Reverso Documents,
The New Generation Document Translation Platform

Théo Hoffenberg and Elodie Segrestan

Reverso, Paris, France

theo@reverso.com

Abstract

Reverso Documents is a widely-adopted translation and post-editing platform that combines advanced machine translation with extensive document format support and layout preservation capabilities. The system features AI-based rephrasing, bilingual dictionaries, and translation memory integration, enabling both professional translators and general users to work efficiently with complex documents. Used by millions globally, it provides API access for workflow integration and batch processing. The upcoming 2025 release will introduce LLM-based translation such as customizable settings with additional context, audio processing and anonymization features. This paper describes the platform's functionality, technical evolution, pricing structure, and competitive advantages in the market.

1 Introduction

Reverso Documents emerged from the European project Flavius (2017-2020, EU Horizon 2020 program, grant agreement No. 779360) as a solution to the growing need for **accessible yet powerful document translation tools**. The system uniquely positions itself between professional translation tools and consumer-grade solutions, offering advanced features while maintaining ease of use. The platform is commercially available and actively maintained by Reverso.

2 System Overview

Reverso Documents supports a wide range of **document formats**, including PDF, Docx, Xlsx, HTML, and XML, while preserving the original

layout. The system enables sentence-level **editing** in a side-by-side interface, maintaining translation memories for post-editing tracking.

The core functionalities include:

- Documents translation supporting 10+ formats (PDF, Docx, ...) with **editable output and a preserved layout**
- **Advanced machine translation** capabilities integrated since 2020, with continuous quality improvements and 25+ languages
- **Online collaborative revision** platform, enabling users to:
  o review the translation directly on the platform, with **side-by-side** segments
  o invite colleagues to review at no extra cost
  o save changes in **translation memory** and apply it to future projects, ensuring terminology consistency
  o import translation memory from another tools, or export translation memory created on the platform
  o use AI-based **rephrasing** system while reviewing, generating full-sentence alternatives
  o access integrated bilingual **dictionaries** directly from the revision platform, for precise word and expression-level work

These features work in concert to provide a comprehensive translation workflow that balances automation with human control, distinguishing it from basic MT services and traditional CAT tools.

3 Technical Evolution

The system's technological advancement has followed a carefully planned trajectory:

- 2020: Integration of **NMT** technologies, marking a significant milestone in translation quality

- 2023: Implementation of AI-based **rephraser**, enhancing the system's ability to suggest alternative translations
- 2025 (Q3 planned release): Introduction of **LLM-based translation** features with customizable translation setting, enhanced named entity identification, audio files processing, content anonymization features and post-editing time reduction. The post-editing time reduction will be accomplished through a combination of improved translation quality from LLM models fine-tuned on specific domains and intelligent contextual suggestions that learn from user edits.

4   Applications and User Base

Reverso Documents serves a diverse and extensive user base that includes **professional translators** seeking efficiency and quality, **students** developing skills with professional-grade tools, **professionals across various domains** needing accurate translations of business documents such as contracts or presentations, and **academic researchers** requiring precise translations of technical content. The system's versatility is enhanced by its **API**, which enables seamless integration into existing workflows and supports batch translation processing. This flexibility makes it equally suitable for individual translators and enterprise-scale operations.

5   Pricing and Licensing Options

Reverso Documents offers a tiered pricing structure to accommodate different user needs:
- **Free trial**: Up to 2,500 words for document translation
- **Premium**: Up to 50,000 words per year, files up to 30Mo, PDF up to 100 pages each + Premium on all Reverso suite
- **Pro**: Premium + Up to 200,000 words per year, files up to 120Mo, PDF up to 250 pages each
- **Enterprise** plans: Tailor-made plan and additional features such as SSO
- **Academic** licenses: 75% discount for educational institutions
- **One-time credits**: Document translation credits for a specific project

All plans include the core translation technology, while **advanced features** like OCR for scanned PDF or API access are reserved for paid plans.

6   Competitive Advantages

In comparison to other translation platforms, Reverso Documents offers several distinct advantages:
- Superior **layout** preservation
- More **user-friendly** interface than professional CAT tools
- Better integration of AI-based **rephrasing** than competitors
- More **affordable pricing** than enterprise-focused solutions
- Emphasis on **data privacy and security**, with no use of users' data to train our models and option to delete documents at any time

7   Future Developments

The roadmap for Reverso Documents includes several significant enhancements to its capabilities:
- Multimedia file processing support, offering new possibilities for content translation including audio transcription and translation
- Enhanced proofreading capabilities providing users with more tools for ensuring translation accuracy
- Advanced rephrasing system offering more nuanced alternatives based on context and domain
- **LLM-based features** including enhanced named entity identification, content anonymization features and post-editing time reduction

8   Conclusion

Reverso Documents represents a significant step forward in **making high-quality, secure machine translation accessible to a broader audience**. Its structured approach and user control features differentiate it from direct LLM use, while its continuous evolution ensures it remains at the forefront of translation technology. The platform's success in serving millions of users while maintaining high standards of translation quality demonstrates the effectiveness of its design philosophy and implementation.

https://documents.reverso.net