

## CFPS 73

(Call for Papers Submission number 73)

# Beyond Lineage-Linked: Functional Requirements

Submitted by: Johnston, Wesley

Created: 2013-05-09

URL: Most recent version: <http://fhiso.org/files/cfp/cfps73.pdf>  
This version: [http://fhiso.org/files/cfp/cfps73\\_v1.pdf](http://fhiso.org/files/cfp/cfps73_v1.pdf)

Description: This paper addresses the high-level functional requirements need for a modern GEDCOM to model the real world in a way that allows a genealogist to share and manipulate (sort, search, aggregate, report or analyze) relationships other than inter-personal lineage-linkage.

Keywords: relationships, community trees, place, merging, sharing, manipulation, sorting, searching, aggregating, reporting, analysis

# Beyond Lineage-Linked: Functional Requirements

---

*by W. Wesley Johnston  
May 9, 2013*

## **Abstract**

*This paper addresses the high-level functional requirements need for a modern GEDCOM to model the real world in a way that allows a genealogist to share and manipulate (sort, search, aggregate, report or analyze) relationships other than inter-personal lineage-linkage. This is a “what” and not a “how” paper. The paper does venture into some aspects of software implementation requirements, beyond the pure representational requirements of GEDCOM.*

## **Context: Original GEDCOM as a Lineage-Linked Database**

GEDCOM stands for Genealogical Data Communication. It had – and still has – two essential purposes:

1. Storage of a genealogical database
2. Sharing a genealogical database, including the merging of multiple databases

The original conception was as a lineage-linked database. People were linked to other people only through two relationships: parent-child and marriage. The paradigm focused on implementing a bare-bones pedigree in the database. Other inter-personal relationships existed explicitly only as they could be defined by these two fundamental relationships.

Events were tied to people. And places and dates were tied to events. But places connected to each other only indirectly and implicitly through connections back to events for the same person or for multiple people.

However, as any good genealogist knows, when you look at a census, you want to look at the prior and next houses. And witnesses at a marriage and sponsors at a baptism in the real world required explicit connections, since the familial relationship might not yet be known or there may have been no familial relationship and yet the relationship was clearly of great importance and needed to be kept in some other way than as implicit in the textual notes about an event. And it was often the reality that the step-siblings of an ancestor figured very significantly in their lives but could only be represented as distant relations through multiple applications of the two basic interpersonal relationships (i.e. child of parent, marriage, marriage, marriage, child of parent). And people went to school together, joined the same organizations, or in other ways had

relationships that figured very significantly in their lives but which were relegated to being implicit within textual notes.

Other genealogists went even a step further, blurring the lines between family and local history in the creation of community trees. Using a comprehensive record set, such as parish registers or consecutive censuses, these genealogists reconstructed the families of the parish or town or neighborhood.

Thus the original conception of a lineage-linked database failed to model the real world in which people lived.

While a modern GEDCOM must support explicit standardized relationships beyond lineage linkage, it must also provide the user the opportunity to create their own explicit relationships. Commercial software developers implementing the original GEDCOM realized that the real world was a lot more than GEDCOM could model. So each of the developers created add-ons to allow their users to have needed functionality that GEDCOM did not support. Many of these additions could not be easily or correctly transferred to another database, undermining the second purpose of GEDCOM (sharing a genealogical database, including the merging of multiple databases).

This paper addresses the high-level functional requirements need for a modern GEDCOM to model the real world in a way that allows a genealogist to share and manipulate (sort, search, aggregate, report or analyze) relationships other than inter-personal lineage-linkage. I include “share” very purposefully, so that the second purpose of GEDCOM has robust support and we do not wind up again in a world of vendor-specific unique additions that cannot be shared.

## **Paradigms to Keep in Mind**

The above discussion touches on some key paradigms that in the real world result in the creation of relationships, inter-personal and otherwise, that a modern GEDCOM should support.

1. Place must go down to the house level and allow connection of one place to another to support easy manipulation, such as listing all the families who lived on the same street or all the people who lived in the same house over the course of time. This extends to robust support for community trees.
2. Key real-world relationships (e.g. marriage witnesses) must have explicit standardized representation, with robust support for sharing and manipulation. Some lesser real-world relationships (e.g. classmate, co-member of an organization, co-worker – the sort of things that show up on LinkedIn) should also be considered for standardized support, though the line must be drawn somewhere to decide what to include or exclude from standardization. GEDCOM should provide strong support for user-defined explicit relationships beyond the standardized ones. The standardized relationships should be established from the outset, and a central authority (FHISO) should manage the

expansion of those standard relationships as usage reveals the need. The need to be able to share and merge databases must be kept in mind in these decisions.

3. Step families are a special case. They must be kept in mind, but I do not have any insight into how to better represent them. The fact that they were sometimes of great importance in a person's life but sometimes of only minor importance complicates this.