Lindsay Robbins  
Spring 2022

INFO 5200.001

Draft 4

**A Collection of Books on German Language Learning:**

**Information Organization System**

**1. Project description**

**1.1. Collection and information objects**

The title of the collection is a collection of books on German language learning, which is housed in a hypothetical retail bookstore in an area with a population that is predominately moderate in socio-economic status. The collection exists in order to provide patrons with a varying selection of books that can be used to aid in learning the German language. There are approximately 1500 books in the collection, and it is estimated to grow by approximately 250-500 titles within the next year. The collection contains physical books only, however some titles may have accompanying download links or learning materials included. Some of the topics included are grammar, travel phrases, pronunciation, syntax, conversational use, regional variations, and vocabulary.

**1.2. Users' demographics and knowledge**

The end user group of this collection consists primarily of middle-class adults from the ages of 18-75. While there are language books for children, these are kept and organized alongside other reference books for their age group in a different area. This user group has a good understanding of the English language, so the collection focuses on books that teach German from the perspective of a native English speaker.

General knowledge is the type of knowledge that comes from life experience and is not specific to any one field of study. The end user group of this collection has a moderate to high level of general knowledge due to their age range, as they are all adults and have also likely completed primary school. Those users that are older may have a higher level of general knowledge due to the amount of life experience they have had. Domain knowledge includes more specific subject area knowledge. In this case, user domain range is low to moderate. Many users are complete beginners of the German language and have absolutely no domain knowledge, though some are slightly more familiar and may seek out more specialized subject matter. None of the users are experts or native speakers. System knowledge relates more specifically to the ability of the user to use the systems of information seeking, which include technology like computers or digital kiosks. As these users come from a population that is moderate in socio-economic status, most have average systems knowledge through ownership of smart devices or computers, though some older users have less familiarity with newer forms of technology. Information seeking knowledge is the user’s ability to find information and problem-solve in the process. The user group of this collection is predominately moderate in their information seeking as most of them have likely used general search engines, databases or catalogues as students or adults.

Taking the knowledge levels of this user group into consideration, this information organization system provides features that allow users to search with moderately sophisticated technology for basic to more specific or specialized topics. Without a system that allows this user group to navigate the collection more specifically, they are going to turn to resources freely available on the internet or from other information centers including competitor retail establishments.

**1.3. Users' problems and questions**

This collection provides a variety of subject matter related to German with the intention of aiding users in learning the language. With users primarily having low to moderate domain knowledge, this means that the questions they have are more general overall, though users with more domain knowledge have a better idea of what topics are greater importance to them. Some users are taking courses and looking for supplemental material in a specific topic such as vocabulary or learning independently and looking for a book that focuses on several major aspects of the language concurrently.

**User question 1:** I’m completely new to German and I want as many easy and cheap books as I can get.

**Object attributes:** Skill level, Cost

**Desired precision:** Low

**Desired recall:** High

**User question 2:** I’m going to Germany for vacation, and I need a short, pocket-sized dictionary.

**Object attributes:** Subject, Format, Size

**Desired precision:** Moderate

**Desired recall:** Low

**User question 3:** I’d like a few different books on how to pronounce umlauts. It would be helpful if there were some illustrations.

**Object attributes:** Illustrations, Subject

**Desired precision:** Moderate - High

**Desired recall:** Moderate

**User question 4:** I need a copy of the German book in the *For Dummies* series.

**Object attributes:** Series, Publisher

**Desired precision:** High

**Desired recall:** Low

The goals of the end users differ depending on personal reasons for learning German. Some are seeking out information for short-term goals such as a vacation while others seek a more in-depth approach that includes more specific breakdowns of topics such as grammar or pronunciation. As for the user seeking a short dictionary illustrates, factors including the format and size of books in the collection are important attributes. Many books on German language include illustrations that aid in teaching a particular subject. The user who is seeking a book on umlauts benefits from a book that contains illustrations of the placement of the tongue in the mouth. As this is a retail establishment, it is crucial to allow users to search for materials based on price as it also ensures that users are able to find not only books that fit their other search criteria, but ones they are more likely to purchase. Bibliographic attributes including publisher, date published, author, illustrator, and editor are crucial in the case that a user is searching for more specific materials, as the user in question four. Since the skill and experience level of users varies, the ability to search by difficulty level or skill level is also an important factor. An individual who has had two semesters of German in college no longer needs materials from a book for an absolute beginner. An additional feature not mentioned in the above examples but is common in modern language learning texts, is the inclusion of additional materials such as flash cards, links to audio on the internet or other learning aids. Many users seek out materials with these types of features, so the inclusion of supplemental or bonus materials is another important attribute of this information system. An ISBN is crucial as a unique identifier and should be considered, as well as a sort of store identifier, publication date and cover image. These are all typical metadata elements that are taken into consideration when reviewing objects.

**2. Representation of information objects**

**2.1. Entity level**

An entity is metadata that describes an object. The entity can describe a part of an object such as a record or book, but it can also describe the entire object, so there may be multiple entities for an object. This can be determined by how users are searching for information; if they are seeking out one chapter in a book, it may be more beneficial to consider the chapter an entity. As this information system considers the whole book, the entity is the whole book. Users are typically looking to locate one or more whole objects rather than a section of a book. The entity level for this database is the entirety of each individual book and each book is represented by one record. Each record for this information system represents on whole book and the metadata that informs the cataloging procedures is primarily on the object itself.

**2.2. Metadata elements and semantics**

This information system has a total of fifteen elements that represent the objects in the collection. These contain both bibliographic and subject related information that cover the major methods of searching conducted by users as indicated in Section 1.3 as many users prefer to find information by Subject, Publisher, or Series.

The FRBR tasks, outlined as Find, Identify, Select, and Obtain, describe the process in which users search for information and are well-known within the library science disciplines. FRBR tasks begin with the Find task, in which a user does an initial search to see if materials exist in a given information system. Users illustrate the Identify task as they review the results of this search for matches. As several records may match the terms the user has searched for, in the next task—select, users choose objects that meet more specific criteria they desire. Lastly, in the Obtain task, users actually locate and take the object or objects that best suit their needs. If this is a physical book located in bookstore, as within this information system, the user locates the item on the shelf and walks out with it at this stage.

Users who are searching for books on the German language encounter books within the entirety of this collection and Identify, Select and Obtain based on more selective criteria. Fifteen different elements have been identified. See Appendix A for element descriptions.

Subject may include more specific topics within German language learning such as specific areas in grammar, conversation or vocabulary which can be broad or narrow in scope. As such, Subject supports both the Find and Identify tasks. As indicated by the example questions provided in section 1.3, one major element that users search for is Skill Level. This element is often explicitly provided on the cover of a book which may be labelled as being for a Beginner, Intermediate or Advanced learner. Objects which cover more specific subjects such as particular grammar points are typically in Intermediate or Advanced categories. This element is elaborated on in Section 2.3 to address more specifically what these definitions mean in the context of this information system. Skill level supports the Find and Identify tasks as users may search for books primarily based on their current knowledge on the topic as the most or second-most important criteria. Format relates to Find and Identify tasks as it is highly related to the goals of the user. As illustrated in section 1.3, users may be searching specifically for a book that follows a dictionary format as a major factor. As of Section 2.2, the Size element has been changed to Travel-Sized to indicate that yes, the object meets the criteria for being travel-sized, or no, it does not. Travel-Sized is somewhat related to Format, as many objects share the same attributes, such as a pocket-sized dictionary which is both small and formatted in a specific manner. As many language books come in a travel-sized version, this is a distinction that supports the Find and Identify tasks like the Format element. Cost supports the Select task as there are likely to be several objects that meet their criteria through the Find and Identify tasks within a range of costs. Question 3 in Section 1.3 suggests that the user is most concerned with the Subject of their results, however they are also seeking out an object that is Illustrated. As this is a more minor concern, it supports the Select task; before it can have or not have illustrations, the object must first include the umlaut or pronunciation as a Subject. User question 4 in Section 1.3 indicates that the primary element they are considering is the Series or Publisher. Both of these elements would provide the particular object they are seeking. Therefore, the Series and Publisher, while not always a priority as indicated by the other question samples, can aid in the Find task depending on the user. Users prioritize other elements above the Author or Editor elements, but it is crucial to include this specific piece of metadata as it helps differentiate and identify the object as well. While no example questions provided indicate a desire for objects with Additional Features such as study aids, this element may be considered concurrently with the Cost element for example in the process of the Select task. If the user finds that two objects meet their other criteria at the same Cost, they may Obtain the object that provides extra Additional Features for the value.

The Store Identifier, which is the area in which the book is located within the physical store, both support the user in the Obtain task as they point them to where to physically locate their desired item and identify it visually. The Store Identifier acts as the Classification of the object. Another element that does not aid in FRBR tasks is the Publication Date. Like ISBN, the Publication Date is a critical piece of metadata that while not typically a search factor, should be included with other primary metadata. The Cover Image is a valuable piece of metadata to include as it helps users in the Obtain task; when users are searching store shelves, the cover image helps them locate their desired object more quickly.

While the ISBN and Publication Date are not elements that most users use to locate an object, they are important to include within the information system as basic metadata elements. This brings the total number of applicable elements to fifteen.

Refer to Appendix A.

**2.3. Record structure and specifications**

The total number of desired fields for the record is fifteen in order to accommodate the desired fields. Specifications for each field determine how they work within the information system and how they can be entered and also utilized by end users. In many cases, these specifications indicate the limitations of how each field can be used within the system and push catalogers to interact with fields in a particular way. The specification for each desired field is discussed below.

The Title field includes primarily textual data. It is mandatory to catalogue as one of the most primary pieces of data about a book in this information system; all books in this system have a title. Only one term is allowed for this field as the books in this system do not have multiple Titles. The field cannot have a controlled vocabulary as there is no way to predict the Title of every object that the collection has. For the same reason, the field cannot have a drop-down list as every Title is unique. The field is searchable, as users who are looking for a specific object use this data as a primary search field.

The Author field supplies the name of the creator or writer of the object and is a required and searchable field due its importance as a piece of identifying metadata. Indicated in Section 1.3 the editor is also an important creator to be included. Here, this element is included with the Author field as they are a contributing creator. Author names typically consist only of text, so data is entered as text. There may be several authors, so the number of allowed entries is three. It is possible that there may be more authors than this, but it is unlikely that users need the names of less significant contributors. There is not a controlled vocabulary as it is not possible to account for every single possible author name. As such, there is also no drop-down as there are too many potential names.

The Publication Date field is not searchable, as users did not indicate that date was an important search feature in Section 1.3. As this is a retail establishment, the collection is mostly comprised of very recent books as well. It is a required field as it is an important piece of metadata and only one entry is allowed as it is not possible for there to be multiple publication dates. Reprints are to be designated as a separate record. There is not a controlled vocabulary as there are too many potential dates and there is not a drop-down for the same reason.

The Publisher field is catalogued through text as Publisher names are typically only provided as text. The cataloger is required to enter this field as all objects in the system have an identified Publisher. As this information system is housed in a retail establishment, the objects in the collection are newer and do not include texts by unknown groups. Only one term is allowed for this field as the objects do not have more than one publisher. The field also does not have a controlled vocabulary. Though many of the objects may share a publisher, there is no way to know every publisher that has and will exist. For this same reason, the Publisher field does not have a drop-down list as there are too many variables. The field is searchable, as indicated by Section 1.3 in which a user requests a book by a specific publisher. With educational or non-fiction books, users may desire one specific or familiar publisher as well.

ISBN is not searchable, as users are relying on other fields to search. This field is not required at the time of cataloging as not all items include an ISBN. Only one entry is allowed, as there is only one identifier. As each ISBN is unique, it cannot have a controlled vocabulary or a drop-down.

The field, Store Identifier, indicates where the object represented by the record is physically located within the store. This follows a simple numerical system that coordinates to shelves beginning with 1 and expanding to encompass the total number of shelves in the retail store. It is not searchable as it is not a field users utilize in searching; it only helps users in the Obtain FRBR task. It can, however, help internally for functions of staff. It is a required field as all objects in the location have a shelf. It does not have a controlled vocabulary, as the number of shelves and identifiers changes over time. It also does not have a drop-down as there are too many identifiers to include.

Format is related to but distinct from Travel-Sized, as many objects that are catalogued as Yes to Travel-Sized may share the same Format, but this is not true for all objects. This data is entered into the field as Text, as the options for the field are all text-based. It is not required because some objects may not match any of the terms. Up to three terms are permitted for this field as it is possible that an object may include a Dictionary portion as well as a Phrase Book portion. The field has a controlled vocabulary including the most common formats for German language learning materials. These include Phrase book, Dictionary, Workbook, Study Guide, and Short Story.

The Subject field is catalogued as text, as the terms are descriptors of a given Subject. The cataloger is required to enter this field as it is a primary method of search as evidenced in Section 1.3, and the field is also searchable due to this. Six terms are permitted for this field as a book in the system may focus on several Subjects. The field does not have a controlled vocabulary or a drop-down list, as the type of Subject can be specific or broad. Only six terms are permitted in order to minimize time for each object and focus on more obvious Subject material that is noted in key areas of the object such as the covers or chapter pages.

As indicated in Section 1.3, cost is a factor users use to inform their selection process and so it is a required and searchable field. The information system is housed in a retail bookstore, not a library or other institution where cost would not normally be a factor. For the field, prices are determined from the MSRP listed on the object and are placed in one of six different numeric, textual categories: “$10 and under”, “$10-20”, “$20-30”, $30-40”, “$40-50” or “$50 and above”. This prevents users from needing to know the exact price of a volume and allows for more flexibility while keeping the number of variants in the field small. Only one entry is allowed; if the price meets an exact number such as $10, it is placed in the category just below. An object priced at $30 exactly is entered in the “$20-30” category to maintain consistency. There are only six categories, so this is included as a drop-down list.

The Illustrated field is a Yes/No field that indicates whether or not the object has illustrations included. This is determined by a quick flip through and if there is more than one illustration inside, the field is marked Yes. Section 1.3 suggests that users may want to search by this field, so this is a searchable field. It has a controlled vocab as there are only two options, Yes/No. These options are listed on a drop-down.

Travel-Sized is a Yes/No field indicating whether the object is travel-sized or not. The cataloger is required to enter this as it is determined by a quick visual judgement and either the object is small in dimensions or not. Most objects indicate whether or not they are intended to be carried as a travel aid; if they are labelled as such, the field should be entered as Yes. If the material is no larger than 6 x 6 x 2, it should also be classified as Travel-Sized. Only one term is permitted in the field as only two options are available, Yes or No. These are also the two options within the controlled vocabulary and are placed in a drop-down list. The field is searchable, as indicated by Section 1.3—some users may prioritize this field.

The field, Skill Level, was previously not addressed in terms of its methods of distinction or importance. It is a somewhat subjective field. Many books in the collection are labelled by creators with the anticipated skill level required to use, however not all are. Three categories are included as controlled vocabulary—Beginner, Intermediate and Advanced. All entries are in one of these three categories. If there is no indication of the Skill Level in the book, it is up to the discretion of the cataloger to decide its classification. Books that have more specific Subjects are not in the Beginner category, as most beginners are not at a point where they are learning about very specific topics. It is possible to include up to two entries in the case that the cataloger is unsure or if the book’s contents are more ambiguous across lines.

The Additional Features field is descriptive. As such, it is entered as text. This is not a searchable field, as it was not indicated as a primary search priority in Section 1.3. Many books use Additional Features as a selling-point and clearly mark whether these are present in the volume. Additional Features is not required as a field, as not all objects in the collection have them. Three terms are allowed in this field, as there can be several Additional Features of an object. There is not a controlled vocabulary for this field. While many objects may have the same or similar features, i.e., study guides, there is too much potential variation in what could be considered an Additional Feature. The field does not include a drop-down as there are too many potential options.

Series is a descriptor, much like a Title, and is entered into the field as through text. The cataloger is not required to enter this as not all objects are a part of a Series. Only one term is allowed in the field as objects are not typically part of more than one series. The field does not have a controlled vocabulary as there are innumerable Series that exist, and it is a field that is subject to change over time. The field also does not include a drop-down for this same reason; there are too many potential Series to include. Series is searchable. As suggested in Section 1.3, users may seek out materials from a specific series. This is important as well for users who are using a Series of learning aids that are sequential and are searching for a later volume in the same Series.

The Cover Image field is not searchable, as users are not using this metadata to find objects. It is not a required field as it is not a primary element that users illustrate preference for, as seen in Section 1.3. Only one entry is allowed, and the entry consists of the front cover only. There is not a drop-down or a controlled vocabulary as it is not possible to know every iteration of this field that is possible.

See field names and associated record structure specifications in Appendix B, table 1.

**2.4. Record content and input rules**

The metadata elements that the cataloger locates come from the chief source(s) of information. In other forms of media this may include the insert for a DVD or a CD jewel case. This collection is comprised entirely of books, so the chief sources of information come primarily from the outside covers, inside covers, title page recto and verso, copyright page, title or chapter pages or index. How the cataloger locates these elements is dictated by content rules. Input rules are the constraints that dictate how a cataloger deals with actually entering a record into the database. Input rules ensure that catalogers are able to enter records into the database quickly and efficiently with little room for error. They also ensure that records are treated consistently.

Refer to Appendices C and G.

**3. Access and authority control**

Authority control is the creation of a controlled vocabulary which contains authorized terms or names; this ensures only approved and recognized instances of a term are used which makes it easier for users to search for terms and less time for the cataloger. This affects spelling and word form of specific words, proper names and more. Authority control dictates the process of indexing on the part of the cataloger and affects users as they interact with an information system by providing standardized vocabulary. There are two major types of authority control—name authority control and subject authority control. Each of these forms directs a cataloger to use accepted terms when entering metadata into a field. Only accepted terms may be used, and these terms come from an outside document that provides a list of pre-approved terminology. These documents may include name authority files or thesauri.

Fields can either describe the bibliographic information or subject of an object. Bibliographic information typically has more to do with the physical object and can include areas such as ISBN or publication date. Subject fields include intellectual content which is determined by subject analysis at the time of cataloging. This can include areas such as plot summary, theme, etc. The subject indicates what the content of an object is; this is what differentiates it from bibliographic fields.

Fields involving subject or intellectual content benefit greatly from subject authority control as they aid users in locating information on a specific topic by providing consistency in terminology. Implementing authority control helps ensure consistency and uniformity in a database as well as consistent search results. Authority control helps direct end users to related terms by demonstrating what terms will yield results, as well as show them related terms they may use for browsing. For instance, a user who searches the term “phonology” will see “umlauts” and “diphthongs” as more narrow terms within their search.

In this information system, Subject is under subject authority control and utilizes a thesaurus while Author, Publisher, and Series would benefit from name authority control as they deal with proper names. The name authority file provides guidelines for the Author field (See Section 5).

**4. Representation of information content**

**4.1. Subject access**

Subject indicates the intellectual content of an object and is frequently a primary method of search for end users in an information system. As indicated in Section 1.3, users prioritize searches in part by subject. Subject authority is the practice of managing subject or intellectual content of records. This is accomplished through the creation, use and management of subject authority files which contain controlled vocabulary terms that represent intellectual content. This is a labor and expertise-intensive process which is managed by professionals in authority work. There are two major types of subject authority files—thesauri and subject heading lists.

Before these files can be developed, first a process of subject analysis must occur.

In order to determine what the subject of an information object is, catalogers begin a process of subject analysis. Subject analysis is the process by which the intellectual content of an information object is determined. When undertaking this process, subject representations are developed that serve as ways to access this content. Subject representations are terms used to represent information objects in a collection and are necessary in order to organize and standardize subjects in an information system for user access. Subject representations are developed in consideration of information seeking behavior, and when the representation matches a search, users are provided with more consistent and accurate results. This also aids users at a variety of different domain knowledge levels to identify what terms are relevant.

Subject analysis consists of several steps in which the entity analyzing the object begins by becoming familiar with it. In the case of a physical book, this may include viewing the front and back cover as well as flipping through the pages. The individual does not read the entire tome; they are simply getting a general overview of its interior contents. The second step, extraction, involves pulling key terms or concepts from the object. Next, in the third step—translation, the cataloger adapts the terms that they have pulled from the object. If there is a controlled vocabulary, the cataloger references this. The final step is formalization where the cataloger enters the information into the information system following any pre-approved formatting guidelines. These representations identified by the cataloger, when matched, work in tandem with the representations of information needs from users to connect users with desired information.

When extracting vocabulary in the subject analysis process, there are two major ways for the cataloger to deal with their results, natural language indexing (NLI) and controlled vocabulary. Natural language indexing is not under authority control, whereas controlled vocabulary indexing is. Natural language indexing involves taking terms from a document directly which means that are more likely to be widely variable from document to document. All indexing begins by extracting vocabulary in a natural language form. When working within a controlled vocabulary, the cataloger finds the closest appropriate controlled vocabulary term and assigns each extracted term to it. Controlled vocabularies have the benefit of being more uniform, whereas natural language indexing can often create issues for users who are searching with terms that do not exactly match those that have been entered into the information system. Controlled vocabularies can make it easier to connect users to records by limiting variation.

Classification systems dictate the location of information objects and are developed with consideration of how users search for information within a system. In this system of physical books, the classification system reflects the location of the object on a shelf. The identifying element of the classification system therefore indicates to the user where to find a specific object in the collection. The majority of classification systems organize first by subject, so the actual physical location of an object on a shelf will be influenced by how the subject of the object has been determined, which directly affects users in obtaining a book from the shelf. Grouping objects together by subject facilitates user browsing as well, as books with related subjects will be nearby.

Subject is the primary field in this information system that contains intellectual content. All of the objects in this collection have German language learning as a common subject so there is a need to create better searchability through more specific subject matter. In this collection, catalogers follow this process in order to identify intellectual content related Subject field, as these are the fields used for subject access. As indicated in Section 1.3, intellectual content that is being requested largely falls into these two fields with the somewhat subjective field of Subject which is intellectual in nature and can be broad or narrow in scope.

**4.2. Thesaurus structure**

Thesauri in the context of subject authority control are lists of approved terms that a cataloger can assign to a record at the time of cataloging and are one of the two major types of subject authority files. They allow for a controlled vocabulary to be used, usually directly related to the subject of a record, and create consistency between records in an information system. The Subject field of this collection is under a subject authority control with a thesaurus due to its direct relationship to intellectual content for objects contained in the information system. Controlled vocabulary in a thesaurus indicates what terminology is permitted to be indexed and what will yield results in a search. Authorized terms are terms that have been approved in the thesaurus and unauthorized terms are those that are not permitted. While thesauri are beneficial to catalogers, they also aid users of an information system as a resource for browsing as well as gaining a better understanding of the range of available topics, rather than having to guess which terms will provide results. It also ensures that items about similar subjects are organized together which aid in the final FRBR task—to obtain.

The other major type of subject authority file is the subject heading. These are lists utilized by catalogers to indicate intellectual content through what are often hierarchical relationships. This information system utilizes a thesaurus; however, it would be possible to introduce a subject heading file as a form of subject authority as many of the subjects represented would fit into a hierarchical structure such as “Verbs” being a subset of “Parts of Speech”. Subject headings allow for items about similar or related topics to be associated together, making it easier for users to locate additional related items to their search. They also ensure that there is order when dealing with elements like synonyms and related terms. For hierarchical terms, this may be represented through terms including “Veterinary science”— “Domestic cats”, illustrating a hierarchical relationship.

The relationship of one term to other related terms in a controlled list of vocabulary is expressed as a syndetic structure. These semantic relationships can be either equivalent, hierarchical, or associative and notated in the thesaurus through the following indicators: UF, USE, BT, NT, and RT. UF indicates a term that is not authorized, where USE indicates a term that may be used instead. Broader terms, narrower terms, related terms. Equivalent relationships are descriptive of terms that are essentially synonyms such as the words “Phonology” and “Phonetics” where both words suggest the manner in which a person creates a sound of speech. However, in this example, only the term “Phonology” is an authorized term which is typical of an equivalent relationship—one term is usually authorized where the equivalent term is not. In the case of equivalent or nearly equivalent terms, one term must be chosen as the preferred term for the controlled vocabulary. A hierarchical relationship includes terms that are more specific subsets of a more broad term, such as how “Adjectives” are a subset of “Grammar” in a hierarchical relationship. Hierarchical relationships may only contain authorized terms. Associative vocabulary includes terms such as “Pronouns” and “Verbs” which are both associated with “Parts of speech” but are not equivocal.

Mandatory reciprocals within a thesaurus ensure that there is consistency in the relationships between terms. If a term such as “umlauts” is indicated to be a narrower term of “Phonology”, the term “Phonology” must have a mandatory reciprocal in order to illustrate the relationship the other way. This means that “umlauts” must be indicated as a narrower term (NT) underneath “Phonology”.

Domain is the overall theme or scope of content that is contained in a thesaurus. In this case, the collection itself, German language learning books is the domain. Scope outlines the limitations of the domain, if applicable. In this system, the scope and domain are identical as they both describe the entire collection and there are no particular restraints on scope. Specificity is medium to high for this thesaurus, as many elements of language can be defined with a fair level of definitiveness such “Nouns” or “Adjectives”, however some subjects are more abstract such as “Regionalism”. As indicated in Section 1.3, some users are seeking information on specific elements of the language, i.e., “Umlauts”, though some are searching more broadly as well. The level of exhaustivity of this thesaurus is somewhat high—as indicated in Appendix B., up to six terms may be entered. This is essential for this information system as many books in the collection feature many subject terms listed in the thesaurus and users may desire books that cover a wide range of topics.

Refer to Appendix D: Sample thesaurus.

**4.3. Classification scheme**

Classification is the way in which objects are organized and summarizes metadata about an object. There are two major types, hierarchical and faceted. Hierarchical classification schemes are similar to an umbrella or tree model, with a more broad term at the top that can be broken down into more specific levels underneath. Hierarchical schemes can be advantageous when dealing with a collection that is heavily influenced by subject in its organization and is unlikely to undergo significant changes over time in its facets. As these schemes typically consider as many concepts as possible, they are more rigid and difficult to adjust once implemented. A hierarchical scheme is likely a beneficial choice if the system contains records which are more varied. Faceted schemes include as many relevant subject areas as possible but do not organize them in a hierarchy, rather they group relevant subjects together as they are associated. They can be easier for the cataloger to use in indexing and work well for collections that do not focus on subject as heavily. This classification scheme is more flexible than a hierarchical scheme and can be adapted over time more easily, so this is the classification scheme this collection utilizes.

While many major subject elements fit logically under a larger umbrella term such as “Parts of speech” with NT “Verbs” and “Nouns”, topical records like those on “Music” may warrant new records in the future.

As a large number of subject terms are related to grammar, the primary facet, the class Grammar (GMR) organizes these books together as being related. Similarly, the Pragmatics (PRGM) class organizes based on topics related to the subject.

The classification code used in this information system consists of three major facets: Subject, skill level and author. As indicated in Section 1.3, skill level and subject are both important to user search needs and author helps to aid in differentiating material on the shelf. After the three facets have been addressed following notation rules in Appendix E., a three digital numerical code is added to the end as a unique identifier for the volume. The unique identifier is determined by the order of acquisition in relation to the other volumes with the same facets. An example is provided below.

Advanced Conversational Idiom in German

W.E. Anderson and H.K. Kägeler

The book is primarily focused on the topic of idioms, so the first part of the classification code IDIO follows the Subject facet name instructions to use the first four letters of a CV term from the thesaurus— “Idioms”. This is capitalized according to notation rules and followed with a period. The A indicates that this is an advanced book which was determined during subject analysis through the title and contents as the book explicitly states that it is for an “advanced” topic. The first author listed is W.E. Anderson; according to notation rules this individual’s last name has been listed in capital letters. This is followed by a period and a unique numerical identifier.

IDIO.A.ANDERSON.001

Refer to Appendix E.

**5. Name authority control**

Name authority control covers proper nouns including names of individual people and organization or groups. For example, in considering a film, this may include featured actors. For books it may include authors, publishers and more. In order to implement and follow a name authority control system, a name authority file must be created and referenced during the cataloging process. A name authority file is a resource separate from the database which provides information regarding authorized forms of names. It is sometimes integrated in such a way that the cataloger is directed to the authorized version of a name, however it also sometimes must be referenced outside of the database in another document. The management of a name authority file is carried out by technical users who regularly update it to reflect any relevant name changes or other issues. The name authority file is structured to contain at least three elements including the authorized form of a name, variations of a name and reference to what sources were used to determine the authorized version. Authorized names are typically chosen based on their frequency of use and more recent usage. Other factors that can affect this are whether or not the name is more well-known compared to a biographical or other name.

Creating a system for name authority control ensures that there is consistency in records which allows catalogers to save time and therefore money. For end users, name authority control systems make it easier to find all records containing the name of one group or individual in particular, even if they go by a pseudonym or alternative spelling. For example, most people would likely not search for “Lewis Carroll” under his real name—Charles Lutwidge Dodson. If some records were only to have “Lewis Carroll” entered into the database record (as they may appear on a book), a user searching for “Charles Lutwidge Dodson” would only find records with the latter entry. Similarly, if a user was to enter “Charles Dodson” when the authorized name is “Charles Lutwidge Dodson” or other variations, recall would be low as well. An ideal implementation of name authority control would ensure that users are able to retrieve records for both iterations of his name. Name authority also controls the method by which a name is entered with regards to spelling, punctuation, and name order.

In Section 3, it was indicated that the fields for Author, Publisher and Series would require name authority control. Appendix F. provides examples of how the Author field is under name authority control. Name authority control is necessary in this information system as there are authors with name variations. For example, Edward Swick’s books in the collection list “Ed Swick” as the author on the cover, however this is not the case for all of his books, other materials, and his information online. The Library of Congress Name Authority File is an authoritative source on proper names and is used as the principal source of authorized name forms in the name authority file for this database. Following this resource, this author’s name is entered into the record as “Edward Swick”. Additionally, it is often preferable to use an individual’s full name—in this case “Edward” instead of “Ed”. Some of the formatting is changed from the way it is displayed through the Library of Congress, however, such as listing author names in order by first then followed by last time in order to match the way in which records appear in Libib. If a name is unavailable through the Library of Congress, the term LCNA (Library of Congress Not Available) is used in order to indicate that no previous record was found. This is followed by the year the file was accessed. By utilizing name authority control for author names, this ensures that users are able to locate all records by the same creator and it saves time on the part of those entering information into the records.

Refer to Appendix F.

**6. System evaluation and development**

### 6.1. SWOT

**SWOT is an acronym that stands for strengths, weaknesses, opportunities, and threats. A SWOT analysis is a tool used that helps organizations in planning for their future. Each section, organized into a 4x4 chart, outlines elements that fit into one of each of the four areas. Strengths are positive aspects and may lend the organization an edge over others. Weaknesses are areas that need improvement and should be considered for potential improvement. Opportunities are positive and Threats are negative but rather than be more actively affecting an organization, both are elements that should be taken into consideration for the future. All four areas are aspects that need to be considered when making important decisions, as any entity should play to their strengths and weaknesses while considering potential opportunities and threats that could affect them.**

While Libib is easy to understand and created in such a way that records can be entered quickly, this information system has many weaknesses due to its reliance on the software. Several fields, as indicated in Appendix B., Table 2, are largely identical to pre-existing Libib fields such as Title, Author, Publication Date, ISBN, and Publisher. Other fields identified for this system including the Subject, Skill Level and others do not translate well into Libib’s pre-existing framework. A more desirable database would allow for several additional fields, rather than on the Tags field in Libib. In entering several different Fields into the Tags field in Libib, it creates a sense of disorder, especially in consideration of the fact that Libib’s Tag system automatically rearranges entries into alphabetical order no matter the order they are originally entered. The ability to filter by Tags would be more beneficial to end users if there was a degree of separation between all of the different fields that have been merged into it. Additionally, when viewing the column to the right-hand side of the interface, the Tags are arranged in alphabetical order in a long list which is difficult to navigate. Fields that are not necessary to the system should not be included, such as Status. This system is built for a retail establishment, and it is unlikely that users would benefit from this feature.

The thesaurus provided in Appendix D. helps ensure that records in the database are consistent and that multiple versions of the same Tag are not used for a Subject, however unless users have access to this document, there is no way for them to know they are searching for an unauthorized term in Libib. For example, as indicated in Appendix D., “spelling” is not an authorized term. The authorized term is “orthography”. “Spelling” is a much more colloquial term that most users would likely use to search over “orthography” but when searching for this through the system, they will not retrieve any results. In this case, it is possible that the user will interpret this to mean that there are no records in the system for this topic, when there are in-fact two.

While it is possible to enter data into a field to indicate the price of an object in the database, there is no ability to filter or search by price which is an incredibly important feature for a retail establishment in order to stay competitive. There is a field for this information, however, so it is a good starting point for potentially integrating this feature in the future. This would be an important addition as users may be deterred by the inability to filter by cost and turn to other establishments that provide this function. Libib seems to have been created for smaller libraries with rental procedures, which may be more beneficial for certain collections including those than contain fiction primarily. In searching for non-fiction books, as indicated by questions in Section 1.3, users are often seeking answers to highly specific questions. Therefore, a system that is somewhat limited like Libib is likely not the best choice to represent this collection.

There is a way to identify records as belonging to the same group, such as a series of titles, however, there is not a meaningful or obvious way to retrieve records based on this information.

Arguably, the only significant benefit to this system is the more specific distinctions made through Tags such as difficulty level or Subject, however this is also the biggest weakness of the system as it is frustrating to navigate.

Authorized name forms adapted from The Library of Congress Name Authority File must be altered in order to fit more readily within the Author field of Libib. Libib requires multiple authors to be separated with a comma and a large majority of authorized names in the Library of Congress Name Authority are formatted as “Last Name, First Name”. Many of these authorized names also include dates following them to indicate active periods by an individual. Entering names in this form causes complications with Libib as it seems the system was created with the intention of providing a first name followed by a last name. If the database entry in Libib was to follow the same format, names would not display properly, as only the last name of the first author is displayed unless the record is viewed in detail.

There are some benefits to hosting this information system through Libib, however. The ability to view items with large displays of the cover illustrations is appealing as a potential buyer. While this feature seems to be intended for tracking copies for lending purposes, the ability to see how many copies of an item are available is ideal for a retail setting. Depending on how quickly it is possible to see the reflection of a book being taken from the record, this could be a very desirable feature for users looking for books to purchase; they are able to see which books are readily available which may influence their purchasing decisions when compared to establishments that do not have this feature. Some features such as the ability to batch process records are potentially very beneficial for catalogers.

**SWOT Analysis Chart:**

|  |  |
| --- | --- |
| **Strengths**   * Easy for beginners to learn to use—both for cataloging and searching * Ability to batch process several records at once makes it easier for catalogers to update information when necessary | **Weaknesses**   * Not enough fields for subject-related metadata * Name authority is not well-integrated * Status field is not necessary, but cannot be removed * Searching by series does not work in keyword search * Authority controlled terms are not always obvious to the user * The user is unable to filter by price |
| **Opportunities**   * Number of available copies of an object could be a valuable tool, especially in a retail setting * Ability to enter exact pricing and update it easily could be helpful as the database for a retail establishment | **Threats**   * Reliance on Tags field could lead to a great amount of confusion for both the catalogers and end users * Retailers with more user-friendly systems will draw away business |

**6.2. Change and development**

Many of the issues that this system faces could be minimized or solved with the ability to separate originally desired fields, rather than having several share the Tags field in Libib. As indicated above, the auto-alphabetization in one of the more harmful aspects of this system. Having the cost of each object represented in its record is important for a retail establishment, however the ability to sort by price or choose from brackets of cost such as $10-$20 would be more beneficial to users who are searching for books to purchase. This could be accomplished by implementing either a sorting feature or a checkbox system for monetary increments.

Many issues could also be resolved simply by including checkboxes like many other systems utilize including journal databases like EBSCO or JSTOR. Checkboxes would save time in entering Fields such as “Travel-Sized” by reducing the number of keystrokes. When implemented for several fields like “Illustrated” and “Additional Features” as well, this could save a good deal of time and money by streamlining the record entry process. Larger databases like EBSCO also provide more advanced search functions. Libib has a basic keyword search, but the only way to search in an alternative method is to sort or filter by Fields such as individual Tags, Publication or Title. A more sophisticated search function would make it easier to search for a specific record; at present, the experience of using the database as an end user feels more like browsing than truly searching for a specific record.

At this time, entering an unauthorized name into the search bar will yield no results. For some names that are abbreviations, the keyword search may still be able to retrieve records. For example, searching “Ed Swick” still yields results with the authorized name “Edward Swick”. In cases where the author uses a pseudonym that is completely different, it is likely that the user will not be able to retrieve any results at all. Many larger systems have a source directly incorporated into the database that will automatically retrieve authorized names, saving time and effort on both the part of the cataloger and the end user. Much more sophisticated in comparison and this may be due to the nature of its use as a primarily academic resource.

**7.0. Project summary**

I chose this specific collection because I am a beginning German language learner and I felt frustrated when searching for books to learn from. I was very specific in my needs but could often not find information in any system that gave me a good indication of what the contents would actually be unless I saw the book physically in person. Several of the traits that became tags reflected my experiences in looking for materials on German including the travel-sized and form considerations. Because of this I am glad that this is the topic I chose for my collection, as I felt it would be a good fit for the requirements as well as give me an opportunity to view more books on a subject I was interested in.

While Libib is not the most ideal system in my personal opinion, I think that the ideas I developed for it within a better system would be helpful for others who are in my position when looking for German books. Most of my challenges stemmed from Libib, however I see the merit in a system that requires very little knowledge of computers. It is also a very easy system to set up and for a very basic database where the metadata is not complicated, this would be an excellent system. This also seems to be a good system for smaller collections. The biggest weakness I think is in having to rely on the tagging system for subject information. There are likely other ways of doing this in Libib, but the tags field seemed the most reasonable to me. In an ideal system, I would prefer having a bit more control over that aspect. It was a bit disheartening as well to see many of my considerations overwritten by Libib such as the automatic alphabetization of tags. If I had realized it did this, I would likely have approached the fields of Skill Level and Format differently. I think that if I had been able to create drop-down lists or checkboxes it would have made the process much better. For some of the elements of metadata I wanted to capture, like whether or not a book was intended for travel use, a checkbox would have been much better as it is a yes or no question. In retrospect, the time it takes to enter this as a typed tag is likely not worth the time for every single book. Now that I am aware of Libib’s quirks, I would be able to build another database from scratch again and more quickly if I had to use this system again. My only concern from this point on is if and when I am in the position where I need to develop a database with different software and experience similar learning curves once again. I do feel that my understanding of the methodology of creating a database is much better, however.

I think overall that my general understanding of the types of practices and considerations that are involved in creating an information system have evolved but I would likely make quite a few changes to my system before I was to ever implement it in a real-life situation. I am still not sure I would be comfortable in working with database software that required more computer-based knowledge, but I think I would be able to create a system again from something like Libib. I am definitely at a point where I could express to an entity creating a system what I would desire from an ideal one, if that was ever asked of me, however.

**Appendix A. Metadata elements and semantics**

|  |  |  |
| --- | --- | --- |
| **No.** | **Element name** | **Semantics** |
| 1 | Title | Name of the object |
| 2 | Author | Writer responsible for the object |
| 3 | Publication Date | Date that the book was published |
| 4 | Publisher | Organization responsible for publication |
| 5 | ISBN | Universal identifier provided for each object |
| 6 | Store Identifier | Unique three-digit numerical code indicating the location of the object within the store |
| 7 | Format | Arrangement of information—i.e., dictionary, phrasebook, narrative, etc. |
| 8 | Subject | Specific or focused area of knowledge—i.e., grammar, conversation, etc. |
| 9 | Cost | Price of the object in USD |
| 10 | Illustrated | Images are/are not provided |
| 11 | Size | Physical dimensions—i.e., classification of an object as travel/pocket-sized or not. |
| 12 | Skill Level | Knowledge level expected of the user: Beginner, intermediate, or advanced. |
| 13 | Additional Features | The object does/does not include special or extra features—i.e., web interactivity, study aids, etc. |
| 14 | Series | Status as part of a series or standalone publication |
| 15 | Cover Image | Photo of the front cover of the book |

**Appendix B. Record structure and specifications**

**1. Record structure specifications**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Field name** | **Field type** | **Searchable** | **Required** | **Number of allowed entries** | **Controlled**  **Vocabulary?** | **Drop Down**  **List?** |
| 1 | Title | Text | Yes | Yes | 1 | No | No |
| 2 | Author | Text | Yes | Yes | 3 | No | No |
| 3 | Publication Date | Date | No | Yes | 1 | No | No |
| 4 | Publisher | Text | Yes | Yes | 1 | No | No |
| 5 | ISBN | Numerals | No | No | 1 | No | No |
| 6 | Store Identifier | Numerals | No | Yes | 1 | No | No |
| 7 | Format | Text | Yes | No | 3 | Yes | Yes |
| 8 | Subject | Text | Yes | Yes | 6 | No | No |
| 9 | Cost | Numerals | Yes | Yes | 1 | Yes | Yes |
| 10 | Illustrated | Text | Yes | Yes | 1 | Yes | Yes |
| 11 | Travel-Sized | Text | Yes | Yes | 1 | Yes | Yes |
| 12 | Skill Level | Text | Yes | Yes | 2 | Yes | Yes |
| 13 | Additional Features | Text | Yes | No | 3 | No | No |
| 14 | Series | Text | Yes | No | 1 | No | No |
| 15 | Cover Image | Image | No | No | 1 | No | No |

**2. Field comparison**

|  |  |  |  |
| --- | --- | --- | --- |
| **No.** | **Desired Field** | **Libib Field** | **Notes** |
| 1 | Title | Title | Identical to Libib |
| 2 | Author | Author | Identical to Libib |
| 3 | Publication Date | Published Date | Identical to Libib |
| 4 | Publisher | Publisher | Identical to Libib |
| 5 | ISBN | ISBN 13 or ISBN 10 | Identical to Libib. If the book has an ISBN 10, the ISBN 10 field is entered while the ISBN 13 field is left blank and vice versa. |
| 6 | Store Identifier | Call # | The Store Identifier is a numeral with the same function as a Call #, so this is the field that is utilized to indicate physical location of the object in the record. |
| 7 | Format | Tags | Controlled vocabulary remains the same, however the Format type is entered as a Tag in Libib |
| 8 | Subject | Tags | There is no Subject field in Libib. Subject is entered as a Tag. |
| 9 | Cost | Tag | There is a cost field in Libib, but it must be added after the record is created and allows only for specific dollar amounts, so this is the manner in which this data is entered. |
| 10 | Illustrated | Tags | There is no drop-down or pre-existing field to indicate whether illustrations are a present element, so the Illustrated field is determined by the Tag “Illustrated” only if there are illustrations present as determined by the cataloger. As such, this field is not required in Libib if there are not illustrations, as was originally conceived in Appendix B part 1. |
| 11 | Travel-Sized | Tags | Libib does not provide drop-down lists, so this field is determined through the Tag “Travel”. This is also not a required field in Libib and is only to be entered if the book meets the requirements of being Travel-Sized. |
| 12 | Skill Level | Tags | There is no pre-existing field in Libib for Skill Level, so entries use the Tags field with text indicating Beginner, Intermediate and Advanced as necessary. |
| 13 | Additional Features | Tags | As there is no field for Additional Features in Libib and no drop-down, check-mark box or other feature, the Additional Feature is added as a textual description of the given Feature as a Tag in Libib. |
| 14 | Series | Series/Group | The Series/Group field in Libib is the closet correlation to Series, however post-SWOT analysis, this may be better suited to Tags. |
| 15 | Cover Image | Cover Image | Identical to Libib |

**Appendix C. Record content and input rules**

**Field #: 1**

**Field Name: Title**

**Semantics:** Name of the object

**Chief Source of Information: 1. Front cover, 2. Copyright page**

**Input Rules: Title is required and only one is permitted. Find the Title on the front cover of the object or on the copyright page and enter it as displayed, following the same punctuation****. Capitalize using title case and ignore special formatting such as italicization or all-caps words.**

**spelling and inclusion of special characters. Include a subtitle if one is provided.**

**Example: Title: Lexicalist Phonology of English and German**

**Field #: 2**

**Field Name:** Author

**Semantics: Individual or individuals who created or wrote the book. This can also include an organization or other entity but should be prominently displayed on the physical item.**

**Chief Source of Information: 1. Front cover, 2. Copyright page**

**Input Rules:** **The Author field is required to have at least one entry but may have up to three. List authors in the same order as they are provided on the object with names followed by a comma and space. Authors after the third do not need to be entered, but if there are multiple, enter the first three when applicable. Names should be entered in first and then last name order. If the author has a middle name or initial, enter it between the first and last names with spaces between each word. As instructed in Libib., last names with spaces should have quotes around them. Initials are followed with a period and space. If names include any punctuation such as a dash, this is to be included as depicted and capitalize the following word.**

**Example: Author: Adriana Borra, Ruth Mader-Koltay**

**Field #:** 3

**Field Name:** Published Date

**Semantics:** Date that the book was published

**Chief Source of Information: 1. Copyright page**

**Input Rules: Enter the date into the fields in Libib as they appear, with the year, month, and day each in separate boxes. This should be in YYYYMMDD format. If there is no publication date of the book provided, leave the field blank and the system automatically enters 01 for the month and day which is acceptable. If only the year is provided, only fill in the year and leave the month and day blank. This field should only include numerals.**

**Example: 2016**

**Field #: 4**

**Field Name:** Publisher

**Semantics: Group or entity that is responsible for the publication of the book.**

**Chief Source of Information: 1. Copyright page**

**Input Rules: Enter the name of the publisher as it appears in the object. Capitalize all individual words, spell them in the same manner as they appear and include any and all punctuation.**

**Example: Publisher: Foris Publications**

**Field #:** 5

**Field Name:** ISBN

**Semantics: Unique numerical identifier provided by the publisher, consisting of thirteen digits followed by the letters ISBN. Typically located near the bar code.**

**Chief Source of Information: 1. Back cover, 2. Copyright page**

**Input Rules: This is not a required field as not all books with have an ISBN, however, if the book does have one listed, it should be entered. Enter the numerals for the ISBN in the order the are displayed from left to write. Do not include any punctuation; the ISBN entry should only include numerals with no spaces. If the ISBN has ten digits, enter it into the ISBN 10 field; if it has thirteen digits, enter it into the ISBN 13 field.**

**Example: ISBN 13: 9780767918602**

**Field #:** 6

**Field Name:** Call #

**Semantics: Classification number of the object.**

**Chief Source of Information: Subject fields, Skill Level fields and Author fields.**

**Input Rules: See Appendix E for schemes and notation rules.**

**Example: IDIO.A.ANDERSON.001**

**Field #:** 7

**Field Name:** Tags

**Semantics:** This database uses the Tags field in a non-standard way in order to accommodate the metadata fields below.

**Desired Field A (Format): Arrangement of information—i.e., dictionary, phrasebook, narrative, etc.**

**Desired Field B (Subject):** Specific or focused areas of knowledge. See appendices D and E for more information.

**Desired Field C (Cost):** Price of the object in USD

**Desired Field D (Illustrated):** Images are/are not provided

**Desired Field E (Travel-Sized): The object is suited for, or intended to use for travel**

**Desired Field F (Skill Level): Knowledge level expected of the user: Beginner, intermediate, or advanced.**

**Desired Field G (Additional Features):** The object does/does not include special or extra features—i.e., web interactivity, study aids, etc.

**Chief Source of Information:**

**Desired Field A (Format): 1. Front cover, 2. Title page, 3. Remainder of the inside of the object**

**Desired Field B (Subject): 1. Front cover, 2. Title page, 3. Contents/Chapter page, 4. Remainder of the inside of the object**

**Desired Field C (Cost): 1. Back cover, 2. Inside front cover, 3. Publisher’s Website**

**Desired Field D (Illustrated): 1. Entirety of the object**

**Desired Field E (Travel-Sized): 1. Entirety of the object**

**Desired Field F (Skill Level): 1. Front cover, 2. Contents/Chapter page, 3. Remainder of the object**

**Desired Field G (Additional Features): 1. Front cover, 2. Back cover, 3. Inside front cover, 4. Inside back cover**

**Input Rules:**

**Desired Field A (Format): This field is not required, as not as all objects fit the controlled vocabulary. Up to three entries are allowed which may include Phrase book,” Dictionary”, “Workbook”, “Study Guide”, and “Short Story”. Each entry should be entered exactly as it appear with each word capitalized and a space in between. Each entry should have a comma followed by a space.**

**Desired Field B (Subject): At least one entry in this field is required, while up to six are permitted. Review the chief sources of information to identify any subjects mentioned by name in the volume such as “Vocabulary” or “Grammar”. Use only terms that are authorized in Appendix D. Each entry is to be capitalized with spaces between words and followed by a comma and space. Enter terms based on their frequency in appearance on the chief sources of information. If a term occurs more than once, it should be prioritized over a term that only appears once.**

**Desired Field C (Cost): The is a field that requires one entry only. Prices are taken from the MSRP listed on the object, which is typically on the back cover or close to the bar code. Enter the price including only numerals and a period as displayed. It is not necessary to include a dollar sign.**

**Desired Field D (Illustrated): The Illustrated field is not required and only one entry is allowed. If there is more than one illustration or image in the entirety of the book, enter the text “Illustrated”. If there are illustrations, the cataloger must enter “Illustrated”. If there are no illustrations, the field is to be left blank.**

**Desired Field E (Travel-Sized): This field is not required. Similar to the Illustrated field, the term “Travel-Sized” is to be entered if the volume explicitly states that it is intended to be carried for travel or if the object measures** is no larger than 6 x 6 x 2. The term “Travel-Sized” is to be entered exactly as it appears here with both words capitalized and with a dash between them. If the object is larger than the provided dimensions or does not state that it is intended for travel use, the field is to remain blank.

**Desired Field F (Skill Level): This field is required and can have up to two entries. Enter no more than two of these terms for each object, following the same spelling and capitalization as follows: “**Beginner”, “Intermediate” and “Advanced”. **This is the most subjective field, and the cataloger has to make their best judgement call based on the content inside and outside the object. If there is doubt about whether the object matches criteria for one or another category, both may be entered. If the concepts introduced at the beginning of the book include very basic such as simple vocabulary or the alphabet, the book should be entered as “Beginner”. If the book focuses on more specific concepts, the book should be entered as “Advanced”.**

**Desired Field G (Additional Features): This is field which is not required and may have a maximum of three entries. Each entry should be typed and followed with a comma and space between each entry and all individual words should be capitalized. Each term is followed with a comma and a space. Additional Features does not have a controlled vocabulary. Survey the object for the chief sources of information. Some Additional Features may include flash cards, other study aids, audio content, etc.**

**Example:**

**Desired Field A (Format): Study Guide**

**Desired Field B (Subject): Music, Art, Vocabulary**

**Desired Field C (Cost): 49.95**

**Desired Field D (Illustrated): Illustrated**

**Desired Field E (Travel-Sized): Travel-Sized**

**Desired Field F (Skill Level): Advanced**

**Desired Field G (Additional Features): Audio, Flash Cards**

**Field #:** 8

**Field Name:** Series/Group

**Semantics:** Name of the series that the book is part of.

**Chief Source of Information: 1. Front cover, 2. Copyright page, 3. Publisher’s webpage**

**Input Rules: This field is not required, as not all books are part of a series. As with the Title field, enter full the name of the Series, capitalizing all nouns and verbs and placing spaces between words. If there is doubt as to if the title is part of a series, the cataloger may reference the publisher’s webpage to confirm.**

**Example: No Nonsense Knowledge**

**Field #:** 9

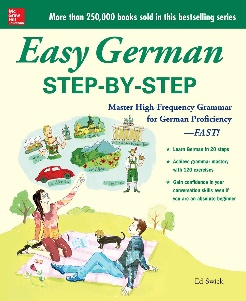
**Field Name:** Cover Image

**Semantics:** Photo of the front cover of the book

**Chief Source of Information: 1. Front cover**

**Input Rules: This field is not required. Take and upload a photo of the front cover of the book in the field as indicated in Libib.**

**Example:**



**Appendix D. Sample thesaurus**

Key: BT = Broader term

NT = Narrower term

RT = Related term

UF = Related term that is not authorized

USE = Term to be used in place of the indicated term

Bolded terms are authorized.

**Adjectives**

BT Parts of Speech

BT Grammar

**Articles**

BT Parts of speech

BT Grammar

**Diphthongs**

BT Phonology

**Film**

**Grammar** (GRMR)

NT Adjectives

Articles

Nouns

Parts of Speech

Prepositions

Pronouns

Syntax

Verbs

**Idioms**

BT Pragmatics

**Music**

**Nouns**

BT Parts of speech

BT Grammar

**Orthography**

UF Spelling

**Parts of speech**

BT Grammar

NT Adjectives

Articles

Nouns

Pronouns

Verbs

Phonetics

USE Phonology

**Phonology**

NT Diphthongs

Umlauts

UF Phonetics

Pronunciation

Speech sounds

**Pragmatics (PRGM)**

NT Regionalism

Idioms

RT Travel

**Prepositions**

BT Parts of speech

BT Grammar

**Pronouns**

BT Parts of Speech

BT Grammar

Pronunciation

USE Phonology

**Regionalism**

BT Pragmatics

RT Travel

Speech Sounds

USE Phonology

Spelling

USE Orthography

**Syntax**

BT Grammar

**Travel**

RT Pragmatics

Regionalism

**Umlauts**

BT Phonology

**Verbs**

BT Parts of speech

BT Grammar

**Vocabulary**

**Appendix E. Classification scheme**

**1. Scheme**

|  |  |  |
| --- | --- | --- |
| Facet 1: Subject | Facet 2: Skill Level | Facet 3: Author |
| Grammar (GRMR) | See rule below | See rule below |
| Pragmatics (PRGM) |  |  |

**2. Notation rules**

**Facet name:** Subject

**Chief source of information:** Subject field

**Notation rules:** Check for a related class first and use the associated code. If there is no class, use the first four letters of one subject term from the thesaurus, all capitalized. Follow the facet with a period.

**Facet name:** Skill Level

**Chief source of information:** Skill Level field

**Notation rules:** Use the following letter codes as applicable—B for beginner, I for intermediate and A for advanced. If more than one Skill Level has been indicated, the classification should list them in order beginning with B, followed by I and then A with no punctuation in between. Letters should all be capitalized. Follow the facet with a period.

**Facet name:** Author

**Chief source of information:** Author field

**Notation rules:** Use the author’s full last name in capitals. If there are multiple authors, use the first authors listed.

**3. Rule for unique number:** The unique number is to be a series of three numerals ordered sequentially from left to right on the shelf where the object is located.

**4. Example**

Advanced Conversational Idiom in German

W.E. Anderson and H.K. Kägeler

The book is primarily focused on the topic of idioms, so the first part of the classification code IDIO follows the Subject facet name instructions to use the first four letters of a CV term from the thesaurus—“Idioms”. This is capitalized according to notation rules and followed with a period. The A indicates that this is an advanced book which was determined during subject analysis through the title and contents as the book explicitly states that it is for an “advanced” topic. The first author listed is W.E. Anderson; according to notation rules this individual’s last name has been listen in capital letters. This is followed by a period and a unique numerical identifier.

IDIO.A.ANDERSON.001

**Appendix F. Name authority file**

**1. Record content and input rules**

**Field #: 1**

**Field name:** AuthorizedName

**Semantics:** Authorized form of a name for use in the database.

**Input rules: Refer to the Library of Congress Name Authority File with the following considerations. Use the most commonly listed and recent name. As indicated in Appendix C., Field #2, only three maximum names are allowed. Variant names take priority over additional authors. Enter pseudonyms directly following their associated name. List variant names exactly as they appear including the same spelling and punctuation—however, ignore capitalization past the first letter of each name and ignore special formatting including italics. Do not invert names and do not include commas between first and last names. Omit credentials including terms such as “Dr.”. Omit years or other dates. If there is no name listed in the Library of Congress Name Authority, copy the name as it appears on the object following the rules above, follow with the term LCNA to indicate no record was available, followed by the year entered.**

**Example: Karl A. Schmidt; LCNA; 2022.**

**Field #: 2**

**Field name:** VariantName

**Semantics:** Other iterations or versions of authorized names.

**Input rules: Refer to the Library of Congress Name Authority File.** Use the same spelling and punctuation as the name is listed. If the variant is not listed, enter the variant following the authorized name using the same format as for the authorized name.

**Example:**

**Field #: 3**

**Field name:** SourcesUsed

**Semantics:** Source referenced in determining authorized form of a name.

**Input rules:** Enter title of resource with the same spelling and punctuation. Ignore special formatting such as italics. Capitalize only the first letter of each word. Follow the title of the source with the LC control number and the year accessed. Separate each element with a semi-colon.

**Example:** Library of Congress Name Authority File; n 2006042519; 2022.

**2. Sample Name authority records** (choose 10 names from your main database)

**Record #1**

**AuthorizedName:** Edward Swick

**VariantNames:** Ed Swick

**SourcesUsed:** Library of Congress Name Authority File; n 98047327; 2022.

**Record #2**

**AuthorizedName:** Adriana Borra

**VariantNames:** Aria Borra; Aria Borra-Olympia

**SourcesUsed:** Library of Congress Name Authority File; n 2006042519; 2022.

**Record #3**

**AuthorizedName:** RuthMader-Koltay

**VariantNames:** Ruth M. Koltay

**SourcesUsed:** Library of Congress Name Authority File; n 2006042524; 2022.

**Record #4**

**AuthorizedName:** Paul G. Graves

**VariantNames:** Paul Graves

**SourcesUsed:** Library of Congress Name Authority File; n 82235481; 2022.

**Record #5**

**AuthorizedName:** Elmer H. Antonsen

**VariantNames:** Elmer Antonsen

**SourcesUsed:** Library of Congress Name Authority File; n 89662920; 2022.

**Record #6**

**AuthorizedName:** Josephine Barber

**VariantNames:** Joséphine Barber

**SourcesUsed:** Library of Congress Name Authority File; n 84238524; 2022.

**Record #7**

**AuthorizedName:** Karl A. Schmidt

**VariantNames:** Karl Schmidt

**SourcesUsed:** Karl A. Schmidt; LCNA; 2022.

**Record #8**

**AuthorizedName:** Arnold Leitner

**VariantNames:** Arnie Leitner

**SourcesUsed:** Library of Congress Name Authority File; n 2006068832; 2022.

**Record #9**

**AuthorizedName:** Elizabeth Bingham

**VariantNames:** Liza Bingham

**SourcesUsed:** Library of Congress Name Authority File; n 2001029271; 2022.

**Record #10**

**AuthorizedName:** Martin Durrell

**VariantNames:** Martin Dodson; German Meister

**SourcesUsed:** Library of Congress Name Authority File; n 90716885; 2022.

**Appendix G. Sample records**

1.

Graphical user interface, text, application, email

Description automatically generated

2.

Graphical user interface, text, application, email

Description automatically generated

3.

Graphical user interface, text, application, email

Description automatically generated

4.

Graphical user interface

Description automatically generated

5.

Graphical user interface, text, application, email

Description automatically generated

6.

Graphical user interface, application

Description automatically generated

7.

Graphical user interface, text, application, email

Description automatically generated

8.

Graphical user interface, text, application, email

Description automatically generated

9.

Graphical user interface, text, application, email

Description automatically generated

10.

Graphical user interface, text, application, email

Description automatically generated

11.

Graphical user interface, text, application

Description automatically generated

12.

Graphical user interface, text, application

Description automatically generated