

# Project Structure

This project is based on a legacy codebase that was originally derived from a GeneWeb system written in OCaml. The folder structure has been carefully designed to ensure the code is both maintainable and scalable, while also preserving compatibility with the original system.

Each folder in the repository **must contain a `README.md` file** describing its purpose, usage, and relevant notes. This helps new contributor and maintainers to quickly understand the role of each module.

## Directory Tree

```
python/  
  README.md      → Main project documentation  
tests/  
  README.md      → Test project documentation  
  config/        → Test configuration files  
  unit/          → Unit test suites  
  common/        → Shared utilities and helpers for tests  
  performance/   → Performance and benchmarking tests  
  golden_master/ → Reference outputs used to verify that two binaries pr  
                   oduce identical results, ensuring backward compatibility and preventing re  
                   gressions.  
  compatibility/ → Compatibility tests (versions, dependencies, formats)  
  reports/       → Generated test reports (logs, HTML, XML...)  
scripts/  
  README.md      → Scripts project documentation  
  consang/  
    README.md    → Docs for consanguinity calculation module  
  gwb2ged/  
    README.md    → Docs for GeneWeb → GEDCOM converter  
  ged2gwb/  
    README.md    → Docs for GEDCOM → GeneWeb converter  
  gwc/
```



## Directory Details

### python/

- **Integration root** containing the GeneWeb binaries, test infrastructure, and utility scripts.
- Contains:
  - `tests/` → Automated validation modules (unit, performance, compatibility, golden master, etc.).
  - `scripts/` → Build, cleanup, and migration utilities.
  - **Binaries** → Core GeneWeb tools.

### tests/

- Core of the automated validation system:
  - **config/** → Test configuration files.
  - **unit/** → Unit test suites.
  - **common/** → Shared test utilities and helpers.
  - **performance/** → Performance benchmarks.
  - **golden\_master/** → Reference "golden files" for regression detection.
  - **compatibility/** → Cross-version and dependency compatibility checks.
  - **reports/** → Generated test reports (logs, HTML, XML).

### scripts/

- General-purpose scripts for maintenance, data generation, and deployment.
- Includes both Python and shell scripts.

### consang/

- Algorithms for **consanguinity (inbreeding) calculations** in genealogical trees.

### gwb2ged/

- Converts **GeneWeb** → **GEDCOM**.
- Useful for exporting genealogical data into standard GEDCOM format.

### ged2gwb/

- Converts **GEDCOM** → **GeneWeb**.
- Essential for importing genealogical data from external sources.

### gwc/

- **GeneWeb Compiler**: manages and creates the genealogical databases.