Gedcom

O termo **Gedcom** refere-se a um formato para o intercâmbio de dados genealógicos.

Foi originalmente desenvolvido pela Igreja dos Mórmons por razões religiosas e mais tarde retomado pelos genealogistas, para trocar dados genealógicos entre diferentes pessoas que não tinham os mesmos sistemas.

É, portanto, um tipo de linguagem.

> A palavra **Gedcom**, que é um acrónimo para **ge**nealogical **d**ata **com**munication, está escrita como uma sigla: GEDCOM. Por derivação metonímica, a palavra designa também um **ficheiro de genealogia em formato Gedcom**. O ficheiro `xxxxx.ged` em que está a trabalhar no Ancestris é um Gedcom..

> Desde meados da década de 1990, com o advento da Internet e a multiplicação dos intercâmbios digitais, a norma Gedcom tornou-se gradualmente um padrão essencial para a maioria dos programas e sítios de genealogia. No entanto, alguns deles não respeitam totalmente o formato e adaptam-no desenvolvendo comandos proprietários. Quanto ao Ancestris, é totalmente compatível com a norma Gedcom (versão 5.5 e 5.5.1), o que significa que permite aos seus utilizadores trocar e partilhar ficheiros genealógicos, perfeitamente fiáveis e sem qualquer risco de perda de dados.

Características de um ficheiro Gedcom

Um **ficheiro em formato Gedcom** é um ficheiro de texto (*.ged), ou seja, um ficheiro que pode ser aberto apenas para leitura ou para modificação a partir de qualquer editor de texto (tais como o Bloco de notas, o Kate, o Kwrite, o Gedit, etc.).

Consequentemente, tal ficheiro pode ser utilizado como está por qualquer programa de genealogia, instalado sob qualquer sistema operativo, sem necessidade de converter nada.

As diferentes informações contidas no ficheiro Gedcom são precedidas por uma **etiqueta**: na norma Gedcom, esta etiqueta é um marcador composto por três ou quatro letras maiúsculas, sempre associado ao mesmo tipo de informação.

Por exemplo, a etiqueta PLAC (= place, i.e. local) indica sempre que a informação que anuncia é um local (local de nascimento, local de óbito, local de uma cerimónia, etc.).

Registos de um ficheiro Gedcom

Um ficheiro Gedcom contém um conjunto de registos, o primeiro e o último dos quais são de um tipo especial:

- o primeiro registo chama-se cabeçalho (etiqueta HEAD);
- o último registo chama-se marcador de fim de ficheiro (etiqueta TRLR).

Cada um dos outros registos pertence a uma categoria de **entidade**, sendo entendido que cada uma destas categorias tem as suas próprias funções e estruturas.

Um ficheiro Gedcom contém 7 categorias de entidades. Os registos que podem ser encontrados num ficheiro Gedcom são, portanto, os seguintes:

- registos que descrevem **indivíduos** (etiqueta INDI) ;
- registos que descrevem **famílias** (etiqueta FAM) ;
- registos que descrevem **notas** (etiqueta NOTE) ;
- registos que descrevem **fontes** (etiqueta SOUR) ;
- registos que descrevem **repositórios** (etiqueta REPO) ;
- registos que descrevem **elementos multimédia** (etiqueta OBJE) ;
- registos que descrevem **fornecedores de informação** (etiqueta SUBM).

A escolha de considerar estas 7 categorias de dados como registos é, naturalmente, arbitrária, mas é o princípio de uma norma.

Poder-se-ia facilmente imaginar outros tipos de registos, tais como locais, por exemplo. O facto de um local não ser uma entidade separada não impede o Ancestris de os gerir.

Árvore de um registo

Cada registo é apresentado numa estrutura em árvore: cada etiqueta pode incluir qualquer número de sub-etiquetas.

As sub-etiquetas são etiquetas dependentes hierarquicamente da etiqueta de nível imediatamente superior e podem, por sua vez, incluir uma ou mais sub-etiquetas, etc.

Cada nível pode assim ser sub-dividido ad infinitum.

Níveis hierárquicos

As hierarquias estão numeradas.

Uma vez que cada linha deve permanecer no seu lugar na hierarquia, a cada linha é atribuído um número correspondente ao nível que ocupa na estrutura em árvore do registo.

Por exemplo, a linha no nível principal de cada registo (ou seja, nível zero) é numerada 0; uma linha no nível imediatamente abaixo é numerada 1; uma linha no nível imediatamente abaixo do nível anterior é numerada 2; e assim por diante.

Identificador e categoria dos níveis zero

O nível *zero* de cada registo (para além dos dois registos extremos, HEAD e TRLR) é composto por dois elementos contíguos:

- o número de identificação da entidade rodeado por dois símbolos @;
- a etiqueta associada à categoria a que a entidade em questão pertence;
  - por exemplo, o cabeçalho @I24@INDI significa que identificador do registo é I24 e que ele pertence à categoria de indivíduos (i.e., indivíduo número I24).

Indentação

Para maior clareza, as diferentes linhas de um registo podem ser indentadas (com um ou mais espaços à esquerda do primeiro carácter) para facilitar a identificação da posição de cada linha na hierarquia.

- Exemplo não indentado:

```
0 @I3@ INDI (etiqueta principal deste disco: individual I3)
1 NAME Jean Martin (nome do indivíduo)
1 SEX M (sexo do indivíduo: masculino)
1 BIRT (nascimento do indivíduo)
2 DATE 16 Abril 1951 (data: 16 de Abril de 1951)
1 FAMC @F5@ (família da qual descende o indivíduo I3: família F5)
```

- o mesmo exemplo, agora indentado:

```
0 @I3@ INDI (etiqueta principal deste disco: individual I3)
  1 NAME Jean Martin (nome do indivíduo)
  1 SEX M (sexo do indivíduo: masculino)
  1 BIRT (nascimento do indivíduo)
    2 DATE 16 Abril 1951 (data: 16 de Abril de 1951)
  1 FAMC @F5@ (família da qual descende o indivíduo I3: família F5)
```

O **editor Gedcom** do Ancestris adopta o estilo indentado, mas não mostra os números de nível. Também adiciona símbolos para mostrar ou ocultar sub-etiquetas dos vários ramos.

O mesmo exemplo indentado, mas sem os números de nível :

```
@I3@ INDI (etiqueta principal deste disco: individual I3)
  NAME Jean Martin (nome do indivíduo)
  SEX M (sexo do indivíduo: masculino)
  BIRT (nascimento do indivíduo)
    DATE 16 Abril 1951 (data: 16 de Abril de 1951)
  FAMC @F5@ (família da qual descende o indivíduo I3: família F5)
```

Composição de uma linha num registo

Linha padrão

Cada linha de um registo contém essencialmente os seguintes elementos:

- o número de nível (de 0 a n);
- a **etiqueta** que indica a natureza da informação contida na linha;
- a informação associada à etiqueta em questão.

Por exemplo:

- Linha `2 DATE 16 Abril 1951` pode ler-se como segue: linha de nível 2, tipo de informação DATA, conteúdo 16 de Abril de 1951

Referência a outra entidade

Algumas linhas também contêm uma **referência a outra entidade**, que consiste num número enquadrado por dois símbolos @. Esta referência é um marcador especial cujo papel é diferente de acordo com o lugar que ocupa em relação à etiqueta da linha:

- uma referência localizada à esquerda da etiqueta indica o número do registo actual (um número que é sempre único na categoria de entidade a que pertence): isto ocorre apenas no nível 0 da linha do registo;
  - exemplo: `0 @I3@ INDI` - linha principal da entidade que é objecto deste registo, número de identificação deste registo - I3, categoria da entidade - indivíduo;
- uma referência localizada à direita da etiqueta indica o número de um outro registo, e refere-se a este último de modo a relacioná-lo com o registo actual;

- exemplo: `1 FAMC @F5@` - linha nível 1, etiqueta FAMC (família da qual o indivíduo actual é descendente) e referência F5 (ou seja, o indivíduo actual é descendente da família F5).

Norma Gedcom

A norma Gedcom refere-se ao conjunto de regras que regem o que pode e não pode ser feito para garantir que todos armazenem informação genealógica de uma certa forma. Esta é a chamada gramática da linguagem Gedcom.

Existem duas normas principais, 5.5 e 5.5.1, sendo a segunda uma ligeira evolução da primeira. Algumas coisas permitidas na primeira já não são permitidas na segunda e vice-versa. Estas diferenças são no entanto, limitadas.

> O Ancestris pode gerir as normas 5.5 e 5.5.1.

Encontrará **ao fundo da página** várias ligações que resumem o conjunto da documentação para a norma Gedcom. Propomos aqui uma lista dos pontos essenciais da norma e sua utilização no **Ancestris**. ***As ligações são para páginas em inglês***.

Norma Gedcom 5.5

Índice de conteúdos

Carta de William S. Harten

Tabela de tipos de dados - Página 1 - Página 2

Introdução

- **Objectivo e conteúdo da norma GEDCOM**
- **Objectivos da versão 5.x**
- **Modificações da versão 5.5 resultantes do estudo da versão (rascunho) 5.4**
- **Adições ou modificações ao rascunho 5.4**
- **Modificações à versão 5.3**

Capítulo 1: Gramática da representação dos dados

- **Conceitos**
- **Gramática**

- Descrição dos componentes gramaticais

Capítulo 2: Gramática ligada à progenitura

- Estruturas de registo dos ficheiros ligados à progenitura
- Infraestrutura dos ficheiros ligados à progenitura
- Elementos-base dos ficheiros ligados à progenitura
- Compatibilidade com outras versões Gedcom
- Estrutura do ficheiro de transmissão Gedcom
- Exemplo da transmissão de GEDCOM ligado à progenitura

Capítulo 3: Utilização de conjuntos de caracteres nos GEDCOM

- 8-Bit ANSEL
- ASCII (Versão USA)
- UNICODE (ISO 10646)

Capítulo 4: registo de um produto GEDCOM

Apêndice A : Definição de etiquetas Gedcom ligadas à progenitura

Apêndice B : Referências cruzadas

- Estrutura das referências cruzadas
- Referências cruzadas de base

Apêndice C : Códigos de Templos SUD

Apêndice D : Conjunto de caracteres ANSEL

- Caracteres gráficos sem espaçamento
- Caracteres gráficos com espaçamento

Apêndice E: Codificar/Descodificar objectos multimédia

- Introdução
- Codificar
- Descodificar

Norma Gedcom 5.5.1

Pode consultar a norma Gedcom 5.5.1 sob o formato pdf em inglês: **Norma Gedcom 5.5.1**. Curiosamente, as duas normas não estão disponíveis no mesmo formato.

Neste mesmo documento encontrará uma comparação entre as duas normas.

Norma Gedcom 7.0.x

> Esta norma foi difundida em 2021.

As especificações para esta norma podem ser encontradas na página **The FamilySearch GEDCOM Specification**.

Ligações úteis

- **Gedcom** : página Gedcom da Wikipédia;
- **The Gedcom Standard Release 5.5** (em inglês);
- **ptexier.pdf** sítio de Patrick TEXIER (PDF), conselhos e explicações da norma Gedcom;
- **The Gedcom Standard Release 5.5 (version HTML)** : o documento de referência (em inglês).

Revision #8
Created 10 August 2020 08:08:00 by pedro
Updated 27 July 2026 04:01:41 by pedro