Project Report: PubMed Paper Fetcher CLI

Project Title: PubMed Paper Fetcher CLI

Developer: Nikhil V

GitHub Repo: https://github.com/Nikhilvh2600/get-papers-nikhilvh

TestPyPI Package: https://test.pypi.org/project/get-papers-nikhilvh/

Objective:

Build a CLI-based Python tool to fetch and filter PubMed research papers based on user queries.

The focus is on identifying papers authored by people affiliated with pharmaceutical or biotech companies.

Key Features:

- Uses the PubMed Entrez API to search and fetch article metadata

- Filters non-academic authors using keyword heuristics and affiliation checks

- Extracts emails using regex

- Outputs to CSV or prints to the console

- Fully packaged using Poetry and published to TestPyPI

Technical Highlights:

Project Report: PubMed Paper Fetcher CLI

- Typed Python using `typing` hints

- Modular code with separation of logic (`fetch.py`, `filter.py`, `output.py`, `cli.py`)

- CLI implementation using `argparse`

- Poetry script entry point: `get-papers-list`

- README with usage instructions, LLM acknowledgment, and TestPyPI instructions

---

Bonus Achievements:

- Published as a TestPyPI package

- LLM assistance properly acknowledged

- Git version controlled with public GitHub repo

Conclusion:

All functional and non-functional requirements were met successfully. The project is clean, modular, and ready for extension or production use.