must match the example in Appendix E, and all codes should follow the rules in Appendix E.

Tasks (Draft 4): Revise and print 10 database records to reflect all final fields, specifications, and rules, including name authority rules and any other revisions. Check records as for Draft 3, plus:

- Although any given record may have empty fields, each field in the record structure must be filled (and therefore displayed) at least once somewhere in the 10 records.
- Data in name fields (e.g., Author, Publisher) must include names from authority records in Appendix F.

Appendix: Inmagic record printout. [Print from Inmagic and insert here; View Inmagic walkthrough for instructions.]