



Project Plan - Entrepreneurship

1. Project team members

Full Name
Abner Silva Barbosa

2. Executive Summary

The project aims to develop an artificial intelligence-based solution for translating documents in the pharmaceutical and clinical research sectors, emphasizing accuracy and speed. The platform will provide contextualized and customized translations, employing a continuous learning system that leverages user feedback. Document security and confidentiality will be ensured through advanced data protection technologies. The business model will be subscription-based, offering dedicated maintenance and support to contracting companies, with the goal of reducing errors, enhancing efficiency, and promoting reliability in document processes.

3. Problem Statement

Currently, the pharmaceutical document translation market faces significant challenges in preserving the original context and meaning, as well as maintaining formatting, indentation, and the accurate placement of images and text as per the original document. This process can take up to four days to complete, causing substantial delays. Furthermore, if the translation contains any errors, corrections may take more than two weeks because the document must undergo all validation stages again, negatively impacting timelines and the efficiency of regulatory processes.

Another major challenge is information security, as these documents contain sensitive data and cannot be processed on public platforms, limiting available automated translation options. Additionally, the translation validation process requires highly qualified staff to manually review each document, making the process even slower and increasing operational costs. This valuable time could be better utilized in other strategic activities, improving overall workflow efficiency.

4. Business Description

The startup proposes an innovative artificial intelligence-based solution for automated and secure translation of technical documents in the pharmaceutical and clinical research sectors. Our platform uses advanced AI models trained to preserve the context, formatting, and integrity of original documents, ensuring accurate, contextualized translations within a significantly reduced timeframe.

We serve pharmaceutical companies, research laboratories, regulatory agencies, and clinics by offering a service that reduces the time and operational costs associated with manual translations. Additionally, the platform ensures data security by avoiding the use of public and unreliable services for processing sensitive information.

Our business model is subscription-based, providing companies continuous access to the tool, along with dedicated support and maintenance. With this solution, we aim to optimize regulatory processes, minimize translation errors, and deliver greater efficiency for industry professionals.

5. Business Objectives

The project aims to develop an artificial intelligence platform specialized in translating pharmaceutical and clinical research documents, ensuring accuracy, speed, and security in processing this information. The startup will focus on serving laboratories, clinics, and regulatory agencies by offering an efficient solution to optimize processes and reduce challenges faced in translating these documents.

Primary Objectives:

1. Automate specialized technical translation

- a. Develop an AI system capable of translating documents while maintaining their original context, terminology, and formatting, thus avoiding common errors found in generic translations.

2. Ensure document security and confidentiality

- a. Implement robust data protection measures, ensuring that sensitive information is not processed on public platforms or exposed to potential data breaches.

3. Reduce translation and validation time

- a. Significantly decrease the time required for document translation, from days to hours, and streamline the validation workflow, assisting staff responsible for reviewing translations.

4. Continuously enhance the solution with user feedback

- a. Implement a continuous learning system where the translation model progressively improves based on user interactions and corrections.

6. Work Schedule

Below is the annual schedule of modules and their deliverables, showing only the main tasks; therefore, it is not possible to view the subtasks.



Módulo 1

- Sprint 1
 - Delivery of business documents

- Definition of project scope
 - GitHub repository or organization setup
- Sprint 2
 - Definition of requirements and user stories
 - Initial wireframe of software screens
 - Initial project architecture
 - Document modification and creation script (level 1)
- Sprint 3
 - Continuation of wireframe
 - Repository setup (frontend | backend | AI)
 - Continuation of the script (level 2)
 - Basic version of the translation model
- Sprint 4
 - Version 2 of the translation model
 - Some frontend screens without integration
 - API to receive and send documents to the translation model
- Sprint 5
 - MVP from the user perspective (not administrator)

Módulo 2

- Sprint 1
 - Initial implementation of model flow monitoring
 - Improve screens figma
- Sprint 2
 - Frontend screens in the code
 - Implemented logs in the backend
 - Initial integration front-end and back-end
- Sprint 3
 - Script for download font in the project
 - Integration of software to save documents in Storage (S3 | MINIO)

- Improve LLM development
- Sprint 4
 - API documentation for built services
 - Document detailing software costs (subscriptions, services, cloud) in the railway
- Sprint 5
 - Document about training process, improvement history, references, and strategies used
 - Implementation of user feedback collection on translations
 - Start implementation edit pdf in the frontend.

Módulo 3

- Sprint 1
 - Document on user research regarding interface and software usage
 - Unit and interface tests
 - Automatic deployment of new service versions
- Sprint 2
 - Service monitoring and metrics definition
 - Email notifications to users upon completion of document translation
- Sprint 3
 - Document on required software security strategies
 - Implementation of security layers
 - Traceability logs
- Sprint 4
 - Notification trigger if model accuracy drops after training
 - Model improvement
 - Document outlining product sales strategy
- Sprint 5
 - Data governance plan (initially LGPD-compliant)
 - UX/UI improvements based on feedback

- Scalability and software expansion plan

Módulo 4

- Sprint 1
 - Document detailing support and maintenance strategy
 - Document outlining market adoption plan
 - Planning of pitch deck and investor presentation
- Sprint 2
 - Document on market analysis (ROI and financial projections)
 - Definition of success KPIs for features
 - Definition of SLA for customers
- Sprint 3
 - Automation of AI training processes
 - Document on future system integrations
- Sprint 4
 - Document on expansion strategies for new languages
 - Model testing and version benchmarking
- Sprint 5
 - Investor presentation
 - Software demo video

7. Scope

The project aims to create an artificial intelligence platform specialized in translating pharmaceutical and clinical research documents, ensuring accuracy, speed, and security. The following deliverables will have a significant impact on the business:

Key Deliverables:

- Web platform with an intuitive interface for sending and receiving translated documents
- AI model specialized in translating technical terms from the pharmaceutical sector
- Integration with secure storage to preserve the formatting and structure of documents
- Robust API for integration with third-party systems
- User feedback system for continuous AI model improvement
- Usage monitoring and metrics for cost control and model performance
- Automation of the translation workflow to reduce turnaround time from days to hours
- Access management and security to ensure compliance with standards such as LGPD and GDPR
- Subscription-based monetization and pricing plan

- 1) **Preliminary Value Proposition:** The solution drastically reduces the time and cost required for translating technical documents in the pharmaceutical field, ensuring translations maintain the original context, formatting, and terminological accuracy.
- a) **Agility:** Translations that previously took days can now be completed in hours
 - b) **Technical Accuracy:** Model trained with sector-specific documents to avoid errors
 - c) **Security:** Private platform ensuring total confidentiality of documents
 - d) **Process Automation:** Reducing the need for lengthy manual revisions
 - e) **Continuous Learning:** AI enhanced through user feedback

By providing a more efficient and secure solution, the project eliminates current bottlenecks faced by pharmaceutical companies and clinics, allowing internal teams to focus on more strategic tasks instead of dealing with manual translation processes.

- 2) **Customer Segments:** The project serves companies and institutions requiring specialized, high-quality translations within the pharmaceutical and clinical research sectors:
- a) **Pharmaceutical Laboratories:** Companies handling regulatory documentation for submissions in different countries
 - b) **Clinics and Hospitals:** Institutions needing translations of medical documents and technical reports

8. General Observations

Challenges and Risks

- **Security and Compliance Assurance:** Due to sensitive documents containing private company information, security must be thoroughly planned and continually monitored.
- **Market Adoption:** Pharmaceutical companies and regulatory agencies often have rigid processes, potentially hindering initial platform adoption.
- **Continuous Model Training:** AI quality depends on the availability of training data, requiring validation of reliable sources for continuous improvement.
- **Scalability and Operational Costs:** Processing large volumes of documents may require robust infrastructure, impacting operational costs that need to be optimized.

Opportunities

- **Strategic Partnerships:** Opportunity to establish partnerships with software companies serving clinics and laboratories, integrating the solution directly into their workflows.
- **Expanding Market:** The demand for specialized technical translations is growing due to pharmaceutical sector globalization and increased international research.

Hypotheses to be Validated

- Can AI maintain the same level of accuracy as a specialized translator?
- Will the subscription pricing model be financially sustainable for companies of different sizes?

- Will translation automation sufficiently reduce time and costs without compromising quality?
- Will the platform maintain the expected security and privacy standards of the pharmaceutical sector?