THE GEDCOM STANDARD

DRAFT Release 5.3

4 November 1993

Prepared by the
Family History Department
The Church of Jesus Christ of Latter-day Saints

Suggestions and Correspondence:
GEDCOM Coordinator - 3T
Family History Department
50 East North Temple
Salt Lake City, UT 84150
USA

Telephone (USA) 801-240-4534
240-5225

TABLE OF CONTENTS

Introduction

GEDCOM was developed by the Family History Department of the Church of Jesus Christ of Latter-day Saints to provide a flexible uniform format for exchanging computerized genealogical data. GEDCOM is an acronym for GEnealogical Data Communication. GEDCOM is provided to foster the sharing of genealogical information and the development of a wide range of inter-operable software products to assist genealogists, historians, and other researchers.

Purpose and Content of This Document

This technical document is written for computer programmers, system developers, and technically sophisticated users.

The chapters in this document contain the following GEDCOM specifications:

* Data Representation Grammar
* Lineage-linked GEDCOM Grammar
* GEDCOM Transmission File
* Values
* Character Sets

This document describes GEDCOM at two different levels. The lower level defines a general-purpose data representation language for representing any kind of structured information in a sequential media. The higher level defines specific content for data to be exchanged between compatible systems.

The lower level is known as the GEDCOM data format and deals with the syntax and identification of structured information in general, but does not deal with the semantic content of any particular kind of data. The lower level GEDCOM format and the basic GEDCOM concepts are presented in chapter 1. This chapter will also be useful to those using GEDCOM for other kinds of data, not just genealogical data.

The higher level is known as a GEDCOM form. A GEDCOM form is defined for each kind of data

that uses the GEDCOM data format. The only GEDCOM form presented in this document is called the Lineage-linked GEDCOM form. Other GEDCOM forms have been used for other kinds of data, including several that are not related to genealogy. The Lineage-linked GEDCOM form is defined in chapter 2 and is the form used by commercial genealogical software systems for exchanging compiled, linked information about individuals with accompanying source citations and evidence records. The other forms of GEDCOM are not publicly exchanged at this time, and are not discussed in this document.

Changes in Version 5.x

Prior versions of The GEDCOM Standard were released in October 1987 (3.0) and August 1989 (4.0). Versions 1 and 2 were drafts for public discussion and were not established as a standard.

This GEDCOM draft version (5.x) includes the first standard definition of the Lineage-linked form of GEDCOM and also includes the first major expansion of the Lineage-linked form since its initial use in GEDCOM 3.0. The existing registered GEDCOM-compatible systems should still be able to exchange most data with newer systems that use this version and will still be considered GEDCOM-compatible for submitting information to the Family History Department. See chapter 2, "Compatibility with previous GEDCOM releases", for compatibility detail. There are several purposes for version 5.x of GEDCOM:

* Re-define the description of the GEDCOM data representation grammar in a shorter, more precise format, for ease of understanding (see chapter 1). The GEDCOM format remains the same, even though the description of it is changed.

* Define the combinations of tags, values, and pointers allowed in the Lineage-linked form (see chapter 2). This is the form of GEDCOM currently exchanged by commercial

genealogical software systems, and it remains unchanged except for new tags and upward-compatible structural extensions listed below. (The Lineage-linked form should not be confused with other forms of GEDCOM, which apply the basic GEDCOM data format with different tag, value, and pointer combinations for other purposes.)

* Define representations for support information such as source citations, and or notes. (See chapter 2 for suggested source citation structure in the Lineage-linked grammar.)

* Define additional EVENt and Role tags.

* Define user-defined ASSOciations with INDIviduals including direct family relationships.

* Require SOURce VERSion (product version) and GEDCom VERSion information in the HEADer record.

* Define DATE modifier (ABT, BEF, AFT, BET) and a more rigorously defined regular date format.

Some changes in Version 5.2 - 5.3 that were not in previous 5.x versions are:

* An address structure was defined to provide consistency to the addresses used in the many different structures. The Phone number is now subordinate to address.

* A new tag for marrital status (MSTAT) at the time of an event was used added to the event structure.

* A mechanism for creating user-defined tags. These are defined in a SCHEMA definition in the header record.

* The inclusion of the Unicode standard (ISO 10646) as an additional character set standard (see chapter 3).

* A MULTI_MEDIA_LINK structure was introduced to provide links to digitized video and sound files.

* The NAME tag used in the SOURCE_STRUCTURE was changed back to the TITLe tag to be used with the title of a book or article.

* The SOURCE_STRUCTURE was changed. Compatibility may affect 5.x systems that

was using the CPLR, XLTR, AUTH, INFT tags in substructures within the source structure. See originator (ORIG) substructure for handling the name of the originator of the source data.

* Relocated all tags from the SUPPORT_INFO structure to the various structures where they specifically apply.

* Added the use of the FORM {FORMAT} tag in both the HEADER and PLACE_STRUCTURE. The FORM tag in the header record subordinate to the PLAC tag indicates that all of the locality names are specified in a consistent hiarchy as specified by the value of the FORM. For example; 2 FORM City, County, State. GEDCOM 5.2 used the TYPE tag subordinate to the PLAC tag for this purpose.

GEDCOM Product Registration

Developers of GEDCOM-compatible products using the Lineage-linked form of GEDCOM (see chapter 2) should register their product by submitting the following information to the GEDCOM

coordinator:

* A diskette containing a small sample of GEDCOM output from the product being registered. This should be data which represents all of the fields managed by your system and that can be used for testing compatibility with other developer's systems.

* A proposed unique SOURce name in the GEDCOM header record to identify the product (not the company). This name can be up to 40 characters long, allowing mixed upper and lower case, with no embedded spaces. Use an underscore (_) to connect multiple words instead of spaces or a combination of upper and lower case letters i.e. FamilyRecords or Family_Records. Family History reserves the right to require uniqueness within the first 10 characters of this name.

* An optional text file containing relevant technical documentation about the product's GEDCOM implementation.

GEDCOM Software Library

A library of unrestricted public domain source code, in the C programming language, is available to help reduce the work required to achieve GEDCOM compatibility.

Chapter 1

DATA REPRESENTATION GRAMMAR

INTRODUCTION

This chapter describes the core GEDCOM data representation language.

The generic data representation language defined in this chapter may be used to represent any form of structured information, not just genealogical data, using a sequential stream of characters.

CONCEPTS

A GEDCOM transmission represents a database in the form of a sequential stream of related records. A record is represented as a sequence of tagged, variable-length lines, arranged in a hierarchy. A line always contains a hierarchical level number, a tag, and an optional value. A line may also contain a cross-reference identifier or a pointer. The GEDCOM-line is terminated by a carriage return, a line feed character, or any combination of these.

The tag in the GEDCOM-line identifies the type of information contained in the line, in the same sense that a field-name identifies a field in a database record. This means that the data is self-defining. Tags allow a field to occur any number of times within a record, including zero times. They also allow the use of different or new fields to be included in the GEDCOM data without introducing incompatibility, because the receiving system will ignore data which it does not understand and process only the data that it does understand.

The hierarchical relationships are indicated by the hierarchical level number. Subordinate lines have a higher level number. The hierarchy allows a line to have sub-lines, which in turn may have their own sub-lines, and so forth. A line and its sub-lines constitute a context or enclosure, that is, a cluster of information pertaining directly to the same thing. This hierarchical arrangement corresponds with the natural hierarchy found in most structured information.

A series of one or more lines constitutes a record. The beginning of a new record is indicated by a line whose level number is 0 (zero).

A GEDCOM receiver system scans the input for expected information by looking for specific tags and processing the associated values. Unrecognized tags (perhaps from a sending system whose database contains some different information) are handled by not processing the associated value nor

its enclosed sub-lines; that is, the entire context is ignored. These are treated as exceptions by printing them in an exception report or saving them in some generic way. Saved exception lines may be recombined when the data is exported.

In addition to hierarchical relationships, GEDCOM defines inter-record relationships which allow a record to be logically related to other records, without introducing redundancy. These relationships are represented by two additional but optional parts of a line: a cross-reference pointer and a cross-reference identifier. The cross-reference pointer "points at" a related record, identified by a required, matching unique cross-reference identifier. The cross-reference identifier is analogous to a primary key in relational database terminology.

GRAMMAR

The grammar for the GEDCOM data format--a data representation language--is defined in this chapter. The grammar is a set of rules that specify what sequences of characters are valid GEDCOM expressions. The rules are expressed as a set of pattern definitions, where each pattern is defined in terms of either a more primitive sub-pattern, or a constant. Pattern definitions consist of the pattern name, a separator (:=), followed by either a constant, a more primitive sub-pattern, or a set of alternatives of these. When a set is used, the alternatives are enclosed in square brackets

[] with the alternatives separated by a vertical bar ([alternative_1 | alternative_2]). Only one is to be selected. The user can read the grammar components of the selected sub-pattern by substituting any sub-patterns until all sub-patterns are resolved.

A GEDCOM transmission consists of a sequence of physical records, each of which consists of a sequence of gedcom_lines, all contained in a sequential file or stream of characters. The following rules pertain to the gedcom_line:

* The beginning of a new physical record is designated by a line whose level number is 0.

* Physical records are intended to be small enough to fit within a memory buffer of typical size, though absolute limits are not established.

* The total length of a GEDCOM-line, including leading white space and terminators, does not exceed 255 characters. Long text can be represented by using CONTinue or CONCatenate tags.

* Leading white space (tabs, spaces, and extra line terminators) preceding a GEDCOM-line should be ignored by the reading system. Systems generating GEDCOM should not place any white space in front of the GEDCOM-line (at least for the near future, see "Compatibility With Previous GEDCOM Versions" at the end of chapter 2).

* Level numbers must not contain leading zeroes which are not significant, for example, level one must be 1, not 01.

* GEDCOM-lines constructed with user defined tags must include a tag definition in the a schema substructure in the transmission header record. The user defined tag must begin with an underscore (_). The schema allows a receiving system to interpret the associated data. (See the User Defined Tags section in chapter 2 for more information).

GRAMMAR SYNTAX

A gedcom_line has the following syntax:

gedcom_line:=
level delim opt_xref_id tag opt_line_value terminator

for example:

1 OCCU Teacher

The components of the sub-patterns above are defined below in alphabetical order. Some of the components are defined in terms of more primitive sub-patterns:

```
alpha:=
 [ (0x41)-(0x5A) | (0x61)-(0x7A) | 0x5F ]
   Any ASCII letter: A-Z, a-z, and (_) underscore

alphanum:=
 [ alpha | digit ]

any_char:=
 [ alpha | digit | otherchar | (#) | ( ) | (@) (@) ]

delim:=
 [ (0x20) ]
```

```
       space_character

    digit:=
     [ (0x30)-(0x39) ]
       One of the digits 0,1,2,3,4,5,6,7,8,9

    escape:=
     [ (@) (#) escape_text (@) non_at ]

    escape_text:=
     [ any_char | escape_text any_char ]
       The escape_text is coded to meet the rules of a particular GEDCOM
form.  For the lineage-
       linked form the definitions are found in Chap. 2.

    level:=
     [ digit | level digit ]
       (Do not use non-significant leading zeroes such as 02.)

    line_item:=
     [ pointer | escape | any_char ]

    line_value:=
     [ line_item |    line_value  line_item ]

    non_at:=
     [ alpha | digit | otherchar | (#) | ( ) ]

    null:=
       () nothing

    opt_line_value:=
     [ null | delim | delim line_value ]

    opt_xref_id:=
     [ null | pointer delim ]

    otherchar:=
     [(0x21)-(0x22) | (0x24)-(0x2F) | (0x3A)-(0x3F) | (0x5B)-(0x5E) | (0x60)
|
      (0x7B)-(0x7E) | (0x80)-(0xFF)]
    Any ASCII character except control characters (0x00 - 0x1F), alphanum,
space ( ), number sign
    (#),  at character (@), and the DEL character (0x7F).

    pointer:=
     [ "@" alphanum pointer_string "@" ]

    pointer_char:=
     [ non_at ]

    pointer_string:=
     [ null | pointer_char | pointer_string pointer_char ]

    tag:=
     [ alphanum | tag alphanum ]
```

```
terminator:=
 [ carriage_return |     line_feed |   carriage_return line_feed |
```

line_feed carriage_return]

USAGE DESCRIPTION:

alpha:=
The alpha characters include the underscore which is used to link word pieces together in forming tag names or tag labels.

any_char:=
Any character except the control characters found in the range of 0x00 - 0x1F. If an @ is desired as part of the line_value, it must be written in GEDCOM as a double @, ie., "3 doz. @ $20.00" must be stored as "3 doz. @@ $20.00".

delim:=
The delim (delimiter), a single space character, terminates both the variable-length level number and the variable-length tag. Note that space characters may also be present in a value.

escape:=
The escape is a sequence in the grammar used to specify special processing, such as switching character sets or calendars for date interpretation, or for indicating an inclusion of a non_GEDCOM data form into the GEDCOM structure. The form of the escape sequence is:

@# escape_text @ non_at.
for example:
@#DJULIAN@.

The non_at after the final at character (@) should be discarded if it is a space ( ). Otherwise, it should be retained as part of the text following the escape. Output systems should always place a space ( ) after the escape sequence.

The specific format of the escape sequence is defined for the specific GEDCOM form being defined. (See chapter 2 for the escape sequence definition for the lineage-linked form).

escape_text:=
The escape_text is defined to meet the requirements of a particular GEDCOM form. For the lineage-linked form the definitions are found in Chap. 2.

level:=
The level number works the same way as the level of indentation in an indented outline, where indented lines provide detail about the item under which they are indented. A line at

any level L is enclosed by and pertains directly to the nearest preceding line at level L-1.

The Level L may increase by 1 at most. Level numbers must not contain leading zeroes

which are not significant, for example level one must be (1), not (01).

The enclosed subordinate lines at level L are said to be in the context of the enclosing

superior line at level L-1. The meaning of a tag (see tag below) is interpreted in the context

of the tags of the enclosing line(s). Take the following record about an individual's birth

and death dates, for example:

```
0 INDI
  1 BIRT
    2 DATE 12 MAY 1920
  1 DEAT
    2 DATE 1960
```

In this example, the expression DATE 12 MAY 1920 is interpreted within the INDI (individual) BIRT (birth) context, representing the Individual's birth date. The second DATE is in the INDI DEAT (death) context. The complete meaning of DATE depends on the context. (Note: the above example is indented according to the level numbers to make the concept more obvious. In the actual GEDCOM data there is no indentation, just level numbers lined up vertically on the left margin).

NOTE: Some existing systems provide an option to produce an indented GEDCOM output for user readability, using space or tab characters between the terminator and the level number of the next line to visibly show the hierarchy. Also, some have suggested allowing extra blank lines to visibly separate physical records. These features may be incorporated into the GEDCOM standard at some future time, but for now, such a change would render some existing systems incompatible. Therefore, we recommend that new systems be prepared to discard extra carriage returns, line feeds, spaces and tabs immediately preceding the level number during input. Output should still be constrained to level numbers without indentation or blank lines, until most receiving systems are prepared to deal with this change.

line_value:=

The line_value identifies an object within the domain of possible values allowed in the context of the tag. The combination of the tag, the line_value, and the hierarchical context of the supporting gedcom_lines provides the understanding of the enclosed values. This domain is defined by a specific grammar for representing a given GEDCOM form (see chapter 2 for Lineage-linked grammar).

Values whose source information contains illegible parts of the value should be indicated by replacing the illegible part with ... (ellipses).

Values are generally not encoded in binary or other abbreviation schemes for reducing space requirements, and they are generally constrained to be understandable by a typical user without decoding. This is intended to reduce the decoding burden on the receiving software. A GEDCOM-optimized data compression standard will be defined in the future to reduce space requirements. Meanwhile, users may agree to compress and decompress GEDCOM

files using any compression system available to both sender and receiver.

The line_value within the context of a tag hierarchy of gedcom_lines represents one piece of information and corresponds to one field in traditional database or file terminology.

opt_xref_id:=
(See pointer.)
The opt_xref_id is formed by any arbitrary combination of characters from the pointer_char set. The first character must be an alpha or a digit. The opt_xref_id is not retained in the receiving system, and may therefore be formed from any convenient combination of identifiers from the sending system. No meaning is attributed by the receiver to any part of the opt_xref_id, other than its unique association with the associated record. The use of the colon (:) character is also reserved.

otherchar:=
[(0x21)-(0x22) | (0x24)-(0x2F) | (0x3A)-(0x3F) | (0x5B)-(0x5E) | (0x60) | (0x7B)-(0x7E) | (0x80)-(0xFF)]
Any ASCII character except control characters (0x00 - 0x1F), alphanum, space ( ), number sign (#), at character (@), and the DEL character (0x7F).

If any of these characters appear in the level, xref_ID, or pointer segments of the GEDCOM line, then that substructure should be written to an exception file. If any of these characters appear in the value segment and the proper escape processing has not been invoked, then

they should be replaced by a (^) (0x5E) character, unless the character is a TAB (0x09)
character which can be replaced with a space (0x20) character. These changes should also
be recorded on an exception file.

pointer:=
A pointer stands in the place of the context identified by the matching xref_id.
Theoretically, a receiving system should be prepared to follow a pointer to find any needed
value in a manner that is transparent to the logic of the subsystem that is looking for specific
tags. This highly-flexible facility will probably be used more in the future. For the time
being, however, the use of pointers is explicitly defined within the GEDCOM form (Such as
defined in Chapter 2).

The pointer represents the association between two objects that usually reside in different
records. There can, however, be an association between objects within the same logical
record. If this condition exists it is indicated in the pointer record composition containing an
(!) character that separates the parent record's cross-reference ID from the specific
substructure's cross-reference ID which is at some subordinate level to the logical at level
zero. The cross-reference ID of the substructure subordinate to a zero level record is always
composed of the Record ID number and the Substructure ID number, such as @I132!1@.
By including the Record Id number in the pointers which associate objects within a record
will allow the GEDCOM processors to build the index only at the record level and then
search sequentially for the appropriate substructure cross reference ID.

Complex logical record structures are divided into small physical records to accommodate
memory constraints, many-to-many relationships, and independent record creation and
deletion.

The pointer must match a corresponding xref_id within the transmission, unless the colon (:)
character is present (future network reference to a permanent file record). A pointer is
given instead of duplicating an object, though the logical result is equivalent. An expanded
traversal of a record tree includes following the pointers to related records to some depth,
and splicing those records (logically) into the resultant expanded tree. Pointers may refer to

either records which have not yet appeared in the transmission (forward reference) or to
records that have already appeared earlier in the transmission (backward reference). This
arrangement usually requires a preliminary pass to construct a look up table to support
random access by xref_id during subsequent passes.

tag:=
A tag consists of a variable length sequence of alphanum characters. All user defined tags,
that is tags used which have not been defined by the GEDCOM standard must begin with an
underscore character. (0x95). All user defined tags must be defined in the SCHEMA
substructure of the HEADer record.

The tag represents the meaning of the line_value within the context of the enclosing lines,
and contributes to the meaning of enclosed subordinate lines. Specific tags are defined in
Appendix A.

Although existing tags are only three or four characters long, systems should prepare to
handle tags of any length. Tags will be unique within the first 15 characters.

Valid combinations of specific tags, line_values, xref_ids, and pointers are constrained by
the GEDCOM form defined for representing a given kind of information (see chapter 2 for
the Lineage-linked form grammar).

terminator:=
The terminator delimits the variable-length line_value and signals the end of the

gedcom_line. The valid terminator characters are:

```
                [     carriage_return |
                      line_feed |
                      carriage_return line_feed |
                      line_feed carriage_return ]
```

Examples:

The following are examples of valid but unrelated GEDCOM-lines:

```
0 @1234@ INDI
. . .
1 AGE 13
. . .
1 CHIL @1234@
. . .
1 NOTE This is a note field that is
2 CONT continued on the next line.
```

The first line has a level number 0, a xref_id of @1234@, an INDI tag, and no value. The second line has a level number 1, no xref_id, an AGE tag, and a value of 13. The third line has a level number 1, no xref_id, a CHIL tag, and a value of a pointer to a xref_id named @1234@.

Chapter 2
LINEAGE-LINKED GRAMMAR

INTRODUCTION

This chapter describes the specific tag, value, and pointer combinations used for exchanging lineage-linked genealogical information in the GEDCOM format. Lineage-linked data pertains to individuals linked in family relationships across multiple generations. The chapter also addresses specific compatibility issues pertaining to previous Lineage-linked GEDCOM releases and contains a sample Lineage-linked GEDCOM transmission.

The Lineage-linked grammar defined in this chapter is based on the general framework of the GEDCOM data representation grammar defined in the Chapter 1. The lineage-linked grammar defines the GEDCOM form used by commercial genealogical software systems to exchange data. Other specialized GEDCOM-based grammars have been created for different uses. These other uses of the general-purpose GEDCOM data representation should not be confused with this specific usage for lineage-linked genealogical data, as defined in this chapter as the only approved form of GEDCOM exchanged by commercial genealogical software systems at this time.

LINEAGE-LINKED GRAMMAR ORGANIZATION

This Lineage-linked GEDCOM grammar is organized into three sections:

* Record structure components
* Substructure patterns (Arranged alphabetically by substructure name)
* Primitive elements (Arranged alphabetically by primitive name)

Structures and substructures are indicated by enclosing the structure name within double angle <<brackets>>. Primitive element patterns are enclosed in single angle <brackets>.

The definition of each structure consists of the structure name, a separator (:=), and the structure's component pattern. This pattern consists of (a) GEDCOM-lines composed of primitive elements, and/or (b) substructures. Some primitive elements consist of two or more alternative sub-pattern choices. These choices are shown by listing the alternative sub-patterns between opening and closing square [brackets] and separating each choice with a vertical bar (|), meaning that exactly one of the alternate substitutions must be selected. Some definitions of primitive elements use the

definition of other primitive elements to complete their definition. This is shown by including the
name of the detailed element type inside angle <brackets> in the definition.

The number of sub-pattern occurrences allowed within a pattern is defined in an occurrence
definition in curly {braces} on each line. This number indicates the minimum and maximum
number of occurrences allowed for a pattern component in the form {minimum:maximum}. Note
that minimum and maximum occurrence limits are defined relative to the enclosing superior line.
This means that a required line (minimum = 1) is not required in an instance where the optional
enclosing line is not given. Similarly, a line occurring only once (maximum = 1) may occur
multiple times as long as each occurs only once under its own multiple-occurring superior line.

The level numbers for any sub-structure are represented as (n), (+1), (+2), and so forth, so that
they may be used in more than one place at different starting level numbers. In these cases, (n)
equals the level number where the pattern first appears, and the (+1) means one level greater than
level n, (+2) means two levels greater than level n, and so forth.

Unless stated otherwise, the only ordering imposed on GEDCOM-lines within an enclosure arises
when multiple opinions or other items are presented for which only one may be expected by a
receiving system. For example, a person may have been known by more than one name, or
evidence may suggest a birth either in 1840 in New York or in 1837 in Pennsylvania. In these
cases, the most credible or preferred information is listed first, followed by less credible or less

preferred items. The QUAY tag may also be used to show the preferred data (see appendix A).
Systems that support only a single field within a context should use the first item in the list.

Conflicting dates or places of an event should be represented in separate event structures to provide
a place for the accompanying source citations, rather than place multiple dates or multiple places
under the same enclosing event.

Even though no other ordering is defined beyond the one described above, some GEDCOM
programming tools optimize performance based on the assumption that tags generally appear in a
typical order. Therefore, sending systems are encouraged to present GEDCOM structures in the
same general order as the one given in these patterns, unless there is a reason to use a different
sequence.

This form uses the tag TYPE as a subordinate tag to names, places, events, etc. The intent of this
tag is meant to further define its superior tag for the viewer only, it is not intended to inform a
computer program how to process the data. The difference between this value and a note value
would be that displaying systems should always display the type value when they display the
associated data. Therefore, cautious consideration should be used in using the TYPE tag.

RECORD STRUCTURES OF THE LINEAGE-LINKED FORM

LINEAGE_LINKED_GEDCOM:=
This is a model of the Lineage-linked GEDCOM structure for submitting data to other
lineage-linked GEDCOM processing systems. A header and a trailer record are required and
they enclose any number of data records.

```
0  <<HEADER>>
{1:1}
0  <<RECORD>>
{0:M}
0 TRLR
{1:1}
```

There are specific subordinate GEDCOM-lines that may be used as subordinate GEDCOM-
lines to other superior GEDCOM-lines. For example:

```
1 BIRT
2 DATE 02 Oct 1937
3 QUAY 1
```

In the above example QUAY at level 3 indicates how reliable or correct the birth date value

is. The QUAY tag applies to any tag that contains a value. This tag is not shown in any of
the structures but the reader and writer of GEDCOM should expect that the QUAY tag
could be present as a subordinate tag to any tag that has an associated value.

HEADER:=
The header structure provides information about the entire transmission. The SOURce system
name identifies which system sent the data. The DESTination system name identifies the
receiving system. Submission to the Family History Department for Ancestral File is
ANSTFILE. For LDS temple submissions it is TempleReady.

```
      n  HEAD
{1:1}
         +1 SOUR <SYSTEM_NAME>
{1:1}
            +2 VERS <VERSION_NUMBER>
{1:1}
            +2 NAME <PRODUCT_NAME>
{0:1}
            +2 CORP <CORPORATE_NAME>
{0:1}
               +3 <<ADDRESS_STRUCTURE>>
{0:1}
            +2 DATA <NAME_OF_SOURCE_DATA>
{0:1}
               +3 DATE <PUBLICATION_DATE>
{0:1}
         +1 DEST <SYSTEM_NAME>
{0:1}
         +1 DATE <TRANSMISSION_DATE>
{0:1}
            +2 TIME <TIME_VALUE>
{0:1}
         +1 SUBM @XREF:SUBM@
{1:1}
         +1 FILE <FILE_NAME>
{0:M}
         +1 COPR <COPYRIGHT_STATEMENT>
{0:1}
            +2 CONT <TEXT>
{0:M}
         +1 SCHEMA
{0:1}
            +2 <<USER_TAG_SCHEMA>>
{1:M}
         +1 GEDC
{1:1}
            +2 VERS <VERSION_NUMBER>
{1:1}
```

```
            +2 FORM <GEDCOM_FORM>
{0:1}
         +1 CHAR <CHARACTER_SET>
{0:1}
            +2 VERS <VERSION_NUMBER>
{0:1}
         +1 LANG <LANGUAGE_OF_TEXT>
{0:1}
         +1 PLAC
{0:1}
            +2 FORM <PLACE_HIERARCHY>
{1:1}

RECORD:=
       [
         n  <<EVENT_RECORD>>
{0:1}
       |
         n  <<FAMILY_RECORD>>
{0:1}
       |
         n  <<INDIVIDUAL_RECORD>>
{0:1}
       |
         n  <<NOTE_RECORD>>
{0:1}
       |
         n  <<REPOSITORY_RECORD>>
{0:1}
       |
         n  <<SOURCE_RECORD>>
{0:1}
       |
         n  <<SUBMITTER_RECORD>>
{1:1}
       ]

FAMILY_RECORD:=
       n  @XREF:FAM@ FAM
{0:1}
         +1 HUSB @XREF:INDI@
{0:1}
         +1 WIFE @XREF:INDI@
{0:1}
         +1 CHIL @XREF:INDI@
{0:M}
         +1 REFN <USER_REFERENCE_NUMBER>
{0:M}
         +1 <FAM_EVNT_TAG>
{0:M}
            +2 TYPE <FAMILY_EVENT_DESCRIPTOR>
{0:1}
            +2 DATE <DATE_VALUE>
{0:1}
            +2 <<PLACE_STRUCTURE>>
{0:1}
```

```
         +1 <DIV_EVNT_TAG>
{0:M}
            +2 TYPE <DIVORCE_DESCRIPTOR>
{0:M}
            +2 DATE <DATE_VALUE>
{0:1}
            +2 <<PLACE_STRUCTURE>
{0:1}
         +1 ASSO @XREF:ANY@
{0:M}
            +2 TYPE <ASSOCIATION_DESCRIPTOR>
{0:1}
         +1 NCHI <COUNT_OF_CHILDREN>
{0:1}
         +1 <<LDS_FAM_ORDINANCE_EVENT>>
{0:M}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}
         +1 <<MULTI_MEDIA_LINK>>
{0:M}
         +1 <<CHANGE_DATE>>
{0:M}
```

```
INDIVIDUAL_RECORD:=
      The occurrence of FAMS and FAMC tags show {0:1}, however; when an
individual is
      referenced in a FAMily record as either a spouse or child, then this
record must include a
      corresponding FAMS and/or FAMC tags.  The association of one individual
to another can be
      represented by using the ASSO tag in the individual record to point to
the record of the
      associated individual.  The relationship or association is shown in the
value field of the
      subordinate TYPE tag.

      n  @XREF:INDI@ INDI
         +1 <<INDIVIDUAL>>
{1:1}
         +1 FAMS @XREF:FAM@
{0:M}
         +1 FAMC @XREF:FAM@
{0:M}
            +2 <<CHILD_FAMILY_EVENT>>
{0:M}
         +1 ASSO @XREF:REC@
{0:M}
            +2 TYPE <ASSOCIATION_DESCRIPTOR>
{0:1}
         +1 <<LDS_INDI_ORDINANCE_EVENT>>
{0:M}
         +1 RFN <PERMANENT_RECORD_FILE_NUMBER>
{0:M}
         +1 REFN <USER_REFERENCE_NUMBER>
{0:M}
         +1 AFN <ANCESTRAL_FILE_NUMBER>
{0:1}
         +1 ALIA @XREF:INDI@
{0:M}
         +1 ANCI @XREF:SUBM@
{0:M}
         +1 DESI @XREF:SUBM@
{0:M}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}
         +1 <<MULTI_MEDIA_LINK>>
{0:M}
         +1 <<CHANGE_DATE>>
{0:M}

EVENT_RECORD:=
      This structure represents event-oriented evidence information that is
claimed as a basis for a
      submitter's opinion expressed in Lineage-linked INDIVIDUAL and FAMILY
records.  Event
      records define an event in terms of a what happened, where and when it
happened, and what
```

individuals are mentioned in the record.

These event records in some cases will be the source for assertions made in compiling lineage-linked data. SOURce pointers to the bibliographic description of where this event information was recorded should be a part of this record.

Evidence records from historical sources are kept separate from opinion records created by the submitter. The information contained in evidence records is not redundant with respect to the information contained in submitter's opinions, even when names, dates, or places are the same, because the authority for asserting the information is different.

Roles of an event which pertain to the event itself are placed subordinate to the event tag.

Roles of individuals mentioned in the event which are relationship roles such as the "husband's father" is placed subordinate to the role tag of the groom. For example, the minister at a wedding's role would be represented by the 0 EVENt-MARRiage-OFFIciator structure. The father of the husband would be represented by the 0 EVENt-MARRiage-HUSBand-FATHer structure.

```
n  @XREF:EVEN@ EVEN
   +1 <<CHANGE_DATE>>
{0:M}
   +1 <EVENT_TAG>
{1:1}
      +2 TYPE <EVENT_DESCRIPTOR>
{0:1}
      +2 DATE <DATE_VALUE>
{0:1}
      +2 <<PLACE_STRUCTURE>>
{0:1}
      +2 PERI <TIME_PERIOD>
{0:M}
```

```
            +2 RELI <RELIGIOUS_AFFILIATION>
{0:1}
            +2 <<MULTI_MEDIA_LINK>>
{0:M}
            +2 <<TEXT_STRUCTURE>>
{0:1}
            +2 <<SOUR_STRUCTURE>>
{0:M}
            +2 <<NOTE_STRUCTURE>>
{0:M}
            +2 <ROLE_TAG>
{0:M}
               +3 TYPE <ROLE_DESCRIPTOR>
{0:1}
               +3 <<INDIVIDUAL>>
{0:1}
               +3 ASSO @XREF:INDI@
{0:M}
                  +4 TYPE <ASSOCIATION_DESCRIPTOR>
{1:1}
               +3 <RELATIONSHIP_ROLE_TAG> [NULL | @XREF:INDI@ ]
{0:M}
                  +4 TYPE <ROLE_DESCRIPTOR>
{0:1}
                  +4 <<INDIVIDUAL>>
{0:1}

NOTE_RECORD:=              /* must contain cross reference ID */
      n  <<NOTE_STRUCTURE>>
{1:1}
         +1 <<CHANGE_DATE>>
{0:M}

REPOSITORY_RECORD:=        /* must contain cross reference ID */
      n  <<REPOSITORY_STRUCTURE>>
{1:1}
         +1 <<CHANGE_DATE>>
{0:M}

SOURCE_RECORD:=            /* must contain cross reference ID */
      n  <<SOURCE_STRUCTURE>>
{1:1}
         +1 <<CHANGE_DATE>>
{0:M}


SUBMITTER_RECORD:=
      The submitter record identifies individuals or organizations that
contributed the opinion
      information contained within the GEDCOM transmission.  All records in
the transmission are
      assumed to be submitted by the SUBMITTER referenced in the HEADer,
unless a SUBMitter
      reference inside a specific record points at a different SUBMITTER.

      n  @XREF:SUBM@ SUBM
{1:1}
```

```
         +1 <<NAME_STRUCTURE>>
{1:1}
         +1 <<ADDRESS_STRUCTURE>>
{0:1}
         +1 LANG <LANGUAGE_PREFERENCE>
{0:3}
         +1 <<CHANGE_DATE>>
{0:M}


SUBSTRUCTURES OF THE LINEAGE-LINKED FORM

ADDRESS_STRUCTURE:=
      n  SITE <SITE_NAME>
{0:1}
      n  ADDR <ADDRESS_LINE>
{0:1}
         +1 CONT <ADDRESS_LINE>
{0:M}
         +1 PHON <PHONE_NUMBER>
{0:3}

BURIAL_STRUCTURE:=
      Used only when cemetery information is managed separately from the
burial place name.  It is
      permissible to include the cemetery name as the low level locality
name; for example,
      Richmond Cemetery, Richmond, Cache, Utah, USA.
      n  CEME <CEMETERY_NAME>
{0:1}
         +1 PLOT <BURIAL_PLOT_ID>
{0:1}

CHANGE_DATE:=
```

```
      n  CHAN
{1:1}
         +1 DATE <CHANGE_DATE>
{1:1}
            +2 TIME <TIME_VALUE>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}

CHILD_FAMILY_EVENT:=
      [
      n  ADOP
{1:1}
         +1 TYPE <CHILD_FAMILY_EVENT_DESCRIPTOR>
{0:1}
         +1 AGE  <AGE_VALUE>
{0:1}
         +1 DATE <DATE_VALUE>
{0:1}
         +1 <<PLACE_STRUCTURE>>
{0:1}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}
      |
      n  <<LDS_CHILD_SEALING_EVENT>>
{0:1}
      ]

CORRECTNESS_ASSESMENT:=
      n  QUAY <QUALITY_OF_DATA>
{0:1}
                          /* used subordinate to any tag containing a value */

EVENT_STRUCTURE:=
      Information about an individual with respect to a specific event, such
as the age, marital
      status, religious affiliation of this individual at time of this event.
Keep in mind that this is
      data specific to the individual owning this event and not the data that
belongs to the source in
      which this data was found.  For instance Immigration and Emigration
events should use a
      reference a source structure to show the SHIP and PORT information
concerning the event.
      Roles of other individuals can be shown using the EVENt record.  A link
to the event record
      can be made by using the SOURce structure to point to the EVENt record.
The event record
      in this case would be an evidence record supporting the assertions made
in creating this event
      structure.

      n  <EVENT_TAG>
{1:1}
```

```
         +1 TYPE <EVENT_DESCRIPTOR>
{0:M}
         +1 DATE <DATE_VALUE>
{0:1}
         +1 <<PLACE_STRUCTURE>>
{0:1}
            +2 <<BURIAL_STRUCTURE>>
{0:1}
         +1 AGE <AGE_VALUE>
{0:1}
         +1 MSTAT <MARITAL_STATUS>
{0:1}
         +1 CAUS <CAUSE_OF_DEATH>
{0:1}
         +1 RELI <RELIGIOUS_AFFILIATION>
{0:1}
         +1 AGNC <GOVERNMENT_AGENCY>
{0:1}
         +1 <<TEXT_STRUCTURE>>
{0:1}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}
         +1 <<CHANGE_DATE>>
{0:M}

INDIVIDUAL:=
      n  <<NAME_STRUCTURE>>
{1:M}
      n  TITL <INDI_TITLE>
{0:M}
      n  SEX <SEX_VALUE>
{0:1}
      n  <<EVENT_STRUCTURE>>
{0:M}
      n  <<ADDRESS_STRUCTURE>>
{0:M}
      n  RELI <RELIGIOUS_AFFILIATION>
{0:M}
      n  NAMR <RELIGIOUS_NAME>
{0:M}
         +1 RELI <RELIGIOUS_AFFILIATION>
{0:1}
      n  EDUC <SCHOLASTIC_ACHIEVEMENT>
{0:M}
      n  OCCU <OCCUPATION>
{0:M}
      n  SSN  <SOCIAL_SECURITY_NUMBER>
{0:M}
      n  IDNO <NATIONAL_ID_NUMBER>
{0:M}
         +1 TYPE <TYPE_OF>
{1:1}
      n  PROP <POSSESSIONS>
{0:M}
```

```
      n  DSCR <PHYSICAL_DESCRIPTION>
{0:M}
         +1 CONT <PHYSICAL_DESCRIPTION>
{0:M}
      n  SIGN <SIGNATURE_INFO>
{0:M}
      n  NMR <COUNT_OF_MARRIAGES>
{0:M}
      n  NCHI <COUNT_OF_CHILDREN>
{0:M}
      n  NATI <NATIONALITY>
{0:M}
      n  CAST <CASTE_NAME>
{0:M}

LDS_CHILD_SEALING_EVENT:=
      n  SLGC
{1:1}
         +1 TYPE <LDS_CHILD_SEALING_DESCRIPTOR>
{0:1}
         +1 DATE <DATE_VALUE>
{0:1}
         +1 TEMP <TEMPLE_VALUE>
{0:1}

LDS_FAM_ORDINANCE_EVENT:=
      n  SLGS
{1:1}
         +1 TYPE <LDS_FAM_ORD_DESCRIPTOR>
{0:1}
         +1 DATE <DATE_VALUE>
{0:1}
         +1 TEMP <TEMPLE_VALUE>
{0:1}

LDS_INDI_ORDINANCE_EVENT:=
      n  <LDS_INDI_ORD>
{1:1}
         +1 TYPE <LDS_INDI_ORD_DESCRIPTOR>
{0:1}
         +1 DATE <DATE_VALUE>
{0:1}
         +1 TEMP <TEMPLE_VALUE>
{0:1}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}

MULTI_MEDIA_LINK:=
      n  AUDIO <ESCAPE_TO_AUXILLARY_PROCESSING>
{0:1}
      n  PHOTO <ESCAPE_TO_AUXILLARY_PROCESSING>
{0:1}
      n  VIDEO <ESCAPE_TO_AUXILLARY_PROCESSING>
{0:1}
```

```
NAME_STRUCTURE:=
      n  NAME <PERSONAL_NAME>
{1:1}
         +1 TYPE <NAME_TYPE_DESCRIPTOR>
{0:1}
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}

NOTE_STRUCTURE:=
      This structure contains information originated by the submitter.
      n  [ @XREF:NOTE@ | NULL ] NOTE [ <SUBMITTER_TEXT> | NULL ]
{1:1}
         +1  CONT <SUBMITTER_TEXT>
{1:M}
         +1  NOTE @XREF:NOTE@
{0:1}

PLACE_STRUCTURE:=
      n  PLAC <PLACE_VALUE>
{1:1}
         +1 FORM <PLACE_HIERARCHY>
{0:1}
         +1 <<ADDRESS_STRUCTURE>>
{0:1}
```

```
         +1 <<SOUR_STRUCTURE>>
{0:1}
         +1 <<NOTE_STRUCTURE>>
{0:1}

REPOSITORY_STRUCTURE:=
      n  [ @XREF:REPO@ | NULL ] REPO
{1:1}
            +2 NAME <NAME_OF_REPOSITORY>
{0:1}
            +2 CNTC <NAME_OF_CONTACT_PERSON>
{0:1}
            +2 <<ADDRESS_STRUCTURE>>
{0:1}
            +2 MEDI <MEDIA_TYPE>
{0:1}
            +2 CALN <SOURCE_CALL_NUMBER>
{0:1}
               +3 ITEM <FILM_ITEM_IDENTIFICATION>
{0:1}
               +3 SHEE <SHEET_NUMBER>
{0:1}
               +3 PAGE <PAGE_NUMBER>
{0:1}
            +2 REFN <MANUAL_FILING_IDENTIFICATION>
{0:1}
            +2 <<NOTE_STRUCTURE>>
{0:1}

SOURCE_STRUCTURE
      The source structure represents the submitter's basis (justification)
for the opinions asserted in
      a lineage linked transmission.  This information is used by other
researchers to (1) determine
      how much confidence to place in the associated assertions, (2) compare
new evidence to old
      evidence from prior research, and (3) locate and examine the evidence
to make an independent
      evaluation of it.  If a source is not explicitly cited for a given
context, the source is by default
      ascribed to be the personal opinion of the submitter, with no further
basis for its credibility.

      The justification takes the form of a description of the source from
which the evidence was
      obtained, and may include a machine-readable representation of the
evidence itself, such as an
      image of a document or an extract of its contents.

      A given source may be the basis for many different assertions.  Thus,
much of the information
      is the same for many different citations of that source, such as the
publisher information; and
      yet, some of the information varies from one citation to the next, such
as the page number for
      a specific item.  Consequently, the SOURCE_STRUCTURE includes a
sophisticated
```

mechanism for sharing general source description information that is common across multiple citations, while at the same time allowing more specific information to be more directly associated with individual citations. All tags within the SOURCE_STRUCTURE participate in this approach.

To implement the mechanism, the SOURCE_STRUCTURE includes a SOURce pointer that refers to another SOURCE_STRUCTURE containing more general information to be included in the citation. This forms a chain of records, beginning within an individual or family record and ending in a source record that does not contain another SOURce pointer.

A given tag may appear in more than one record along the chain. In this case, the tag occurring in one link (source record) of the chain is said to shadow or supersede the same tag found in subsequent records of the chain. A program looking for a particular tag (or tags) in the citation starts looking in the first record of the chain and continues looking in each subsequent record in the chain for the appropriate tag, succeeding when the tag is found or failing when the end of the chain is reached. In effect, a complete logical source citation is the set of all tags of all records within the source chain, excluding shadowed tags.

The chain may consist of only one SOURCE_STRUCTURE contained entirely inside an individual or family record, with no SOURce pointer leading out from the individual or family record. More typically, the chain will begin in the individual or family record and end in an ordinary source description record. Occasionally, a multiple volume source may be represented using a record in the middle of the chain for specific information about the volume.

For example, in a multiple volume source where each volume covered a range of years, a volume description would contain the PERIod covered by the volume, and the more general description of the set of volumes would contain the PERIod covered by the entire set of volumes. In assembling the complete source citation, the program would stop searching for the PERIod as soon as it found a PERIod tag, which in this case would be in the volume description. In a multiple volume source where each volume covered a specific place as part of a larger grouping of places, the program would find the PLACE_STRUCTURE information in the intermediate volume description, and it would find the PERIod information in the final, more general description of the set of volumes.

We encourage data entry systems to develop flexible entry screens which will prompt their users for information which will meet the minimum standards for citing sources. At the minimum there should be an entry form for published sources and one for unpublished sources. The elements below are marked if they were recommended by the National Genealogical Society as being a help in citing puplished (p) or unpublished (u) sources.

```
SOURCE_STRUCTURE:=
                                /****** TYPE OF SOURCE ******/
      n  [ @XREF:SOUR@ | NULL ] SOUR [ <TEXT> | NULL ]
         +1 [ CONT | CONC ] <TEXT>
{0:1}
         +1 CLAS <SOURCE_CLASSIFICATION_CODE>
{1:1}up
         +1 EVEN  <EVENT_CLASSIFICATION_CODE>
{0:1}
         +1 PERI <TIME_PERIOD_COVERED>
{0:M}up
                                /****** CITATION SPECIFIC INFO ******/
         +1 TITL [<DESCRIPTIVE_TITLE> | @XREF:SOUR@]
{0:1}up
         +1 SOUR [ @XREF:SOUR@ | @XREF:EVEN ]
{0:M}up
         +1 PAGE <PAGE_DESCRIPTION>
{0:1}up
         +1 DATE <ENTRY_RECORDED_DATE>
{0:1}u
         +1 CENS
{0:1}
            +2 DATE <CENSUS_DATE>
{0:1}u
            +2 LINE <LINE_NUMBER>
{0:1}u
            +2 DWEL <DWELLING_NUMBER>
{0:1}u
```

```
            +2 FAMN <FAMILY_NUMBER>
{0:1}u
            +2 <<NOTE_STRUCTURE>>
{0:1}
                                /****** WHO CREATED IT ******/
      +1 ORIG
{0:M}
            +2 NAME <ORIGINATOR_NAME>
{0:1}up
            +2 TYPE <ORIGINATOR_TYPE>
{1:1}up
            +2 <<NOTE_STRUCTURE>>
{0:1}

                                /****** PUBLICATION INFO ******/
         +1 PUBL
{0:1}
            +2 TYPE <PUBLICATION_TYPE>
{1:1}up
            +2 NAME <NAME_OF_PUBLICATION>
{0:1}p
            +2 PUBR <PUBLISHER_NAME>
{0:1}p
            +2 <<ADDRESS_STRUCTURE>
{0:1}
            +2 DATE <PUBLICATION_DATE>
{0:1}up
            +2 EDTN <PUBLICATION_EDITION>
{0:1}p
            +2 SERS <SERIES_VOLUME_DESCRIPTION>
{0:1}p
            +2 ISSU <PERIODICAL_ISSUE_NUMBER>
{0:1}p
            +2 LCCN <LIBRARY_CONGRESS_CALL_NUMBER>
{0:1}

                                /****** WHERE IS IT STORED ******/
         +1 <<REPOSITORY_STRUCTURE>>
{0:1}up

                                /****** IMMIGRATION/EMIGRATION ***/
```

```
            +2 NAME <NAME_OF_VESSEL>
{0:1}
            +2 PORT
{0:1}
               +3 ARVL
{0:1}
                  +4 DATE <ARRIVAL_DATE>
{0:1}
                  +4 PLAC <ARRIVAL_PLACE>
{0:1}
               +3 DPRT
{0:1}
                  +4 DATE <DEPARTURE_DATE>
{0:1}
                  +4 PLAC <DEPARTURE_PLACE>
{0:1}
            +2 <<TEXT_STRUCTURE>>
{0:1}
            +2 <<NOTE_STRUCTURE>>
{0:1}
                                /****** SUPPORT DATA ******/
         +1 <<TEXT_STRUCTURE>>
{0:1}
         +1 <<MULTI_MEDIA_LINK>>
{0:M}
         +1 <<NOTE_STRUCTURE>>
{0:1}
         +1 STAT <SEARCH_STATUS>
{0:1}
            +2 DATE <SEARCH_STATUS_DATE>
{0:1}
         +1 REFS @XREF:SOUR@    /* REFERENCED SOURCE */
{0:1}
         +1 FIDE <SOURCE_FIDELITY_CODE>
{0:1}
         +1 QUAY <QUALITY_OF_DATA>
{0:1}

TEXT_STRUCTURE:=
      This structure contains information from the source document.
      n  TEXT <SOURCE_TEXT>
{1:1}
         +1  [ CONT | CONC ] <SOURCE_TEXT>
{1:M}
         +1  <<NOTE_STRUCTURE>>
{0:1}

USER_TAG_IN_CONTEXT:=
      A context structure which represents all of the superior level numbers
and associated tags
      from level zero to the level of the new user tag.  All user tag names
must start with and
      underscore (_).

      0  <OLD_TAG_1>
{1:1}
```

```
         1 <OLD_TAG_2>
{0:M}
            2 _<NEW_TAG>
{0:M}

                /* always start user tag name with an underscore (_).*/
```

For example, two new user tags are to be defined as _HOSP and _NURS and placed subordinate to an individual's birth. The user tag in context would be: (Example only)

```
n  INDI
   +1 BIRT
      +2 _HOSP
      +2 _NURS
```

The resulting USER_TAG_SCHEMA, to be included in the HEADer record, would then look like the following:

(Example only)

```
n  SCHEMA
   +1 INDI
      +2 BIRT
         +3 _HOSP
            +4 LABL <FULL_TAG_NAME>
            +4 DEFN <USER_TAG-DEFINITION>
            +4 ISA  <IS_A_KIND_OF_TAG>
         +3 _NURSE
            +4 LABL <FULL_TAG_NAME>
            +4 DEFN <USER_TAG-DEFINITION>
            +4 ISA  <IS_A_KIND_OF_TAG>
```

See User Defined Tag section at the end of chapter 2 for additional information.

```
USER_TAG_SCHEMA:=
      n  <<USER_TAG_IN_CONTEXT>>
{1:M}
         +m LABL <FULL_TAG_NAME>
{1:1}
         +m DEFN <USER_TAG_DEFINITION>
{1:1}
         +m ISA  <IS_A_KIND_OF_TAG>
{1:1}
           /*  +m represents the first subordinate level to the new user
defined tag level.  (See
               example shown under the substructure definition for
USER_TAG_IN_CONTEXT). */
```

PRIMITIVE ELEMENTS OF THE LINEAGE-LINKED FORM

The fields sizes are to show the minimum recommended field length within a database that is
constrained to fixed length fields. GEDCOM lines are limited to 255 characters. However, data of
any length can be included in GEDCOM by using the CONCatenation or CONTinuation tag to
expand a field beyond the 255 limit. These two tags are being used to extend text type messages
rather than extending, for example, a name line. Text lines are used in ADDR, DSCR, NOTE,
SOUR, TEXT, etc.

ADDRESS_LINE:= {Size=1:40}
Address information that, when combined with NAME and CONTinuation lines, meets
requirements for sending communications through the mail.

AGE_VALUE:= {Size=1:30}
A number that indicates the age in years, months, and/or days. Any labels must come after their
corresponding number, for example; 4 yr 8 mo 10 da. The year is required, and listed first,
even if it is 0 (zero).

ANCESTRAL_FILE_NUMBER:= {Size=1:8}
A unique permanent record number of an individual record contained in the LDS Ancestral File.

ARRIVAL_DATE:= {Size=1:90}
<DATE_VALUE>
A date associated with an arrival event, such as the arrival of a ship into a port.

ARRIVAL_PLACE:= {Size=1:120}
<PLACE_VALUE>
The place from which travel terminated, such as the locality name of a port of arrival, such as
Ellis Island, New York, New York.

ASSOCIATION_DESCRIPTOR:= {Size=1:90}
A word or phrase that describes the association between this person and another person identified
by a pointer. (For example, n ASSO great grandfather @XREF:SUBM@ would be read, this
person is a great-grandfather of the person defined in the submitter record.)

AUXILLARY_FILE_REFERENCE:= {Size=1:30}
A full file reference to the auxillary data to be linked to the GEDCOM context.

AUXILLARY_SET_FORMAT:= {Size=1:10}
[OLE | GIF | TIF | WPG | etc.]
Indicates the format of the data that is being linked to the GEDCOM context. This will allow

the GEDCOM processor to determine whether they are able to process the auxillary data. The
auxillary file should contain a header record with data required, by the indicated format, to
process the file data.

CALENDAR_ESCAPE_SEQUENCE:= {Size=4:15}
[@#DHEBREW@ | @#DROMAN@ | @#DFRENCH R@ | @#DGREGORIAN@ |
@#DJULIAN@ | @#DUNKNOWN@]
An escape sequence that allows dates from one of the indicated calendars to be represented. The
default calendar is the Gregorian calendar.

CASTE_NAME:= {Size=1:90}
A name assigned to a particular group that this person was associated with, such as a particular
racial group, religious group, or a group with an inherited status.

CAUSE_OF_DEATH:= {Size=1:90}
The cause of death of this person. This should be the same cause as listed on the death
certificate if known. (A medical history structure may be developed for a future GEDCOM

release.)

CEMETERY_NAME:= {Size=1:90}
The name of the cemetery where a person was buried.

CHANGE_DATE:= {Size=10:11}
<DATE_EXACT>
The date that this data was last changed.

CHARACTER_SET:= {Size=1:8}
A code value that represents the character set to be used to interpret this data. The default
character set is ANSEL which includes ASCII as a subset. UNICODE is also will be allowed.
See chapter 3.

CHILD_FAMILY_EVENT_DESCRIPTOR:= {Size=1:90}
A word or phrase that describes or modifies the adoption event being reported.

CONCATENATED_DATA:= {Size=1:247}
Adds new data to the end of the data in the preceding context.

CONTACT_PERSON:= {Size=1:120}
<PERSONAL_NAME>
The name of the person to whom communications should be addressed.

CONTINUED_DATA:= {Size=1:247}
A new line which logically is included in the preceding line. This may be used in specified
situations where the value length exceeds the maximum allowed length for the line.

COPYRIGHT_STATEMENT:= {Size=1:90}
A copyright statement needed to protect the rights of the owner of this data.

CORPORATE_NAME:= {Size=1:90}
The company, corporate or government agency name.

COUNT_OF_CHILDREN:= {Size=1:3, Type=NUMBER}
The number of children of this individual from all marriages or of this family, regardless of
whether the associated children are represented in the GEDCOM file.

COUNT_OF_MARRIAGES:= {Size=1:3, Type=NUMBER}
The number of different families that this person was known to have been a member of as a
spouse or parent, regardless of whether the associated families are represented in the GEDCOM
file.

DATE_DUAL:= {Size=1:90}
<DATE_REGULAR/<YEAR_ALTERNATIVE>

A date which shows the possible date alternatives arising from a calendar change, for example,
15 Dec 1752/3.

DATE_EXACT:= {Size=10:11}
<DAY> <MONTH> <YEAR>
A formatted date with one space between the day and the month and one space between the
month and the year.

DATE_MODIFIER:= {Size=3:15}
[ABT | AFT | BEF | EST | <CALENDAR_ESCAPE_SEQUENCE>]
Qualifies the meaning of a date.

```
       ABT = About
       AFT = After
       BEF = Before
       EST = Estimated

DATE_PHRASE:=         {Size=1:90}
    <text>
    Any statement offered as a date when the specific year is not known, but
which gives
    information about when an event occurred.

DATE_RANGE:=          {Size=17:31}
    [ BET <DATE_REGULAR> AND <DATE_REGULAR> ]

DATE_REGULAR:=                                                {Size=4:35}
    [ <DATE_MODIFIER | blank ] [ <DATE_EXACT> | <MONTH> <YEAR> |
    <YEAR> ]

DATE_VALUE:=          {Size=1:90}
    [ <DATE_REGULAR> | <DATE_PHRASE> | <DATE_RANGE> |
    <DATE_WITH_BC> |
     <DATE_DUAL> | <DATE_MODIFIER> <DATE_REGULAR> ]
    Examples:
       15 JUN 1990
       2 days after easter 1790
       BET NOV 1830 AND 25 DEC 1830
       600 B.C.
       ABT 1 JAN 1440
       @#DFRENCH R@28 NIVOSE AN09

DATE_WITH_BC:=                                                {Size=1:90}
    [ <DATE_PHRASE> <YEAR> B.C. ]
    A date of an event that occurred before Christ.

DAY:=                                                         {Size=1:2,
Type=NUMBER}
    dd
    Day of the month, where dd is a numeric digit whose value is within the
valid range of the days
    for the associated month.
DEPARTURE_DATE:=                                              {Size=1:90}
    <DATE_VALUE>
    A date associated with an departure event, such as the departure of a
ship from a port.

DEPARTURE_PLACE:=                                             {Size=1:120}
    <PLACE_VALUE>
    The place from which travel began, such as the locality name of a port of
departure, such as
    Pier 37, San Francisco, California.

DESCRIPTIVE_TITLE:=                                           {Size=1:247}
    A descriptive title of the information source, such as a description of:
       *        A title of an article published in a periodical.
       *        A letter including the date, the sender and the receiver.
       *        A transaction between a buyer and seller including their
names and date of
```

transaction.
* A Family Bible containing genealogical information including past and present
owners and a physical description of the book.
* A personal interview.

DIVORCE_DESCRIPTOR:= {Size=1:90}
A word or phrase that commonly describes the kind of separation, such as "divorce" or
"separated", that took place between husband and wife. The separation descriptor should use the
same word or phrase and in the same language, whenever possible, that was used by the
recorder of the event.

DIV_EVNT_TAG:= {Size=3:4}
[ANUL | DIV | DIVF] (See Appendix B for additional Tags)
A family event tag which describes the event of separation.

ENTRY_RECORDING_DATE:= {Size=1:90}
<DATE_VALUE>
The date that the entry was entered into the source record by the recorder.

ESCAPE_TO_AUXILLARY_PROCESSING:= {Size=1:30}
[@#A<AUXILLARY_FILE_REFERENCE> <AUXILLARY_SET_FORMAT>
An escape sequence which allows for alternate data formats to be linked to a specific context
within the GEDCOM file. The linked data referenced is for special processing and is tied to the
context in which the escape was issued. For instance, data specific to Window's Object linking
and embedding servers would be referenced in this manner. See Chapter 6, "Microsoft
Windows Programmer's Reference" for the format of the standard OLE data stream. This
allows the transmission of images, sounds, or other auxillary processing associated with the
enclosing context. The format of the escape sequence has only been designed for including data
by referencing a specific file name. This means that there will be an unique auxillary data file
for each link. In the future we may adopt a method of including all of the auxillary data in a
single auxillary transmission file. Other auxillary process formats may also be defined in later
GEDCOM versions.

EVENT_CLASSIFICATION_CODE:= {Size=1:90}
[<IND_EVNT_TAG> | <EVENT_DESCRIPTOR>]
A code that classifies the principal event that caused this source record to be created.
EVENT_DESCRIPTOR:= {Size=1:90}
A descriptor that should be used whenever the EVEN tag is used to define the event being cited.
For example, if the event was a purchase of a residence, the EVEN tag would be followed by

the phrase "Purchased Residence." When this descriptor is used with any of the defined event tags, it modifies the basic definition of the associated tag. For example the BIRT tag could be used in connection with an EVENT_DESCRIPTOR of "Stillborn" to modify the birth event as a stillborn birth. An EVENT_DESCRIPTOR of "DEAD" shows a person is dead but the death date is not known. The event descriptor should use the same word or phrase and in the same language, when possible, that was used by the recorder of the event. Systems that display data from the GEDCOM form should be able to display the descriptor value in their screen or printed output.

EVENT_TAG:= {Size=3:4}
[<IND_EVNT_TAG> | <FAM_EVNT_TAG> | <DIV_EVNT_TAG>]
An event tag chosen from the tags identifying either individual or family events, including the EVEN tag with an event descriptor.

FAMILY_EVENT_DESCRIPTOR:= {Size=1:90}
A word or phrase that best describes the circumstances that created this family. The marriage descriptor should use the same word or phrase and in the same language, when possible, that was used by the recorder of the event. Possible descriptor values include "Childbirth-unmarried," "Common Law," "Tribal Custom," for example. Systems that display data from the GEDCOM form should be able to display the descriptor value in their screen or printed output. (See also <DIV_EVNT_TAG>.)

FAM_EVNT_TAG:= {Size=3:4}
[CENS | MARR | MARB | MARC | MARL | MARS | ENGA | EVEN]
(See Appendix B for additional Tags)
An event tag indicating the reason for defining a family.

FILE_NAME:= {Size=1:90}
The name of the GEDCOM transmission file on the source operating system. It includes the path, file name, and file extension. The path may optionally include the drive letter.

FILM_ITEM_IDENTIFICATION:= {Size=1:90}
A particular book or unit of material that may have been filmed with other books or units on the same microfilm. The convention used in the Family History Department microfilms is to include a separator frame with a sequential item number to separate multiple books on a single film.

FULL_TAG_NAME:= {Size=1:15}

The long name of a user defined GEDCOM tag. For example, HOSP tag would have a long
name of HOSPITAL. This name should be a name that could be used as a field label for reports
and screens. The name may include underscore characters (_).

GEDCOM_FORM:= {Size=1:15}
[LINEAGE-LINKED | (others to be registered)]
The GEDCOM form used to construct this transmission.

GOVERNMENT_AGENCY:= {Size=1:90}
The name of the branch of government associated with this event or data.

IND_EVNT_TAG:= {Size=3:4}
[ADOP | BIRT | BAPM | BARM | BASM | BLES | BURI | CENS | CHR | CHRA |
CONF | DEAT | EVEN | EMIG | GRAD | IMMI | MARR | NATU | ORDN | RETI |
PROB | WILL]

An individual event tag. The EVEN tag must be followed by a TYPE and an
<EVENT_DESCRIPTOR>. The <EVENT_DESCRIPTOR> is optional for the defined event
tags, for example:

```
1 EVEN
2 TYPE Farley Family Reunion
1 BIRT
2 TYPE illegitimate
```

(See Appendix A for tag definitions or see Appendix B for proposed Tags. These proposed tags
have not been standardized. They may be used as a value for the TYPE tag under the EVEN
tag or under the appropriate approved event tags. Appropriate means that the event should be
processed the same as the selected superior tag)

INDI_TITLE:= {Size=1:90}
A formal designation used by an individual in connection with the individuals name, for
example, (Captain) John Smith.

INFORMANTS_NAME:= {Size=1:90}
<PERSONAL_NAME>
The name of a person who contributed evidence information.

INTERVIEWERS_NAME:= {Size=1:90}
<PERSONAL_NAME>
The name of the person who conducted the interview for information.

IS_A_KIND_OF_TAG:= {Size=1:25}
[<LANGUAGE_TABLE>]
The human language in which the data in the transmission is normally read or written. It is used
primarily by programs to select language-specific sorting sequences and phonetic name matching
algorithms.

LANGUAGE_PREFERENCE:= {Size=1:90}
[<LANGUAGE_TABLE>]
The language in which a person prefers to communicate. Multiple language preference is shown
by using multiple occurrences in order of priority.

LANGUAGE_TABLE:= {Size=1:25}
A table of valid language codes. This table of valid languages may be found in the
Encyclopedia Britannica 1989 Book of the Year.

LDS_CHILD_SEALING_DESCRIPTOR:= {Size=1:20}
<LDS_ORDINANCE_DESCRIPTOR>
A descriptor that describes the disposition of this ordinance. The appropriate descriptor is one
of the choices defined by <LDS_ORDINANCE_DESCRIPTOR>.

LDS_FAM_ORD_DESCRIPTOR:= {Size=1:20}
<LDS_ORDINANCE_DESCRIPTOR>
A descriptor that describes the disposition of this ordinance. The appropriate descriptor is one
of the choices defined by <LDS_ORDINANCE_DESCRIPTOR>.

LDS_INDI_ORD:= {Size=3:4}
[BAPL | CONL | WAC | ENDL]
A tag that represents an individual's religious event associated with The Church of Jesus Christ
of Latter-day Saints. (See Appendix A for a definition of these tags.)

LDS_INDI_ORD_DESCRIPTOR:= {Size=1:90}
<LDS_ORDINANCE_DESCRIPTOR>
A descriptor that specifies the disposition of this ordinance. The appropriate descriptor is one of
the choices defined by <LDS_ORDINANCE_DESCRIPTOR>.

LDS_ORDINANCE_DESCRIPTOR:= {Size=1:20}
[BIC | CANCELED | COMPLETED | DNS | DONE | INFANT | STILLBORN |
SUBMITTED]
A code indicating the status of an LDS ordinance.
BIC = This person was born in the covenant, meaning that he or she automatically
receives the blessing of 'child to parent' sealing.
COMPLETED= This ordinances has been completed but the date is not known.
DNS = This record is not being submitted for this temple ordinances.
DONE = This ordinance has been completed but the date is not known.

INFANT = This person died before eight years old.
STILLBORN = This person was stillborn.
SUBMITTED = This ordinance was previously submitted.

LIBRARY_CONGRESS_CALL_NUMBER:= {Size=1:20}
The call number assigned to this item by the U.S. Library of Congress.

MANUAL_FILING_IDENTIFICATION:= {Size=1:90}
A description of where the source is manually filed at this repository or personal collection.
Personal genealogical collections should be organized and filed so that items can be specifically
identified and retrieved. For example, "Probate file Drawer 83, File D, Number 18", or "Box

3, Smith Folder".

MARITAL_STATUS:= {Size=1:20}
[D | S | W | _<TEXT>]
The marital status at the time of the associated event. Status values are:

D = Single but legally Divorced at time of event.
M = Married at time of event.
S = Single, never married at time of event.
W = Single because of the death of a spouse.
_ = If other information about marital status is to be shown add the appropriate text preceded by an underscore "_".

MEDIA_TYPE:= {Size=1:15}
[AUDIO | BOOK | CARD | ELECTRONIC | FICHE | FILM | MAGAZINE | MANUSCRIPT | MAP | NEWSPAPER | PHOTO | TOMBSTONE | VIDEO]
A code, selected from one of the media classifications choices above that indicates the type of material in which the referenced source is stored.

MONTH:= {Size=3:3}
[JAN | FEB | MAR | APR | MAY | JUN |
JUL | AUG | SEP | OCT | NOV | DEC]
A month name abbreviation selected from the choices above, used in forming dates.

NAME_OF_SOURCE_DATA:= {Size=1:90}
The name of the electronic data source that was used to obtain the data in this transmission. For example, the data may have been obtained from a CD-ROM disc that was named "U.S. 1880 CENSUS CD-ROM vol. 13."

NAME_OF_VESSEL:= {Size=1:90}
A name of the ship, air ship, or commercial vehicle used for travel, immigration, emigration, etc.

NATIONALITY:= {Size=1:90}
The person's national origin in common usage. Examples: Irish, Native American, Swede, and so forth.

NATIONAL_ID_NUMBER:= {Size=1:30}
A nationally-controlled number assigned to an individual. Commonly known national numbers should be assigned their own tag, such as SSN for U.S. Social Security Number. The use of the IDNO tag requires a subordinate TYPE tag to identify what kind of number is being stored. For example:

n IDNO 43-456-1899
+1 TYPE Canadian Health Registration

NEW_TAG:= {Size=3:15}

A user defined tag that is contained in the GEDCOM current transmission. This tag must be
defined within the SCHEMA context in the HEADer record and its name must begin with an
underscore (_). The SCHEMA context defines the data associated with this new tag. (See tags
LABL, DEFN, and ISA).

NULL:= {Size=0:0}
convention that indicates the absence of any characters in the value including
A the null character (0x00) which is prohibited.

OCCUPATION:= {Size=1:90}
The kind of activity that an individual does for a job, profession, or principal activity.

OLD_TAG_1:= {Size=3:15}
    This is any tag defined by the GEDCOM standard and is used in the SCHEMA context of the
    HEADer record to show the context in which a new user defined tag is being used. This tag
    always represents a tag which was used at level 0.

OLD_TAG_2:= {Size=3:15}
    This is any tag defined by the GEDCOM standard and is used in the SCHEMA context of the
    HEADer record to show the context in which a new user defined tag is being used. Old_TAG_2
    represents any tag at any level between level 1 and the level in which the new user defined tag
    resides. For example,

```
   n SCHEMA
    +1 INDI  (zero level)
       +2 BURI
         +3 PLAC
                +4 CEME
                  +5 _PLOT (new user tag)
```

ORD_BY_PATRON_CODE:= {Size=1:1}
    [Y | N]
    A code that identifies whether the patron will provide proxies for the cleared ordinances
    specified by the associated tag.
       Y = Patron will provide proxies for the associated cleared ordinance.
       N = Temple is to provide proxies for the associated cleared ordinance.

ORIGINATOR_NAME:= {Size=1:120}
    [<PERSONAL_NAME> | <CORPORATE_NAME>]
    The name of the person or organization that created this source.

ORIGINATOR_TYPE:= {Size=3:15}
    [AUTHOR | COMPILER | TRANSCRIBER | ABSTRACTOR | EDITOR |
      INFORMANT | INTERVIEWER | GOVERNMENT | BUSINESS | ORGANIZATION]
    A classification of the type of the person or entity that created this source.

PAGE_DESCRIPTION:= {Size=1:90}
    A field that identifies the page within the source. This may be a page number range, a specific
    page number, or another way of defining how to find the specified information within the
    source.

PERIODICAL_ISSUE_NUMBER:= {Size=1:90}
    The number or description of the specific periodical publication.

PERMANENT_RECORD_FILE_NUMBER:= {Size=1:18}
    <REGISTERED_RESOURCE_IDENTIFIER>:<RECORD_IDENTIFIER>
    The record number that uniquely identifies this record within a registered network resource.
    The number will be usable as a cross-reference pointer. The use of the colon (:) is reserved to

indicate the separation of the 'registered resource identifier'(precedes the colon) and the unique
'record identifier' within that resource (follows the colon). In cases where the colon is used,
implementations that check pointers should not expect to find a matching cross reference
identifier in the transmission but would find them in the indicated database within a network.
Making resource files available to a public network is a future implementation.

```
PERSONAL_NAME:=                                          {Size=1:120}
    [
     <TEXT> |
    /<TEXT>/ |
```

```
     <TEXT> /<TEXT>/ |
    /<TEXT>/ <TEXT> |
     <TEXT> /<TEXT>/ <TEXT>
    ]
```
    The surname of an individual, if known, is enclosed between two slash (/) characters. The order
    of the name parts should be the order that the person would customarily have used when giving
    it to a recorder. If part of name is illegible, that part is indicated by ... (ellipses).

    Examples:
       William Lee
       /Parry/
       William Lee /Parry/
       William /Lee/ Parry
       William Lee /Pa.../

PHONE_NUMBER:= {Size=1:25}
    A phone number.

PHYSICAL_DESCRIPTION:= {Size=1:247}
    A comma delimited, unstructured list of the attributes that describe the physical characteristics of
    a person, place, or object.
    Example:
       1 DSCR Hair Brown, Eyes Brown, Height 5 ft 8 in

PLACE_VALUE:= {Size=1:120}
```
    [
    <TEXT> |
    <TEXT>, <PLACE_VALUE>
    ]
```
    The jurisdictional name of the place where the event took place. Jurisdictions are separated by
    commas, that is, town, county, state or village, parish, country. Receiving systems cannot
    assume that the nth locality position is necessarily a specific level of jurisdiction. Some systems
    may include a PLAC context in the HEADer record which will specify the jurisdictional levels
    to the place names. Missing intermediate jurisdictions is represented by adjacent placeholder
    commas. If FORM value within the PLACe context of the HEADer record is present, then all
    levels of jurisdiction must be accounted in this way. For example if the following was included
    in the header record:
```
       0 HEAD
        1 PLAC
          2 FORM city, county, state, country
```
    Then each place name would be expected to account for the four levels by using appropriately
    placed commas.

    A FORM tag showing a change to this default assumption shown in the HEADer record can be

used subordinate to an individual place structure to show the variant jurisdictional levels.

A place of origin that is not necessarily a birth place is shown by preceding the place name with
the word "of." Missing or illegible characters within a place name are indicated by ...
(ellipses).

POSSESSIONS:= {Size=1:247}
A list of possessions (real estate or other property) belonging to this individual, separated by
commas.

PRODUCT_NAME:= {Size=1:90}
The name of the software product that produced this transmission.

PUBLICATION_DATE:= {Size=1:90}
    <DATE_REGULAR>
    The date this source was published or compiled.

PUBLICATION_EDITION:= {Size=1:90}
    A description of the specific version of the publication which is being referenced.

PUBLICATION_NAME:= {Size=1:90}
    The name of a publication such as a book, pamphlet, periodical, newspaper, or other
    monographic publication.

PUBLICATION_PLACE:= {Size=1:120}
    <PLACE_VALUE>
    The name of the place (city, state) where an item was published or the location of the publisher's
    main office.

PUBLICATION_TYPE:= {Size=4:12}
 [BOOK | PERIODICAL | NEWSPAPER | UNPUBLISHED | ELECTRONIC]

PUBLISHER_NAME:= {Size=1:90}
    The name of the publisher of the referenced publication.

QUALITY_OF_DATA:= {Size=1:1, Type=NUMBER}
    [0 | 1 | 2 | 3]
    The submitter's assessment of the reliability of the information for the associated fact:
      0 = Unreliable evidence or data was estimated.
      1 = Direct or primary evidence with some question of reliability
          or potential for bias for example, an autobiography).
      2 = Secondary evidence.
      3 = Direct and primary evidence used, or by dominance of the evidence.

RECORD_IDENTIFIER:= {Size=1:18}
    An identification number assigned to each record within a specific data base. If this identifier is
    associated with a preceding colon (:), then it is the record number within the registered resource
    identified by the data that precedes the (:) else it is a specific reference to a record within the
    current database if no registered resource identifier precedes the (:). If the colon is not present
    it is the identification of a record within the current GEDCOM transmission file.

REGISTERED_RESOURCE_IDENTIFIER:= {Size=1:18}
    This is an identifier assigned to a resource data base which is available through access to an
    available network. (Future plans.)

RELATIONSHIP_ROLE_TAG:= {Size=1:90}
   [BROT | CHIL | FATH | HEIR | HUSB | MOTH | PARE | PHUS | PWIF | SIBL |
     SIST | WIFE]

RELIGIOUS_AFFILIATION:= {Size=1:90}
A name of the religion with which this person or record was affiliated.

RELIGIOUS_NAME:= {Size=1:120}
A name given to a person to be used in connection with a religion.

REPOSITORY_NAME:= {Size=1:90}
The official name of the archive in which the stated source material is stored.

ROLE_DESCRIPTOR:= {Size=1:90}
A word or phrase that identifies the role of each person in the event being described. This should be the same word or phrase, and in the same language, that the recorder used to define the role in the actual record. This is used in connection with the ROLE_TAG.

ROLE_TAG:= {Size=1:20}
[BUYR | CHIL | FATH | GODP | HDOH | HDOG | HEIR | HFAT | HMOT | HUSB | INFT | LEGA | MEMBER| MOTH | OFFI | PARE | PHUS | PWIF | RECO | REL | ROLE | SELR | TXPY | WFAT | WIFE | WITN | WMOT | INDI]
A tag that indicates the role of the individuals mentioned in a source event record. If the above list does not include the role being cited, use the ROLE_TAG followed by a ROLE_DESCRIPTOR to define the role. (See appendix A for the definition of these tags and Appendix B for additional ROLEs which have been proposed as GEDCOM tags). Names of individuals mentioned in the event but their role was not mentioned, should be identified by using the INDI role tag. Any associations between others of known roles and this individual can be shown by using the ASSOciation pointer.

SCHOLASTIC_ACHIEVEMENT:= {Size=1:247}
A description of a scholastic or educational achievement or pursuit.

SEARCH_STATUS:= {Size=1:90}
[ACTIVE | FOUND | NO | ORDERED | PLANNED | PROVED]
A field that shows the research status with respect to the cited source. Where:
ACTIVE = This source is currently being searched.
FOUND = Part or all of the expected information has been found.
NO = This source is no longer in use because the information could not be found.
ORDERD = A request for this source has been sent to the Repository.
PLANNED= This source is to be examined.
PROVED = This source has been reconciled with the data in this record.

SEARCH_STATUS_DATE:= {Size=1:90}
<DATE_EXACT>
The date on which the current SEARCH_STATUS was set.

SERIES_VOLUME_DESCRIPTION:= {Size=1:247}
A description of a successive publication. The description should identify the timing of the
publication, for example, Spring, Summer, Fall, Winter. The description should also state the
volume number of periodicals or of multi-volume books.

SEX_VALUE:= {Size=1:7}
A code that indicates the sex of the individual:
M = Male
F = Female

SIGNATURE_INFO:= {Size=1:90}
A description of the capabilities of this person to sign documents, the symbol used in signing,
did they know how to sign, did they use a model to produce a signature.

SITE_NAME:= {Size=1:90}
The name of a specific site associated with an event, address, or place.

SOCIAL_SECURITY_NUMBER:= {Size=9:11}
A social security identification number assigned to this person.

SOURCE_CALL_NUMBER:= {Size=1:90}
An identification number used to file and retrieve items from the holdings of a repository.

SOURCE_CLASS_DESCRIPTOR:= {Size=1:25}

A descriptive word or phrase that classifies the type of source being cited. This descriptor is used only when none of the classifications defined under the <SOURCE_CLASSIFICATION_CODE> fit this source type. Systems that display data from the GEDCOM form should be able to display the descriptor value in their screen or printed output.

SOURCE_CLASSIFICATION_CODE:= {Size=7:90}

[BOOK | CENSUS | CHURCH | COURT | HISTORY | INTERVIEW | JOURNAL | LAND | LETTER | MILITARY | NEWSPAPER | PERIODICAL | PERSONAL | RECITED | TRADITION | VITAL | OTHER!<SOURCE_CLASS_DESCRIPTOR>]

A code which classifies the source which contained the evidence data. Where:

BOOK = A published work including biographies and genealogies.
CENSUS = A official census.
CHURCH = A church record.
COURT = A record from a court, both criminal and civil.
HISTORY = A published historical account.
INTERVIEW = An interview.
JOURNAL = A personal record or diary.
LAND = A record of land holdings or transactions, both federal and state.
LETTER = A letter or other written communication.
MILITARY = A military record.
NEWSPAPER = A newspaper account.
PERIODICAL = A work that is published at certain intervals, such as monthly, quarterly, or yearly.
PERSONAL = A source that was compiled from accounts given from a person's memory.
RECITED = A recited genealogy, such as a tribal or clan genealogy.
TRADITION = A source that was compiled from accounts communicated by word-of-mouth from one generation to another.
VITAL = A vital record created by a government agency of vital records such as births, marriages, and divorces.
OTHER! = Other sources can be identified by using (OTHER!) followed by <SOURCE_CLASS_DESCRIPTOR>.

Systems that display data from the GEDCOM form should be able to display the descriptor value in their output.

SOURCE_FIDELITY_CODE:= {Size=7:17}

[ORIGINAL | PHOTOCOPY | TRANSCRIPT | EXTRACT]

A code is a selected from the above choices that provides an assessment of the fidelity (the exactness) of this source material.

ORIGINAL = This source is the original record being cited.

PHOTOCOPY = This source is a photocopy of the original record.
TRANSCRIPT = This source is a complete transcription of the original record.
EXTRACT = This source is an abridgement, subset, and/or interpretation.

SOURCE_FILM_NUMBER:= {Size=1:15}
A unique number assigned by the repository to identify the specific microfilm containing
information about the event of interest.

SOURCE_JURISDICTION_PLACE:= {Size=1:120}
<PLACE_VALUE>
The name of the lowest jurisdiction that encompasses all lower-level places named in this source.
For example, "Franklin, Idaho" would be used as a source jurisdiction place for events
occurring in the various towns within Franklin county but "Idaho" would be used as a source

jurisdiction place if the source records referenced other counties in Idaho besides Franklin
county.

SOURCE_TEXT:= {Size=1:247}
<TEXT>
A verbatim copy of any description contained within the source. This indicates notes that are
actually contained in the source document, not the submitter's opinion about the source.

SUBMITTER_TEXT:= {Size=1:247}
Comments or opinions from the submitter.

SYSTEM_NAME:= {Size=1:20}
The name of the sending or receiving GEDCOM-compatible product. The system name for the
sending system was obtained when the product was registered as a GEDCOM-compatible
product. All GEDCOM transmissions must be so identified. The system name used with the
DESTination tag should be:

* "ANSTFILE" when sending to the ancestral file.
* "TempleReady" when submitting for temple ordinances.
* The same DESTination system name as was used with the SOURce tag is used when the
destination is unknown.

TEMPLE_VALUE:= {Size=5:5}
A 5-character abbreviation of the temple in which LDS temple ordinances are performed.
(Contact the GEDCOM Coordinator for a table of valid abbreviations)

TEXT:= {Size=1:247}
A string composed of any valid character or string of characters in the GEDCOM character set.

TIME_PERIOD:= {Size=1:90}
[FROM <DATE_REGULAR> TO <DATE_REGULAR> |
FROM <DATE_REGULAR> |
TO <DATE_REGULAR>]
The range in time of an event or set of events, inclusive. The choice FROM
<DATE_REGULAR> indicates a range from a beginning date to an indefinite future date.
This differs from the date range notation in that the date range is to indicate that an event took
place on a given date within the range. The time period date indicates that the event or events
cover or happened over the time period specified.

The choice TO <DATE_REGULAR> indicates from an indefinite beginning to a specified
date.
Examples:
FROM 1904 to 1915
FROM 1904
TO 1905

TIME_VALUE:= {Size=1:10}
[hh:mm:ss.fs]
The time of a specific event, usually a computer-timed event, where:
hh = hours on a 24 hour clock
mm = minutes
ss = seconds, (optional)
fs = decimal fraction of a second, (optional)

TRANSMISSION_DATE:= {Size=10:11}
<DATE_EXACT>
The date that this transmission was created.

TYPE_OF:= {Size=1:20}
A user-defined number or text that the submitter uses to identify this record. For instance, it
may be a record number within the submitter's automated or manual system, or it may be a page
and position number on a pedigree chart.

USER_TAG_DEFINITION:=
A formal description of the user defined tag. This description can be used by the receiving
system to give meaning to the user defined tags. (See Chapter 2, User Defined Tags section.)

VERSION_NUMBER:= {Size=1:15}
An identifier that represents the version level assigned to the associated product. It is defined
and changed by the creators of the product.

XREF:= {Size=1:15}
Either a pointer or a cross-reference identifier. If this element appears before the tag in a
GEDCOM-line, then it is a cross-reference identifier. If it appears after the tag in a GEDCOM-
line, then it is a pointer. The method of delimiting a pointer or cross-reference identifier is to
enclose the pointer or cross reference identifier within at-signs (@), for example, @I123@. A
XREF may not begin with a number sign (#). This is to avoid confusion with an escape
sequence prefix (@#). The use of a colon (:) in the XREF is reserved for creating future
network cross-references.

XREF:ANY:= {Size=1:15}
<XREF>
A universal pointer. It may point to any other cross-reference identifier type.

XREF:EVEN:= {Size=1:15}
<XREF>
A pointer to or a cross reference identifier of a source event record.

XREF:FACT:= {Size=1:15}
<XREF>

A pointer to or a cross reference identifier of a facts record.

XREF:FAM:= {Size=1:15}
<XREF>
A pointer to or a cross reference identifier of a family record.

XREF:INDI:= {Size=1:15}
<XREF>
A pointer to or a cross reference identifier of an individual record.

XREF:NOTE:= {Size=1:15}
<XREF>
A pointer to or a cross reference identifier of a note record.

XREF:REPO:= {Size=1:15}
<XREF>
Either a pointer to a REPOsitory, a SUBMitter, or an INDIvidual record, or a cross reference
identifier of a repository record.

XREF:REC!ID:= {Size=1:15}
[<FILE:REC!ID> | <REC!ID> | <!ID>]
Enclosed in at-signs (@), this is a pointer to a context within a record. Normally the pointer
will only be used to point to role contexts within the current event record but the principle
should allow the reference to a context within a specific record within a specific file. The

following are valid ways of representing this pointer:

@FILE:REC!ID@ = A pointer to a specific context <!ID>, within a specific record
<REC> within a specific file <FILE:>, that logically replaces the
context containing the cross reference pointer. (Future.)

@REC!ID@ = A pointer to a specific context <!ID> within a specific record within the
current GEDCOM transmission.

not valid:

@!ID@ = A pointer to a specific context <!ID> within the current record of this
GEDCOM transmission must also contain the record level pointer, such
as @I13!3@.

XREF:SOUR:= {Size=1:15}
<XREF>
Either a pointer to a SOURce, a SUBMitter, or an INDIvidual record, or a cross reference
identifier of a source record.

XREF:SUBM:= {Size=1:15}
<XREF>
Either a pointer to a SUBMitter, or an INDIvidual record, or a cross reference identifier of a
submitter record.

YEAR:= {Size=3:4, Type=NUMBER}
A numeric representation of the calendar year in which an event occurred.

YEAR_ALTERNATIVE:= {Size=1:1, Type=NUMBER}
A year modifier which shows the possible date alternatives for pre-1752 date brought about by a
calendar change, for example, 15 Dec 1752/3.

COMPATIBILITY WITH OTHER GEDCOM VERSIONS

Products based on GEDCOM 5.3 are generally compatible with products based on prior GEDCOM
versions. However, there are four issues related to specific products that introduce incompatibilities
which can be accommodated by programming to handle the information in both the standard and the
non-standard way. Compatibility with prior implementations may be maintained by doing the
following:

1. Treat a TITL tag found at level 0 as if it were a SOUR record, including its subordinate
structure. Roots III points from a SOUR structure in an INDI record to a 0 TITL source

record in this manner. Likewise, the TITL tag must be used instead of the SOUR tag in the level 0 SOUR record to send source information to Roots III.

2. The structure for LDS sealing of child to parents was changed in the standard from the FAM-CHIL-SLGC structure to the INDI-FAMC-SLGC structure to conform with the more natural access path to this information. PAF 2.1 reads the sealing date in the FAM-CHIL-SLGC structure, while other products read it in the INDI-FAMC-SLGC structure. To accommodate all implementations, systems handling the LDS ordinance events should look for the child sealing information in either place. Systems should also write the child sealing information in both structures when preparing a transmission. Other child events were also moved to the INDI-FAMC structure, namely ADOPtion, which should receive the same treatment.

3. When an individual has multiple names, GEDCOM 5.x requires listing the preferred instance first, followed by less-preferred names. However, PAF and other products take only the last instance during a transmission, causing the preferred name to be dropped when more than one name is present. The same happens with all multiple-instance tags where only one instance is

received. When writing to GEDCOM 4.0 (or earlier) compatible systems you should only
output the preferred name under the name tag and export the also-known-as name in a note
field.

We anticipate a future change to allow use of indentation to make GEDCOM files easier to read.
To make this transition easier, beginning with GEDCOM 5.3, leading white space in a GEDCOM
line should be handled by receiving systems by ignoring it. Indentation should NOT be transmitted
in GEDCOM files until this change is established in a future version of The GEDCOM Standard.

PACKAGING THE GEDCOM TRANSMISSION FILE

The GEDCOM transmission is normally created on a DOS or Macintosh compatible diskette. The
DOS filename extension is (.GED). Macintosh filenames do not use file extensions.

When the GEDCOM file is too large to fit on a single diskette, the file is divided after any whole-
line (last character is the terminator), and the DOS filename extension becomes (G##) where (##) is
(00) for the second disk, (01) for the third, and so forth. For Macintosh filenames, append the two
digits to the subsequent filenames in parentheses. (See example below.) This allows the receiving
software to ensure that disks are read in the correct sequence.

Given that the user-supplied portion of the file name is SMITH, then the complete filenames for a
three-disk transmission would be:

Disk	DOS Filename	Macintosh Filename
1	SMITH.GED	SMITH
2	SMITH.G00	SMITH(00)
3	SMITH.G01	SMITH(01)

The required GEDCOM HEADer record appears only on the first disk and the required TRLR
(trailer) record appears only on the last disk and must be followed by the terminator.

USER DEFINED TAGS
Data stored in different systems within a user defined context will not be easy to share between
other systems. GEDCOM defines a schema that can be included within the HEADer record which
will give receiving systems the information to assist them in interpreting the user defined data.
Utmost care should be taken when defining User tags. The primary use would be for transmitting
data between the same software driven system, system developers are encouraged to find ways of

supporting user defined tags, but GEDCOM only provides a way to express the data, it usage is left
to the receiving software.

This schema is designed to show:

a. The context within which the new tag appears in the records.
b. The name of the new tag, which must start with an underscore (_).
c. The definition of the new tag.
d. The label or long name of the new tag, if different from the tag name.
e. The kind of data that this new tag represents in terms of a predefined standard GEDCOM tag. For Example, if HOSPital was being defined as a user tag, then we would use the SITE tag to show that hospital is a kind of SITE.

In the Sample Lineage-linked GEDCOM Transmission example below is the SCHEMA required for
defining a new user defined tag "_HOSP" which is intended to show the name of the name of the
hospital where a birth took place.

Included in the schema context is:

1. The LABL tag to define a longer tag name that can be used as a field label.
2. The DEFN tag which allows sharing of the definition of the new tag.

3. The ISA tag to show that this tag is a kind of another standardized tag. In this case
_HOSPital is a kind of SITE.

ESCAPE SEQUENCE FORMAT FOR THE LINEAGE-LINKED FORM

The Lineage-linked form utilizes the escape sequence feature provided in the GEDCOM grammar in
the following way:

* An escape sequence in the HEADer structure invokes variant processing for the entire
transmission.

* An escape sequence that appears in subsequent structures affect only the line on which the
escape sequence appears unless that line has subordinate CONTinuation or CONCatenation
lines. In this case the variant processing applies to the subordinate CONTinuation and
CONCatenation substructure lines as well.

* The form of the escape sequence is @# escape_type_code escape_text @ where the
escape_type_code indicates that:

A = A auxillary data format or processing is being referenced. Auxillary data formats include such forms as images, sound, or other data requiring auxillary processing. (See primitive element ESCAPE_TO_AUXILLARY_PROCESSING above in this chapter).

C = Character set processing is being invoked.

D = Date processing for special calendar is being invoked. (see primitive element CALENDAR_ESCAPE_SEQUENCE above in this chapter).

The escape_text specifies the specific processing to be done within that particular type, for
example, @#DJULIAN@ indicates julian date processing.

SAMPLE LINEAGE-LINKED GEDCOM TRANSMISSION

The example below shows how some of these value types appear in a valid GEDCOM Lineage-
linked transmission. The example is a sample transmission of genealogical information about three
individuals who are members of the same family--husband, wife, and child. In the example,
"Joe/Williams/" is the value specified by the tag NAME under the INDI tag for the record (@3@).
Other values in other lines, such as the birth date and place, provide additional information about

Joe Williams. The value (@4@) specified by the FAMC tag is a pointer to the FAMily record
(@4@) of which Joe Williams is a child. Included also in this transmission example are three other
record types: a source record, a submitter record, and a repository record. These records are
pointed to from within other records in the transmission. This shows how pointer values can be
used in creating the GEDCOM Lineage-linked form.

Example: (Indentation is for readability only.)

```
 0 HEAD
   1 SOUR PAF
     2 VERS 2.1
   1 DEST ANSTFILE
   1 SUBM @5@
   1 GEDC
     2 VERS 5.2
   1 SCHEMA
     2 INDI
         3 BIRT
           4 _HOSP
```

```
            5 LABL HOSPITAL
            5 DEFN The name of a hospital
            5 ISA  SITE
0 @1@ INDI
  1 NAME Robert Eugene/Williams/
  1 SEX M
  1 BIRT
    2 DATE 02 OCT 1822
    2 PLAC Weston, Madison, Connecticut
    2 _HOSP St. Marks
    2 SOUR @6@
  1 DEAT
    2 DATE 14 APR 1905
    2 PLAC Stamford, Fairfield, CT
    2 QUAY 2
  1 BURI
    2 PLAC Stamford, CT
      3 CEME Spring Hill Cemetery
  1 OCCU Publisher
  1 FAMS @4@
0 @2@ INDI
  1 NAME Mary Ann/Wilson/
  1 SEX F
  1 BIRT
    2 DATE BEF 1828
    2 PLAC Connecticut
  1 FAMS @4@
0 @3@ INDI
  1 NAME Joe/Williams/
  1 SEX M
  1 BIRT
    2 DATE 11 JUN 1861
    2 PLAC Idaho Falls, Bonneville, Idaho
  1 FAMC @4@
0 @4@ FAM
  1 HUSB @1@
  1 WIFE @2@
  1 CHIL @3@
  1 MARR
    2 DATE DEC 1859
0 @5@ SUBM
  1 NAME Reldon /Poulson/
  1 ADDR 1900 43rd Street West
    2 CONT Billings, MT 68051
    2 PHON (406) 555-1232
0 @6@ SOUR
  1 TYPE VITAL
  1 EVEN BIRT
  1 TITL County Birth Records
  1 PERI FROM 1820 TO 1825
  1 PLAC ,Madison, Connecticut
  1 RECO CIVIL
  1 FIDE PHOTOCOPY
  1 REPO @7@
    2 MEDI FILM
    2 CALN 13B-1234.01
```

```
0 @7@ REPO
  1 NAME Family History Library
  1 ADDR 35 N West Temple Street
    2 CONT Salt Lake City, UT 84150
0 TRLR
```

SAMPLE EVENT_RECORD

This example shows how the Evidence_Record format might be used to store an extraction of a
christening record:

```
0 @EV13@ EVEN
 1 TYPE CHR
    2 DATE 17 NOV 1830
    2 PLAC Littlehampton, West Sussex, England
       3 ADDR 9 Chiltern Close
         4 CONT East Preston
    2 @EV13!1@ CHIL
       3 NAME Jason \Wilde\
       3 AGE 4 yrs
    2 @EV13!2@ MOTH
       3 NAME Wilma \Wilson\
       3 BIRT
         4 DATE 15 MAY 1810
         4 PLAC Nottingham, England
    2  @EV13!3@ FATH
       3 NAME William \Wilde\
       3 BIRT
         4 DATE 15 OCT 1805
         4 PLAC Nottingham, England
       3 ASSO @EV13!4@
         4 TYPE BROTHER
    2 @EV13!4@ GODF
       3 NAME David \Wilde\
```

Chapter 3

USING CHARACTER SETS IN GEDCOM

INTRODUCTION

GEDCOM needs to be designed to accommodate different character sets to facilitate the sharing of genealogical data in different languages. In order to minimize the number of differing standards to accomplish this, we have chosen to have each system convert their usage to ANSEL and eventually UNICODE. In January of 1991 a Unicode Consortium was founded to promote the use of the Unicode standard which accommodates all characters in one character set (see the section on Unicode below). Unicode Consortium has agreed with the ISO 10646 standard to merge and Unicode will be a subset of the ISO 10646 international character encoding standard. The difficulty is in handling the two character code sequences. Therefore, until the multi-byte handling becomes more common, the usage of ANSEL to represent the latin-based international characters will be the standard.

The GEDCOM specification does not address the implementation methods for multilingual processing, such as keyboard arrangements, sorting sequences, or character and graphic representations (font styles, proportional spacing, and so forth) on the CRT or printers, however, Unicode standard has defined formatting characters which will indicate the direction of the text presentation as well as other text formatting character codes.

Most of the genealogy systems developed so far utilize either ASCII or ANSEL, or both. ANSEL accommodates the set of Latin-based languages, as explained below.

8-Bit ANSEL

The 8-Bit ANSEL (American National Standard for Extended Latin Alphabet Coded Character Set for Bibliographic Use, Z39.47, 1985 copyright) is the default character set for GEDCOM. It is used for all transmissions of information unless another character set is specified. The use of this character set standard makes it possible to preserve the full integrity of the language by providing a method of using the standard ASCII character set and supplementing it with both non-spacing character modifiers (diacritic) as well as spacing special characters. Non-spacing means that the diacritic is printed without advancing the device's print position. The character being modified is then printed in the same position, resulting in a combined image of both the character and the

diacritic(s). The storage of ANSEL requires storing the non-spacing graphic character(s) preceding
the ASCII character that the diacritic is to modify. The ANSEL standard specifies an extended 8-bit
configuration (above 128) to represent the spacing and non-spacing graphic characters that make up
most of the Latin based languages. ANSEL is a super-set of ASCII. The standard ASCII
characters including the control characters are preserved.

ANSEL is known by two other names: (1) ANSI Z39.47-1985) and (2) the American Library
Association character set, used in library systems worldwide, including the MARC (MAchine-
Readable Catalog) format.

A description of the codes for the ANSEL character set has been reproduced with permission and is
included with the printed version of The GEDCOM Standard. The description of ANSEL codes is
not included in the electronic version. This description may be purchased from the American
National Standards Institute at 1430 Broadway, New York, N.Y. 10018.
The description of the ANSEL character set standard includes the following:

* An 8-Bit Code Table showing the ASCII and extended ANSEL codes
* An explanation or legend of these codes
* A chart that identifies the ANSEL Non-spacing Graphic Characters
* A chart that identifies the ASCII Control Characters
* A chart that identifies the ASCII Graphic Characters

Character-set codes 0 through 127 are the same for 8-Bit ANSEL and 8-Bit ASCII (USA version--

ANSI 8-Bit). Character-set codes 128 through 255 are unique to the ANSEL character set.

ASCII (USA version)

When there isn't a need for diacritics or other special characters, and if you are not transmitting binary data, you will find it convenient to use ASCII (8-bit USA version) if your computer already supports it. This is a standard of the American National Standards Institute (ANSI). Most of the basic printable characters of ANSEL and ASCII (USA version--ANSI 8-Bit) are identical.

Binary Character Set

Binary formats for representing photographs and other bit-mapped graphics should use the escape sequence "escape_to_supplementary_processing" for linking supplementary files to the GEDCOM context (see chapter 2).

UNICODE (ISO 10646)

The Unicode standard is a new character code designed to encode text for storage in computer files. It is a subset of the upcoming ISO 10646 standard. The design of the Unicode standard is based on the simplicity and consistency of today's prevalent character code set, extended ASCII code set, but goes far beyond ASCII's limited ability to encode only the Latin alphabet: the Unicode encoding provides the capacity to encode all of the characters used for written languages throughout the world. In order to accommodate the many thousands of characters used in the international text, the Unicode standard uses a 16-bit code set instead of extended ASCII's 8-bit code set. This expansion provides codes for more than 65,000 characters. The Unicode standard assigns each character a unique 16-bit value, and does not use complex modes or escape codes to specify modified characters or special cases. The text representation of the Unicode 16-bit numbers is U+0041 which is assigned to the letter A, 65 decimal. The Unicode standard includes the Latin alphabet used for English, the Cyrillic alphabet used for Russian, the Greek, Hebrew, and Arabic alphabets. Other alphabets used in countries across Europe, Africa, the Indian subcontinent, and Asia, such as Japanese Kana, Korean Hangul, and Chinese Bopomofo are included. The largest part of the Unicode standard is devoted to thousands of unified character codes for Chinese, Japanese, and Korean ideographs. (See "The Unicode standard", vol. 1 and 2, published by Addison-Wesley Publishing, for character code standards.)

The Unicode character set environment, which contains a character set for all languages, minimizes
previous GEDCOM requirements to provide escape_sequences for moving from one character set to
another. If the Unicode environment is used to produce a GEDCOM transmission, the header
record would also be in Unicode, requiring receiving systems to determine whether the transmission
is Unicode or ASCII before they could interpret the GEDCOM header. This would be done by
reading the first two bytes of the transmission. If the first two bytes are 0x30 and 0x20 then the
transmission will be in either ASCII or ANSEL as determined by the header record. If the first two
bytes are 0x30 and 0x00 then the transmission should be processed as a Unicode transmission.
(Different platforms may reverse the position of the null byte, in which case the test would be for
0x00 and 0x30.)

How to change character sets

The character set for an entire transmission is specified in the character-set line of the header
record.

The example below shows the specification in the header record.

Example:

Lvl	Tag	Value
0	HEAD	
1	SOUR	PAF
2	VERS	2.1
1	DEST	ANSTFILE
1	CHAR	ANSEL

The character-set change remains in effect until the TRLR record is encountered at the end of the
transmission.

The lineage_linked form no longer makes use of the character escape_sequence to change a
character set context inside of the transmission. Unicode does not require shifting from character
set to character set and we should encourage its use for multi-language support.

For more information about character sets, see the following:

* Extended Latin Alphabet Coded Character Set for Bibliographic Use. American National
Standards (ANSI), Z39.47, 1985.

* "8-Bit ASCII--Structure and Rules." American National Standards (ANSI) X3.134.1-198x.

* "7-Bit and 8-Bit ASCII Supplemental Multilingual Graphic Character Set (ASCII Multilingual Set)" (manuscript). American National Standards (ANSI), X3.134.2-198x.

* "The Unicode standard", vol. 1 and 2, published by Addison-Wesley Publishing.

Appendix A
LINEAGE-LINKED GEDCOM TAG DEFINITION

Introduction

Appendix A is a glossary of the tags approved for use with Lineage-linked GEDCOM. (See chapter 2 for an example of the tags in context that describes the Lineage-linked structure.) Every tag must be used within the context shown to ensure that all information transmitted by means of GEDCOM is uniformly identified.

The tags vary in type, depending on their role or use in a transmission. They are used to identify individuals, families, names, dates, places, events, roles, sex, sources, relationships, control codes and other kinds of data for computers, computer programs, and computer systems.

Generally, the definition for each tag is broad enough to cover all uses of the tag. Any new tag needed to extend the Lineage-linked form can be used for by a system that generates GEDCOM output may be used and will not violate the Lineage-linked GEDCOM standard as long as the context for the Lineage-linked GEDCOM grammar is not violated. System builders using new tags should register them and their definitions with the GEDCOM Coordinator at the address listed on the title page of this document. The Coordinator will evaluate the feasibility of including them as a part of the next release of the standard. Suggestions and proposed additions are welcome.

Lineage-Linked GEDCOM Tag Definitions

This section provides the definition of the standardized GEDCOM tags and shows the formal name of the tag inside of {braces}.

ADDR {ADDRESS}:=
The contemporary place, usually required for postal purposes, of an individual, a submitter of information, a repository, a business, a school, or a company.

ADOP {ADOPTION}:=
The event of a legal creation of the child-parent relationship that does not exist biologically.

AFN {AFN}:=
A unique permanent record file number of an individual record stored in the Ancestral File.

AGE {AGE}:=
The age of the individual at the time an event occurred, or the age listed in the document.

AGNC {AGENCY}:=
The name of the branch of government.

ALIA {ALIAS}:=
A pointer to which indicates that another record is suspected of being the same person.
When the suspicions are confirmed, drop the alias line, combine all data into one record, and
delete the other record. Alias should NOT be used to record alternate names for the same
person. (See Name tag definition.)

ANCI {ANCES_INTEREST}:=
Indicates an individual in which the submitter has interest in additional research for ancestors
of this individual. (See also DESI)

ANUL {ANNULMENT}:=
An event declaring a marriage void from the beginning (never existed).

ARVL {ARRIVAL}:=

An event declaring the arrival or reaching of a destination.

ASSO {ASSOCIATES}:=
Identifies friends, neighbors, or associates of an individual.

AUTH {AUTHOR}:=
The name of the individual who created or compiled information.

BAPL {BAPTISM-LDS}:=
The event of baptism performed at age eight or later by priesthood authority of The Church
of Jesus Christ of Latter-day Saints. (See also BAPM.)

BAPM {BAPTISM}:=
The event of baptism (not LDS), performed in infancy or later. (See also BAPL and CHR.

BARM {BAR_MITZVAH}:=
The ceremonial event held when a Jewish boy reaches age 13.

BASM {BAS_MITZVAH}:=
The ceremonial event held when a Jewish girl reaches age 12, also known as "Bat Mitzvah".

BIRT {BIRTH}:=
The event of entering into life.

BLES {BLESSING}:=
A religious event of bestowing divine care or intercession.

BROT {BROTHER}:=
A male sibling.

BURI {BURIAL}:=
The event of the proper disposing of the mortal remains of a deceased person.

BUYR {BUYER}:=
A person who purchased or purchases from another.

CALN {CALL_NUMBER}:=
The number used by a repository to identify the specific items in its collections.

CAST {CASTE}:=
The name of an individual's rank or status in society, based on racial or religious differences, or differences in wealth, inherited rank, profession, occupation, etc.

CAUS {CAUSE}:=
A description of the cause of the associated event or fact, such as the cause of death.

CEME {CEMETERY}:=
The name of the cemetery or other resting place where an individual is buried.

CENS {CENSUS}:=
The event of the periodic count of the population for a designated locality, such as a national
or state Census.

CHAN {CHANGE}:=
Indicates a change, correction, or modification. Typically used in connection with a DATE to

specify when a change in information occurred.

CHAR {CHARACTER}:=
An indicator of the character set used in writing this automated information.

CHIL {CHILD}:=
The natural, adopted, or sealed (LDS) child of a father and a mother.

CHR {CHRISTENING}:=
The religious event (not LDS) of baptizing and/or naming a child.

CHRA {ADULT_CHRISTNG}:=
The religious event (not LDS) of baptizing and/or naming an adult person.

CLAS {CLASSIFICATION}:=
A classification name given to identify objects because they posses a set of similar attributes
or characteristics.

CNTC {CONTACT_PERSON}:=
The name of a person that is listed as the contact person at an institution such as a
repository, college, business, etc.

CONC {CONCATENATION}:=
An indicator that the additional value information follows and is to be connected to the value
of the superior preceding line without a new line.

CONF {CONFIRMATION}:=
The religious event (not LDS) of conferring the gift of the Holy Ghost and, among
protestants, full church membership.

CONL {CONFIRMATION_L}:=
The religious event by which a person receives membership in The Church of Jesus Christ of
Latter-day Saints.

CONT {CONTINUED}:=
An indicator that additional value information follows and is to be connected with the value
of the superior preceding line as a new line.

COPR {COPYRIGHT}:=
A statement that accompanies data to protect it from unlawful duplication and distribution.

CORP {CORPORATE}:=
A name of an institution, agency, corporation, or company.

CPLR {COMPILER}:=
The name of the person that compiled writings of others.

DATA {DATA}:=

Pertaining to stored automated information.

DATE {DATE}:=
The time of an event in calendar days.

DEAT {DEATH}:=
The event when mortal life terminates.

DEFN {DEFINITION}:=
A textual description of something.

DESI {DESCENDANT_INT}:=
Indicates the submitter that has interest in research to identify additional descendants of this
individual. (See also ANCI.)

DEST {DESTINATION}:=
A system receiving data.

DIV {DIVORCE}:=
An event of dissolving a marriage through civil action.

DIVF {DIVORCE_FILED}:=
An event of filing for a divorce by a spouse.

DPRT {DEPARTURE}:=
An event declaring the departure or leaving for another destination.

DSCR {PHY_DESCRIPTION}:=
The physical characteristics of a person, place, or thing.

EDTR {EDITOR}:=
The name of a person who edited information.

EDUC {EDUCATION}:=
Indicates the education attained.

ENDL {ENDOWMENT}:=
A religious event where an endowment ordinance for an individual was performed by
priesthood authority in an LDS Temple.

ENGA {ENGAGEMENT}:=
An event of recording or announcing an agreement between two people to become married.

EMIG {EMIGRATION}:=
An event of leaving one's homeland with the intent of residing elsewhere.

EVEN {EVENT}:=
A noteworthy event related to an individual, a group, or an organization.

FAM {FAMILY}:=
Identifies a legal, common law, or other customary relationship of husband and wife and

their children, if any, or a family created by virtue of the birth of a child to its biological
father and mother.

FAMC {FAMILY_CHILD}:=
Identifies the family in which an individual appears as a child.

FAMS {FAMILY_SPOUSE}:=
Identifies the family in which an individual appears as a spouse.

FATH {FATHER}:=
Identifies the male parent in a family. In the Lineage-linked form this tag is used only in the
EVENT_RECORD role tag structure (See Chapter 2). Direct parent relationships are
represented using the HUSBand and WIFE tags as part of the FAMILY_RECORD.

FIDE {FIDELITY}:=
A description of the state of originality of the record to permit an assessment of the potential
for accuracy or errors due to the use of a copy of the record.

FILE {FILE}:=
An information storage place that is ordered and arranged for preservation and reference.

FILM {FILM_NUMBER}:=
An assigned, unique number used to identify a reel of film.

FORM {FORMAT}:=
An assigned name given to a consistent format in which information can be conveyed.

GEDC {GEDCOM}:=
Information about the use of GEDCOM in a transmission.

GODP {GODPARENT}:=
A sponsor at a religious rite (baptism).

GRAD {GRADUATION}:=
An event of awarding educational diplomas or degrees to individuals.

HDOH {HEAD_HOUSEHOLD}:=
Identifies a person whose role was recorded as "head of household" for an event such as a
census.

HEAD {HEADER}:=
Identifies information pertaining to an entire GEDCOM transmission.

HEIR {HEIR}:=
A role of an individual who inherited or is entitled to inherit an estate.

HFAT {HUSB_FATHER}:=
A role of an individual acting as the husband's father for a cited event.

HMOT {HUSB_MOTHER}:=
A role of an individual acting as the husband's mother for a cited event.

HUSB {HUSBAND}:=
An individual in the family role of a married man or father.

IDNO {IDENT_NUMBER}:=
A number assigned to identify a person within some significant external system.

IMMI {IMMIGRATION}:=
An event of entering into a new locality with the intent of residing there.

INDI {INDIVIDUAL}:=
A person.

INDX {INDEXED}:=
Specifies information about an index to simplify finding information in a source.

INFT {INFORMANT}:=
An individual who reported facts concerning an event.

INTV {INTERVIEWER}:=
The person who facilitated, recorded, and obtained information during an interview.

ISA {IS_A_KIND_OF}:=
Indicates the tag of an object of which this object inherits its characteristics from.

ISSUE {ISSUE}:=
An identifier used to differentiate one giving out from another, such as a number
differentiating one periodical publication from another.

ITEM {ITEM}:=
Refers to a unit within a set of things that belong together. The unit itself may be made up
of other objects but collectively they are referred to as an unit (item) of the set. A group of
papers filmed together under one header page is referred to as an item on a film.

LABL {LABEL}:=
A name assigned to a field or product which helps to identify it.

LANG {LANGUAGE}:=
The name of the language used in a communication or transmission of information.

LCCN {LIB_CONGRS_CALL}:=
The number assigned by the U.S. Library of Congress to a document, book, etc.

LGTE {LEGATEE}:=
A role of an individual acting as a person receiving a bequest or legal devise.

MARB {MARRIAGE_BANN}:=
An event of an official public notice given that two people intend to marry.

MARC {MARR_CONTRACT}:=
An event of recording a formal agreement of marriage, including the prenuptial agreement in
which marriage partners reach agreement about the property rights of one or both, securing
property to their children.

MARL {MARR_LICENSE}:=
An event of obtaining a legal license to marry.

MARR {MARRIAGE}:=
A legal, common-law, or customary event of creating a family unit of a man and a woman as
husband and wife.

MARS {MARR_SETTLEMENT}:=

An event of creating an agreement between two people contemplating marriage, at which
time they agree to release or modify property rights that would otherwise arise from the
marriage.

MEDI {MEDIA}:=
The medium used to store or transmit information.

MBR {MEMBER}:=
Identifies an individual (element) belonging to a group (set).

MOTH {MOTHER}:=
Identifies the female parent in a family. In the Lineage-linked form this tag is used only in

the EVENT_RECORD role tag structure (See Chapter 2). Parent relationships are
represented using the HUSBand and WIFE tags as part of the FAMILY_RECORD.

NAME {NAME}:=
A word or combination of words used to help identify an individual, title, or other item.
More than one NAME line should be used for people who were known by multiple names.

NAMR {NAME_RELIGIOUS }:=
A name given to an individual to be used in association with one's religion.

NAMS {NAME_SAKE}:=
Identifies the person that an individual is named after to perpetuate the person's name.

NATI {NATIONALITY}:=
The national heritage of an individual.

NATU {NATURALIZATION}:=
The event of obtaining citizenship.

NCHI {CHILDREN_COUNT}:=
The number of children that this person is known to be the parent of (all marriages), or that
belong to this family.

NMR {MARRIAGE_COUNT}:=
The number of times this person has participated in a family as a spouse or parent.

NOTE {NOTE}:=
Additional information provided by the submitter for understanding the enclosing data.

OCCU {OCCUPATION}:=
The type of work or profession of an individual.

OFFI {OFFICIATOR}:=
A person acting in an authorized capacity as voice in performing an ordinance or ceremony.

ORDN {ORDINATION}:=
A religious event of receiving authority to act in religious matters.

ORIG {ORIGINATION}:=
Pertains to the creation or root of an object.

OWNR {OWNER}:=
The name of the person who is the owner of the associated item or property.

PAGE {PAGE}:=
A number or description to identify the page in a document.

PERI {PERIOD}:=
Indicates the range of time during which an event took place.

PHON {PHONE}:=
A unique number assigned to dial a specific telephone.

PHOTO {PHOTO}:=
A photograph (graphic image) of a person, place, or thing, depending on the enclosing
context.

PHUS {PREV_HUSB}:=
       An individual in the role of the principal's previous husband for a cited event.

PLAC {PLACE}:=
       A jurisdictional name to identify the place or location of an event.
PORT {PORT}:=
       A site identifier of entering or leaving, such as an air port, harbor, port of entry, or a data
       port where data enters or leaves a system.

PROB {PROBATE}:=
       An event of judicial determination of the validity of a will. May indicate several related
       court activities over several dates.

PROP {PROPERTY}:=
       The name of land and/or other properties possessed by this individual.

PUBL {PUBLICATION}:=
       A published work.

PUBR {PUBLISHER}:=
       The name of the company or individual who published a work.

PWIF {PREV_WIFE}:=
       An individual in the role of the principal's previous wife for a cited event.

QUAY {QUALITY_OF_DATA}:=
       An assessment of the reliability of the evidence to support the conclusion drawn from the
       evidence.

RECO {RECORDER}:=
       A person responsible for recording information about an event, place, or person.

REFN {REFERENCE}:=
       A description or number used to identify an item for filing, storage, or other reference
       purposes.

REFS {REFERENCED_SOUR}:=
       A source that was referenced by the cited source but was not examined by the submitter.
       Examined sources are listed using a SOUR tag.

RELI {RELIGION}:=
       A religious denomination to which a person is affiliated or for which a record applies.

REPO {REPOSITORY}:=
       An institution that has the specified item as part of its collection(s).

RETI {RETIREMENT}:=
An event of exiting an occupational relationship with an employer after a qualifying time
period.

RFN {REC_FILE_NUMBER}:=
A permanent number assigned to a record that uniquely identifies it within a known file.

ROLE {ROLE}:=
A name given to a role played by an individual in connection with an event.

SCHEMA {SCHEMA}:=
A context pattern definition that specifies the meaning and the valid context(s) of a user
defined tag. See the SCHEMA_STRUCTURE substructure definition.

SELR {SELLER}:=
A person who sold or sells to another.

SEQU {SEQUENCE}:=
Indicates the sequence or order of an event or information.

SERS {SERIES}:=
Designates the volume within a series in which a given work is a part.

SEX {SEX}:=
Indicates the sex of an individual--male or female. No SEX line is present if the sex is
unknown.

SIBL {SIBLING}:=
A male or female child of a common parent.

SIGN {SIGNATURE}:=
Used to identify information about an individual's signature.

SIST {SISTER}:=
A female sibling.

SITE {SITE}:=
The name of the specific location, building, etc. that is in connection with the address or
place value, such as, "Shriners Hospital" or "The Church of the Epiphany".

SLGC {SEALING_CHILD}:=
A religious event pertaining to the sealing of a child to his or her parents in an LDS temple
ceremony.

SLGS {SEALING_SPOUSE}:=
A religious event pertaining to the sealing of a husband and wife in an LDS temple
ceremony.

SOUND {SOUND}:=
A collection of sound bits pertaining to the enclosed context.

SOUR {SOURCE}:=
The initial or original material from which information was obtained.

SPOU {SPOUSE}:=
A husband or wife of a person.

SSN {SOC_SEC_NUMBER}:=
A number assigned by the United States Social Security Administration. Used for tax
identification purposes.

STAT {STATUS}:=
An assessment of the state or condition of something.

SUBM {SUBMITTER}:=
An individual or organization who contributes genealogical data to a file or transfers it to
someone else.

TEMP {TEMPLE}:=
The name or code that represents the name of a temple of The Church of Jesus Christ of
Latter-day Saints.

TEXT {TEXT}:=
The exact wording found in an original source document.

TIME {TIME}:=
A time value in a 24-hour clock format, including hours, minutes, and optional seconds,
separated by a colon ":". Fractions of seconds are shown in decimal notation.

TITL {TITLE}:=
A descriptive description of a specific writing, such as the title of a book when used in a
source context, or a formal designation used by an individual in connection with individual's
name, such as Captain.

TRLR {TRAILER}:=
At level 0, specifies the end of a GEDCOM transmission.

TXPY {TAXPAYER}:=
A role of a person who has been assessed a tax.

TYPE {TYPE}:=
A further qualification to the meaning of the associated superior tag. The value does not
have any computer processing reliability. It is more in the form of a short one or two word
note that should be displayed any time the associated data is displayed.

VERS {VERSION}:=
Indicates which version of a product, item, or publication is being used or referenced.

WFAT {WIFE_FATHER}:=
A role of an individual acting as the wife's father for a cited event.

WIFE {WIFE}:=
An individual in the family role of a married woman or mother.

WILL {WILL}:=
A legal document treated as an event, by which a person disposes of his or her estate, to take
effect after death. The event date is the date the will was signed while the person was alive.
(See also PROBate.)

WITN {WITNESS}:=
An individual who attested that he or she saw an event take place.

WMOT {WIFE_MOTHER}:=
A role of an individual acting as the wife's mother for a cited event.

XLTR {TRANSLATOR}:=
The name of a person who translated words from one language to another.

Appendix B

PROPOSED EVENT AND ROLE TAG DEFINITIONS

The additional event and roll tags below have not yet been standardized. They are shown here in
this draft form to obtain opinions as well as definitions. We will standardize as many as makes
sense by the time the draft is finalized. The underscore '_' in front of the tags indicate the tags
which have not been standardized and should be structured as user defined tags complete with your
own definition and classification using the ISA tag. The other tags, the ones with the asterisk '*'
have been standardized and defined in the 5.x Appendix A. Tags not appearing in Appendix A are
not used in any of the lineage-linked structures of 5.x and were therefore dropped from the standard
approved list.

Events:

TAG:	TAG NAME	DEFINITION
_ABJUR	Abjuration	
_ABSOL	Absolution	
ADOP	Adoption*	
_APPRN	Apprenticeship	
BAPM	Baptism*	
BIRT	Birth*	
CENS	Census*	
_CHARTR	Charter	
CHR	Christening*	
_CITZN	Citizenship	
_CIVIL	Court Civil	
_CNFSCTN	Confiscation	
_COMUN	Communion	
CONF	Confirmation*	
_CRIME	Court Criminal	
_CRTULRY	Cartulary	
DEAT	Death*	
_DEAT_NOTE	Death_Notice	
DIV	Divorce*	
_DIV_ANUL	Divorce_Annulment	
_DIV_SEP	Divorce_Separation	
_DOWRY	Dowry	
_DPORTN	Deportation	
EDUC	Education*	
EMIG	Emigration*	
_EMPLYMT	Employment	
_ENRLMNT	Enrollment - matriculation	
_EXCUTN	Execution	
_F_COMM	First_Communion	
_FUNRL_HOME	Funeral Home	

Events: (cont')

TAG:	TAG NAME	DEFINITION
_GALLEY	Galley	
GRAD	Graduation*	
IMMI	Immigration*	
_INTRO	Introduction	
_LAND	Land	
_LND_LEAS	Land_Lease	
_LND_PURC	Land_Purchase	
_MARR_BTRO	Marriage_Betrothal	
_MARR_CMLAW	Marriage Common Law	
_MARR_CNSNU	Marriage_Consanguinity	- marriage to blood relatives
_MARR_CNTRC	Marriage_Contract	
_MARR_DIMIS	Marriage_Dimissorial	- permission to get married in another jurisdiction
_MARR_DISPN	Marriage_Dispensations	
_MARR_ENGA	Marriage_Engagements	
_MARR_INTNT	Marriage_Intention	
_MARR_REHAB	Marriage_Rehabilitation	
_MARR_BANN	Marriage_Banns - Announcements	
MARR	Marriage*	
MILI	Military*	
_MILI-INDU	Military_Induction	
_MILI_DIS	Military_Discharge	
_MISS_PRSN	Missing Person	
_NAME_CHNG	Name Change	
NATU	Naturalization*	
ORDN	Ordination*	
_PASL	Passenger_List	
_PASP	Passport	
_POLI_RPT	Police_Reports	
_POPL_REG	Population_Register	
_POOR_LAW	Poor_law	
PROB	Probate*	
_ROSTR	Roster	
_S_COMM	Solemn_Communion	
_SASINE	Sasine	
_SEPRTN	Separation	
_SLAVE	Slavery	

Events: (cont')

TAG:	TAG NAME	DEFINITION
TXPY	Tax_payer*	
_TSTMNT	Testament	
_VOTE_REG	Voting_Registration	
_VOW	Vow	
WILL	Will*	

Roles:
The following are roles which could be used to describe participants in events. The status of these
tags are the same as those listed for the event tags listed above.

TAG:	TAG NAME	DEFINITION
_ANCE	Ancestor	
_APLCNT	Applicant	
_APPRN	Apprentice	
_APRSR	Appraiser	
_AUNT	Aunt	
_BISHP	Bishop	
_BOARDR	Boarder	
_BOROWR	Borrower	
_BRID	Bride	
_BRO	Brother	
BUYR	Buyer*	
_CAPT	Captain	
CHIL	Child*	
_CLRGY	Clergymen	
_CMDR	Commander	
_COUSN	Cousins	
_CREW	Crew	
_DEAD	Deceased	
_DESC	Descendant	
_EMPLYR	Employer	
_EXCUTR	Executor	
FATH	Father*	
_FIANCE	Fiance	
_FREND	Friend	

TAG:	TAG NAME	DEFINITION
_GODF	Godfather	
_GODM	Godmother	
GODP	Godparent*	
_GR_AUNT	Grand_Aunt	
_GR_FATH	Grand_Father	
_GR_MOTH	Grand Mother	
_GR_UNCL	Grand_Uncle	
_GROO	Groom	
_GUARDN	Guardian	
HDOH	Head_of_house*	
_HEIR	Heir	
HUSB	Husband*	
INFT	Informant*	
_INSTR	Instructor	
_JRNYMN	Journeyman	
_JUDGE	Judge	
_LENDR	Lender	
_M_WIFE	Midwife	
_MNSTR	Minister	
_MONK	Monk	
MOTH	Mother*	
_MSTR	Master	
_NIECE	Niece	
_NEPH	Nephew	
_NLAW	In_law	
_NLAW_BRO	Brother_in_law	
_NLAW_DAU	Daughter_in_law	
_NLAW_FATH	Father_in_law	
_NLAW_MOTH	Mother_in_law	
_NLAW_SIS	Sister_in_law	
_NLAW_SON	Son_in_law	
_NOTRY	Notary	
_NUN	Nun	
_NURS	Nurse	
OFFI	Official*	
_ORPHN	Orphan	
_PHYSN	Physician	
_PROF	Professor	
_PRISNR	Prisoner	
_PATIENT	Patient	
_PASNGR	Passenger	

TAG:	TAG NAME	DEFINITION
RECO	Recorder*	
REL	Relative*	
_RNTR	Renter	
_RSDNT	Resident	
_SASSIER	Sassier	
_SBLNG	Sibling	
SELR	Seller*	
_SIS	Sister	
_SLAV	Slave	
_SOLDR	Soldier	
SPOU	Spouse*	
_SERVNT	Servant	
_STEWRT	Stewart	
_STUD	Student	
_TEACHR	Teacher	
_TENANT	Tenant	
_UNCL	Uncle	
_WARD	Ward	
WIFE	Wife*	
WITN	Witness*	

Appendix C

ANSEL CHARACTER SET

The following tables show the spacing and non-spacing diacritic characters that are contained in the
ANSEL set. This table was added to give help to those receiving the machine version to the
GEDCOM standard. The graphic characters shown are not always accurate, however the name of
the diacritic and the decimal equivalent should agree with the ANSEL standard.

C/R column refers to the column and row of the American National Standard Z39.47-
1985 showing the ANSEL character graphic and its 8 bit binary representation.

wpcode column shows the Wordperfect (code page , character number) (1,2) chosen as the
closest representation of the diacritic as shown in Wordperfect Appendix P. of version
(5.1)

Dec column shows to the decimal equivalent for that diacritic as is used in the ANSEL
character set.

Name column gives the english name of the diacritic.

example of use column shows an example of words using this diacritic. For the non-
spacing diacritic, this mark appears before the character in which it should be
superimposed.

ANSEL Non-spacing graphic characters
8-bit

C/R	wpcode	Dec	Graphic	Name	example of use
14/0	2,4	224		low rising tone mark	c□ui
14/1	1,0	225		grave accent	r□egle
14/2	1,6	226		acute accent	est□a
14/3	1,3	227		circumflex accent	m□eme
14/4	1,2	228		tilde	ni□no
14/5	1,8	229		macron	g□aj□ejs
14/6	1,22	230		breve	alt□a
14/7	1,15	231		dot above	□zaba
14/8	1,7	232		umlaut (diaeresis)	□oppna

14/9	1,19	233	◆	hacek	v◆zdy
14/10	1,14	234	◆	circle above (angstrm)	h◆ar

ANSEL Non-spacing graphic characters
8-bit

C/R	wpcode	Dec	Graphic	Name	example of use
14/11	2,11	235		ligature, left half	akademii��a
14/12	2,12	236		ligature, right hlf	akademii��a
14/13	1,10	237		high comma, off center	rozde�lovac
14/14	1,16	238		double acute accent	id�oszaki
14/15	2,25	239		candrabindu	Ali�iev
15/0	2,15	240		cedilla	�ca
15/1	2,17	241		right hook	viet�a
15/2	2,0	242		dot below	te�da
15/3	2,1	243		double dot below	khu�tbah
15/4	2,3	244		circle below	Mahar�sicaritam�rtam
15/5	2,6	245		double underscore	�Ghulam
15/6	2,7	246		underscore	samar
15/7	2,16	247		left hook	darzi�na
15/8	2,14	248		right cedilla	kh�ong
15/9	2,9	249		half circle below	�humantu�s
15/10		250		double tilde, left half	�ngalan
15/11		251		double tilde, right hlf	�ngalan
15/12	1,5	252		diacritic slash through char (LDS extension)	
15/13					
15/14	1,9	254		high comma, centered	g�eotermika

ANSEL Spacing graphic characters
8-bit

C/R	wpcode	Dec	Graphic	Name	example of use
10/0					
10/1	1,152	161		slash L - uppercase	��d|
10/2	1,80	162		slash O - uppercase	|st
10/3	1,78	163		slash D - uppercase	|uro
10/4	1,88	164		thorn - uppercase	|ann
10/5	1,36	165		ligature AE - uppercase	|gir
10/6	1,166	166		ligature OE - uppercase	|uvre
10/7	1,6	167		miagkii znak	fakul|tet
10/8	1,1	168		middle dot	novel|la
10/9	5,28	169		musical flat	B|
10/10	4,22	170		patent mark	ABC|
10/11	6,1	171		plus or minus	A|B
10/12	1,230	172		hook O - uppercase	B|
10/13	1,232	173		hook U - uppercase	X|A
10/14	1,11	174		alif	Un|yusho
10/15		175			reserved - future
11/0	2,11	176		ayn	fa|il
11/1	1,153	177		slash l - lowercase	rozbi|
11/2	1,81	178		slash o - lowercase	h|j
11/3	1,79	179		slash d - lowercase	|avola
11/4	1,89	180		thorn - lowercase	|ann

ANSEL Spacing graphic characters (cont.)

C/R	wpcode	Dec	Graphic	Name	example of use
11/5	1,37	181		ligature ae - lowercase	sk|g
11/6	1,167	182		ligature oe - lowercase	|uvre
11/7	1,16	183		tverdyi znak	ob|iavlenie
11/8	1,24	184		dotless i - lowercase	masal|
11/9	4,11	185		British pound	|5.00
11/10		186		eth	
11/11		187		reserved - future	
11/12	1,231	188		hook o - lowercase	S|
11/13	1,233	189		hook u - lowercase	T| D|c
11/14		190		empty box (LDS-extension)	
11/15		191		black box (LDS-extension)	
12/0	6,33	192		degree sign	10| C
12/1	6,49	193		script l	25|.
12/2	4,71	194		phonograph cpyright mrk	Decca|
12/3	4,23	195		copyright mark	|1993
12/4	5,27	196		musical sharp	D|
12/5	4,8	197		inverted question mark	|Que
12/6	4,7	198		inverted exclamtn mrk	|Esta
12/13		205		e in middle of line (LDS extension)	
12/14		206		o in middle of line (LDS extension)	
12/15	1,23	207		Es Zet	Preu|en