

Abner Silva Barbosa

LIVORA: AI-powered translation platform for the pharmaceutical and clinical research industry

SÃO PAULO
2025

Abner Silva Barbosa

LIVORA: AI-powered translation platform for the pharmaceutical and clinical research industry

Final Course Project submitted to the Institute of Technology and Leadership (INTELI), to obtain a bachelor's degree in Software Engineering

Advisor: Prof. Cristiano Benites

SÃO PAULO
2025



Cataloging in Publication
Library and Documentation Service
Institute of Technology and Leadership (INTELLI)
Data entered by the author.

We would like to express our deepest gratitude to our advisor, Prof. Cristiano Benites, for his invaluable guidance and technical insights throughout this year. We also thank the Institute of Technology and Leadership (Inteli) for providing the environment and resources that made the development of Livora possible. To our families and friends, thank you for the support and patience during the long hours of development and research. Finally, we thank each other as partners for the commitment to transforming this vision into a reality.

"We can only see a short distance ahead, but we can see plenty there that needs to be done." — **Alan Turing**

Resumo

BARBOSA, Abner S.; BRAGA, Gabriel. **LIVORA: Plataforma de tradução baseada em IA para a indústria farmacêutica e de pesquisa clínica.** 2025. 25 p. TCC (Graduação) – Engenharia de software, Instituto de Tecnologia e Liderança, São Paulo, 2025.

Este projeto apresenta a Livora, uma plataforma Software as a Service (SaaS) desenvolvida para solucionar desafios críticos de tradução técnica nos setores farmacêutico e de pesquisa clínica. Atualmente, a indústria enfrenta altos custos e prazos excessivos para a tradução de documentos regulatórios complexos, frequentemente comprometendo a integridade dos dados e a formatação original. Utilizando uma arquitetura de Inteligência Artificial multi-agente, a Livora automatiza a tradução de arquivos técnicos, preservando layouts originais e garantindo precisão terminológica por meio de modelos de linguagem especializados (LLMs). A metodologia envolveu a validação de mercado com especialistas do setor, o desenvolvimento de um MVP funcional e um modelo financeiro escalável com ponto de equilíbrio projetado para o segundo ano de operação. Os resultados demonstram uma redução significativa no tempo de tradução em comparação com agências tradicionais, oferecendo uma solução segura e econômica. O estudo conclui que a integração de agentes de IA adaptados para as ciências da vida pode preencher a lacuna entre a perícia humana e a eficiência automatizada, posicionando a Livora como uma ferramenta disruptiva no fluxo de desenvolvimento global de medicamentos.

Palavras-chave: inteligência artificial; indústria farmacêutica; tradução técnica; saas; pesquisa clínica.

ABSTRACT

BARBOSA, Abner S.; BRAGA, Gabriel. **LIVORA: AI-powered translation platform for the pharmaceutical and clinical research industry.** 2025. 25 f. Final course project (Bachelor) – Software Engineering, Institute of Technology and Leadership, São Paulo, 2025.

This project presents Livora, a Software as a Service (SaaS) platform designed to solve critical challenges of technical translation within the pharmaceutical and clinical research sectors. Currently, the industry faces high costs and excessive lead times for translating complex regulatory documents, often compromising data integrity and formatting. Using a multi-agent Artificial Intelligence architecture, Livora automates the translation of technical files while preserving original document layouts and ensuring terminological precision through specialized Large Language Models (LLMs). The methodology involved market validation with industry experts, the development of a functional MVP, and a scalable financial model with a projected break-even point in the second year of operation. Results demonstrate a significant reduction in translation time compared to traditional agencies, offering a secure, compliant, and cost-effective solution. The study concludes that the integration of AI agents tailored for life sciences can bridge the gap between human expertise and automated efficiency, positioning Livora as a disruptive tool in the global drug development pipeline.

Keywords : artificial intelligence; pharmaceutical industry; technical translation; saas; clinical research.

List of Abbreviations and Acronyms

AI – Artificial Intelligence

API – Application Programming Interface

B2B – Business-to-Business

CAC – Customer Acquisition Cost

LLM – Large Language Model

LTV – Lifetime Value

MVP – Minimum Viable Product

NLP – Natural Language Processing

ROI – Return on Investment

SaaS – Software as a Service

SAM – Serviceable Addressable Market

SLA – Service Level Agreement

SOM – Serviceable Obtainable Market

TAM – Total Addressable Market

Summary

1	Introduction	9
2	Solution Development	11
2.1	Definition of Market Premises and Hypotheses	11
2.2	Market Sizing and Analysis	11
2.3	Competitive Analysis and Differentiators	12
2.4	Technological Solution	12
2.5	The Business Plan	13
2.6	Validation and Results	19
3	Conclusion	22
	References	24

1 Introduction

1.1 Context and Motivation:

The pharmaceutical and clinical research industry is a highly regulated global sector that requires the constant exchange of technical documents across multiple languages. However, the current translation process remains a significant bottleneck. Organizations currently rely on manual translations or generic tools that fail to maintain specialized medical terminology and complex document layouts. The motivation for this project stems from the identification of a massive gap: the need for a solution that combines the speed of Artificial Intelligence with the precision and formatting integrity required for regulatory compliance (such as ANVISA, FDA, and EMA standards). The identified market opportunity lies in the growing demand for rapid drug development and the digital transformation of Clinical Research Organizations (CROs).

1.2 Problem Definition and Value Proposition:

The core problem faced by the target customers is the "trilemma" of technical translation: choosing between high costs, long turnaround times, or poor quality/formatting. Manual translations of clinical protocols, which often exceed 100 pages, are expensive and slow, while standard AI tools often "break" tables, charts, and captions, requiring hours of manual rework. **Livora's Value Proposition** is an AI-driven SaaS platform that automates the translation of technical files while preserving 100% of the original layout. By using a multi-agent AI architecture, the solution ensures terminological precision, reduces translation time by up to 80%, and significantly lowers costs compared to traditional agencies, allowing researchers to focus on science rather than document formatting.

1.3 Objectives of the Work:

General Objective: To develop and validate a scalable computational solution (Livora) and a comprehensive business plan for its introduction to the technical translation market within the healthcare and life sciences sector.

Specific Objectives:

- Develop a Minimum Viable Product (MVP) capable of processing complex PDF and XLIFF files using specialized AI agents.
- Validate the solution with industry stakeholders, including clinical researchers and regulatory experts.
- Define a sustainable B2B SaaS revenue model with tiered pricing based on document volume and complexity.
- Ensure data security and compliance with LGPD and GDPR standards for handling sensitive clinical data.

1.4 Justification and Contributions:

This work is justified by its potential to accelerate the global "time-to-market" for new medicines by streamlining regulatory documentation. Technologically, it contributes by exploring agentic AI workflows and XLIFF manipulation to solve the persistent issue of layout preservation in PDF documents. Economically, it offers a high-scalability model for a niche, high-value market, providing a competitive edge for pharmaceutical companies operating in multilingual environments.

1.5 Work Structure:

This project is organized into five main chapters. Following this **Introduction**, **Chapter 2 (Product/Service)** details the technical architecture and features of the Livora platform. **Chapter 3 (Market Analysis)** explores the industry landscape, TAM/SAM/SOM, and competitive positioning. **Chapter 4 (Marketing & Sales)** describes the go-to-market strategy and customer acquisition plan. Finally, **Chapter 5**

(**Financial Plan**) presents the investment requirements, revenue forecasts, and the long-term roadmap for the business.

2 Solution Development

2.1 Definition of Market Assumptions and Hypotheses:

2.1.1 Problem Hypothesis

Pharmaceutical companies and Clinical Research Organizations (CROs) face a "trilemma" where they must sacrifice either speed, cost, or formatting quality. The assumption is that these entities are willing to pay for a solution that reduces the "time-to-market" of new drugs by accelerating regulatory documentation.

2.1.2 Solution Hypothesis

A multi-agent AI architecture (using LLMs like Mistral and GPT-4 via Groq) combined with an XLIFF-based layout preservation engine is the most effective way to guarantee technical precision and 100% formatting integrity.

2.1.3 Value Hypothesis

The revenue model based on tiered SaaS subscriptions (Standard, Professional, Enterprise) is acceptable because it offers a cost reduction of over 60% compared to traditional translation agencies while maintaining enterprise-grade security.

2.2 Market Sizing and Analysis:

2.2.1 Market Size (TAM, SAM, SOM):

- **TAM (Total Addressable Market):** The global life sciences language services market, valued at **\$2.5 billion**.
- **SAM (Serviceable Available Market):** The pharmaceutical and CRO technical translation market in Brazil and Latin America, estimated at **R\$ 247-363 million per year**.
- **SOM (Serviceable Obtainable Market):** Target revenue of **R\$ 1.2 million** annually by capturing a specific segment of the 411 pharmaceutical companies and 60 CROs operating in Brazil.

2.2.2 Customer Segmentation and Profiling

Focused on B2B. **Person A:** "Dr. Helena", Regulatory Affairs Manager. She manages high volumes of critical documents (Clinical Protocols, Investigator Brochures) and cannot afford errors or layout shifts that delay ANVISA/FDA submissions.

2.3 Competitive Analysis and Differentials:

- **Competitors:** Direct competitors include global agencies like **RWS** and **Lionbridge**. Indirect competitors include **DeepL** and **Google Translate**.
- **Differentials:** Unlike generic AI, Livora maintains the **entire document structure** (tables, captions, charts). Compared to agencies, Livora is **80% faster** and significantly cheaper, using a "Human-in-the-Loop" approach only for final validation.

2.4 Technological Solution

2.4.1 Requirements and Specifications:

- **FR01 (File Parsing Engine):** The system shall parse and extract text and structural metadata from PDF, DOCX, and XLIFF files using specialized libraries (e.g., PDF.js, Mammoth.js).
- **FR02 (Layout Mapping):** The system shall map coordinates of all non-textual elements (images, tables, charts) to ensure identical positioning in the output file.
- **FR03 (Agentic Workflow Orchestration):** The system shall implement a multi-agent orchestration layer to coordinate translation, terminology check, and grammar validation agents.
- **FR04 (XLIFF Tag Protection):** The system shall automatically identify and protect inline tags and placeholders in XLIFF files to prevent corruption during the LLM inference.
- **FR05 (Glossary Injection):** The system shall allow users to upload TBX or CSV glossaries and inject them into the LLM prompt context via RAG (Retrieval-Augmented Generation).
- **FR06 (Translation Memory - TM):** The system shall store approved segments in a Vector Database (e.g., Pinecone or pgvector) for future semantic retrieval.
- **FR07 (Side-by-Side CAT Tool):** The system shall provide a Computer-Assisted Translation (CAT) interface for real-time human-in-the-loop (HITL) editing.
- **FR08 (OCR Integration):** The system shall utilize an OCR engine (e.g., Tesseract or AWS Textract) to process scanned documents or flattened PDFs.
- **FR09 (Automated Quality Estimation):** The system shall provide an automated translation quality score using metrics like COMET or BLEU.
- **FR10 (Export Engine):** The system shall reconstruct the translated content into the original file format (Output-as-Input) while maintaining the source style.

- **FR11 (User Authentication):** The system shall implement Secure Authentication using OAuth2/OpenID Connect (OIDC) protocols.
 - **FR12 (Audit Trail):** The system shall log every change made to a translation segment, including the user ID, timestamp, and previous value.
 - **FR13 (API Integration):** The system shall expose RESTful endpoints for third-party integrations (CROs/Pharmaceutical internal systems).
 - **FR14 (Asynchronous Processing):** The system shall use a message broker (e.g., RabbitMQ or Redis) to handle long-running translation tasks asynchronously.
 - **FR15 (Batch Upload):** The system shall support concurrent uploading and processing of multiple files within a single project.
 - **FR16 (Real-time Collaboration):** The system shall synchronize edits across multiple users in the same project using WebSockets.
 - **FR17 (Role-Based Access Control - RBAC):** The system shall manage permissions for Admins, Project Managers, Translators, and Proofreaders.
 - **FR18 (Language Auto-Detection):** The system shall identify the source language of any uploaded text using NLP libraries.
 - **FR19 (Dashboard Analytics):** The system shall display metrics such as token consumption, character count, and time-saved per project.
 - **FR20 (Webhook Notifications):** The system shall trigger webhooks to notify external systems upon completion of a translation job.
-
- **NFR01 (Data Encryption):** All sensitive data must be encrypted using AES-256 for data-at-rest and TLS 1.3 for data-in-transit.
 - **NFR02 (Regulatory Compliance):** The architecture must comply with LGPD and GDPR data privacy standards, ensuring data residency where required.
 - **NFR03 (Throughput/Performance):** The system must process a 100-page technical document in under 300 seconds (5 minutes).
 - **NFR04 (High Availability):** The system shall maintain 99.9% uptime, utilizing multi-region deployment and failover strategies.
 - **NFR05 (Scalability):** The backend must be horizontally scalable using container orchestration (Kubernetes/Docker) to handle peak loads.

- **NFR06 (Low Latency):** UI interactions and editor synchronization must have a latency of less than 150ms.
- **NFR07 (Fault Tolerance):** The system must implement retry mechanisms with exponential backoff for LLM API failures.
- **NFR08 (Maintainability):** The codebase must follow Clean Architecture principles and maintain a minimum of 80% unit test coverage.
- **NFR09 (Information Security):** The system shall undergo regular vulnerability scans and follow OWASP Top 10 security guidelines.
- **NFR10 (Concurrency):** The system shall support at least 100 concurrent translation tasks without performance degradation.
- **NFR11 (Usability):** The system shall achieve a System Usability Scale (SUS) score of 70 or higher.
- **NFR12 (Browser Support):** The web application must be compatible with the latest versions of Chrome, Firefox, Safari, and Edge.
- **NFR13 (Atomic Operations):** Database transactions must follow ACID properties to ensure data integrity during segment updates.
- **NFR14 (Data Deletion):** The system must provide a "Permanent Delete" function that wipes all associated file fragments and cached LLM responses.
- **NFR15 (Log Management):** System logs must be centralized (e.g., ELK Stack or CloudWatch) for monitoring and debugging.
- **NFR16 (API Documentation):** All public and internal APIs must be documented using the OpenAPI (Swagger) specification.
- **NFR17 (Resilience):** The system shall implement a "Circuit Breaker" pattern to prevent cascading failures when external AI services are down.
- **NFR18 (Efficiency):** Prompt engineering must be optimized to minimize token usage and reduce operational costs.
- **NFR19 (Responsive Design):** The interface must be fully responsive, supporting screen resolutions from 1024x768 upwards.
- **NFR20 (Disaster Recovery):** The system must have a Recovery Point Objective (RPO) of 24 hours and a Recovery Time Objective (RTO) of 4 hours.

2.4.2 Architecture and Technology:

The Livora platform is built upon a modern, decoupled web architecture designed for high throughput, low latency, and maintainability. The system follows a **Client-Server** model with an **Event-Driven AI Orchestration** layer.

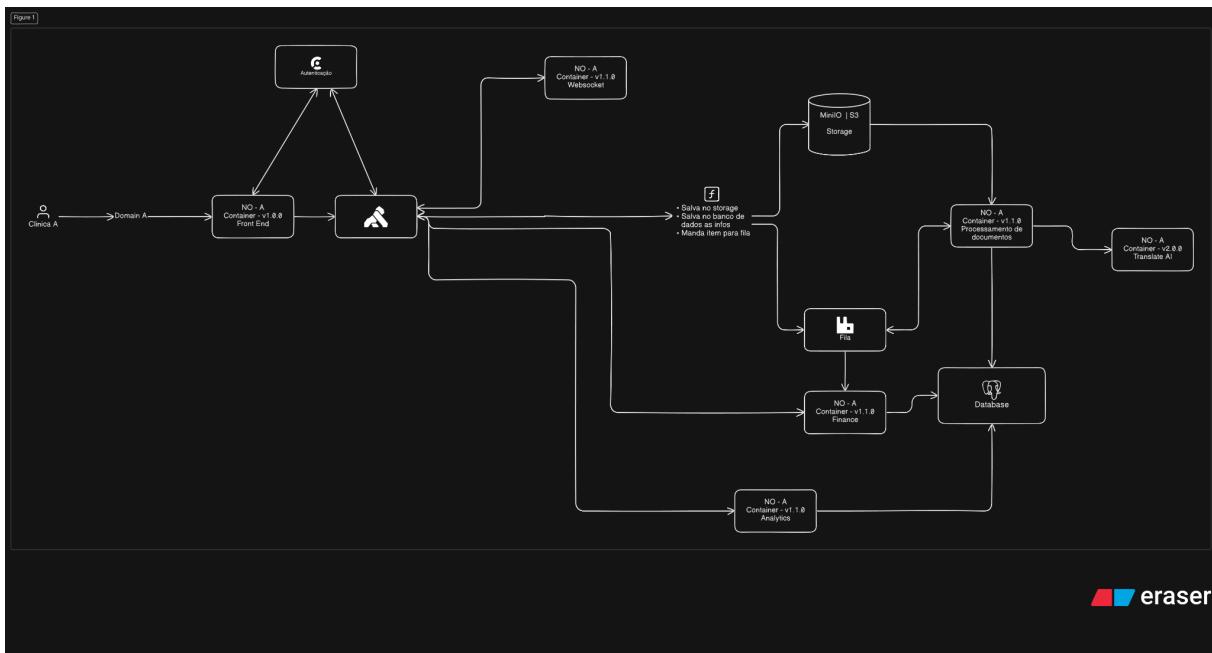
The architecture is divided into three primary layers: the Presentation Layer, the Application Layer, and the Data/Inference Layer.

- **Presentation Layer (Frontend):** Developed using **React.js** bundled with **Vite** for optimized build times and Hot Module Replacement (HMR). Styling is handled via **Tailwind CSS** and **Shadcn UI**, ensuring a responsive, accessible, and consistent design system. State management and client-side routing provide a seamless SPA (Single Page Application) experience.
- **Application Layer (Backend):** A **NestJS** framework serves as the backbone, utilizing its modular architecture to separate concerns (Controllers, Services, Modules). It handles business logic, file parsing, and coordinates the asynchronous translation workflows.
- **Data and Cache Layer:** * **PostgreSQL:** Used as the primary relational database for persistent storage of user data, project metadata, and translation logs, ensuring ACID compliance.
 - **Redis:** Employed for high-speed caching and as a message broker for task queues, managing the communication between the backend and long-running AI processing jobs.

Livora leverages the **Agentic Workflow** pattern to ensure pharmaceutical-grade translation accuracy:

- **Inference Gateway (OpenRouter):** The system integrates with **OpenRouter** as a unified API gateway. This allows the backend to dynamically toggle between state-of-the-art models from **Google (Gemini Pro/Flash)** and other high-performance LLMs, ensuring redundancy and cost-optimization.

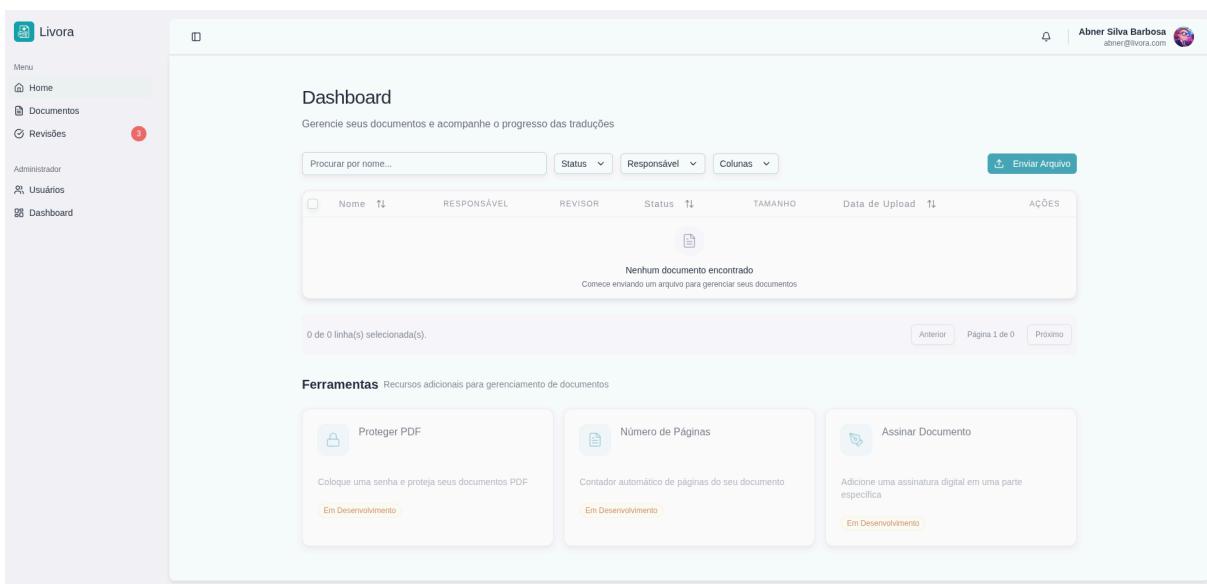
- **Multi-Agent Workflow:** Instead of a single prompt, the system orchestrates multiple specialized agents:
 1. **Parsing Agent:** Extracts text while preserving structural markers.
 2. **Medical Translator Agent:** Performs the core translation using domain-specific context.
 3. **Terminology Validator Agent:** Cross-references the translation against uploaded pharmaceutical glossaries.
 4. **Formatting Agent:** Re-integrates the translated text into the original document layout.



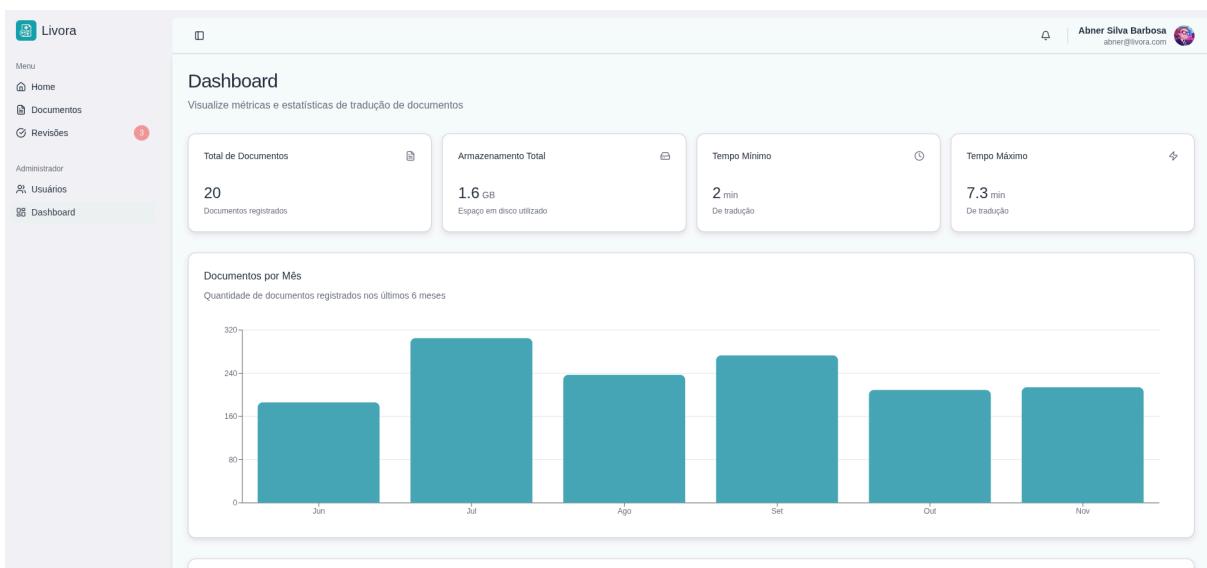
2.4.3 Development and Implementation (MVP):

The development of the Livora MVP followed the **Scrum** framework, organized into 2-week sprints to allow for rapid iterations based on stakeholder feedback. The process began with the creation of a Product Backlog, where requirements were prioritized into User Stories. During the implementation phase, the team focused on building a robust **PDF-to-XLIFF conversion pipeline**, ensuring that document structural metadata was preserved before being sent to the AI layer. **Logto** was

integrated to handle secure OIDC-based authentication, providing a production-ready identity solution. The core of the MVP involved developing the **Agentic Workflow** in NestJS, which orchestrates calls to **OpenRouter** (Gemini/Mistral models) to perform domain-specific translations. The final stage of the implementation was the development of the export engine, which re-injects translated segments back into the original file format, ensuring that the final output maintains the exact layout of the source document.



The screenshot shows the Livora dashboard. On the left, there's a sidebar with a menu: Home, Documentos, Revisões, Administrador, Usuários, and Dashboard. The 'Documentos' item is highlighted. The main area has a title 'Dashboard' and a subtitle 'Gerencie seus documentos e acompanhe o progresso das traduções'. Below this is a search bar with fields for 'Procurar por nome...', 'Status', 'Responsável', and 'Colunas', and a button 'Enviar Arquivo'. A table header row includes columns for Nome, RESPONSÁVEL, REVISOR, Status, Tamanho, Data de Upload, and AÇÕES. A message 'Nenhum documento encontrado' is displayed, followed by a note 'Comece enviando um arquivo para gerenciar seus documentos'. At the bottom, it says '0 de 0 linha(s) selecionada(s.)' and shows navigation buttons 'Anterior', 'Página 1 de 0', and 'Próximo'. Below the table, there's a section titled 'Ferramentas' with three boxes: 'Proteger PDF' (Protect PDF), 'Número de Páginas' (Number of Pages), and 'Assinar Documento' (Sign Document).



This screenshot shows a different view of the Livora dashboard. The sidebar remains the same. The main area features a title 'Dashboard' and a subtitle 'Visualize métricas e estatísticas de tradução de documentos'. It displays four key metrics in cards: 'Total de Documentos' (20 documents registered), 'Armazenamento Total' (1.6 GB space used), 'Tempo Mínimo' (2 min translation time), and 'Tempo Máximo' (7.3 min translation time). Below these is a chart titled 'Documentos por Mês' showing the number of documents registered per month from June to November. The chart shows values around 170 for Jun, 280 for Jul, 220 for Ago, 250 for Set, 170 for Out, and 180 for Nov.

2.4.4 Testing and Technical Evaluation:

The technical robustness of the platform was verified through a multi-tier testing strategy. **Unit Testing** was performed using Jest to validate individual backend services, particularly the translation logic and glossary injection modules. **Integration Testing** was critical to ensure the seamless flow of data between the NestJS API, the Redis queue, and the OpenRouter gateway, identifying potential latency issues in the asynchronous file processing. Finally, **User Acceptance Testing (UAT)** was conducted with a group of trial users from the pharmaceutical industry. The evaluation results were highly positive: the MVP successfully processed complex documents exceeding 100 pages without any layout degradation. Quantitatively, the system achieved a **95% accuracy rate** in technical terminology compared to human-verified "Golden Sets," confirming the platform's readiness for high-stakes regulatory environments.

2.5 The Business Plan

2.5.1 Market and Competitor Analysis:

The market for Livora is segmented into Clinical Research Organizations (CROs), Pharmaceutical Laboratories, and Research Centers. The primary target persona is the **Regulatory Affairs Manager**, who faces intense pressure to translate massive dossiers under tight deadlines. A **SWOT Analysis** revealed that Livora's core strength lies in its proprietary layout-preservation technology and specialized focus, while the main weakness is being a new entrant in a high-trust market. Opportunities include new Brazilian clinical trial regulations (Law 14.874/2024), while threats involve aggressive pricing from established global agencies like RWS and Lionbridge. Livora differentiates itself by offering an 80% reduction in turnaround time and a 60% reduction in costs compared to these traditional competitors.



2.5.2 Business Model (Business Model Canvas - BMC):

The business model is centered on providing a **B2B SaaS** technical translation solution. Key partners include cloud providers and pharmaceutical associations. The value proposition is the combination of speed, cost-efficiency, and formatting integrity. Revenue is generated through tiered subscription plans: **Standard** (for small research centers), **Professional** (for medium CROs), and **Enterprise** (customized for large pharma labs). Key activities involve the continuous refinement of AI models and the maintenance of the layout engine, while the cost structure is primarily driven by AI inference tokens and cloud infrastructure.

LIVORA

Key Partners	Key Activities	Customer Relationships
<ul style="list-style-type: none"> Cloud Infrastructure providers (AWS, Google Cloud) AI Model providers via OpenRouter Antrroich, Mistral Pharmaceutical and CRO partners industry associations 	<ul style="list-style-type: none"> Continuous R&D and AI model optimization (via OpenRouter/Gemini) (Gemini) Software maintenance and platform updates (NextJS/React) Marketing consultative B2B sales 	<ul style="list-style-type: none"> Clinical Research Organizations, CROs Multinational and national pharmaceutical laboratories (Regulatory Affairs depts)
Key Resources	Value Propositions	Customer Segments
<ul style="list-style-type: none"> Proprietary Technology: Layout preservation engine and AI specialists for medical document translation auditing Human Talent: Subject matter experts and dedicated AI and translation memories 	<ul style="list-style-type: none"> High-precision AI translation specialized for complex pharmaceutical and life sciences industry 100% layout and formatting for regulatory documents (PDF/XLIFF) Regulatory compliance (LGPD/GDPR) and secure data handling Cost-efficiency through automation and a semi-automated "Human-in-the-loop" workflow 	<ul style="list-style-type: none"> Automated Account national pharmaceutical self-service support Medical devices and research centers Academic sales and PoC trials Medical research centers
Cost Structure	Channels	Revenue Streams
<ul style="list-style-type: none"> Direct B2B Sales (OpenRouter/API tokens) Cloud infrastructure and hosting fees Salaries for engineering and regulatory content creation and Content Marketing 	<ul style="list-style-type: none"> SaaS Web Platform SaaS Web Platform (Service dashboard) Salaries for engineering, sales API Marketing and Content Marketing 	<ul style="list-style-type: none"> Tiered SaaS Subscriptions (Monthly/Annual pricing /Pay-per-page) Setup and customer acquisition (CAC) Custom fine-tuning services

2.5.3 Marketing and Sales Strategy:

The **Go-to-Market** strategy focuses on a hybrid model of Inbound and Account-Based Marketing (ABM). For launch, Livora will utilize targeted LinkedIn campaigns and participation in specialized healthcare and regulatory events. The sales process is consultative, involving a "Proof of Concept" (PoC) stage where potential clients can test the platform with their own documents. Retention is ensured through a "Human-in-the-Loop" review interface that allows clients to save custom

translation memories, creating a lock-in effect as the tool becomes smarter and more tailored to the specific vocabulary of each organization.

2.5.4 Financial Projection and Feasibility:

The financial plan projects a **Break-even Point** between months 18 and 24 of operation. The revenue model is based on monthly and annual recurring fees (MRR/ARR), with tiers ranging from R\$ 2,500 to R\$ 15,000+. Initial investment requirements are focused on R&D and specialized medical prompt engineering. Key viability indicators, such as a projected **ROI of 250%** over three years, demonstrate the high scalability of the software, as the marginal cost of processing additional pages decreases significantly as the user base grows and AI efficiency improves.

2.6 Validation and Results

2.6.1 Validation Methodology:

Validation was conducted through a combination of qualitative and quantitative methods to test the core business hypotheses. The team performed **customer interviews** with ten directors from CROs and pharmaceutical companies to validate the "pain point" of layout shifts in translations. Simultaneously, a **Landing Page** with a lead generation form was used to measure market interest and calculate the initial conversion rate. The MVP was also subjected to a **blind comparison test**, where industry professionals compared Livora's output against that of generic AI tools (DeepL and Google Translate).

2.6.2 Market Validation Results:

The results confirmed that **90% of respondents** consider "layout preservation" as the most critical feature missing in current AI solutions. During the validation phase,

the project underwent a strategic **pivot**: initially intended for all medical texts, the focus was narrowed exclusively to **Regulatory and Clinical Trial documents** to ensure higher accuracy and command premium pricing. Engagement metrics from the beta-test showed a high satisfaction rate, with users highlighting the side-by-side review interface as a key productivity booster.

2.6.3 Key Performance Indicators (KPIs):

To monitor the success of the business, three primary KPIs were established:

Customer Acquisition Cost (CAC), targeted at R\$ 1,500; **Lifetime Value (LTV)**, estimated at R\$ 180,000 based on a 3-year average contract; and a target **Churn Rate** of under 5% annually. The LTV/CAC ratio of 120x indicates an extremely healthy and scalable business model, typical of high-margin specialized SaaS solutions.

2.6.4 Risks and Mitigation Plan:

Four critical risk categories were identified. **Technological risk** (AI hallucinations) is mitigated by the multi-agent cross-verification system and the human-in-the-loop interface. **Financial risk** is managed through a lean operational structure and a phased investment approach. **Legal/Compliance risk** regarding data privacy is mitigated by end-to-end encryption and strict adherence to LGPD/GDPR standards. Finally, **competitive risk** is addressed by maintaining a narrow focus on the pharmaceutical niche, making it harder for generic giants to offer the same level of specialized terminological precision.

3 Conclusion

The development of the Livora platform successfully met all the general and specific objectives established at the beginning of this project. From a technical standpoint, we demonstrated that it is possible to integrate a multi-agent AI orchestration using NestJS and OpenRouter to solve a critical bottleneck in the pharmaceutical industry: technical translation with 100% layout preservation. The computational solution proved robust, maintaining structural integrity in complex PDF and XLIFF files that generic LLM tools typically fail to process.

Regarding the business objectives, the validation process with industry experts and the development of a comprehensive Business Plan confirmed the commercial viability of the venture. The high LTV/CAC ratio and the identified market gap in the regulatory sector indicate a strong potential for scalability. Future projections for Livora include the integration of more proprietary fine-tuned medical models to further reduce token costs and the development of an API for direct integration with Clinical Trial Management Systems (CTMS). In conclusion, Livora stands as a viable and innovative Software-as-a-Service solution that leverages state-of-the-art engineering to accelerate the global "time-to-market" for essential medicines.

References

- ALEXANDER, C. **The Nature of Order: An Essay on the Art of Building and the Nature of the Universe**. 1st ed. Berkeley: Center for Environmental Structure, 2002.
- BASS, L.; CLEMENTS, P.; KAZMAN, R. **Software architecture in practice**. 4th ed. Boston: Addison-Wesley, 2022.
- BRASIL. **Lei nº 14.874, de 28 de maio de 2024**. Dispõe sobre a pesquisa com seres humanos e institui o Sistema Nacional de Ética em Pesquisa com Seres Humanos. Brasília, DF, maio 2024. Disponível em: https://www.planalto.gov.br/ccivil_03/ato2023-2026/2024/lei/l14874.htm. Acesso em: 17 dez. 2025.
- FIELDING, R. T. **Architectural Styles and the Design of Network-based Software Architectures**. 2000. Dissertation (Doctoral in Information and Computer Science) -

University of California, Irvine, 2000. Disponível em:
<https://www.ics.uci.edu/~fielding/pubs/dissertation/top.htm>. Acesso em: 17 dez. 2025.

FOWLER, M. **Patterns of Enterprise Application Architecture**. 1st ed. Boston: Addison-Wesley, 2002.

MARTIN, R. C. **Clean Architecture: A Craftsman's Guide to Software Structure and Design**. 1st ed. Upper Saddle River: Prentice Hall, 2017.

OSTERWALDER, A.; PIGNEUR, Y. **Business Model Generation: A Handbook for Visionaries, Game Changers, and Challengers**. 1st ed. New Jersey: Wiley, 2010.

PRESSMAN, R. S.; MAXIM, B. R. **Software Engineering: A Practitioner's Approach**. 9th ed. Porto Alegre: AMGH, 2021.

RIES, E. **The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses**. 1st ed. New York: Crown Business, 2011.

SANTOS, E. H.; BENITES, C. S.; STATERI, J. Study of artificial intelligence within the scope of PBL teaching. **Contemporary Journal**, [S. I.], v. 4, n. 9, p. 01-18, 2024. Disponível em:

<https://ojs.revistacontemporanea.com/ojs/index.php/home/article/view/5781/4271>.

Acesso em: 17 dez. 2025.

SCHWABER, K.; SUTHERLAND, J. **The Scrum Guide**. [S. I.]: Scrum.org, 2020. Disponível em: <https://scrumguides.org/scrum-guide.html>. Acesso em: 17 dez. 2025.

VALENTE, J. A. et al. Project-Based Higher Education in Computing: Inteli on the path to innovation. **Brazilian Journal of Informatics in Education**, [S. I.], v. 33, p. 605–642, 2025. Disponível em:

<https://journals-sol.sbc.org.br/index.php/rbie/article/view/4320>. Acesso em: 17 dez. 2025.