Gedcom

Le terme **Gedcom** désigne un format d'échange de données généalogiques.

Il a été développé à l'origine par l'Église des Mormons pour des raisons religieuses, ensuite récupéré par les généalogistes, pour échanger des données généalogiques entre différentes personnes qui n'avaient pas les mêmes systèmes.

C'est donc en quelque sorte un langage.

> Le mot **Gedcom**, qui est un acronyme de **ge**nealogical **d**ata **com**munication (communication de données généalogiques), s'écrit comme un sigle : GEDCOM. Par dérivation métonymique, le mot désigne également un **fichier de généalogie au format Gedcom**. Le fichier `xxxxx.ged` sur lequel vous travaillez dans Ancestris, est un **Gedcom**.

> Depuis le milieu des années '90, avec l'avènement d'Internet et la multiplication des échanges numériques, la spécification Gedcom est progressivement devenue une norme incontournable pour la plupart des logiciels et sites de généalogie. Cependant, certains d'entre eux ne respectent pas totalement le format et adaptent celui-ci en développant des commandes propriétaires. Ancestris quant à lui est totalement compatible Gedcom (version 5.5 et 5.5.1), ce qui signifie qu'il permet à ses utilisateurs d'échanger et partager des fichiers de généalogie parfaitement fiables, sans risque de pertes de données.

Caractéristiques d'un fichier Gedcom

Un **fichier au format Gedcom** est un fichier texte (*.ged), c'est-à-dire un fichier pouvant être ouvert pour lecture seule ou pour modification depuis n'importe quel éditeur de texte (tel que Notepad, Kate, Kwrite, Gedit, etc.).

Par voie de conséquence, un tel fichier peut être utilisé tel quel par n'importe quel logiciel de généalogie, installé sous n'importe quel système d'exploitation, sans avoir besoin de convertir quoi que ce soit.

Les différentes informations contenues dans le fichier Gedcom sont précédées d'une ***étiquette*** (**tag** en anglais): dans la norme Gedcom, ce tag est un marqueur composé de trois ou quatre lettres capitales, toujours associé au même type d'information.

- Par exemple, le tag PLAC (= *place*, c'est-à-dire *lieu*) indique toujours que l'information qu'il annonce est un *lieu* (lieu de naissance, lieu de décès, lieu d'une cérémonie, etc.)

Enregistrements d'un fichier Gedcom

Un fichier Gedcom contient un ensemble d'enregistrements, dont le premier et le dernier sont d'un type particulier :

- Le premier enregistrement s'appelle l'*en-tête* (tag HEAD) ;
- Le dernier enregistrement s'appelle le *marqueur de fin de fichier* (tag TRLR).

Chacun des autres enregistrements appartient à une catégorie d'**entité**, étant bien entendu que chacune de ces catégories possède ses fonctions et ses structures propres.

Un fichier Gedcom contient 7 catégories d'entité. Les enregistrements qu'on peut trouver dans un fichier Gedcom sont donc les suivants :

- Des enregistrements décrivant des **individus** (tag INDI) ;
- Des enregistrements décrivant des **familles** (tag FAM) ;
- Des enregistrements décrivant des **notes** (tag NOTE) ;
- Des enregistrements décrivant des **sources** (tag SOUR) ;
- Des enregistrements décrivant des **dépôts** (tag REPO) ;
- Des enregistrements décrivant des **éléments multimédias** (tag OBJE) ;
- Des enregistrements décrivant des **fournisseurs d'informations** (tag SUBM).

Le choix de considérer ces 7 catégories de donnée comme étant des enregistrements est arbitraire bien sûr, mais c'est le principe d'une norme.

On pourrait facilement imaginer d'autres types d'enregistrements, comme les lieux par exemple. Le fait qu'un lieu ne soit pas une entité à part n'empêche pas Ancestris de les gérer.

Arborescence d'un enregistrement

Chaque enregistrement se présente de manière arborescente : chaque *tag* peut comprendre un nombre quelconque de *sous-tags.*

Des sous-tags sont des tags hiérarchiquement dépendants du tag de niveau immédiatement supérieur, et susceptibles de comprendre à leur tour un ou plusieurs sous-tags, etc.

Chaque niveau peut ainsi se subdiviser à l'infini.

Niveaux hiérarchiques

Les hiérarchie sont numérotées.

Comme chaque ligne doit rester impérativement à sa place du point de vue de la hiérarchie, chacune d'elles est affectée d'un numéro correspondant au niveau qu'elle occupe dans l'arborescence de l'enregistrement.

C'est ainsi que la ligne du niveau principal de chaque enregistrement (c'est-à-dire le niveau *zéro*) porte le numéro 0 ; une ligne située au niveau immédiatement inférieur porte le numéro 1 ; une ligne située au niveau immédiatement inférieur au niveau précédent porte le numéro 2 ; et ainsi de suite.

Identifiant et catégorie des niveaux zéros

Le niveau *zéro* de chaque enregistrement (autre que les deux enregistrements extrêmes, HEAD et TRLR) est composé de deux éléments accolés :

- Le numéro ID de l'entité encadré de deux arobases (@),
- Le tag associé à la catégorie à laquelle appartient l'entité en question.
  - Par exemple, l'en-tête @I24@INDI signifie que l'enregistrement a pour numéro ID le I24, et qu'il appartient à la catégorie des individus (autrement dit, *individu numéro I24*).

Indentation

Pour plus de clarté, les différentes lignes d'un enregistrement peuvent être **indentées** (affectées d'un ou plusieurs espaces à gauche du premier caractère), de manière à repérer plus facilement la place qu'occupe chacune d'elles dans la hiérarchie.

- Exemple non indenté :

```
0 @I3@ INDI (tag principal de cet enregistrement : individu I3)
1 NAME Jean Martin (nom de l'individu)
1 SEX M (sexe de l'individu : masculin)
1 BIRT (naissance de l'individu)
2 DATE 16 avril 1951 (date : 16 avril 1951)
1 FAMC @F5@ (famille dont descend l'individu I3 : famille F5)
```

- Le même exemple indenté :

```
0 @I3@ INDI (tag principal de cet enregistrement : individu I3)
  1 NAME Jean Martin (nom de l'individu)
```

```
  1 SEX M (sexe de l'individu : masculin)
  1 BIRT (naissance de l'individu)
    2 DATE 16 avril 1951 (date : 16 avril 1951)
  1 FAMC @F5@ (famille dont descend l'individu I3 : famille F5)
```

L'**éditeur Gedcom** d'Ancestris adopte un affichage indenté, mais ne fait pas apparaître les numéros de niveau. Il ajoute en outre des poignées permettant d'afficher ou de masquer les sous-tags de telle ou telle branche.

- Le même exemple indenté, mais sans les numéraux de niveaux :

```
@I3@ INDI (tag principal de cet enregistrement : individu I3)
  NAME Jean Martin (nom de l'individu)
  SEX M (sexe de l'individu : masculin)
  BIRT (naissance de l'individu)
    DATE 16 avril 1951 (date : 16 avril 1951)
  FAMC @F5@ (famille dont descend l'individu I3 : famille F5)
```

Composition d'une ligne dans un enregistrement

Ligne standard

Chaque ligne d'un enregistrement contient essentiellement les éléments suivants :

- Le numéro de niveau (de 0 à n),
- Le **tag** indiquant la nature des informations contenues sur la ligne,
- Les informations associées au tag en question.

Exemple :

- La ligne `2 DATE 16 avril 1951` peut se lire ainsi : ligne de niveau *2*, information de type *DATE*, et de contenu *16 avril 1951*

Référence à une autre entité

Certaines lignes contiennent en outre une **référence à une autre entité**, laquelle consiste en un numéro encadré par deux arobases (@). Cette référence constitue un marqueur spécial dont le rôle

est différent selon la place qu'il occupe par rapport au tag de la ligne.

- Une référence située à gauche du tag indique le numéro de l'enregistrement courant (numéro toujours unique dans la catégorie d'entité dont il relève) : ce cas de figure ne se produit que sur la ligne de niveau 0 de l'enregistrement. Exemple :
  - 0 @I3@ INDI : ligne principale de l'entité faisant l'objet de cet enregistrement, numéro ID de cet enregistrement : *I3*, catégorie d'entité : *individu*.
- Une référence située à droite du tag, indique le numéro d'un autre enregistrement, et renvoie à ce dernier afin de le mettre en relation avec l'enregistrement courant. Exemple :
  - 1 FAMC @F5@ : ligne de niveau 1, tag *FAMC* (famille dont descend l'individu courant) et référence *F5* (autrement dit : l'individu courant descend de la famille F5)

Norme Gedcom

La norme Gedcom désigne l'ensemble des règles qui régissent ce qu'il est possible de faire et ne pas faire pour que tout le monde range les informations généalogiques d'une certaine façon. C'est donc la grammaire du langage Gedcom.

Deux normes principales existent, 5.5 et 5.5.1, la seconde étant une légère évolution de la première. Des choses permises dans la première ne le sont plus dans la seconde, et vice-versa. Ces différences sont néanmoins limitées.

> Ancestris sait gérer les normes 5.5 et 5.5.1.

Vous trouverez **en bas de page** plusieurs liens qui rassemblent l'ensemble de la documentation que l'on a trouvé sur les normes Gedcom.

Nous vous proposons ici une traduction des points essentiels de la norme et leur utilisation dans **Ancestris**.

Norme Gedcom 5.5

Vous trouverez ici le détail de toute la norme 5.5 sous forme de liens web.

Tables des matières

Lettre de William S. Harten

Tableau type des données - Page 1 - Page 2

Introduction

- **Objet et Contenu du Standard GEDCOM**

- Fonctionnalités de la Version 5.x
- Modifications de la Version 5.5 résultant de l’Étude de la Version (d'essai) 5.4
- Ajouts et Modifications par rapport à la Version d'Essai 5.4
- Modifications par rapport à à la Version 5.3

Chapter 1: Grammaire de la Représentation des données

- Concepts
- Grammaire
- Description des Composants Grammaticaux

Chapter 2: Grammaire Liée à la Parenté

- Structures d'Enregistrement du Fichier Lié à la Parenté
- Infrastructure du Fichier Lié à la Parenté
- Éléments de Base du Fichier Lié à la Parenté
- Compatibilité avec les Autres Versions Gedcom
- Structure du Fichier de Transmission Gedcom
- Exemple de la Transmission de GEDCOM lié à la parenté

Chapter 3: Utilisation des jeux de caractères dans GEDCOM

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

Chapter 4: Enregistrement de la Production GEDCOM

Appendice A : Définition du Tag Gedcom Lié à la Parenté

Appendice B : Références Croisées

- Structure des Références Croisées
- Références Croisées de Base

Appendice C : Codes LDS Temple

Appendice D : Jeu de Caractères ANSEL

- **Caractères graphiques sans déplacement**
- **Caractères graphiques avec déplacement**

Appendie E : Encoder/Décoder Objets Multimedia

- **Introduction**
- **Encoder**
- **Décoder**

Norme Gedcom 5.5.1

Vous pouvez aussi consulter la norme Gedcom 5.5.1 diffusée en 1999, disponible ici sous forme de fichier pdf en anglais : **Norme Gedcom 5.5.1**.

Vous trouverez dans ce même document un comparatif entre les deux normes.

Norme Gedcom 7.0.18

> Cette norme a été diffusée en 2021.

Les spécifications de cette norme se trouve sur la page **The FamilySearch GEDCOM Specification**.

Revision #26
Created 20 October 2019 10:32:10
Updated 26 July 2026 12:10:43 by frederic