



Uliege Guest Lecture -
05/05/25

Who We Are – Sagacify

We're a **service-oriented AI company** that partners with clients to design and deliver custom AI solutions. Our goal is to solve concrete business problems using machine learning and modern software tools.

- We're a **team of ~30** people
- **End-to-end** ML Solutions
- We don't sell products; we design highly customized solutions





Project organization

ML Team	Dev Team	Ops Team	Sales, PM, ...
<ul style="list-style-type: none">- ML Development- Containerization- No security- No interface	<ul style="list-style-type: none">- Security- Client interface / connection- Data preparation for ML service- Automation	<ul style="list-style-type: none">- Terraform- Optimization of the performance (multiple deployment, ...)	<ul style="list-style-type: none">- Client relation- Time management- Budget organizer



Challenges of Putting LLMs in Production

- Latency + compute overhead
- Hallucination risks
- Lack of observability
- Scaling complexity

Our answer: RAG + MLOps discipline + Monitoring tools



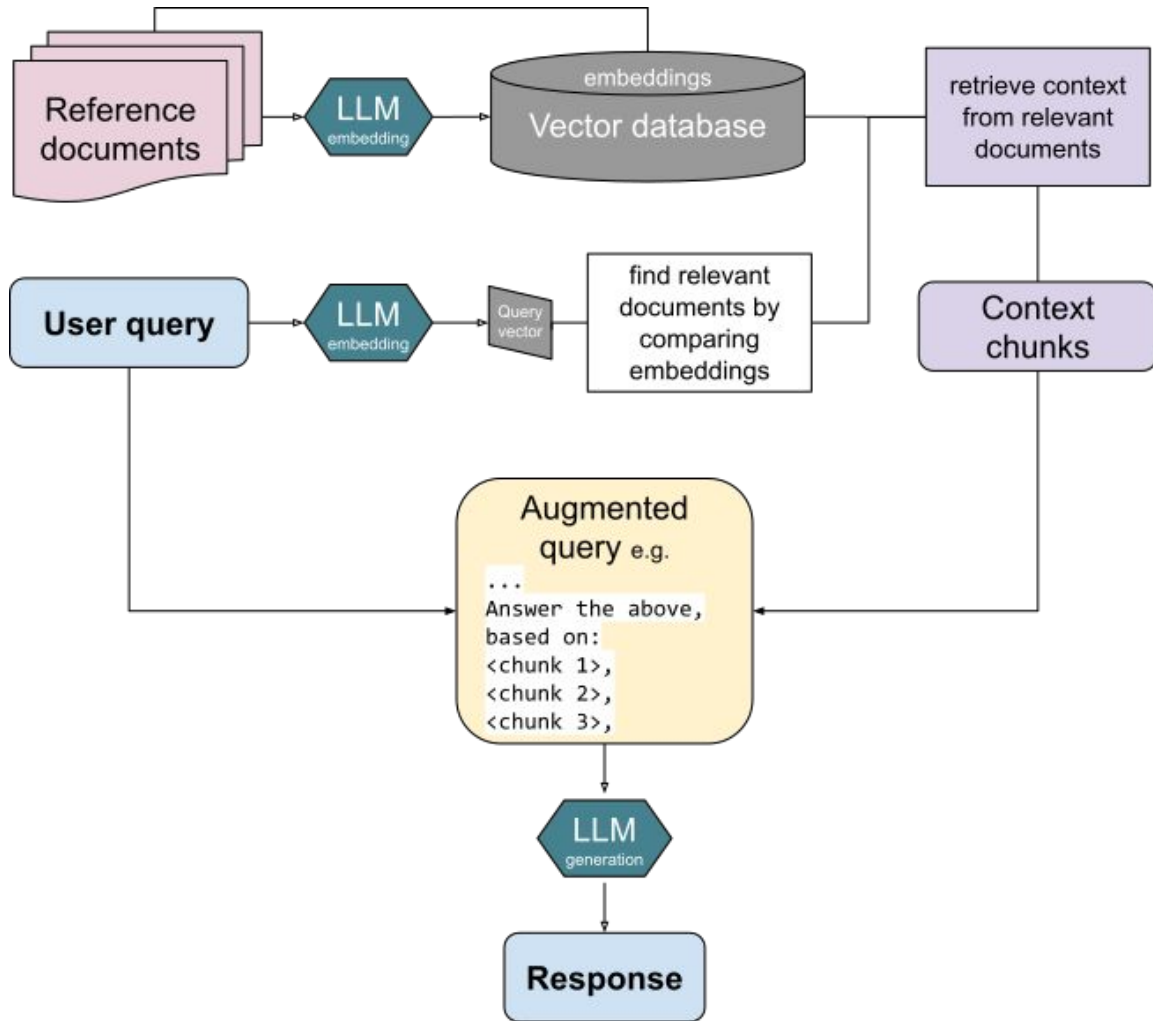


External Knowledge via RAG

RAG: Retrieve & ground LLMs with domain-specific context

Benefits:

- Reduced hallucinations
- Custom domain vocab
- Higher accuracy on user tasks





LLM Deployment Pipeline

1. Model Choice: Azure OpenAI (cloud) or Ollama (local, open-source)
2. Architectures: One-shot LLMs vs. Agent-based systems
3. DevOps Tools:
 - Docker (for isolation)
 - Poetry (for dependency management)
 - GitHub Actions (for CI/CD)





Container Strategy

- 2 Virtual environments:
 - `dev`: includes linting, testing
 - `prod`: lightweight, minimal
- Internal artifacts stored via DockerHub
- Fully reproducible environments





CI/CD in Practice

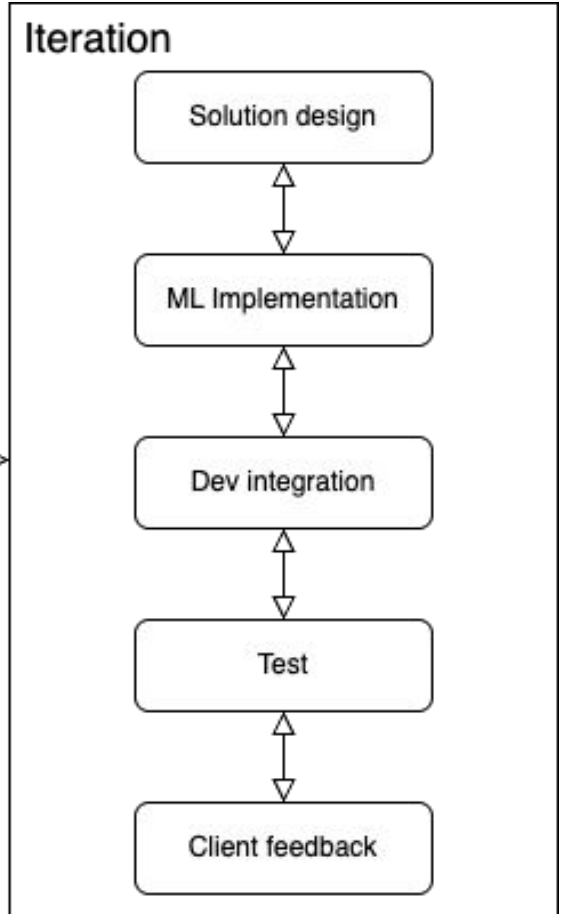
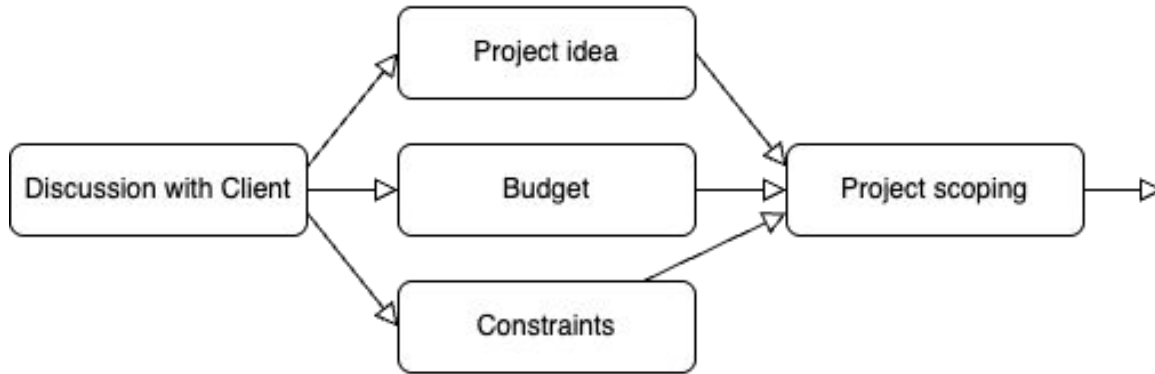
- GitHub Actions automates:
 - Commit checks (`pylint`, `black`, `pytest`)
 - Package versioning
 - Container builds & publishing

Ensures consistency from development to deployment.

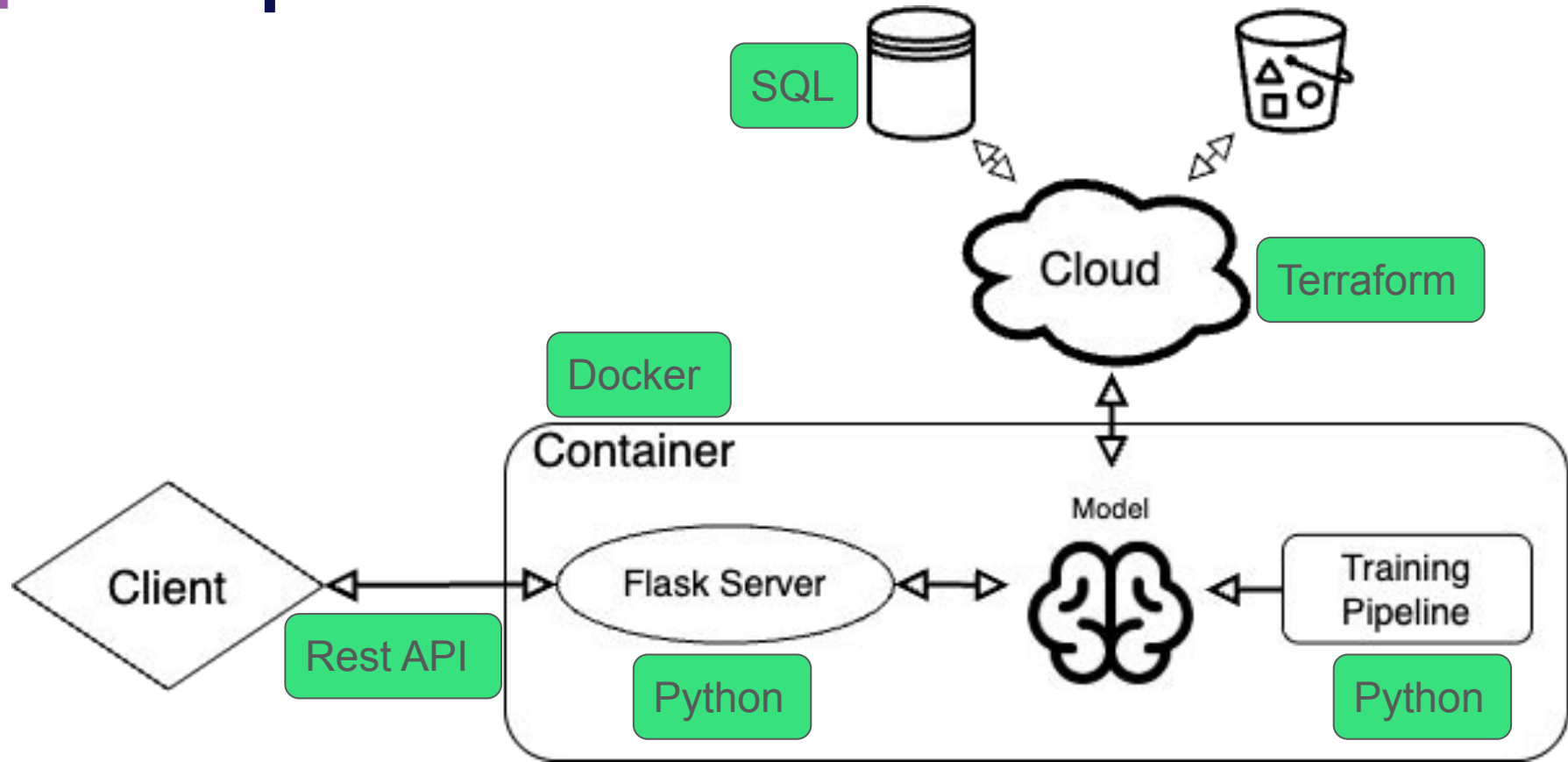




Project Lifecycle



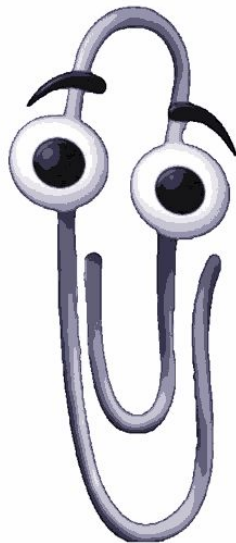
ML Implementation





Project “Clip-it” assistant

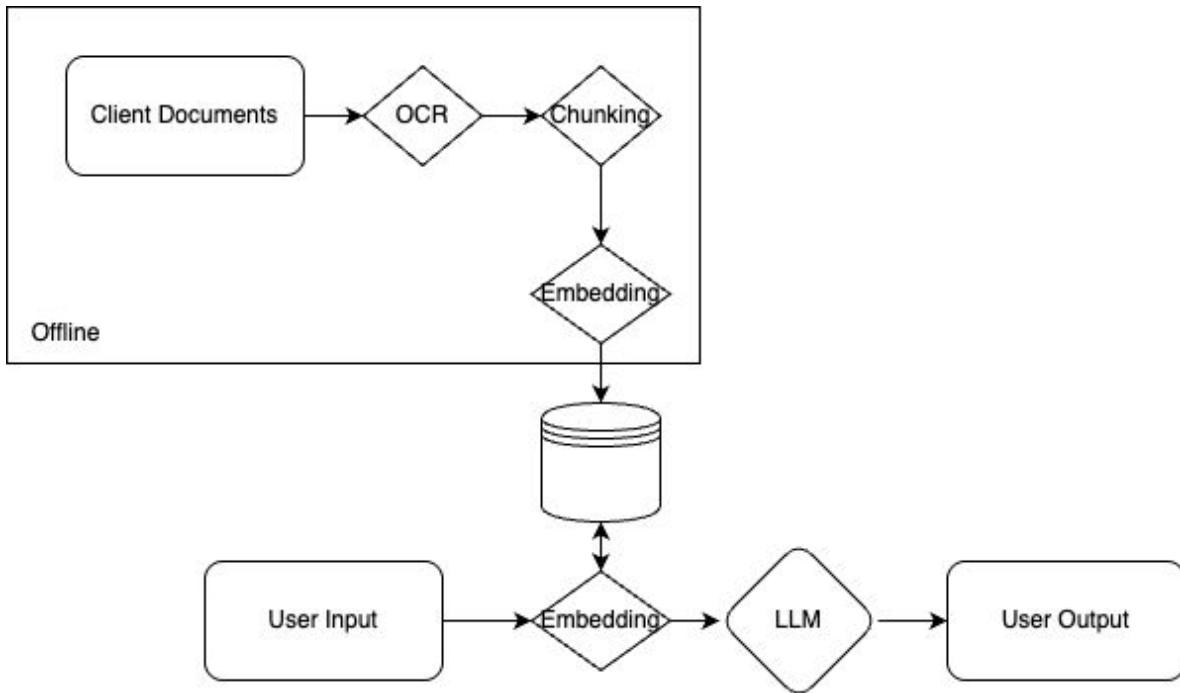
Scope : Create an internal assistant with business knowledge





Project “Clip-it” assistant

Scope : Create an internal assistant with business knowledge



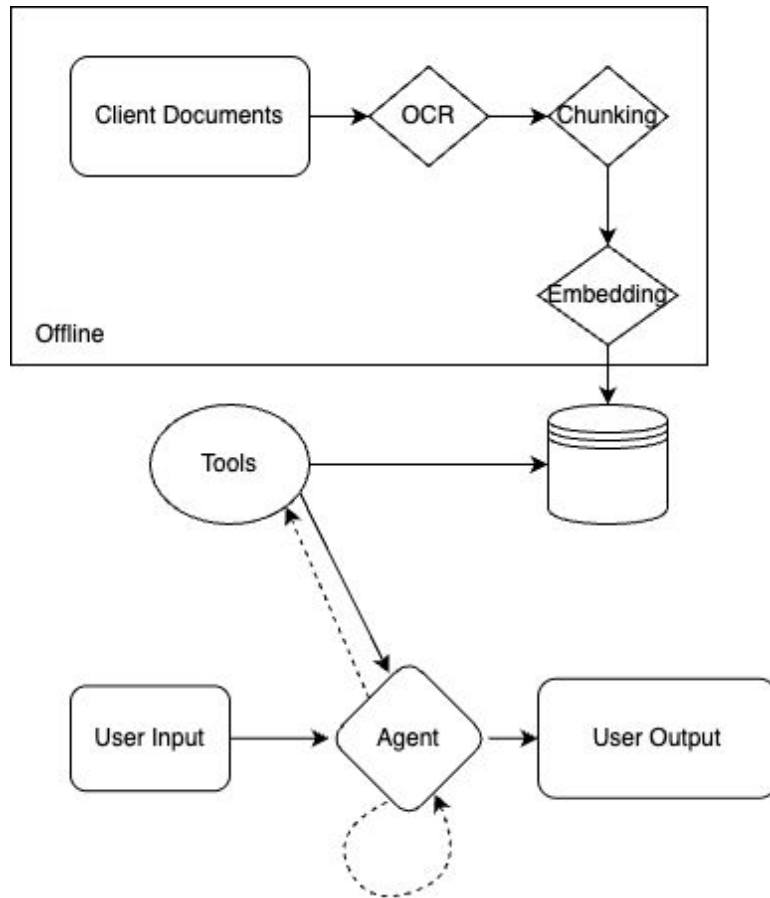


Project “Clip-it” assistant

Scope : Create an internal assistant with business knowledge

Adding agent architecture, enabling tool usage and reflection

Problem : The workflow can no longer operate synchronously

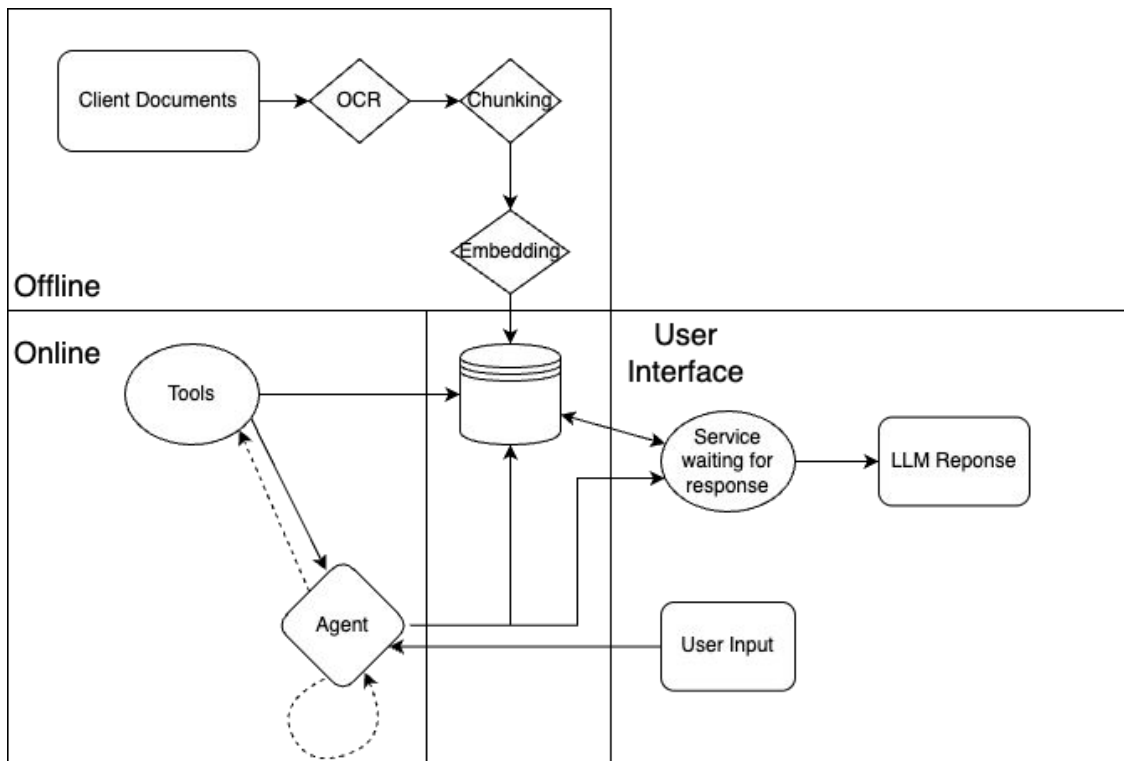




Project “Clip-it” assistant

Scope : Create an internal assistant with business knowledge

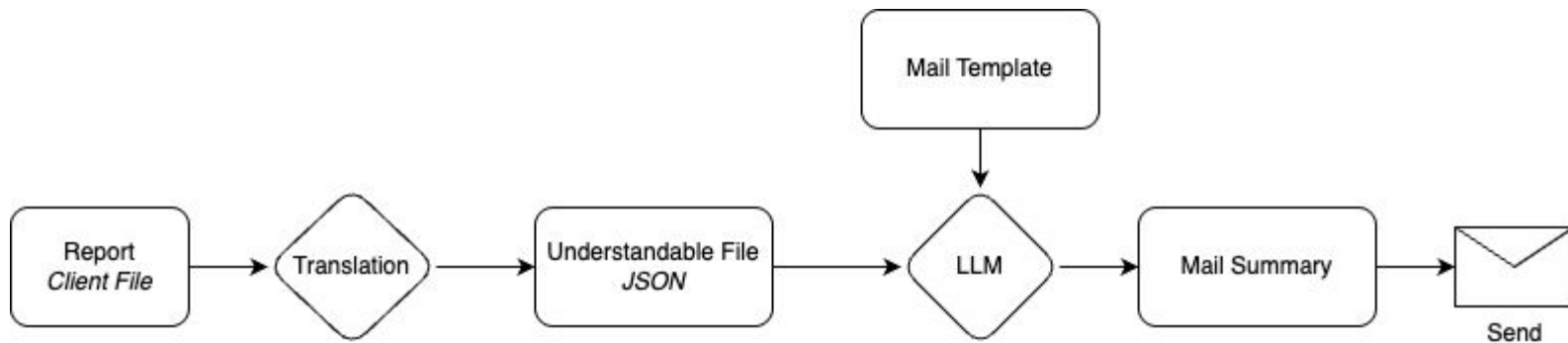
Adding agent architecture, enabling tool usage and reflection enabling asynchronous operation





Project tank sealing report

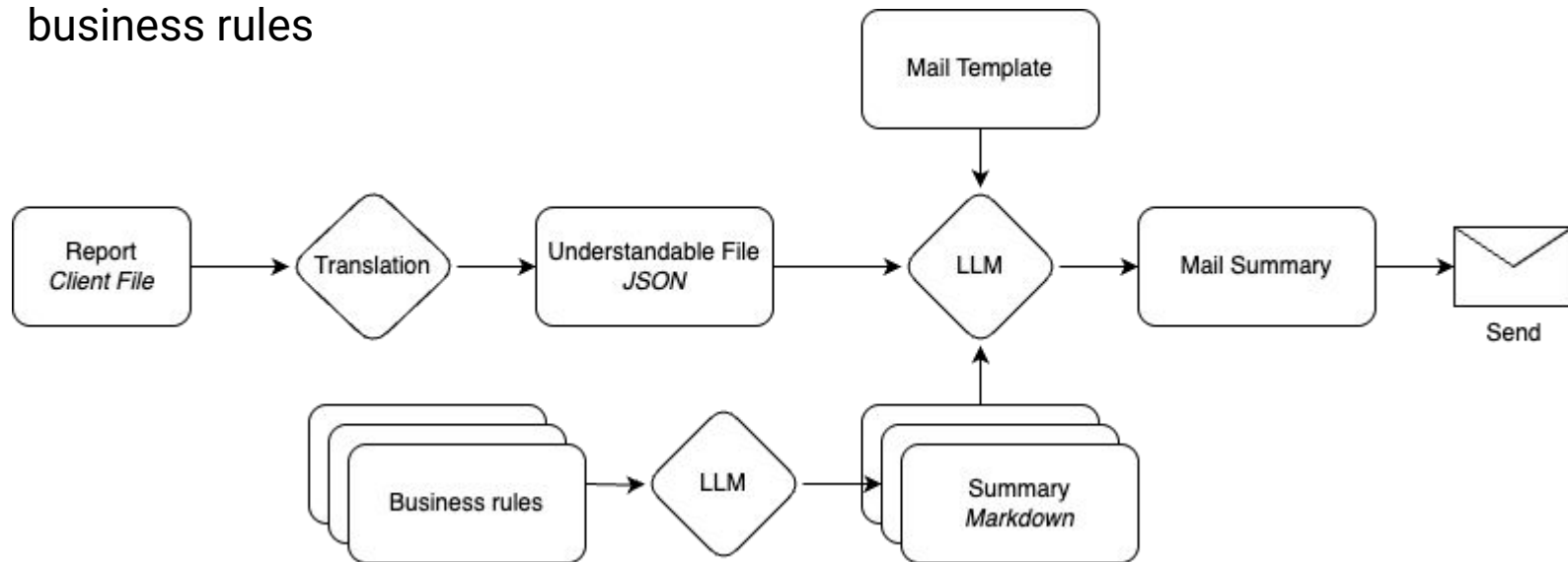
Initial scope : convert a sealing report into a summary mail for the client





Project tank sealing report

Updated scope : convert a sealing report into a summary mail for the client using business rules





Project tank sealing report

Further updated scope : convert a sealing report into a summary mail for the client using business rules and and added a mechanism for updating business rules.

