

# Project File Map

[Русский](#) · [Reference](#) · [Install](#)

Where everything lives. This document is for whoever maintains the project; a user does not need it.

Working folders (`input`, `output`, `logs`, `report`, `release`, `workspace`, `runtime`, `wheelhouse`) are described by role rather than file by file — their contents are produced by runs and rebuilt.

## Root

| File                                 | What it is                              |
|--------------------------------------|-----------------------------------------|
| <code>Start.exe</code>               | Compiled window launcher                |
| <code>launcher_gui.cmd</code>        | Window launcher                         |
| <code>launcher_project.cmd</code>    | Project menu, English                   |
| <code>launcher_project_ru.cmd</code> | Project menu, Russian                   |
| <code>launcher_tools.cmd</code>      | Service menu: doctor, licenses, release |
| <code>builder_main.cmd</code>        | Portable environment build              |
| <code>cleanup_project.cmd</code>     | Cleanup of generated artifacts          |
| <code>pytest.ini</code>              | Test settings                           |
| <code>release_manifest.json</code>   | Manifest of the last release build      |
| <code>README.md</code>               | Public project description              |

CMD files must stay UTF-8 without BOM with CRLF line endings. After editing, run `install\Check-CmdEncoding.cmd`.

---

## config — Settings

---

| File                                      | What it is                                                                             |
|-------------------------------------------|----------------------------------------------------------------------------------------|
| <code>project.yaml</code>                 | Engine settings: city, reference column, target columns, linked columns, safety limits |
| <code>gui_settings.yaml</code>            | Language, theme, window behaviour                                                      |
| <code>tool_manifest.yaml</code>           | Every command and field of the window; the interface is built from it                  |
| <code>version.json</code>                 | Version and the language of the root README                                            |
| <code>ui_colors.yaml</code>               | Interface colours                                                                      |
| <code>path_history.json</code>            | History and pins of source and target paths                                            |
| <code>oktmo_current_keys.txt</code>       | Current OKTMO search keys                                                              |
| <code>oktmo_region_pins.json</code>       | Pinned regions                                                                         |
| <code>oktmo_municipality_pins.json</code> | Pinned municipalities                                                                  |
| <code>oktmo_pins_bundle.json</code>       | The combined pin set                                                                   |

---

## system\_core — Code

---

### Top Level

| File                                | What it does                                                         |
|-------------------------------------|----------------------------------------------------------------------|
| <code>Ultimate_GT_Aligner.py</code> | The matching engine: address parsing, passes, linked column transfer |
| <code>safe_table_join.py</code>     | Comparison of two workbooks and column transfer through Excel        |
| <code>excel_layout.py</code>        | Result workbook formatting: column widths, text wrap, fill           |
| <code>remove_empty_rows.py</code>   | Removal of entirely empty rows                                       |
| <code>Parser_Diagnostic.py</code>   | Address parsing diagnostics                                          |

| File                              | What it does                           |
|-----------------------------------|----------------------------------------|
| <code>inspect_md_tables.py</code> | Markdown table inspection              |
| <code>doctor.py</code>            | Environment and import checks          |
| <code>main.py</code>              | Direct matching run without the window |
| <code>fzf.exe</code>              | Menu for the CMD launchers             |

## address\_engine — The Address Engine

| File                                  | What it does                                                       |
|---------------------------------------|--------------------------------------------------------------------|
| <code>audion_address_core.py</code>   | Address decomposition, recognition of address-like text            |
| <code>address_slots.py</code>         | Slots: definition, filling, address line assembly                  |
| <code>slot_fill.py</code>             | Filling slots from evidence                                        |
| <code>models.py</code>                | Data structures: observation, address parts, fingerprint, cluster  |
| <code>source_intake.py</code>         | Reading sources in every supported format                          |
| <code>source_table_assembly.py</code> | Collecting addresses from sources into a workbook                  |
| <code>source_cleanup.py</code>        | Filtering out Office temporary files                               |
| <code>reference_builder.py</code>     | Building the reference directory; slot and column definitions      |
| <code>reference_processor.py</code>   | Sorting and decomposing the directory by slots                     |
| <code>oktmo_lookup.py</code>          | OKTMO index, disk cache, region / municipality / settlement lookup |
| <code>oktmo_user_keys.py</code>       | The project key file and pins                                      |
| <code>rosstat_update.py</code>        | Downloading and updating the Rosstat registry                      |
| <code>region_reference.py</code>      | Region reference                                                   |
| <code>russian_morphology.py</code>    | Grammatical forms of names                                         |
| <code>universal_address.py</code>     | The common address normalizer                                      |
| <code>normalizer_runtime.py</code>    | Configuring the normalizer per run                                 |

| File                                    | What it does        |
|-----------------------------------------|---------------------|
| <code>address_diagnostics.py</code>     | Diagnostic extracts |
| <code>legacy_office_converter.py</code> | DOC/XLS conversion  |

## services — The Service Layer

`address_aligner_service.py` is the only file. Every window command points to a function here: the service takes the parameters, calls the engine, and writes the log, the report, and the artifact list.

## ui\_nicegui — The Window

| File                                                                     | What it does                                              |
|--------------------------------------------------------------------------|-----------------------------------------------------------|
| <code>window.py</code>                                                   | Window wrapper: port, icon, WebView2 engine               |
| <code>app.py</code>                                                      | The interface itself: commands, fields, log, OKTMO panels |
| <code>workbench.py</code>                                                | Source and target, path history, pins                     |
| <code>i18n.py</code>                                                     | Language switching                                        |
| <code>tokens.css</code> , <code>base.css</code> , <code>theme.css</code> | Styling                                                   |

## core — Core

| File                          | What it does                               |
|-------------------------------|--------------------------------------------|
| <code>jobs.py</code>          | Running operations, cancellation, progress |
| <code>manifest.py</code>      | Parsing <code>tool_manifest.yaml</code>    |
| <code>paths.py</code>         | Managed project folders                    |
| <code>config.py</code>        | Reading settings                           |
| <code>ui_settings.py</code>   | Window settings                            |
| <code>ansi_terminal.py</code> | Terminal output into the log               |
| <code>cmd_encoding.py</code>  | CMD encoding checks                        |
| <code>output_decode.py</code> | Decoding subprocess output                 |

| File                          | What it does |
|-------------------------------|--------------|
| <code>logging_utils.py</code> | Logging      |

## Other

| Path                         | What it is                                                     |
|------------------------------|----------------------------------------------------------------|
| <code>license/</code>        | Collection, deduplication, and pruning of third-party licenses |
| <code>icons/</code>          | Window icons                                                   |
| <code>start_launcher/</code> | Source of <code>Start.exe</code>                               |
| <code>powershell/</code>     | Helper scripts                                                 |

## install — Install and Build

| File                                                                     | What it does                                   |
|--------------------------------------------------------------------------|------------------------------------------------|
| <code>Build_Portable_Env.cmd</code> / <code>.ps1</code>                  | Portable environment build                     |
| <code>install_portable_offline.cmd</code>                                | Install from the local <code>wheelhouse</code> |
| <code>verify_portable_env.cmd</code>                                     | Verify the built environment                   |
| <code>init_folders.cmd</code>                                            | Create the working folders                     |
| <code>Check-CmdEncoding.cmd</code> , <code>Repair-CmdEncoding.ps1</code> | Check and repair CMD encodings                 |
| <code>Clean-Install-Cache.cmd</code> / <code>.ps1</code>                 | Clean the build cache                          |
| <code>Update-Requirements-Lock.cmd</code>                                | Update the dependency lock file                |
| <code>make_release_archive.cmd</code>                                    | Build the release archive                      |
| <code>Build-StartLauncher.cmd</code> / <code>.ps1</code>                 | Build <code>Start.exe</code>                   |
| <code>launcher-tools-update_fzf.cmd</code>                               | Update <code>fzf.exe</code>                    |
| <code>requirements_full.in</code>                                        | Dependency list                                |

| File                           | What it does  |
|--------------------------------|---------------|
| <code>README_INSTALL.md</code> | Install notes |

## data — Data

| Path                                         | What it is                                               |
|----------------------------------------------|----------------------------------------------------------|
| <code>regions_reference.yaml</code>          | Region reference                                         |
| <code>rosstat/data-*.csv</code>              | The OKTMO registry, downloaded by a button in the window |
| <code>rosstat/oktmo_lookup_cache.json</code> | The parsed OKTMO index                                   |

## docs — Documentation

| File                                                          | What it is                                    |
|---------------------------------------------------------------|-----------------------------------------------|
| <code>README_RU.md</code> , <code>README_EN.md</code>         | Project description, the entry point          |
| <code>USER_GUIDE_RU.md</code> , <code>USER_GUIDE_EN.md</code> | User guide                                    |
| <code>OKTMO_RU.md</code> , <code>OKTMO_EN.md</code>           | The OKTMO registry, keys, and pins            |
| <code>REFERENCE_RU.md</code> , <code>REFERENCE_EN.md</code>   | Technical reference                           |
| <code>INSTALL_RU.md</code> , <code>INSTALL_EN.md</code>       | Install and launch                            |
| <code>tools/PROJECT_FILES_GUIDE_*.md</code>                   | This document                                 |
| <code>PDF/</code>                                             | Rendered PDFs of every document in two themes |

PDFs are built by a separate tool from every `.md` in `docs`, nested folders included. The folder structure is preserved.

## tests — Tests

| File                                            | What it covers             |
|-------------------------------------------------|----------------------------|
| <code>test_address_aligner_regression.py</code> | The matching passes        |
| <code>test_safe_table_join.py</code>            | Moving columns by address  |
| <code>test_excel_layout.py</code>               | Result workbook formatting |
| <code>test_gui_render.py</code>                 | Window rendering           |
| <code>test_workbench_route_regression.py</code> | Source and target          |
| <code>test_ansi_terminal.py</code>              | Terminal output            |

The window is tested with NiceGUI's headless fixture, parsing the rendered page with BeautifulSoup; selenium is deliberately not used.

## licenses — Licenses

License texts of third-party components, `THIRD_PARTY_NOTICES.txt`, `COMPONENTS.json`, the scan report, and the pack manifest. Produced by the tools in `system_core/license/`.