

Application of Geographical Gazetteer Standards to Named Time Periods

Prepared by Melanie Feinberg in consultation with Ruth Mostern, Susan Stone, and Michael Buckland

Introduction

The categorization of information by time period is a ubiquitous organizational device, especially for historical data. Yet instead of referring to specific years, we often use period terms to suggest calendar dates. Time period terms distinguish *neolithic* ruins, *Elizabethan* drama, and the *Napoleonic* wars. Indeed, these terms often carry a stronger meaning than calendar dates, because they implicate a subject, time, and place together. For example, the French term *le grand siècle* encapsulates a place (France), time (seventeenth century), and subject (a flowering of arts and culture, and the height of absolute rule). In contrast, the term *seventeenth century* is markedly less informative.

Yet if time is a significant organizational attribute, and time period terms are often used to convey date information, the terms can be ambiguous and imprecise. For example, neolithic ruins in China and in South America are from different periods. The ideas of the Italian Renaissance moved through Europe at different times, so that what we might call the Renaissance in France and the Renaissance in the Netherlands are not contemporaneous. Even if I refer strictly to the Italian Renaissance, limiting my usage to a period in music or a period in fine arts, instead of to a broader cultural movement, can also change the dates associated with the period name. Furthermore, the term *renaissance* can indicate a period in the twelfth century in Europe, the 1920s Harlem renaissance in the United States, and many another period. Time period terms also vary culturally: the term *Great Patriotic War* is used in Russia to refer to Soviet involvement in World War II, but this period is just part of World War II from the American perspective.

Similar problems occur in the realm of geography, where there is a Paris in Texas as well as in France, where Erie is both a canal and a lake, where the location formerly known as British Honduras is now Belize, and so on. In geography, a device known as a *gazetteer* disambiguates place names by linking them to the exactness of longitude and latitude. By matching a name with its coordinates, I can clarify that I mean the city of Long Beach in California and not the Long Beach peninsula in Washington, for example.

Recent work by the Alexandria Digital Library (ADL) and the Electronic Cultural Atlas Initiative (ECAI) describes specifications for digital gazetteer content.¹ In a digital context, gazetteer information can facilitate searching, sorting, and display of data sets, in addition to disambiguating place names. For example, if a gazetteer includes information about geographical feature type, such as canal, lake, or town, one can search a database for all the rivers

¹ Alexandria Digital Library Project, gazetteer content standard, version 3. <http://www.alexandria.ucsb.edu/gazetteer/ContentStandard/version3/GCS3-guide.htm> (last accessed September 2003).
Electronic Cultural Atlas gazetteer project. <http://ecai.org/projects/gazetteer/> (last accessed September 2003).

in Switzerland, in addition to all the locations named Springfield or all the named places within a set of geographic coordinates.

Gazetteer information can also be used to display data from multiple sources on a map. One can create subsets of data by using a time bar to restrict display to certain periods (to only show, for example, the data from 500 to 1000 CE). Using this structure, one can plot, for example, the adoption of different musical instruments along the Silk Road over time.² By making the spatial and temporal dimensions of historical information more apparent through grouping and visualization, one can emphasize connections or discontinuities between data sets. Such mappings supplement the traditional monograph and provide an alternate means of presenting historical information to multiple audiences.³ For example, the visualizations made possible by linking a gazetteer with a historical data set enable one to answer the following questions more easily:

- What happened in a particular geographic location (at a particular time)?
- How does data about a location (such as population) shift over time?
- How does data about a location compare to data about other locations? (For example, what locations exceeded the population of the first location?)

Project Objectives

This project attempts to see if gazetteer standards for place names can be adapted for time period descriptions. Just as a gazetteer matches place names to coordinates, a time period directory could match time period terms to date ranges, location, and other information that characterizes the period.

Project Motivation

A time period directory might enable data searches such as:

- When was the Uruk period of ancient Sumer?
- When the Italian Renaissance was underway in Florence, what cultural movements were prevalent in China?
- What wars were being fought in the year 800?

To enable searches of the type described, a time period directory would need to incorporate not only time period names (such as Uruk period and Italian Renaissance), locations (such as Sumer and China), and dates (year 800), but also period types, or categories that describe certain kinds of time period (such as cultural movements and wars).

² See Electronic Cultural Atlas Initiative Silk Road Cultural Atlas, <http://ecai.org/silkroad/> (last accessed September 2003)

³ Mostern, Ruth, "Digital Gazetteer Standards for History and Culture," 2002. Electronic Cultural Atlas gazetteer project. <http://ecai.org/projects/gazetteer/> (last accessed September 2003).

The ADL content standard for digital gazetteers incorporates similar data elements. The directory would also need to model temporal uncertainty; the dates for the Uruk period, for example, are by no means clear.

The information produced by a time period directory could be used as input for library catalog searches. If, for example, a time period directory search produced a list of wars going on in the year 800, the names and places associated with those time periods could be used to find appropriate texts. The results of the time period directory search (period names and places) could also be displayed on a map. Alternatively, a search for data related to periods of rule in France from 1500 to 1600 could produce an instant timeline.

Although these examples refer to historical data in the traditional sense, one could see specific time period directories being developed for information referring to government documents, or for historical data within a company or other organization. One could search for research and development projects within the first quarter of 2001, for example, even if documents were keyed by code name, instead of dates, or look for NASA projects concurrent with the Apollo program.

Initial Goals

The primary goal of this project was to describe, in a general way, what a time period directory would include, and how directory entries would be structured. Borrowing from XML and database models, we call this description a *content schema*, although our description is looser than a formal data model. Our purpose is to provide a first blueprint for time period directory implementations and a basis for further discussion on this subject. An implementation could use this schema to create, for example, a set of database tables and populate them with time period data.

In addition to the content schema, we created an initial period type list, to segment the broad universe of time period terms into a framework of categories.

Project Scope and Methodology

We began the project by collecting an extensive list of time periods culled from a variety of reference sources.⁴ Using the ADL version 3 content standard⁵ as a starting point, we designed a metadata structure that would accommodate the range of time period data that we had gathered.

⁴ We extracted time period terms used in library records from the University of California MELVYL catalog, isolating the terms used in subfield Y of the 650 area for subject headings. We also scanned a wide variety of Web-based sources, including these examples:

- Wikipedia timelines, particularly those on science and technology (http://www.wikipedia.org/wiki/List_of_timelines)
- Asian history timelines from Asia Source (<http://www.asiasource.org/features/timelines.cfm>)
- British history timelines from the BBC (<http://www.bbc.co.uk/history/timelines/>)
- Art history timelines from the Metropolitan Museum of Art (<http://www.metmuseum.org/>)

In developing the schema, we included only those elements that served the goals of disambiguation, search facilitation, and visualization. Therefore, we decided against including elements, for example, to contain specific data about a time period (such as the number of casualties from a war). One could use a directory of the sort that we propose to acquire search terms and find documents that include additional information, but we do not see the directory as a vehicle to hold supplementary information itself.

During the development process, we made the following decisions regarding the directory's scope:

- The schema does not incorporate a definition of *named time period* (put another way, the schema does not provide criteria to determine acceptable entries).
- The schema does not distinguish between eras (ranges) and events (points).
- The schema does not allow entries for relative time designations (such as *summer*) or other time modifiers (such as *early*), unless these are combined with an absolute time designation and given precise dates.

However, although the content schema is flexible, individual implementers of time period directories may find it useful to limit entries based on their own criteria. For this reason, we discuss various points that implementers might wish to consider.

Limiting Entries to Named Time Periods

It appears straightforward that a time period directory should include commonly accepted terms, such as the Warring States period of Chinese history. However, the situation grows more complex with periods that do not have conventionally agreed-upon names, and yet have historical significance. For example, the domestication of the camel took place over a lengthy period. However, while this era can be easily described, it does not have a conventional name. In other words, this era has no predefined vocabulary control.

Certainly, the domestication of the camel constitutes a significant historical epoch, even if it is not consistently referred to by a specific name. Timelines, the most common visualization device for chronological data, often include periods expressed by descriptive phrases, and not just named terms. Sometimes the timeline itself is set up

These are just a few of the sources used. Our goal in this process was breadth (the range of terms that might be expressed, and the means of their expression), not scholarly accuracy.

⁵ Alexandria Digital Library Project, gazetteer content standard, version 3.

<http://www.alexandria.ucsb.edu/gazetteer/ContentStandard/version3/GCS3-guide.htm> (last accessed September 2003).

as a sequence of descriptive phrases.⁶ Other timelines include descriptively defined periods along with named eras, to avoid gaps in the narrative being expressed.⁷ Leaving out such periods results in an incomplete directory.

However, a conventionalized name implies an agreement that, for a certain group, the period can be concretely defined in a manner acceptable to most members of the discourse community. (Similarly, dictionaries include only words and phrases whose definitions have become solidified through consistent usage.) If one includes descriptively defined periods in a period directory, it is difficult to enforce vocabulary control. There are many ways to express “domestication of the camel”; there are relatively few to express the Khmer Rouge regime. Implementers of time period directories will need to take this into consideration.

This issue does not seem as prevalent in geographic gazetteer development. If we don’t give proper names to every location, we do have systems of addressing (such as 200 Green Street) that can substitute as names (as opposed to relative, uncontrolled descriptions like “my house” or “the purple building near the park”). One reason for this is that geographic divisions often need to be expressed in government records, for purposes of rule, representation, taxation, and so forth, which require control and consistency of naming. In contrast, we use calendar dates to record time for administrative purposes, not time period terms.

Making Distinctions Between Eras and Events

The tendency to use a subject approach to describe time periods, specifically that of describing an era based on surrounding events, highlights another issue: the definition of *period*. A timeline, for example, can present a subject as a sequence of events, such as the signing of the Magna Carta, in addition to eras of significant duration, such as the rule of the House of Bourbon. When we think of something as an event, however, the duration seems unimportant, whereas for an era, the sense of duration is part of its meaning. For example, an era such as the Egyptian Middle Kingdom is partly defined by its duration: it represents the time in which certain dynasties ruled. In contrast, while the assassination of John F. Kennedy has a duration, primary significance attaches to what happened, and not to the time it took. By this definition, events seem less appropriate as time period directory entries.

However, we sometimes use events such as the Kennedy assassination as markers around which we define eras, so that the event becomes noteworthy for its ability to call up a certain time. Mention of Kennedy’s death often evokes a greater context, or, by extension, creates an era around the event. Also, in certain contexts, the duration of events

⁶ For example, the timelines in Wikipedia, the “free encyclopedia,” are set up this way. Wikipedia includes many scientific and technical timelines, such as the timeline of solar astronomy and the timeline of antibiotics, for which named eras are seldom defined. (See Wikipedia, http://www.wikipedia.org/wiki/List_of_themed_timelines, last accessed September 2003.)

⁷ An example is the Asian history timeline found at <http://www.wwnorton.com/college/history/worldciv/referenc/asiatime.htm>, which includes named eras such as the Akkadian Empire along with descriptive phrases such as “introduction of the horse in W. Asia,” and “development of the alphabet.” (Timeline created by Thomas Percy and Mary Dickson to supplement the textbook *World Civilizations*, by Phillip Ralph, Robert E. Lerner; Standish Meacham, Alan T. Wood, Richard W. Hull, and Edward McNall Burns.)

does matter. The Kennedy assassination is again a good example. Historians, journalists, conspiracy theorists, and others have constructed detailed narratives and second-by-second timelines that transform that particular event into a kind of micro-era.

Furthermore, it can be difficult to determine whether a particular happening is an event or an era. It's pretty clear that, for example, the Time of Troubles in Russian history is not an event, because no single happening distinguishes it; the period is a time of unsettled rule, but it can't be characterized as a clear conflict between identified combatants or anything so neatly delineated. Yet often categorizations are not so easy. For example, is the Great Depression an event or an era? Although one may hear recollections of "growing up in the Depression," for example, usage that describes an era, a depression is an economic event, of the same ilk as recessions, panics, and crashes. How would we categorize the South Sea Bubble or the Panic of 1837?

As with the Kennedy example, the answer seems to depend on the context in which the time period term is used. In constructing a detailed history of the South Sea Bubble, one might see it as a period, while to someone creating a general timeline of the eighteenth century, it is an event.

For these reasons, it does not seem useful to formulate universal criteria for defining time periods based on duration. However, such criteria might be possible for a particular gazetteer instance, if the context remains stable. The implementer will need to consider whether events should be included in a directory, and, if so, whether to devise a set of criteria to differentiate between eras and events.

Using Relative Time Designations and Other Modifiers

Relative time designations use descriptions of cyclic events in relation to an absolute time (typically whenever the reference is made). Examples include *next Tuesday*, *baseball season*, *winter*, *the full moon*, *Ramadan*, and so on. Our proposed content schema only supports relative time designations that modify absolute dates. For example, *baseball season 2003* has relatively precise beginning and ending dates, and can be specified in the content schema, while *baseball season* alone cannot.

Cyclic time periods are but one type of relative time designation. Modifiers such as *pre*, *post*, *classic*, *neo*, *early*, *late*, and so on are often used with time periods: examples include *pre-war Manhattan apartments*, *early Baroque concertos*, and *post-September 11 world*. Sometimes the modifier is used to create an era using a precise date as an anchor for the modifier; *post-September 11* falls into that camp. While the beginning date, in this case, is defined, the end date is not specified, and is presently ongoing. Other times, the modifier carves out a fuzzy subset within a defined era; *early Baroque* is in that category.

As with cyclic time designations, the content schema only supports these as defined periods in themselves. For example, one could define a period named *early Baroque* and assign dates to it, but the schema does not support a standalone definition of *early* that would enable one to add it to Baroque or other time periods.

Content Schema Description

The content schema includes these primary elements:

- Entry ID.
- Name.
- Descriptive notes.
- Time scale.
- Dates.
- Period type.
- Location.
- Relationship.
- Administrative details.
- Directory-wide administrative details.

The description is at a general, design level, not an implementation level; the schema could be implemented in XML or as a database, but that level of detail is not addressed here. Instead, we focus only on what elements should be included in the schema, which elements are optional and which required, and which elements are repeatable.

A table-format summary of the content schema, with example entries, appears the end of this document.

Entry ID

Each entry requires a unique ID. Implementations can determine the appropriate format.

Name

Each entry in a time period directory must have at least one name, and can have any number of alternate names. For each name, the following elements are required:

- Name.
- Language and script (such as the Roman alphabet, Cyrillic alphabet, or Devanagari) of the name.
Ideally, the language would be represented using a standard scheme, such as the ISO 639 standard, which defines two-letter or three-letter codes to represent languages. If necessary, specification of a transliteration scheme would also be required.

Optionally, the following elements could also be included:

- Source data.

While documentation is important for scholarly work, not all applications require source information. In addition to standard citations, this element could also contain URLs to source data.

- Notes.

Notes might include information about when and where the name is used; for example, an entry on the Dark Ages might note that referring to the medieval period as “dark” was not considered accurate at the time of entry. Similarly, the term *Great Patriotic War* reflects a particular (Soviet) perspective and function; a more prosaic term, such as *German-Soviet War*, reflects a different understanding of the time period and its significance. Notes could also indicate etymology or pronunciation information as appropriate.

The proposed name information for time periods omits several elements included in the ADL content schema for geographical gazetteers, including items for abbreviated forms of the name, an authority that sanctions use of the name, a confidence note and time period for the name, and separate items for etymology and pronunciation. Because time period names are more generically descriptive than place names, official and abbreviated forms are not typically applicable. While occasionally the time period when a particular name is used might be of some interest, that secondary period would not be germane to the rest of the entry, as it would with geographical names. For example, the period when St. Petersburg was officially named Leningrad is a subset of the period of the existence of a particular city, variously known as St. Petersburg and Leningrad. However, the period of time when the medieval period was referred to as the Dark Ages (say, the eighteenth and nineteenth centuries) has no relation to the duration of the time period represented by that name (say the sixth through fifteenth centuries).

Similarly, etymology and pronunciation are often interesting or unusual for geographical names, but much less so for time period names. In the relatively rare case that a time period name is especially interesting, this information can be conveyed through the note element in the name area, as opposed to separate elements.

In an implementation of the content schema, one would need to specify the default name, if multiple names are specified; for example, to determine which name to display in a list of search results. However, criteria to determine the display name might differ according to different uses of the directory; for example, the same directory might be used in multiple locations, but the display name would need to be in Hindi for one use, and English for another use. Because of this, criteria to determine the display name are not specified in the content schema itself.

Descriptive Notes

This optional item applies to all names and is envisioned as a brief description of the time period. This element should clarify any basic information not apparent from the other content elements. For example, in an entry for the

American Revolution, other elements in the content schema would indicate the name, the type of time period (a period of conflict, for example), and the location where the time period occurred (the American colonies). However, none of these elements would clarify that the combatants were the British and American colonists, for example. The Descriptive Notes element could do so.

As another example, because the terms used to convey time period data may also imply other types of information, names by themselves can be difficult to interpret precisely. For example, the term *Ming* indicates a Chinese dynasty that ran from 1368 to 1644. But when applied to porcelain, *Ming* sometimes connotes expensive, fine wares. Yet while porcelain from imperial workshops in the Ming period is indeed sought after, the wares of the common folk are not as rare or valuable. This aspect of meaning is not likely to be expressed through the period type element (discussed later in this section). Instead, one could use the Descriptive Notes element to clarify that one was using *Ming* to represent the time period when imperial Ming was produced, or that one was using *Ming* to represent any wares produced during that particular dynasty.

Time Scale

Required elements include:

- Calendar system.
An entry must specify a calendar system, such as the Gregorian calendar, Hebrew calendar, Balinese calendar, geologic time scale (which runs in MYA, or millions of years ago, and thus defines the year zero as the current date), and so on. A calendar system defines units of time, such as years, months, and days, and provides a starting point for counting its base units (typically years).
Optionally, an entry can provide notes on the calendar system or source data for it.
- Date format.
An entry must specify the date format, such as the ISO 8601 standard. A date format specifies how the units defined by the calendar system (years, months, days, and so on) are represented, such as 06-08-1970 for August 6, 1970.

Dates

Some time period dates are known with extreme precision; others are mere guesses, with multiple plausible alternatives. All entries need at least one begin date and one end date, but can include multiple dates of each type. Any begin date and any end date provide a possible range. Date elements include these items:

- Date information, expressed according to the calendar system and format previously specified.
- Date type.
Date types define possible ranges. They include:

- Earliest.
- Latest.
- Most likely.
- Advocated by a particular source.
- Ongoing.
- Date type scheme.

The date types listed in the preceding paragraph are a basic set of types. By specifying a type scheme, implementers can develop their own sets of date types.
- Source data.
- Notes.

As an example date scenario, in an entry defining the lifespan of a person, there might be two possible birthdates, but a confirmed date of death. This entry would have three date elements: two begin dates and one end date. One begin date might be identified as the earliest, and the other begin date latest.

If a time period directory includes events, defined here as time markers where the duration does not matter for the purpose of the entry (for example, an entry for the death of Mansa Musa, ancient ruler of Mali), we use the same date for begin and end dates. This way, events and eras can both be included according to the same basic date arrangement.

In cases where only a single date is known with certainty, but the intent is to extrapolate a range from that date, we use the known date as a basis for creating possible begin and end dates, and record the known date in the notes element. For example, if a Chinese scholar produced some documents in the year 536, but no other dates are known about this scholar, a historical source might express the scholar's lifespan as "flourished 536," or "circa 536," and so on. To keep the date structure consistent, the schema only provides for begin and end dates, and not for dates inside a range. Therefore, an entry must be created out of a likely guess based on the known information. This is essentially what the "flourished" note indicates in an ad-hoc way, but it must be spelled out in the directory entry.

The schema does not provide for unknown dates. If a begin date is known, but no end date, a likely guess is required, and the circumstances explained in the notes. This is also similar to the "flourished" case. However, instead of requiring a reader to extrapolate that a Florentine merchant born in 1622 was likely dead by 1722, the directory entry makes this explicit, for the sake of internal consistency. In cases of this nature, one would most likely use the Latest Date type for the guess.

Continuing periods at the time of entry form a somewhat special case. To keep the schema consistent, an end date is required, but guessing a future date would be nonsensical. For these situations, we have recommended a date type of Ongoing. For example, an entry created today for the administration of George W. Bush as president of the United

States would be ongoing. In an implementation, a date with a type of Ongoing could be configured so that an actual date cannot be placed in the Date element; the Date element would also read Ongoing.

Period Type

The period type describes the category of the entry, such as a period of conflict, a cultural movement, or the lifespan of an individual. Period types are useful both for disambiguating periods with similar names and for searching for similar periods. In the ADL geographical gazetteer schema, the feature type is analogous to our period type. ADL feature types include such geographical features as bridges, administrative districts, and mountains.

The required period type element includes:

- Name of the period type.
- Classification scheme that defines the period type.

Following the description of the content schema in this document, we define a short, generic classification scheme for period types, to be extended by the creators of individual time period directories. The period type list that we define is a broad set of categories, meant to serve as an extensible framework through which specific thesauri, created by specialists in particular areas, might be combined. Most implementations will not use our classification scheme; instead, they will extend categories defined in our scheme as necessary. We selected this approach because we recognized that it would be impossible for any single organization to create a master thesaurus of time period types. Instead of attempting to anticipate the needs of the many specialists who might make use of a time period directory, our approach creates a general framework to simplify the union of specialist vocabularies. For example, one of the general categories in our proposed period type list is *period of conflict*. A military historian might extend this category with a more appropriate, specific thesaurus, including such terms as *war*, *battle*, *rebellion*, *mutiny*, and so forth. An intellectual historian might do the same with such general categories as *cultural movement* and *intellectual movement*. In developing this framework, we relied on the previous work of ECAI, which used this reasoning to create a feature type list for geographical gazetteers.⁸

With some entries, multiple period types might be necessary. For example, the civil rights movement in the United States had both social and political aspects. Instead of forcing a choice between social movement and political movement period types, or including two identical entries to describe the same period, one could include both types. As another example, the Italian Renaissance could be described as both an intellectual movement and a cultural movement, if the time period were being defined in a general way. This example is interesting, because a more specific time period directory might make distinctions between the Renaissance as a fine arts movement, the

⁸ Electronic Cultural Atlas gazetteer project. <http://ecai.org/projects/gazetteer/> (last accessed September 2003).

Renaissance as an intellectual movement based on rediscovery of classical texts, and so on, but a more general directory might not.

This multiple assignment applies only when no single period type adequately describes a single time period, and not when two periods have the same name and duration. For example, the Regency period in England refers to both a period of rule and to a period within architecture. The description of each Regency period, as well as the period type, is different. Here, two period types apply to two different time periods. In the case of the civil rights movement, however, the description is the same; two period types apply to one time period. For the Renaissance example, multiple feature types would only be appropriate if a directory took a broad view of the term. If one wanted to associate different date ranges and locations with the different feature types, one would create multiple entries for Renaissance, instead of a single entry.

As a geographic analogy, the single location represented by the name *San Francisco* is both a city and a county. The geographic boundaries are the same for the city and county, city and county government is intertwined, and so forth. One location has two geographical feature types, and a single gazetteer entry is appropriate. In contrast, while Los Angeles is the name of both a city and a county, the county encompasses a larger area than the city of Los Angeles, and governmental functions are separate. In this case, Los Angeles is the name of one entity with the feature type of city and another entity with the feature type of county. Two entries are required.

Location

The location that applies to a time period is often necessary for disambiguation. Even with date information, common terms such as *civil war* can be clarified only when the location is specified. Location information also enables time period data to be plotted on a map.

A time period directory could be used in conjunction with a geographical gazetteer, and so location information could be incorporated by reference. To account for this possibility, the schema defines two methods for including location data. If a geographical gazetteer is available, the following elements are required:

- Gazetteer entry ID.
- Citation information.

For an online gazetteer, this could include gazetteer name and version, or latest date of gazetteer modification. For a printed gazetteer, this would be a standard bibliographical citation.

- Persistent URL to gazetteer, or other location data.

Multiple gazetteer references can be used to incorporate multiple locations, if necessary. As a matter of good practice, if referencing an online gazetteer, one might want to ensure that a local copy is available for users of the time period directory.

However, a geographical gazetteer might not always be available, or the location required by a time period might not always be adequately represented by a gazetteer entry. For example, in an entry for the Battle of Gettysburg, the actual location was on the outskirts of the town, not in the town itself. The area where the battle was fought would not necessarily be included in a gazetteer. For such situations, this schema recommends the inclusion of the minimum elements required to define a particular location. These are:

- Location name.
Multiple names can be specified if necessary. The language and, if necessary, transliteration scheme must also be specified.
- Spatial data.
Spatial data can be represented in various ways, including a point, bounding box, or shape file, that defines the maximum extent of the location.
- Source data.
- Notes.

Determining the appropriate location information to include is not always a straightforward task. For example, the Americans participated in World War I, but the war did not occur in the United States. Should the United States be included in location information that represents where World War I was fought? Implementers should rely on a consensus of historical sources or clearly articulated criteria to make such decisions. As another point to consider, location information may change over the course of a time period. For example, during the Neo-Babylonian empire, the extent of lands held by the Babylonians shifted often, typically with each ruler. There are various ways to handle this. One could include location information for the largest extent of the empire or for an average extent, or include multiple location elements to indicate different locations during the period. Ideally, an individual directory would use a consistent approach, and the location elements would be documented in the note areas.

Location data is assumed to represent the location during the time period being described, not during the period of entry. An entry for the American Civil War would define the location as the United States from 1861 to 1865, and not as the current definition of the United States.

Relationship

A relationship element seems necessary to distinguish periods that are contained by other periods. For example, one would want the battle of the Somme to be recognized as being part of World War I. But not all battles that took place during this period were part of the war. Major internal conflicts, including revolution in Russia and unrest in Ireland, were concurrent with the war, but not ordinarily considered part of it. A `partOf` relationship between the battle of the Somme and World War I resolves this possible ambiguity.

One could also resolve this kind of ambiguity by using more specific period types. A period type scheme that differentiated between battles related to internal unrest and battles between a group and outsiders would avoid the possible problem noted in the World War I example. However, the use of specific period types can also be problematic. A period type distinction between types of battles, for example, would require both data enterers and searchers to likewise make this distinction.

In addition to the partOf (and corollary Contains) relationship, other relationship types might be useful for specific purposes. For example, the Reconstruction period that followed the American Civil War was not part of the war, and yet it emerged directly out of the war. Undoubtedly these periods are linked. Yet it's difficult to define just how: one might say that in one sense the Civil War caused Reconstruction, or that Reconstruction evolved from the Civil War, but it's difficult to define the relationship precisely. As another example, one might want to show how the literary style of sentimentalism prefigured romanticism. Yet the precise nature of this relationship could be a matter for much scholarly debate.

While the introduction of such relationships between time periods might enrich historical data and enable useful visualizations of related information, only the partOf/Contains relationship appears to be required for disambiguation. The difficulty of defining other relationship types precisely, and the multitude of possible relationships that might exist between periods, make additional types problematic. The goal of the time period directory is not to encapsulate all that is known about a time period, but to identify, define, and disambiguate time periods efficiently.

For this reason, this schema includes only the partOf/Contains relationship type. However, creators of individual time period directories may choose to create their own relationship type lists and include additional relationships. One possibility includes a Series type of relationship that links recurring events, such as the Olympic Games. Another possible relationship would be an Eponymous type, that links eras that bear a person's name (such as the Edwardian period in England) to the lifespan of that person. Still another type might link a person's lifespan to the artistic movement with which he or she is identified (for example, Duchamp to Dada).

Administrative Details

Information about the circumstances of the entry is necessary to evaluate older entries, track changes, and provide accountability. Required elements include:

- Creator contact information.
Includes name, organization, and contact information for the enterer.
- Date of entry.
- Modification history.
- Notes.

Directory-wide Administrative Details

A brief set of information is required to identify the directory as a whole.

- Directory name and version number.
- Contact information.

The name and contact information for the person or organization responsible for the directory as a whole, as opposed to the creator of an individual entry.

- Encoding scheme.

The encoding scheme, most likely Unicode, specifies how letters are represented in a computer. The encoding scheme would not be unique per entry.

Future Work

This paper describes the metadata elements that we believe would make an effective time period directory. The next step is to create a formal data model, for example as a set of relational database tables, and implement a prototype.

We are currently pursuing this goal. We also hope that others will provide comments or alternative schemas, and attempt their own implementations.

Appendix A: Period Type List

As described earlier in this document, a period type assigns a category to a time period, to aid in disambiguation and to assist searching of similar kinds of time periods.

The draft period type list that we have created is not meant to define a thesaurus of all possible time period categories. To produce such a thesaurus would be a major undertaking, and even so it would be unlikely to suit all needs. It is impossible to predict the variety of time period categories that might be needed by specialists in a particular area, from art history to evolutionary biology. Instead, our approach aims to provide a broad range of categories, to produce a general-purpose, high-level framework, from which a specialist might expand subcategories to any level of detail, or declare and use a separate specialized thesaurus of named time periods.

Period Type List Structure

Because the set of categories described by the period type list is so general, the list differs from a traditional thesaurus in that it defines containers for similar terms, as opposed to defining preferred terms from equal synonyms.

The period type list was created using a hierarchical approach. There are six top-level categories, with subcategories within. The primary categories are:

- Groups of people.
- Trends in thought and expression.
- Natural events.
- Acts of creation or discovery.
- Biography.
- Cycles and intervals.

A brief set of subcategories is defined for each category. The period type list is presented in full at the end of this section, following a summary of the feature areas.

The period type list does not currently include an area for mythic time. While most time periods relating to mythical or supernatural events would fit into one of the previous categories, there are no obvious slots for communication between human beings and deities, for apocalyptic events relating to previous instantiations of the earth, and so on. While time period names relating to mythic events may be of scholarly or religious interest, it is unclear whether mythic time periods have the same requirements as historical time periods, and further exploration of this topic was deferred.

Groups of People

This period type area includes time periods relating to the history of groups, including peoples; nations; smaller governmental units, such as provinces, counties, and municipalities; schools; churches; and companies. In initial versions of the period type list, we included different categories for nations, for peoples, and for other types of organizations, but the categories within these areas were clearly similar, and it seemed redundant to keep these areas separate.

Trends in Thought or Expression

This area includes eras in the history of ideas, beyond the existence of any particular organized group that might have promulgated a particular movement. For example, while the women's suffrage movement in the United States might have involved a number of organized groups, the movement itself transcended the groups involved within it.

Natural Events

This category includes time periods defined by geologic change, climate events, evolution of species, planetary movements, and anything else not defined by human history.

Acts of Creation or Discovery

This set of types includes periods defined by acts of human endeavor, some of which might be part of a trend in thought or expression. For example, the painting of the Mona Lisa was an act of creation that was part of the cultural movement of the Italian Renaissance.

Biography

This period type area includes events relating to a single person's life. It can be considered a special case of the Groups of People category.

Cycles and Intervals

The final top-level category defines categories relating to cyclic events, such as weather patterns and calendar divisions. As explained earlier in this document, in the content schema discussion, cyclic descriptions describe relative time, and not absolute time. Still, because they are so common, it seemed useful to provide a set of categories for them as well.

Period Type List Development Process

To create the period type list, we first scoured a variety of sources to get a sense of the range of time period terms used in general discourse.⁹ In addition, we reviewed the Library of Congress rules for creating subject headings, to see if we could extrapolate rules for how they create subject headings using time periods. There did not appear to be a fruitful connection there.

At the end of this first phase, we had an extensive list of time period names, supplemented by lists of synonymous terms that encapsulated some of specific examples (such as *war*, *battle*, *uprising*, and so forth). We then began successive rounds of creating categories to fit the collected examples. The first attempt at categorization included 7 general categories, 63 subcategories, and many uncategorized examples. Through a series of subsequent drafts, the list was slimmed down to 6 top-level categories (quite different from the initial primary categories), and 34 subcategories. In each revision, the overall goal was to reduce the number of categories, to achieve the minimum number required so that any specific time period term could be placed in a category.

⁹ See note 4 for a partial list of sources.

Period Type List

Category	Subcategory	Examples	Includes These Concepts ¹⁰
Groups of people	Existence	<ul style="list-style-type: none"> • La Tene culture • Standard Oil in existence • British East India Tea Company • Jesuit Order • United Nations • Progressive political party 	(Existence of) People, culture, civilization, state, kingdom, province, municipality, organization, company, church, school
Groups of people	Stages of development	<ul style="list-style-type: none"> • Pax Romana • Early Minoan period • New Kingdom of Egypt • Golden age of Gupta empire • AT&T broken up into regional phone companies 	Height, decline, fall
Groups of people	Population or territory changes	<ul style="list-style-type: none"> • Settlement of Slovene territory by Celtic Tene culture • Black Death in Europe • Louisiana Purchase • Reunion of Tibet • Irish potato famine 	Immigration, migration, settlement, colonization, annexation, assimilation, expulsion
Groups of people	Period of control	<ul style="list-style-type: none"> • Shang dynasty in China • Commune in France • Communist system in USSR • Rule of Mansa Musa in Mali • Hyksos control of Egypt • Warring States period in China • Earl Warren as Chief Justice of the United States 	Reign, rule, dynasty, term (of office), occupation
Groups of people	Period of conflict	<ul style="list-style-type: none"> • Time of Troubles in Russia • Thirty Years War in Europe • Warsaw Ghetto Uprising • Haymarket riots in Chicago • Intifada • Great Schism 	War, battle, campaign, uprising, insurrection, rebellion, revolt, massacre, persecution
Groups of people	Periods of (non-aggressive) relations between groups	<ul style="list-style-type: none"> • NATO alliance • Era of Good Feelings • Common Market 	Alliance, partnership, coalition, detente
Groups of people	Decisions and policies	<ul style="list-style-type: none"> • Spanish Inquisition • Prohibition • New Deal 	Law, act, policy, edict, decree
Groups of people	Economic fluctuations	<ul style="list-style-type: none"> • Great Depression • South Sea Bubble 	Depression, panic, recession, bubble, prosperity

¹⁰ These descriptions and term lists are not meant to be all-inclusive, but merely to account for the range of topics that might be placed under a particular subcategory.

Category	Subcategory	Examples	Includes These Concepts ¹⁰
Groups of people	Activities and products	<ul style="list-style-type: none"> • Model T produced • NASA Gemini program • Human sacrifices in Aztec culture 	Production, program
Trends in thought and expression	Political movements	<ul style="list-style-type: none"> • Women's suffrage movement in United States • Federalism in United States 	Periods defined by thinking regarding government and interactions between states
Trends in thought and expression	Economic movements	Decline of guild system in Europe	Periods defined by means of production and exchange
Trends in thought and expression	Intellectual movements	Enlightenment	Periods defined by ways of thinking
Trends in thought and expression	Social movements	<ul style="list-style-type: none"> • Civil rights movement in United States • Women's liberation movement 	Periods defined by thinking regarding groups of people and their placement within societies
Trends in thought and expression	Cultural movements	<ul style="list-style-type: none"> • Classical era in music • Sentimentalism in literature • Renaissance in Italy 	Fine arts, music, literature, film, architecture, fads and fashion
Trends in thought and expression	Evolution of languages	<ul style="list-style-type: none"> • Old English • Middle High German • Great Vowel Shift • Cuneiform becomes common language of West Asia 	Periods defined by use of language
Trends in thought and expression	Religious movements	<ul style="list-style-type: none"> • Reformation • Great Awakening (revival movement in colonial America) 	Periods defined by types of religious thinking (beyond internal workings of specific churches)
Trends in thought and expression	Scientific and technical movements	Industrial Revolution	Periods defined by technical advancement or scientific theory
Trends in thought and expression	Spirit of the times (zeitgeist)	<ul style="list-style-type: none"> • Gilded Age • Roaring Twenties 	Periods that reflect perceived overall sentiment or attitude
Natural events	Geological	Mesozoic Era	Volcanic eruptions, earthquakes, plate shift
Natural events	Climate	Pleistocene Ice Age	Ice ages, storms
Natural events	Astronomy	Big Bang	Meteor shower, eclipse, comet
Natural events	Biological	<ul style="list-style-type: none"> • Black Death • Cambrian Explosion • Cro-Magnon man • Plants evolve seeds 	Epidemic, outbreak, evolution of species, evolution of traits
Acts of creation or discovery	Exploration and journeys	<ul style="list-style-type: none"> • Phoenicians circumnavigate Africa • Moon landing 	Events regarding discovery of previously unknown lands or peoples
Acts of creation or discovery	Discoveries, inventions, and investigations	<ul style="list-style-type: none"> • Camels domesticated • First Roman coins • Tuskegee syphilis experiment 	Studies, experiments, inventions, discoveries
Acts of creation or discovery	Creation of single works and dissemination of ideas	<ul style="list-style-type: none"> • Enuma Elish schematized on cuneiform tablets • Mona Lisa painted • Buffy the Vampire Slayer on television 	Publications, works of art

Category	Subcategory	Examples	Includes These Concepts ¹⁰
Acts of creation or discovery	Performances, demonstrations, and public speech	<ul style="list-style-type: none"> Fisk University Singers Jubilee Tour FDR fireside chats on radio March on Washington Urinetown runs on Broadway 	Lectures, concerts, broadcasts, exhibitions, fairs, protests, marches, talks, seminars, readings
Biography	Lifespan	Marcus Aurelius, 121-180	Period defined by the time a person lived
Biography	Stages of development	<ul style="list-style-type: none"> Beethoven becomes deaf Picasso's blue period 	External characterizations of periods in a subject's life; periods defined by a subject's passive participation
Biography	Activities and relationships	<ul style="list-style-type: none"> Henry VIII marriage to Anne Boleyn Italian Jesuit priest Desideri studies and teaches in Lhasa Einstein employed in patent office 	Periods defined by the subject's active participation
Cycles and intervals	Weather patterns	<ul style="list-style-type: none"> El Nino Monsoon season Winter 	Seasons
Cycles and intervals	Planetary cycles	<ul style="list-style-type: none"> Full moon Halley's comet Northern Lights 	Lunar cycle, comet cycles, planetary orbits, rotation of galaxies
Cycles and intervals	Calendar divisions	<ul style="list-style-type: none"> January Tuesday 	Months, weeks, years, calendars
Cycles and intervals	Human-organized periods and events	<ul style="list-style-type: none"> Baseball season Academy Awards Ramadan My birthday Tax time 	Holidays
Cycles and intervals	Life and chemical cycles	<ul style="list-style-type: none"> Mouse gestation period Half-life of carbon-14 	Standard cycles that occur in the lifespan of a species or substance

Appendix B: Content Schema Summary and Examples

The following table summarizes the content schema and provides simple examples. The examples are not structured as completely as they would be in an actual implementation.

Content Schema Area	Required Elements	Optional Elements	Example 1	Example 2
ID	<ul style="list-style-type: none"> Entry ID 	None	XXXX	XXXY
Name	<ul style="list-style-type: none"> Name Language (If applicable) Transliteration scheme Multiple names allowed.	<ul style="list-style-type: none"> Source data Notes 	Spanish Inquisition English	Fourteenth Dynasty English
Descriptive Notes	None	Descriptive notes	Court to find and punish heresies and insincere conversions to the Catholic Church. Invoked by the Spanish government, not the church.	Obscure Egyptian dynasty that ruled from the western delta.
Time Scale	<ul style="list-style-type: none"> Calendar system Date format 	<ul style="list-style-type: none"> Notes 	Gregorian calendar. ISO format.	Gregorian calendar. ISO format.

Content Schema Area	Required Elements	Optional Elements	Example 1	Example 2
Dates	<ul style="list-style-type: none"> Date Date period type Date period type scheme At least one begin date and one end date required.	<ul style="list-style-type: none"> Source data Notes 	Date 1 Begin Date 1478 Date 2 End Date 1834	Date 1 Most Likely Begin Date 1640 BC Date 2 Most Likely End Date 1540 BC
Period type	<ul style="list-style-type: none"> Name of period type Period type classification scheme Multiple period types allowed if applicable.	None	Groups of people–decisions and policies. Proposed period type scheme.	Groups of people–period of rule
Location	<ul style="list-style-type: none"> Location name Spatial data or <ul style="list-style-type: none"> Geographical gazetteer entry ID Geographical gazetteer version information Link to gazetteer Multiple locations allowed.	<ul style="list-style-type: none"> Source data Notes 	Spain and its possessions.	Egypt
Relationship	None	<ul style="list-style-type: none"> Relationship type Relationship scheme Entry ID to related period 	N/A	Part of Second Intermediate Period
Administrative details	<ul style="list-style-type: none"> Creator contact information Date of entry 	<ul style="list-style-type: none"> Modification history Notes 	Junko Doe University of Anonymity Spanish History Project doe@unanonymous.com Entered 8/22/03	Juanita Doe University of Anonymity Egyptian History Project doe2@unanonymous.com Entered 8/22/03