PRESS RELEASE – GEDCOM 5.5.5 IS A BETTER GEDCOM

keywords: GEDCOM, genealogy, software, standards, gedcom.org

LEIDEN – 2 October 2019
GEDCOM version 5.5.5 is the first new version of GEDCOM in twenty years.
GEDCOM 5.5.1 was introduced on 2 October 1999. Today, exactly twenty years later,
GEDCOM 5.5.5 is available from gedcom.org; the GEDCOM site is back.

Maintenance Release

GEDCOM 5.5.5 isn't just another new version of GEDCOM. Previous versions of GEDCOM were feature releases; new versions that introduced new features. GEDCOM 5.5.5 is a maintenance release; it addresses issues with GEDCOM itself, both long-standing issues as well as those that result from the twenty year hiatus.
GEDCOM 5.5.5 does not introduce any major new features. On the contrary, GEDCOM 5.5.5 removes dated, obsolete, deprecated and duplicate features to create a leaner and meaner GEDCOM no longer burdened by features nobody needs.

High Quality GEDCOM

The introduction of the first GEDCOM version in twenty years is a major event, yet the version number of the new version, merely 5.5.5 instead of 5.6 or 6.0, indicates a minor revision.

The best way to understand GEDCOM 5.5.5 is to *think of GEDCOM 5.5.5 files as high quality GEDCOM 5.5.1 files*. GEDCOM 5.5.5 files are so compatible with GEDCOM 5.5.1 files, that many existing genealogy applications will read GEDCOM 5.5.5 files without any problem.
It even is a deliberate feature of GEDCOM 5.5.5 (and GEDCOM 5.5.5 only) than you can manually change the version number from 5.5.5 to 5.5.1 to have it imported by applications that do not support GEDCOM 5.5.5 yet.
That manual version number change only works in that direction; you can downgrade a GEDCOM 5.5.5 file to a GEDCOM 5.5.1 file that way, but you cannot upgrade a GEDCOM 5.5.1 file to a GEDCOM 5.5.5 file that way; most GEDCOM 5.5.1 files simply do not meet the higher quality standards of GEDCOM 5.5.5.

Easier to Import

Developers will welcome technical changes that it make it easier to import GEDCOM files, such as
- Simplified GEDCOM grammar
- Clarification that tags are case-sensitive
- A well-defined version number format
- All GEDCOM identifiers must be alphanumerical now
- The LF/CR combination is no longer a legal line terminator
- no CONC or CONT in GEDCOM header
- The Byte Order Mark (BOM) is mandatory
- Clear & simple rules for dealing with white space
  Simple CONC & CONT rules
- and many other simplifications

Significant Simplification

Despite not introducing any major features, the What's New section of the GEDCOM 5.5.5 specification is about twenty pages long. GEDCOM 5.5.5 features such things as more accurate, consistent and current terminology, new sections, and more bonus chapters, but the key change that makes GEDCOM 5.5.5 a better GEDCOM is significant simplification.

One File per GEDCOM File

GEDCOM 5.5.1 allows a GEDCOM file to be split among multiple files, called volumes, GEDCOM 5.5.5 demands the simplicity of a single file.

One GEDCOM Version per File

GEDCOM 5.5.1 demands supports for some GEDCOM 5.5 record in GEDCOM 5.5.1 files. GEDCOM 5.5.5 does not allow such oddities. GEDCOM 5.5.5 demands that all records be GEDCOM 5.5.5 records.

One Line Terminator per File

The GEDCOM 5.5.1 grammar allows each GEDCOM line to have a different line terminator. The GEDCOM 5.5.5 grammar explicitly demands one line terminator per file.

One Encoding per GEDCOM File

GEDCOM 5.5.1 demands use of ANSEL in the GEDCOM header, even if that header specifies the use of another encoding. GEDCOM 5.5.5 demands use of the specified encoding for the entire file.

One Character Set for GEDCOM

GEDCOM 5.5.5 no longer allows characters sets that may lose information. GEDCOM 5.5.5 is Unicode-only.

One Submitter per GEDCOM File

GEDCOM 5.5.1 allows each individual and each family group to have its own submitter. This is a complication that remains widely unsupported. GEDCOM 5.5.5 simplifies GEDCOM by codifying existing practice, and demanding exactly one submitter per GEDCOM file.

One Place Name Format

GEDCOM 5.5.1 has a feature that allows a GEDCOM file to define its own place name format, and even use multiple, different place name formats in the same file. It remains widely unsupported; the one application that tried to support it makes it mess of your place names.
GEDCOM 5.5.5 simplifies GEDCOM an increases compatibility between applications by removing this feature, and simply supporting the one place name format all users know; jurisdictional levels, ordered small to large, separated by a comma and a space.

One Way of Doing Things: One Record per Thing

GEDCOM 5.5.1 has various records that duplicate functionality already provided by other records. GEDCOM 5.5.5 gets rid of the duplicates. The What's New section has a table showing which record to use instead.

One Way of Doing Things: One Record Format

Some records have both an old, deprecated unstructured format and a newer, preferred structured format. GEDCOM 5.5.5 finishes the transitions started in GEDCOM 5.5.1 and earlier by demanding the preferred structured format.

Improved Data Transfer

One Multimedia Record Per File, One File Per Multimedia Record

GEDCOM 5.5.1 allows a single multimedia record to link to multiple files, but only a few application support that. So, if you export from one of these applications to another, links may be lost. GEDCOM 5.5.5 simplifies GEDCOM and ends this data loss by demanding one multimedia record per file, one file per multimedia record.

One Set of Religious Records

GEDCOM used to have one set of religious record, until FamilySearch introduced separate records for LDS (mormon) events; they added an LDS Baptism record, while GEDCOM already had a baptism record. Some genealogy applications support these duplicate events, most do not, so if you transfer your data from one to the other, you are practically sure to lose that data.
GEDCOM 5.5.5 simplifies GEDCOM and makes sure LDS events aren't lost by returning to one set of religious records for all religions.

One Character Set: Unicode

GEDCOM 5.5.1 allows ASCII, ANSEL and Unicode. What's more, in practice many vendors still use code pages such as Windows ANSI, while they've never been legal.
During the past two decades, the genealogy industry has switched to Unicode, and allowing export to anything but Unicode is allowing loss of characters. GEDCOM 5.5.5 prevents such loss by demanding Unicode.

Simple CONC & CONT rules

One aspect of GEDCOM that has presented problems for users is the use of CONC & CONT records, because some applications do not handle these correctly. The GEDCOM 5.5.1 Annotated Edition provided simple guidelines for handling CONC & CONT correctly.
GEDCOM 5.5.5 provides the simplest CONC & CONT rules ever, making it near impossible to mess up.

A Faster GEDCOM

A welcome benefit of all the simplifications and tightened rules is that import of GEDCOM 5.5.5 files should be noticeably faster than import of GEDCOM 5.5.1 files.

New Features

GEDCOM 5.5.5 is mostly an improved and simplified GEDCOM 5.5.1, but the creators did sneak in a few new features that did not require any new syntax.

Same-Sex Marriage

GEDCOM 5.5.5 explicitly supports same-sex marriage. On a first reading of GEDCOM 5.5.1, the specification seems to prohibit same-sex marriage, but many vendors already figured out that GEDCOM 5.5.1 actually supports same-sex marriage without any GEDCOM extensions. The

GEDCOM 5.5.5 specification clarifies that what they are doing conforms to the existing specification.

Relationships

GEDCOM 5.5.5 explicitly explains how to support relationships other than marriage, and provides a list of relationships to use. This significantly increases cross-product compatibility.

Intersex and Not Recorded

GEDCOM 5.5.5 adds two new sex values: X for Intersex, and N for Not Recorded.

Birthdays

GEDCOM 5.5.5 explicitly allows what many applications were already allowing: a date without a year.

Centenarian Support

GEDCOM 5.5.1 allows recording someone's age at an event, but its AGE_AT_EVENT definition restricts the number of years to only two digits. GEDCOM 5.5.5 allows three digits.

Industry Experts

GEDCOM 5.5.5 has been created by an international group of industry experts, many of whom also contributed to the *GEDCOM 5.5.1 Annotated Edition* released last year, which effectively replaced the original GEDCOM 5.5.1 specification.

The GEDCOM 5.5.5 specification was edited by renowned genealogy technologist Tamura Jones, working together with GEDCOM experts Bob Coret (Genealogy Online, Open Archives), Diedrich Hesmer (Our Family Book, GEDCOM Service Programs), Andrew Hoyle (Chronoplex My Family Tree, Chronoplex GEDCOM Validator), Kari Kujansuu (Genealogy Society of Finland's isotammi.net), Louis Kessler (Behold, GEDCOM File Finder, Double Match Triangulator), Stanley Mitchell (ezGED Viewer) and Nigel Munro Parker (GED-inline GEDCOM validator).

MyHeritage CEO Gilad Japhet bought and donated the gedcom.org domain.

A Better GEDCOM with Annotations

GEDCOM 5.5.5 is a much better GEDCOM than GEDCOM 5.5.1, but does not fix two decades of overdue maintenance with a single release. The GEDCOM 5.5.5 specification is based on the *GEDCOM 5.5.1 Annotated Edition*, and keeps the annotations for issues not addressed by this release.

Availability

The GEDCOM 5.5.5 specification is available today from gedcom.org. The gedcom.org website offers a GEDCOM FAQ for end users, but is largely aimed at developers. It offers GEDCOM 5.5.5, the current specification, as well as GEDCOM 5.5.1, the previous specification, and the Second Revision of the GEDCOM 5.5.1 Annotated Edition. These specifications are complemented by an overview of GEDCOM validators, and a few GEDCOM 5.5.5 sample files to help developers get started with the new standard.

gedcom.org: https://www.gedcom.org

NOTES TO THE EDITOR

illustration

There is a simple transparent PNG "GEDCOM 5.5.5" image you may use as an illustration at www.gedcom.org/media.html

abbreviations

GEDCOM = Originally an abbreviation of **G**enealogical **D**ata **COM**munications, but that is a misnomer. Now considered an ALL-CAPS name instead of an abbreviation.

LDS = The Church of Jesus Christ of of Latter-Day Saints (the mormons)