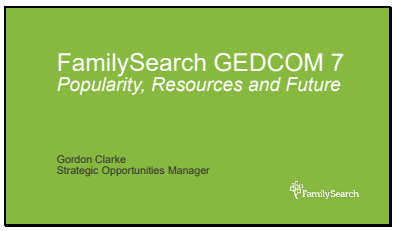
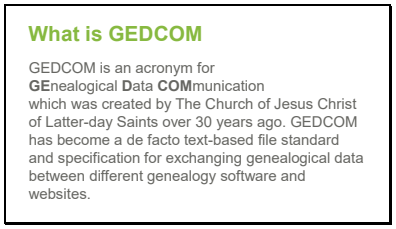
Slide 1



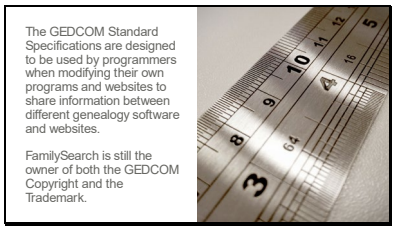
Hello, this is Gordon Clarke, FamilySearch's Strategic Opportunities Manager.
Welcome to the video where you will learning everything you need to know about why and how the newest version of GEDCOM was created. Which companies have implemented it and our making it popular. The many resources that are available to help in software integration. And finally what's ahead for the next couple of years

Slide 2



GEDCOM is an acronym for **GE**nealogical **D**ata **COM**munication which was created by The Church of Jesus Christ of Latter-Day Saints over 30 years ago. GEDCOM has become a de facto text-based file specification for exchanging genealogical data between different software applications and websites.

Slide 3



The GEDCOM File Specifications are design by programmers for programmers. It's the best measurement for the quality of exchanging genealogy information between software companies for the owners of the data.

https://pixabay.com/photos/metal-ruler-metal-measures-flexible-2765212/

Slide 4

Brief History of GEDCOM

	1987 - 1995	1999	2001	2019	2020	2024
Version	3.0, 4.0, 5.0, 5.5	5.5.1	6.0 XML	5.5.1	New Version Started	FamilySearch 7.0.13
Copyright	yes	Draft limited rights	Incomplete	yes	yes	yes
Participation	Private	Private	Private	Private	Public Github	Public Github

Though there was a lot of preparatory work in 2019, the kickoff was 4 years ago at RootsTech 2020. All the major companies, software provider, and industry influencers attended as we got things rolling.

FamilySearch GEDCOM 7 was released in May of 2021. And since then many additions and changes had been made. Various tools to support development have now been released. Various companies have completed there ability to read and write version 7.

Slide 5

Why Continue GEDCOM

1. Popular valuable file format for storing, exchanging and presenting genealogical data
2. Continued responsibility, ownership and improvement by FamilySearch
3. Continue and improve importing into FamilySearch.org
4. Eliminate any question of copywrite and trademark ownership

GEDCOM lives on because the industry wants a Standard File Format for personally controllable genealogical information that's shareable by a branch or the whole tree.

Share what you want with friends, family, genealogy consultants and FamilySearch.

Slide 6

Why Continue GEDCOM

5. Used in all major countries and languages
6. Personal private backup of family tree information and control for editing, copying and sharing
7. Hundreds of website and desktop applications that support the reading and writing of GEDCOM files
8. Many valuable applications need a file for virtual reality, visual tree (engraving, sculpting, machine embroidery), tree analysis and charting solutions

GEDCOM can be used in all the major countries and languages. Copies and backups are just a file to put on local storage or in the cloud. You data can be moved in and out of all the popular desktop and cloud tree solutions.

Slide 7

But GEDCOM 5.5.1 was Outdated. There were many ambiguities, grammar, and spelling errors. The tags and character sets were outdated. Some of the examples didn't even match the specification itself.
Transfer-speed and storage can now handle image and other files associated with the GEDCOM 7 file.

https://pixabay.com/illustrations/smartphone-hand-photomontage-faces-1445489/

Slide 8

The rapid advances in technology allowed the developers of genealogical software and websites to incorporate digitized documents, photographs and sophisticated search options.

Websites and local media can incorporate digitized documents, photographs, and more sophisticated search options.

https://pixabay.com/illustrations/analytics-information-innovation-3088958/

Slide 9

FamilySearch GEDCOM 7

A standard genealogical file format allowing for

the ***creation***
the ***preservation***
the ***collaboration***

importing and exporting data and images
with different applications and cloud services

while maintaining control of the original copy

This statement says it all.

Slide 10

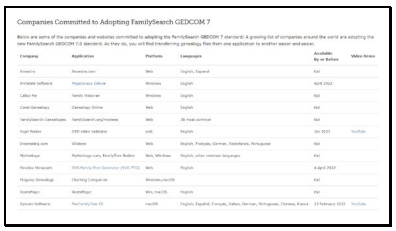
FamilySearch GEDCOM 7 Contributors

Many long-standing veterans of the family history marketplace have contributed to the groundswell that is making the new version and continuous maintenance of GEDCOM possible.

MAJOR WEBSITES	DESKTOP PROVIDERS	STEERING COMMITTEE
Ancestry.com	Ancestral Quest	Albert Emmerich
FamilySearch.org	Family Historian	Dave Thaler
MyHeritage	Family Tree Maker	Gordon Clarke
Puzzilla	Legacy Family Tree	Seth Green
	Mac FamilyTree	James Tanner
	Roots Magic	Luther Tychonievich
		Russell Lynch

Thanks goes to many companies, organizations, and contributors.

Slide 11



Learn about the many companies planning on or already importing FamilySearch GEDCOM 7 at GEDCOM.info.

Slide 12

Videos from Application Supporting FamilySearch GEDCOM 7

Synium Software, Mac FamilyTree
Calico Pie, Family Historian
GEDMatch
Nigel Parker, GEDinline Validator
Joseph Schnieder, GES 2000

These companies have videos showcasing their implementation of GEDCOM 7.

Slide 13

2023 FamilySearch Import to GEDCOM 7

GEDCOM 7 files can now be imported from FamilySearch Family Tree:

- Read and Conversion Code is available at: https://github.com/gedcom7code/xto7
- Sample Application available at: https://gedcom.surge.sh/

In mid-2023 the ability to download and import GEDCOM 7 from Familysearch Family Tree was made possible.

Slide 14

Questions have been asked about GEDCOM X data format which is the backbone for the FamilySearch website. There are many common data elements with FamilySearch GEDCOM but the purposes are different. The overlap will gradually increase but I don't every see one replacing the other. **GEDCOM-X is the file specification use underneath FamilySearch Family Tree's and is widely implemented by software that interoperates with FamilySearch Family Tree. However, GEDCOM remains more common as a file format for sharing files between software platforms.**

Slide 15

We now have three very experience presenters that will share screens and comment about:
Our websites, technical features, and the technical improvements in 2023.

Slide 16

Overview of Websites

Russell Lynch

Slide 17

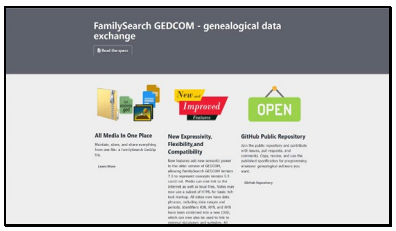
FamilySearch GEDCOM 7.0 release with valuable new enhancements

Hello, I am Russell Lynch, a member of the GECOM Steering Committee.
We released GEDCOM 7.0 in May of 2021.
The updated version of the specification was just the beginning.
We have been preparing for the rollout by developing tools, guides and other programming resources for many months.

Slide 18

This is the webpage for GEDCOM.info
It is a welcome page for public users with links to pages for developers.

Slide 19



This is the GEDCOM.io webpage. I will spend most of my time explaining the resources available on this website.

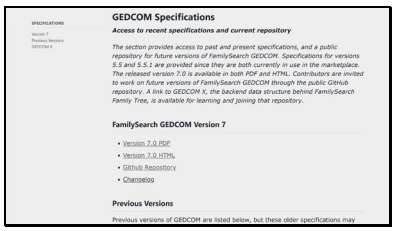
Slide 20

The website for developers has three main sections that we thought were the most important: 1) GedZip to combine all media in one place, 2) new features of GEDCOM explained, and 3) the GitHub Public Repository because we want an open-source approach to the GEDCOM standard

Slide 21

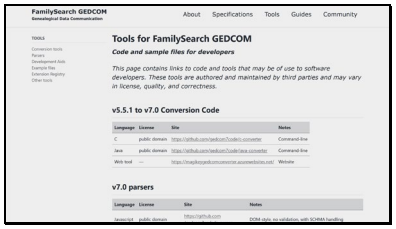
A GEDZIP is a GEDCOM document with a set of external files. It is the archive with the photographs and multi-media files referenced in the GEDCOM.

Slide 22



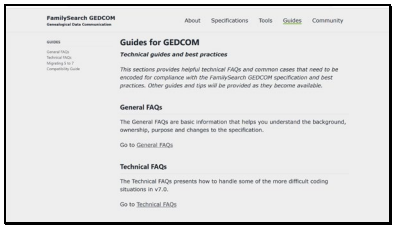
The Specifications tab has the current version of GEDCOM and previous versions for reference purposes.

Slide 23

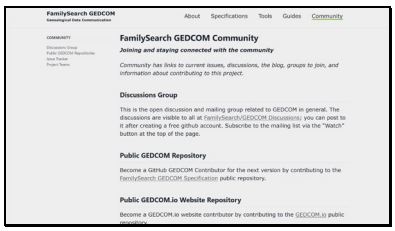


This is the Tools tab. The GEDCOM Steering Committee wants to provide tools for developers. To this end we offer conversion code from version 5.5.1 to 7.0; a parser, a validator, and example files of both excellent use of 7.0 GEDCOM and poor use of 7.0. We will continue to add resources to this page.

Slide 24

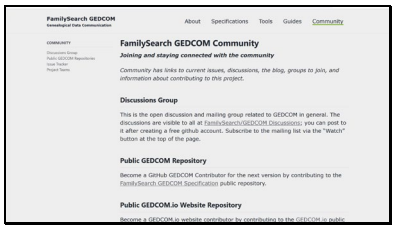


Slide 25



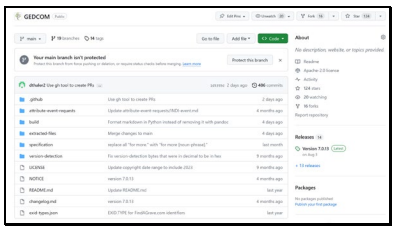
As Gordon Clarke mentioned, in the past few years we have worked to strengthen the GEDCOM Community. We know there are elements of data communication that need enhancements. These will require major revisions to GEDCOM. Please see our session called “What’s Next for GEDCOM” to learn about our plans to improve structures for places, names, calendars, relationships, source citations, and hypotheses.

Slide 26



As Gordon Clarke mentioned, in the past few years we have worked to strengthen the GEDCOM Community. We know there are elements of data communication that need enhancements. These will require major revisions to GEDCOM. Please see our session called “What’s Next for GEDCOM” to learn about our plans to improve structures for places, names, calendars, relationships, source citations, and hypotheses.

Slide 27

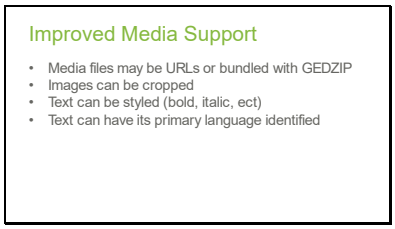


And last, but not least, I will mention our GitHub repository where developers can propose changes to the GEDCOM standard. The GitHub pages for GEDCOM has issue tracking, discussions, and projects. It is an open repo for GEDCOM development.

Slide 28

Let's look at some of the technical features introduced with Family Search GEDCOM 7

Slide 29



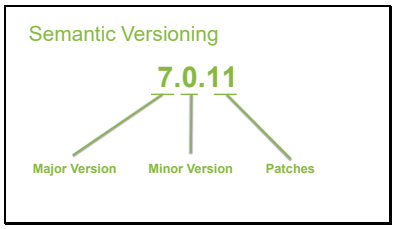
One of the most visible new features is improved support for media. Media files like images, videos, and audio have been linkable from GEDCOM files for many years. Family Search GEDCOM 7 adds the ability to bundle those media files with the GEDCOM a new GEDZIP file format. It also allows those image files to be externally linked via URLs from the web. In addition, images can now have cropping information. So if you have a large image and want to refer to only a small part in some context, you can do that in the GEDCOM file itself instead of needing to edit the image file. Text is also a form of media that has been improved: text can now have basic markup with bold and italics and other basic formatting information, and text can have its primary language identified per text block. So if you have a GEDCOM file with text in multiple languages, that information is now storable as well.

Slide 30

In addition to media, there is new support for research.

One of the biggest additions is the NO tag which allows you to express the idea that some event did not occur in some individual or family. An information was commonly stored in notes, but often overlooked because of that location, and is now stored in data just like the assertion that events did occur.

Slide 31



FamilySearch GEDCOM 7 introduces the use of semantic versioning. The version number can now tell you something about the changes between different versions of FamilySearch GEDCOM.

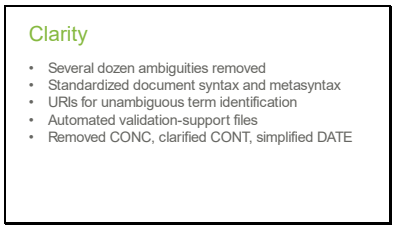
Semantic versioning includes three numbers.

The patch number, the farthest on the right, is used to represent a change in the specification that does not change the underlying files. 7.0.0 and 7.0.6, for example, represented the exact same data, just more clearly specified in 7.0.6.

The middle number, the minor version represents an addition of information without the removal or change of any of the existing information. So version 7.1 will have features that 7.0 doesn't have, but will also have all of the features (implemented the same way) that 7.0 does. So a 7.1 compatible application can open both 7.0 and 7.1 files.

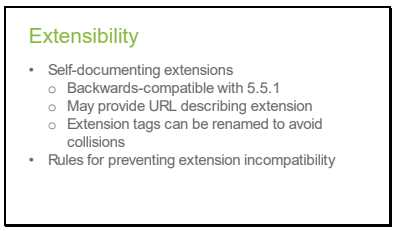
The major version, the 7, is changed only when we're changing how something is done or removing some functionality. So an 8.0 compatible system might not be able to open 7.0 files. And that's part of the reason we picked 7: because we are not compatible either with 5.5.1 or with the draft version 6 that was circulated a few years after that. There are converters that allow conversion between them, but a 5.5.1 file in general is not a valid 7.0 file.

Slide 32



Just a semantic versioning makes it clear what the meaning of a version is, so to other clarity has been added to the specification. We remove dozens of ambiguities from the previous specification; standardized the way the document presents information; provided URIs to give unambiguous meaning for terms that were used in multiple contexts mean different things in previous versions of the specification. We've designed the specification a way that we can automatically extract validation data files so validators can be built to work for all future versions as well as the current version. We've also gone into some specific troubling tags that were present in previous versions and adjusted or remove those to make the specification clearer and simpler to implement.

Slide 33

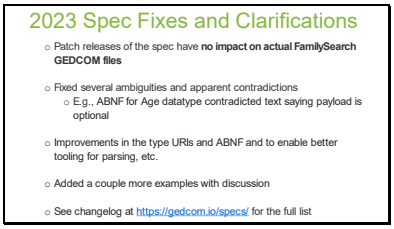


A major strength of all versions of GEDCOM is their support for extensions, where applications can support all of the standards, tags and structures of the specification, and also add additional features that they understand and store them in the same files. We've looked into that in version 7.0 was self-documenting extensions that are backwards compatible with the extensions of previous versions, but can also have a URL whereby implementers can find a definition of what this extension means, and allows extensions created by different developers that happen to use the same name to be renamed to avoid ambiguity and confusion in what they're meaning. You've also created a new rule text describing what extensions can and can't be used for. So the extensions will not break interoperability between applications

Slide 34

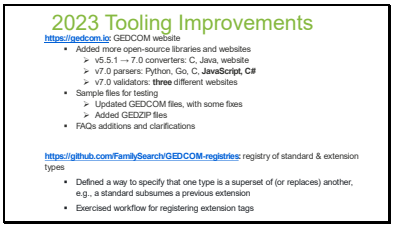
Hello, I'm Dave Thaler, I'm a member of the GEDCOM Steering Committee, and I'm going to tell you about some of the improvements we've made during 2023.

Slide 35



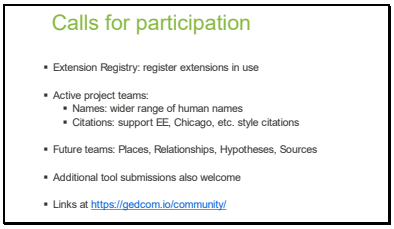
The first thing to keep in mind is that any of the changes that we've made in the GEDCOM specification do not affect GEDCOM files themselves, they only affect the specification. So for example, we fixed a number of ambiguities and apparent contradictions, such as where a formal specification was not consistent with the text explaining it. We also made a number of improvements in ***how*** types are defined, in order to enable better tooling such as parsers and validators, which I'll talk more about on the next slide. Finally, we added some more examples. You can find the full changelog at gedcom.io/specs

Slide 36



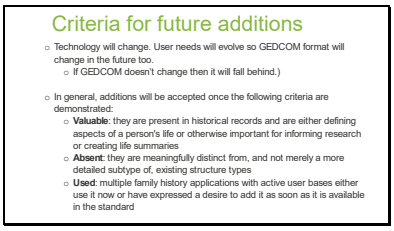
Since we improved the ability for tools to parse and validate GEDCOM, we saw additional open source libraries and websites be developed during 2023, which are now listed on our website at gedcom.io/tools. These include open source parser libraries in two additional programming languages. And the number of free websites implementing GEDCOM validators is now up to three. We also updated the sample GEDCOM files listed on the website, incorporated fixes and additions, and recently added some sample GEDZIP files as well. The committee also created a repository to catalog known GEDCOM types, including both standard types, and extension types, whether those extensions are from existing applications or websites, or whether they are proposed for future standardization. We've defined a way in this registry to denote that one type is a superset of, or replaces, another type, such as when an extension is later incorporated into the standard and replaced by a standard type. And we've just started registering a few extension tags to verify that we're ready to accept other registrations.

Slide 37



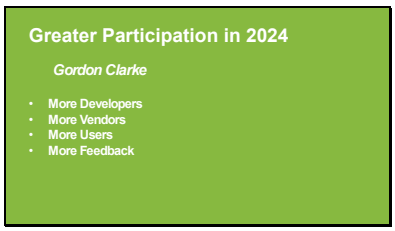
As we discuss the concept of "extensions", it's important to keep in mind that we live in an evolving world where technology and user needs change over time. This means that GEDCOM as a format must be able to evolve as well to meet emerging requirements. As we think about potential future additions to the standard, the GEDCOM Steering Committee has defined three criteria that must be met to be considered for standardization. First, the addition must be shown to be valuable, such as by being present in historical records, or by being important for tracking research. Second, there must not already be some way to represent it in standard GEDCOM today. And third, there should be *multiple* applications or websites that either already use the functionality as an extension or that have expressed a desire to use it if incorporated into the standard.

Slide 38



So that brings us to the Calls for Participation! First, we invite those who are using GEDCOM extensions today to register them in the extension registry, to inform the rest of the community and help motivate future standards work. Second, we welcome interested volunteers and experts in various topics to join GEDCOM project teams, two of which have started, with others potentially starting in the future. Finally, if there are other open source libraries and developer websites that are not yet listed on the tools page, let us know and we are happy to add them to the list. You can find links to all of these at gedcom.io/community.

Slide 39



The GEDCOM 7 Steering committee has be discussing when and how we start working on version 7.1 and what marketing efforts will increase adoption of GEDCOM 7 for many months. We anticipate many more developers, vendors, and users joining the FamilySearch GEDCOM 7 effort in 2024.

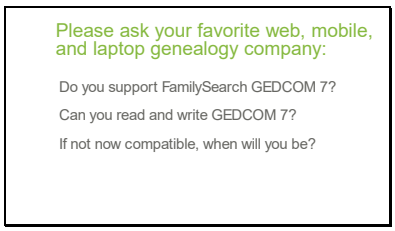
Slide 40

Finally, If your are a software developer or genealogical technician don't forget to join one of these important teams by sending an email to GEDCOM@familysearch.org and put "Project: {project name} " in the subject.

Slide 41

Also Please join any or all GEDCOM 7 related repositories.

Slide 42



If you a consumer or genealogist. Please communicate with your favorite software provider to find out when they will provide support for FamilySearch GEDCOM Version 7.

Slide 43

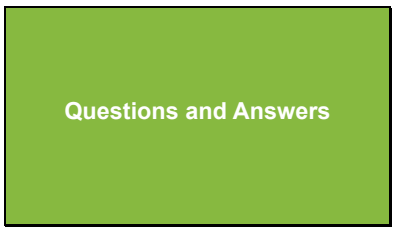
Also Please go to http://GEDCOM.info or http://FamilySearch.org/GEDCOM for general information.

For the actual specification, programmatic tools and guides. please go to http://GEDCOM.io

Ongoing technical communications happen at the new public repository.

Any inquiries can be emailed directly to GEDCOM@FamilySearch.org .

Slide 44



Slide 45