Contributions

Below is a summary of the main contributors for each part of the Eprice project, based on the current project structure.

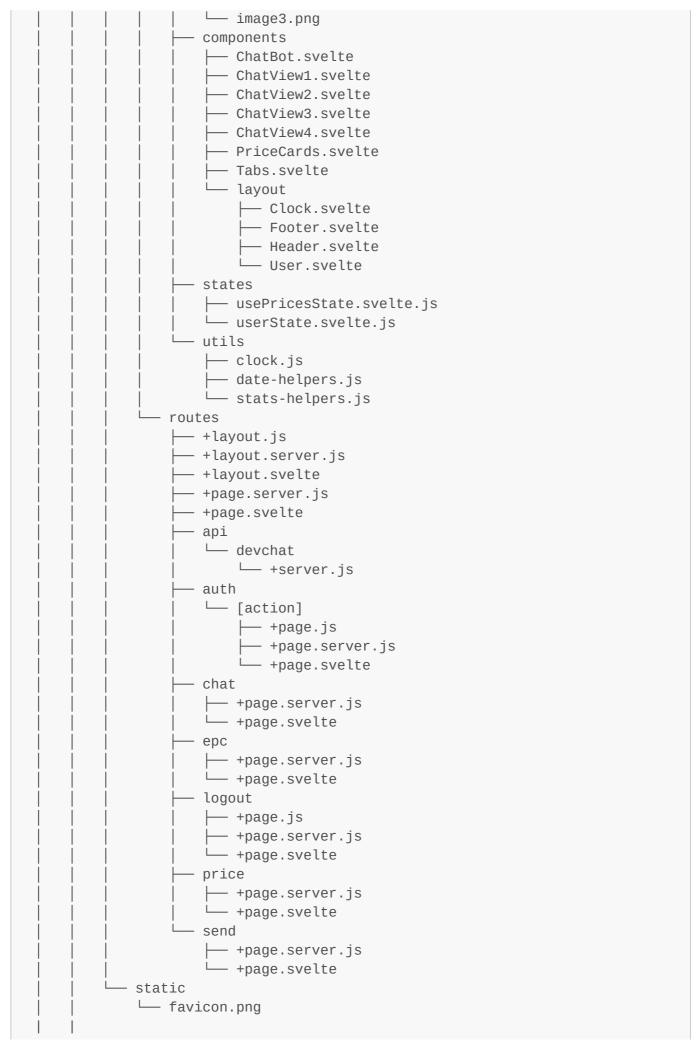
```
— LICENSE
  - README.md

    App

    - README.md
    — compose.yaml (anything with containers, Paavo)
    pgdata.tar.gz - database snapshot (Paavo)
    — project.env (anything with project and container level
configurations, Paavo)
     — backend-tests (setup, Paavo)
        Dockerfile
         — README.md
        └─ tests
            test_auth_controller.py (Juho & Paavo)

    test_data_controller.py (Juho)

     — chat-engine (Paavo)
        ─ Dockerfile
         — README.md
         — agent_app.py
        — agent_manager.py
        — app.py
        ├─ autoloading_uml.py
        — chat_app.py
         — chat_manager_with_tools.py
        ├── file_app.py
         — diagrams
        └─ utils
            ├─ __init__.py
            ├─ db_calls.py
              - helpers.py
            └─ tools.py
      - client (Paavo)
        ─ Dockerfile
         — README.md
          - src
            ├─ app.css
             — app.html
            ├─ hooks.server.js
              - lib
                — apis
                  └─ data-api.js
                  - assets
                    ├─ image12.jpg
```



```
- data-preparation (Paavo)
  - README.md
    scripts
      ├─ Dockerfile
      - README.md
      clean_porssisahko.py
       populate_porssisahko.py
      └─ retrieve_porssisahko_update.sh
— database-migrations (Paavo)
  ├─ V10__code_constraint.sql
  ├─ V11__documents.sql
  ├─ V12__files.sql
  ├─ V13__fingrid.sql
  ├─ V14__fingrid_load_entries.sql
  ├─ V1__users.sql
  ├─ V2__porssisahko.sql
  ├─ V3__timezone.sql
  ├─ V4__users_add_role.sql
  ├─ V5__porssisahko_load_entries.sql
  ├─ V6__users_add_isverified.sql
  ├─ V8__extension_vector.sql
  └─ V9__code.sql
– e2e-tests (setup, Paavo)
  ├─ Dockerfile
   tests
     └─ home.spec.js (Paavo)
- python-server (Juho & Paavo)
  ├─ Dockerfile
   — README.md
   -- main.py
   requirements.txt
   — confiq
     — __init__.py
     └─ secrets.py
   - controllers
      — auth_controller.py (Juho & Paavo)
     └─ data_controller.py (Juho)
   — ext_apis
     — models
     custom_exception.py (Juho)
        - data_model.py (Juho)
     └─ user_model.py (Paavo)
   — repositories
      fingrid_repository.py (Juho)
        porssisahko_repository.py (Paavo)
      └─ user_repository.py (Paavo)
    scheduled_tasks
     porssisahko_scheduler.py (Paavo)
    - services
      ─ auth_service.py (Paavo)
```

```
└─ data_service.py (Juho)
        - utils
          — email_tools.py (Paavo)
           — fingrid_service_tools.py (Juho)
            - porssisahko_service_tools.py (Juho)
           porssisahko_tools.py (Paavo)
- Documents
  — CONTRIBUTIONS.md

    CONTRIBUTIONS.pdf

  — KONTRIBUUTIOT.md

    KONTRIBUUTIOT.pdf

   — README.md
   backend_design.md
   frontend_description.md
   openapi_endpoint_descriptions.md
  — project_description.md
  — project_description.pdf
   project_directory_structure.txt
   – projektin_kuvaus.md
   projektin_kuvaus.pdf
   — diagrams
     └─ sources
          authentication_call_sequence_diagram.wsd

    authentication_class_diagram.wsd

           authentication_use_case.wsd

    data_access_call_sequence_diagram.wsd

           frontend_structure.wsd
            llm_retrieval.wsd
           services_diagram.wsd
            tool_calling.wsd
           use_case.wsd
```

Notebooks

Additionally, the project/developer chat's retrieval mechanism requires some amount of parsing/cleaning/loading (all project code and documents), which is done offline using scripts and document_loading.ipynb inside Eprice/Notebooks/. These are at least in some parts based on Paavo's (heh, me in 3rd person...) other previous and contemporary projects/works. Not everything is included in this repo, but will be later made available in https://github.com/PaavoReinikka.

To aid in the documentation of the project, there is also a UML diagram generator dashboard in Eprice/Notebooks/ (using PlantUML syntax). It was written with Python (see Notebooks README for details and instructions).

Developer chat

The developer chat is based on OpenAI's gtp-4o-mini, which is basically free for research (and such) purposes. There are also a lot of open-source generative text models (chat and completion) in HF, and also good datasets for fine-tuning. Some of the experiments, with various open models, during this project will eventually be published in https://github.com/PaavoReinikka.

The embedding model used here, is open-source from HuggingFace. There are many good models available at HF, and a there is a leader board for embedding models, which is a good place to start looking (https://huggingface.co/spaces/mteb/leaderboard).

Documentation

Most of the documentation was written by Paavo Reinikka, with some help from Juho Ahopelto, and with a lot of help from Copilot.