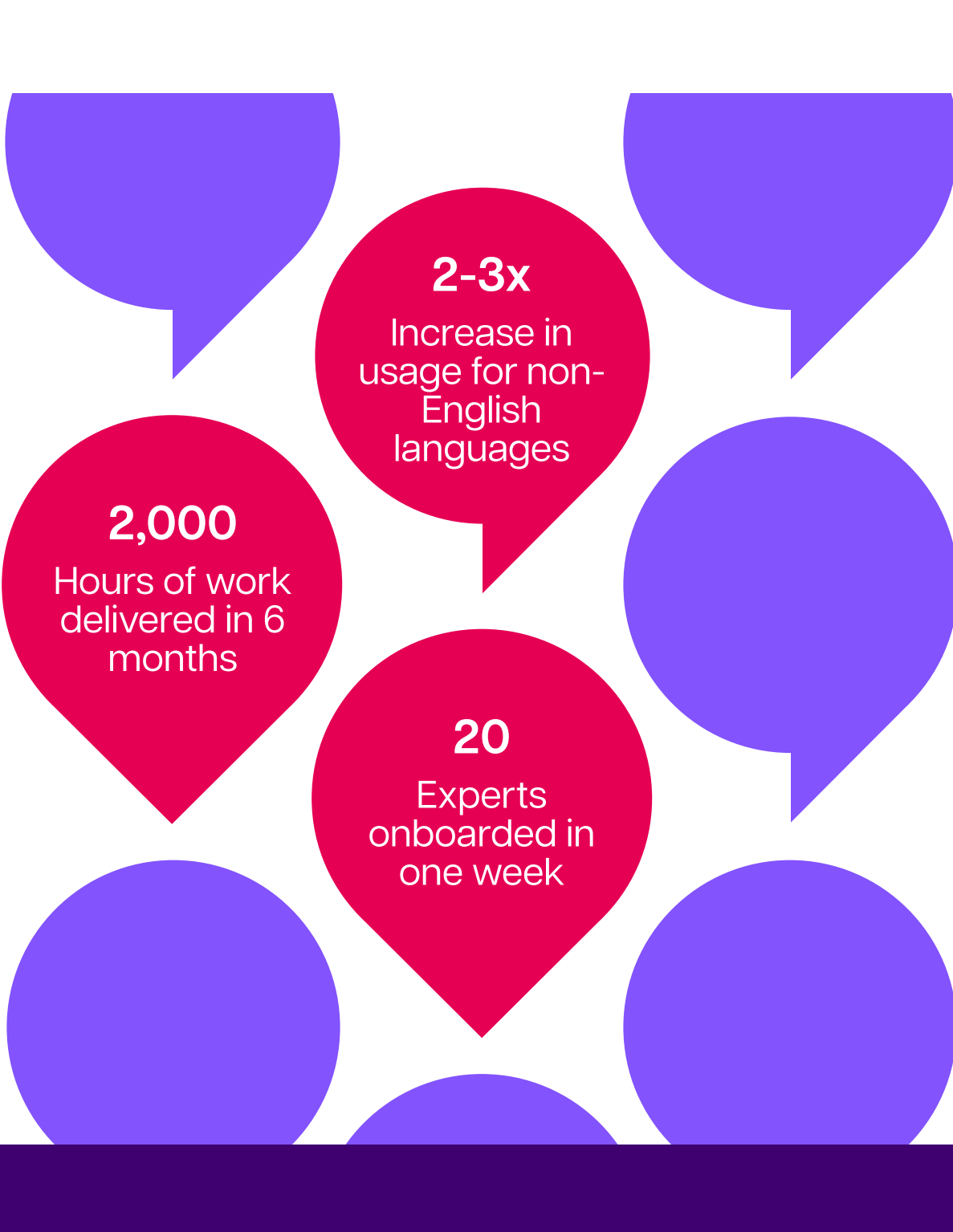


Multinational PDF company unlocks millions of GenAI users with TrainAI

One of the world's top PDF software companies set out to make its GenAI assistant features work for everyone, regardless of location or profession. To succeed, it needed specialized domain knowledge and native language expertise to ensure its GenAI assistant delivered accurate summarization, reliable responses and scalable performance across global markets. By partnering with TrainAI, the company was able to rapidly localize, train and fine-tune its GenAI assistant, unlocking tens of millions of new users worldwide.

Contact us



Key benefits

- 20 language and domain experts onboarded in one week
- 2,000 hours of work delivered within six months
- 2-3x increase in GenAI assistant usage in non-English languages
- Tens of millions of users across new markets

One of the world's top PDF software companies set out to solve a key challenge: make its GenAI assistant features work for everyone, regardless of location or profession.

They decided to partner with RWS and leverage our TrainAI data services to deliver AI functionality across multiple languages and specialized domains.

The result? A successful expansion that brought tens of millions of new users into the company's GenAI ecosystem.

Here's how we approached this client's challenges, solved them and brought their vision to life.

The global AI expansion

The client, who already relies on RWS as a trusted partner for localization services, had successfully rolled out its GenAI assistant features for English-speaking users. But integrating those features into the core user experience across global markets and domains required expertise they just didn't have in-house. To start, they needed to scale their GenAI assistant features across five specialized knowledge areas (academic, healthcare, legal, finance, tech and sales) and two languages (French and German).

The technical requirements were demanding. To succeed, they needed capabilities like:

- AI data specialists with [domain-specific experience](#)
- Ongoing model fine-tuning with real-time feedback for continuous improvement
- Linguistic expertise to evaluate non-English AI outputs

The client lacked the internal capacity to fulfill these requirements themselves, making TrainAI by RWS critical to their expansion strategy.

Challenges

- Scale GenAI assistant across multiple languages and domains
- Meet highly technical requirements including summarizing of documents up to 600 pages long, responding accurately to user inputs, attributing answers to reliable sources, avoiding hallucination and more
- Recruit native language speakers with the appropriate domain expertise to support training and fine-tuning RLHF efforts
- Balance speed of deployment with quality assurance

Solutions

- [TrainAI](#) from RWS
- [Generative AI data services](#)
 - Linguistic and domain expertise
 - Scalable processes
 - Custom technology solutions
 - Human-in-the-loop inputs to support reinforcement learning from human feedback (RLHF)
 - Long-term advisory AI data partnership

Results

Onboarded 20 language and domain experts in one week

Delivered 2,000+ hours of expert work within six months

Increased GenAI assistant usage by 2-3x increase in non-English languages

Gained tens of millions of users in new markets

Developed a scalable process enabling further expansion into additional markets

Solution: vetted experts, rapid scaling and advisory services

RWS's TrainAI team addressed these challenges through a comprehensive approach that combined:

- Linguistic and domain expertise
- Scalable processes
- Custom technology solutions
- Reinforcement learning from human feedback (RLHF)
- Long-term partnership

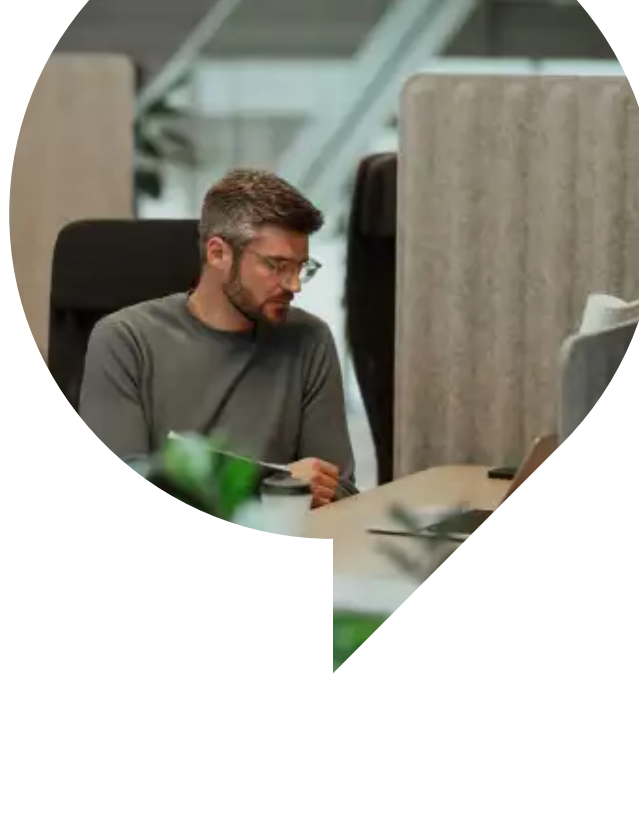
Our deep [AI training data experience](#) informed our project approach, and the client trusted us to deliver in each of these areas.

Linguistic and domain expertise

We tapped into our TrainAI community of AI data specialists, linguists and domain experts to handpick and onboard 20 subject matter experts across five specialized fields and two languages in just one week. Each candidate completed domain vetting, a machine learning assessment, language proficiency testing and training developed specifically for the project that evolved from feedback loop coaching into a comprehensive training program. By the time they joined the project, they were ready to start evaluating.

Scalable processes

Our vetting process is based on our vast experience on previous AI data projects, and we continually reassess our expert pool to keep them project-ready. That's how we were able to successfully ramp up resources for this client's project without sacrificing quality. They appreciated the scalability of our approach, which gave them confidence that we could handle even larger projects down the line.



Challenges: multiple layers of complexity

The project presented several interconnected challenges that made it particularly complex compared to typical AI training initiatives.

First time working with a third-party AI data services provider

The client's team had never worked with external AI data vendors in this way before. This made it challenging to define the right workflows or determine what exactly they needed us to deliver to help them achieve their project objectives. They also wanted to deploy AI updates as quickly as possible. This was difficult, since they were still trying to determine their specific needs.

High technical and quality assurance needs

The client wanted its AI tool to excel at four types of tasks, and each needed to be evaluated for multiple languages and domains:

- Summarizing PDF documents up to 600 pages long
- Answering user questions and attributing answers to reliable sources
- Avoiding AI hallucinations while being able to justify responses by citing source texts
- Responding accurately and appropriately to user inputs

To achieve their goals, the client needed a customized approach to evaluating and scoring the AI's outputs.

Native speakers and linguistic expertise

Finally, the evaluation process called for native language speakers with expertise in AI linguistics. These experts would not only rate responses but also provide comprehensive feedback to identify patterns and areas for improvement.

The client initially planned to use groups of judges to evaluate AI responses and debate their quality. This process proved too time-consuming. They needed a more efficient way to assess output without sacrificing data quality.

2-3x increase in non-English languages

Onboarded 20 language and domain experts in one week

Launched GenAI assistant to French and German users

TrainAI by RWS provided the expertise, scale and operational precision needed to help a global PDF software leader unlock new growth opportunities and transform how millions of users interact with PDF documents.



Custom technology solutions

Custom technology development played a crucial role on this project. It also happens to be an area we excel at.

The TrainAI team created tailored automation tools that could adapt to the client's changing requirements and unique project specifications. This flexibility was essential as the project evolved. We were able to make quick modifications based on client requests and any opportunities for improvement we identified.

Reinforcement learning from human feedback (RLHF)

TrainAI's domain and language experts scored the client's GenAI assistant's responses across languages and domains on a 1-5 quality scale. Domain expertise was required to understand PDF content and rate the assistant's responses and summaries on a range of criteria, including accuracy, completeness, readability, conciseness and more.

The human-in-the-loop input we provided was then used by the client to guide the GenAI assistant's learning through RLHF and improve its performance. The TrainAI team proactively fine-tuned workflows over time to maximize project efficiency and results, effectively demonstrating our capabilities as strategic partners and advisors to the client.

Long-term advisory partnership development

Our collaboration with the client's team evolved into an advisory partnership rather than a simple data vendor relationship. For example, the client initially proposed real-time feedback mechanisms, but our TrainAI experts identified that the time and cost involved could hurt the project's goals.

We recommended consensus-based evaluation, where an odd number of reviewers would evaluate whether an output is acceptable or needs improvement. In this case, [panels of three human experts reviewed the data](#), allowing us to build consensus quickly and keep the project moving while maintaining data quality.

Results: expanding global reach

The project delivered measurable success across multiple areas, and the client had pretty nice things to say about it, too.

Tens of millions of new users

The client successfully launched its GenAI assistant tool to French and German users, directly enabling access for their 20+ million existing users in those markets. They experienced an uptick in usage of +2-3x in non-English languages. More importantly, the client gained tens of millions of new users in these markets. As a result, not only were more people using their GenAI assistant, but more people were now also using their core product.

Efficient and repeatable processes to enter new markets

Leveraging the TrainAI team's data solutions expertise, we developed and optimized a multilingual data localization workflow for the client. The project required 2,000+ hours of expert input within the first six months, and we delivered that work through a scalable process that could handle multiple requests and languages in parallel.

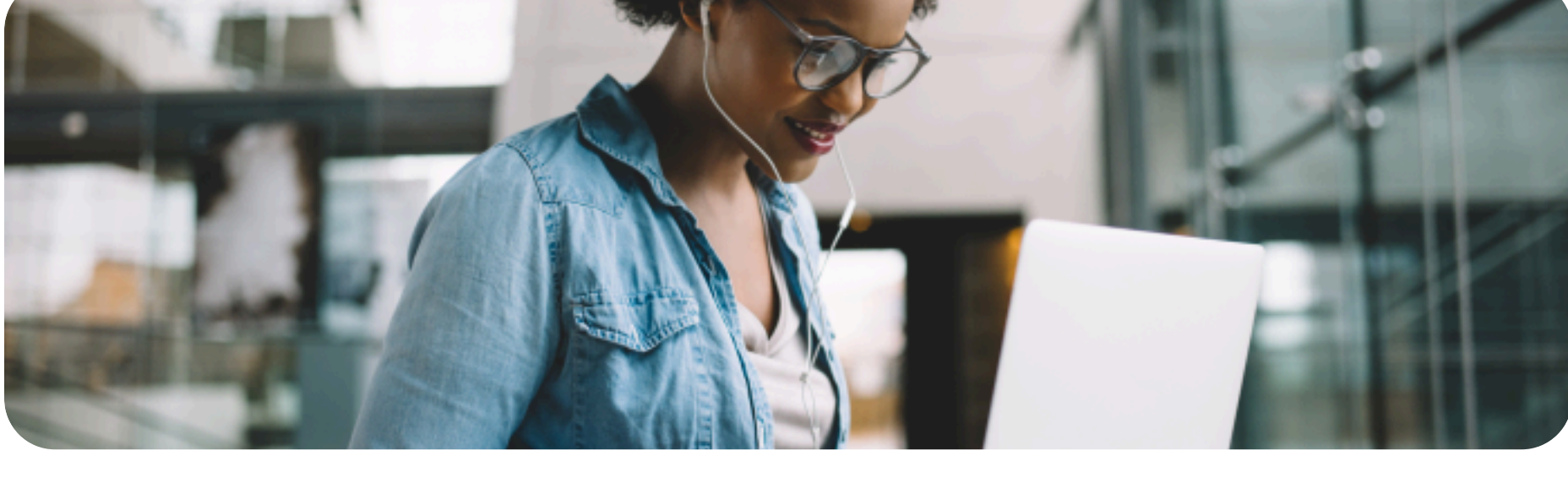
The automated processes and QA workflows we designed also made it easier for the client to expand into additional languages. They've since continued working with us on new languages, including Japanese, Italian, Spanish and Portuguese.

They were so happy with the results that they now acknowledge TrainAI by RWS as their primary AI training and fine-tuning data partner.

A more effective means of improvement and quality assurance

While the client was pleased with the quality of their GenAI assistant's localized capabilities, they gained something even more valuable: a repeatable process for training and fine-tuning it across languages and locales around the world.

The three-person consensus model we implemented provided reliable scoring throughout the project and generated detailed, actionable feedback that could be used right away. The client can now identify patterns across evaluator comments and pinpoint areas for improvement across languages. This approach has become essential for refining the model over time, and will continue to serve the company as it enters new markets.



TrainAI by RWS: accelerating GenAI growth through scalable processes and vetted expertise

The TrainAI team brought their full range of skills, expertise and ingenuity to this task, transforming an enormous technical challenge into a growth opportunity for the client's global expansion. The multinational PDF client appreciated our multilingual AI training and fine-tuning data services, scalable evaluation processes and ongoing AI advisory support. Thanks to the team's work, the client is transforming how people interact with PDF documents across languages, cultures and domains.

Here's what the client shared with our team after their GenAI language and domain expansion brought in 10 million new users:

"Thanks a lot to the RWS [TrainAI] team who worked on testing the support for GenAI!"

"Congrats due also for the TrainAI team that was immersed in validating these new GenAI features."

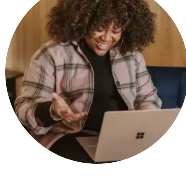
Have a challenging GenAI training or fine-tuning project? We can help.

If you're working on a similar GenAI project, don't underestimate the value of specialized domain and multilingual expertise in the AI training and fine-tuning process. With TrainAI, you get access to the language skills, domain knowledge and adaptive tech solutions you need to overcome your GenAI data challenges and scale faster.

Ready to expand your GenAI capabilities? [Contact us today](#) to learn how TrainAI's data services can support your next phase of AI growth.

Discover more about TrainAI by RWS

[rws.com/trainai](#)



Contact us

We provide a range of specialized services and advanced technologies to help you take global further.

Business E-mail*

First Name*

Last Name*

Company Name*

Job Title*

Country or Region*

Additional message*

Your personal data will only be used to provide you with relevant product and services information. We use the services of third parties both in and outside the EU that have agreed that your personal data will be processed in accordance with the relevant data protection laws. By entering your personal data you are affirmatively consenting to our use of your personal data and our [privacy policy](#).

Submit

Dive deeper

IT leader boosts OCR AI accuracy with 3.5M transcriptions and 30k image annotations in 32 languages

IT leader turns to TrainAI by RWS to help enhance the accuracy and language capabilities of its OCR engine.

[Read case study →](#)

Tech conglomerate enhances GenAI chatbot with TrainAI

To strengthen their competitive advantage, a global technology company wanted to improve the accuracy and capabilities of their generative AI chatbot. However, they quickly realized they needed help to find sufficient technical, data and linguistic resources to evaluate, review and refine prompt-response pairs. TrainAI by RWS improved the client's AI chatbot by harnessing the expertise of content writers, AI data specialists and linguists across the globe.

[Read case study →](#)

TrainAI powers tech leader's AI Data safety benchmarking

A leading technology company turned to TrainAI to benchmark AI data safety across vendors and nine languages. Mobilizing 285+ specialists and delivering up to 10,000 hours per month, TrainAI conducted multilingual audits that surfaced safety risks, vendor performance gaps, and new benchmarks for evaluating LLM training data. The result gave the client objective insights to guide vendor selection and strengthen global AI reliability.

[Read case study →](#)