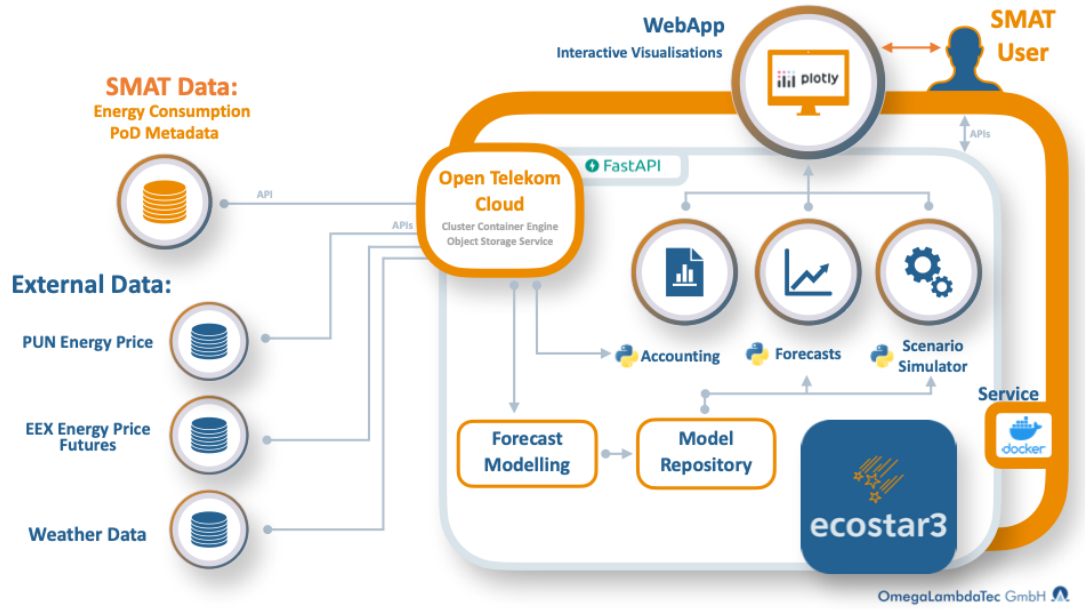
1. Technical Specification Double-side Page

2. TECHNICAL SCOPE

We introduce *ecostar3*—the Energy COST Analyzer, Forecaster, and Simulator designed specifically for the water utility SMAT. This tool addresses the dual challenges of monitoring energy costs and projecting future budgets by leveraging advanced in-house algorithms for energy management and forecasting. Our offering also includes a specialized scenario simulator, enabling long term risk-assessment capabilities. We provide a fully operational MVP via a user-friendly web interface that delivers real-time, interactive visualizations. The dashboard data is up-to-date, sourced directly from SMAT databases and enriched through Machine Learning pipelines. The current implementation provides an easy to access platform to analyse in depth the past energy costs of the utility and to estimate with accuracy the upcoming costs of the future months. Our solution not only meets all data provider specifications but also introduces value-added features like the scenario simulator, which allows for in-depth exploration of various factors influencing future energy costs, such as energy price projections, efficiency improvements, and renewable energy self-production.



3. ALGORITHMS, TOOLS AND CONCLUSIONS

The primary objective is to create an energy cost accounting tool. The input consists of time-series data for energy consumption across more than 1500 Points of Distribution (PoD) and corresponding energy price series. These series differ in time resolution—15 minutes for consumption and one hour for pricing—as well as units (kWh for consumption and MWh for pricing). To tackle this, we implemented an aggregation algorithm that resamples price data to a 15-minute resolution, subsequently calculating energy costs at 15-minute intervals for each PoD. This enables SMAT to validate incoming bills against actual consumption data. The processed data can be further segmented by PoD type, specific time frames, and aggregated to determine the overall cost.

The second objective involved developing a tool for future cost forecasting. To achieve this, we incorporated our in-house time-series forecasting library into *ecostar3*. This comprehensive library supports multiple Machine Learning frameworks, including Scikit-Learn, XGBoost, and Prophet. For our prototype, we employed an LGBM model to project future energy consumption and pricing, complete with 10% and 90% forecast uncertainty quantiles. Our tool provides forecasts up to one year in the future.

The backend is engineered in Python and employs the FastAPI library to expose our algorithmic features via APIs. These algorithms leverage state-of-the-art scientific and ML libraries, including scipy, pandas, and Scikit-Learn. Our prototype dashboard is crafted with Plotly Dash and relies on Flask. Deployment is streamlined via Docker containers, ensuring cloud compatibility. The live prototype is hosted on Open Telekom Cloud's Kubernetes engine, our cloud service provider known for rigorous, GDPR-compliant data security. Access is gated through an Elastic Load Balancer and necessitates login credentials. Data is sourced from various origins using both internal and external APIs and is securely housed in Open Telekom Cloud's Object Storage Service, which offers detailed operational logs.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951981

4. SCALABILITY AND FLEXIBILITY OF THE SOLUTION

The anticipated dataset from SMAT includes 1,500 Points of Distribution (PoDs), a figure typical for utility grids. *ecostar3* employs algorithms that have undergone rigorous testing and integration in multiple projects with German energy utilities and are capable of handling even larger datasets.

Modularity and complementarity ensure scalability. Each tool—accounting, forecasting, and scenario simulation—operates independently, but they can create an added value when used in combination. Within each tool, single PoDs are processed in parallel for efficiency. For forecasting, ML models are trained initially and retrained either on-demand or periodically. These models and data preprocessing pipelines are version-controlled and readily accessible.

In terms of adaptability, we have a proven track record of successful projects across German utilities specializing in water, electricity, natural gas, and district heating. The tool is data-agnostic and can easily incorporate additional variables specific to customer needs, such as demographic changes or climate impacts. Our goal is to develop a flexible solution that will be useful for all utilities in Europe.

Implementation-wise, the core lies in our flexible algorithms. These can be seamlessly integrated into existing customer infrastructure, specialized tools like the REACH Big Data Infrastructure, or the most common cloud infrastructure providers. We are also developing an Algorithm-as-a-Service model, an evolution of the traditional SaaS, aiming to put our algorithms in the spotlight and make them more accessible to a wider customer base.

Within the standard DVC (data acquisition -> data anonymisation -> data secure transfer -> data pre-processing -> data storage -> data analysis -> data visualisation) *ecostar3* can be readily inserted at the data analysis stage. As a value-add, we also offer optional services for data preprocessing and visualization.

5. DATA GOVERNANCE AND LEGAL COMPLIANCE

Energy data processed by ecostar3 predominantly originate from utility companies and are semi-aggregated to ensure individual personal data are not exposed. This ensures that no individual can be personally identified, fulfilling all requirements set forth by the General Data Protection Regulation (GDPR) and other applicable authentication policies.

Data processed and analyzed by *ecostar3* is securely stored on the Open Telekom Cloud, a cloud service provider known for its stringent data protection measures that comply with GDPR standards. This secure storage infrastructure employs multiple layers of encryption and access controls to ensure that only authorized personnel can access the data. Additionally, detailed logs of all data access and operations are maintained. These logs serve as an audit trail, allowing for comprehensive monitoring and traceability of data access, thereby further enhancing the security and integrity of the data throughout its lifecycle.

If we are also responsible for presenting the analyzed data through a web-based visualization tool, multiple layers of security are implemented. Access to the tool is restricted by a secure login mechanism; each user is given a unique username and password, which are securely stored and encrypted. The web connection itself is further secured through SSL/TLS encryption, verified by a third-party certificate authority. This ensures that the data transmission between the user and the web application is encrypted, adding an additional layer of security against potential data breaches or unauthorized access. This multi-layered approach to security ensures that both the data and the results of the analysis are protected at all stages, from initial data ingestion to final visualization.

6. QUALITY ASSURANCE AND RISK MANAGEMENT

Each algorithm within *ecostar3* undergoes rigorous individual testing in both development and operational environments. Post-implementation, these algorithms are further validated using historical data specific to the project. Our infrastructure code benefits from automated unit and integration tests, and we conduct exhaustive system testing with real-world data before any deployment. To ensure resilience, data sanity and quality checks are built into the system to handle scenarios where incoming data may be corrupted or unavailable.

Although *ecostar3* algorithms have proven successful in specific B2B solutions, adapting them to new use-cases will require additional customization, integration, and testing. This is where the bulk of our resources will be allocated. We recognize these challenges as familiar risks, having encountered and successfully mitigated them in past projects. Consequently, we have well-defined strategies in place to navigate these complexities effectively.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951981

Access to the web application

URL: https://ecostar3.omegalambdatec.com

Password: `reach_incubator_2023`

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 951981