

# Guide Détaillé des Options gwdiff

Date de création : 11 septembre 2025 Outil : gwdiff - Comparaison entre deux bases GeneWeb Version analysée : master

---

## Table des matières

1. Vue d'ensemble
  2. Principe de comparaison
  3. Options et syntaxe
  4. Algorithmes et critères de compatibilité
  5. Sortie et interprétation des résultats
  6. Exemples
  7. Performance et mémoire
  8. Dépannage et messages d'erreur
  9. Bonnes pratiques
- 

## 1. Vue d'ensemble

gwdiff aide à cibler les différences entre deux bases GeneWeb. L'outil compare des individus et leurs familles selon deux modes: - Mode descendants (-d): compare les descendants (et aussi les conjoints et leurs parents) des deux individus racines, un dans chaque base. - Mode ascendants puis descendants (-ad): trouve les ascendants communs les plus anciens dans les deux bases, puis compare leurs descendants.

Le rapport est produit en texte (stdout) ou en HTML (option -html) avec des liens vers les individus dans chaque base.

---

## 2. Principe de comparaison

Pour un couple d'individus de référence (un dans base1, un dans base2), gwdiff: - compare les informations de la personne (prenoms, noms, sexe, naissance, décès, profession), - compare les familles (mariage, divorce, lieu de mariage), - tente d'apparier conjoints et enfants selon des compatibilités de nom et d'événements, - descend récursivement sur la descendance (et, en mode -ad, à partir d'ancêtres communs).

Les différents écarts identifiés sont affichés sous forme de messages.

---

### 3. Options et syntaxe

#### Synopsis

`gwdiff [options] base1 base2`

- `base1`: chemin vers la base GeneWeb de reference
- `base2`: chemin vers la base GeneWeb a comparer

#### Options

- `-1 <fn> <occ> <sn>`: (obligatoire) personne de depart dans base1
  - `<fn>`: prenom
  - `<occ>`: occurrence (entier, 0 si premiere occurrence)
  - `<sn>`: nom de famille
- `-2 <fn> <occ> <sn>`: (obligatoire) personne de depart dans base2
- `-d`: mode descendants (default si `-ad` n'est pas fourni)
- `-ad`: mode descendants des ascendants communs
- `-html <root>`: sortie HTML; construit des liens vers individus, prefixes par `<root>`
  - Exemple de lien genere: `<root><base>_w?i=<id>`
- `-mem`: economise la memoire (plus lent);
  - charge uniquement les tableaux necessaires (ascendants, strings; familles/couples/descendants charges selon besoin reduit)

#### Rappels d'utilisation et erreurs communes

- Les deux options `-1` et `-2` sont obligatoires et doivent recevoir 3 valeurs chacune.
  - Les deux bases doivent etre fournies et accessibles.
  - En cas d'argument manquant, `gwdiff` affiche l'usage et quitte avec un code d'erreur.
- 

### 4. Algorithmes et criteres de compatibilite

#### 4.1 Compatibilite des noms

- Prenoms: le prenom de `base1` doit apparaitre parmi le prenom ou les alias de prenom de l'individu candidate dans `base2` (comparaison insensible a la casse, normalisee).
- Noms: de meme, le nom de `base1` doit apparaitre parmi le nom ou ses alias dans `base2`.
- Messages possibles: `first name, surname`.

#### 4.2 Compatibilite des dates

- Dates manipulees sous forme de `cdate/date` (Def.Date).

- Compatibilite de deux dates:
  - Comparaison calendaire (changement de calendrier tolere dans certains cas, ex. gregorien vs julien/francais selon code).
  - Gestion des imprecisions (About, Maybe, Before, After, OrYear, YearInt), traduites en intervalles SDN et comparees par inclusion.
- Messages possibles: `birth date`, `death (status or date)`, `marriage date`.

#### 4.3 Compatibilite des lieux et champs texte

- Lieu de naissance, deces, mariage: si le champ est defini dans base1, il doit etre defini dans base2 (quelle que soit la valeur pour etre considere compatible).
- Profession: meme regle.
- Messages possibles: `birth place`, `death place`, `marriage place`, `occupation`.

#### 4.4 Compatibilite du deces

- Etats consideres compatibles si equivalence faible (ex. inconnu vs inconnu) ou si deux dates compatibles en cas de deces avec date.
- Messages possibles: `death (status or date)`.

#### 4.5 Compatibilite des unions et familles

- Appariement des conjoints par compatibilite de noms (leger).
- Mariage: compatibilite de la date de mariage, du divorce, et du lieu de mariage.
- Messages possibles: `divorce`, `marriage date`, `marriage place`.

#### 4.6 Descendance et recursivite

- En mode `-d`: pour chaque union apparaitée, les enfants de base1 sont recherches dans la liste des enfants correspondants de base2 par compatibilite (noms -> puis verification complete). En cas d'ambiguite, messages:
  - `child missing: <personne>` si aucun enfant compatible trouve.
  - `can not isolate one child match: <personne>` si aucun meilleur appariement clair.
  - `more than one child match: <personne>` si plusieurs candidats potentiels.
- En parallele, les conjoints manquants ou multiples declenchent:
  - `spouse missing: <personne>` ou `more than one spouse match: <personne>`.
- Les parents des individus compares sont verifiees (si disponibles) avec messages possibles:
  - `parents missing` si l'un des couples parentaux est absent dans base2.

#### 4.7 Protection contre re-comparaisons

- Marquage des couples déjà comparés pour éviter des boucles ou répétitions (table de marqueurs sur iper1 avec liste des iper2 déjà visités).
- 

### 5. Sortie et interpretation des resultats

- Format texte (default): lignes indiquant les individus/familles comparés suivies des messages de différences indentifiées.
  - Format HTML (-html): memes informations avec liens cliquables vers les individus dans chaque base (utilise <root> et le nom de la base pour construire l'URL).
  - Exemple de tête d'une comparaison de personnes:
    - **First.Last / Other.Last** puis une liste de messages: **birth date**, **occupation**, etc.
  - Exemple de tête d'une comparaison de couples:
    - **Father x Mother (base1) / Father x Mother (base2)**, puis messages comme **marriage date**.
- 

## 6. Exemples

### 6.1 Mode descendants

```
gwdiff -d -1 Jean 0 DUPONT -2 Jean 3 DUPONT b1 b2
```

- Compare les descendants de **Jean.0 DUPONT** (base b1) avec ceux de **Jean.3 DUPONT** (base b2), y compris conjoints et parents des conjoints.

### 6.2 Mode ascendants puis descendants

```
gwdiff -ad -1 Jean 0 DUPONT -2 Jean 3 DUPONT b1 b2
```

- Recherche d'abord les ancêtres communs les plus anciens et compare leurs descendants.

### 6.3 Rapport HTML

```
gwdiff -d -html http://localhost:2317/ -1 Marie 0 MARTIN -2 Marie 1 MARTIN b1 b2
```

- Génère des liens cliquables vers chaque individu comparé dans les deux bases, avec le préfixe fourni.
-

## 7. Performance et memoire

- `-mem` active un mode economique en memoire (moins de tableaux charges completement), ce qui peut ralentir la comparaison sur certaines bases mais reduit l'empreinte memoire.
  - L'algorithme evite de revisiter des couples deja compares via un marquage, ce qui limite les couts en profondeur.
- 

## 8. Depannage et messages d'erreur

- `Missing reference data base` ou `Missing destination data base`: un des deux chemins de base est manquant.
- `-1 parameter is mandatory` / `-2 parameter is mandatory`: la definition d'au moins une personne de depart est absente ou incomplete.
- `Cannot find person ... in reference/destination base`: la cle personne n'existe pas dans la base.
- `Too much arguments`: plus de deux bases passees en arguments.

Conseils: - Verifier les prenom et noms (casse, accents), l'occurrence (0 si premiere occurrence). - S'assurer que les bases sont bien accessibles et chargeables par GeneWeb.

---

## 9. Bonnes pratiques

- Commencer par un individu dont l'identite est certaine dans les deux bases.
- Utiliser le mode `-ad` pour explorer des ecarts genealogiques plus larges a partir des ancetres communs.
- Activer `-html` pour une revue visuelle rapide avec navigation.
- Utiliser `-mem` sur des machines a memoire limitee ou des bases tres volumineuses.