

PDFMathTranslate: Scientific Document Translation Preserving Layouts

Rongxin Ouyang^{1,†}, Chang Chu^{2,*†}, Zhikuang Xin³, Xiangyao Ma⁴

¹National University of Singapore. ²Tsinghua University.

³University of Chinese Academy of Sciences ⁴Funstory.ai Limited.

*Correspondence: chuc23@mails.tsinghua.edu.cn.

† These authors contribute equally to this work.

Abstract

Language barriers in scientific documents hinder the diffusion and development of science and technologies. However, prior efforts in translating such documents largely overlooked the information in layouts. To bridge the gap, we introduce **PDFMathTranslate**, the world’s first open-source software for translating scientific documents while preserving layouts. Leveraging the most recent advances in large language models and precise layout detection, we contribute to the community with key improvements in precision, flexibility, and efficiency. The work has been open-sourced at <https://github.com/byaidu/pdfmathtranslate> with more than 22k downloads.

1 Introduction

The advances in science and technologies spread and evolve in different languages ([Von Gizecki; Montgomery](#)). Yet, language barriers hinder the diffusion and development of scientific progress ([Ramírez-Castañeda, 2020](#); [Hwang, 2005](#); [Ammon, 2012](#)). For example, while 98% publications in science are written in English ([Liu; Ammon](#)), the language only has 7.3% native speakers and no more than 20% speakers worldwide ([Bahji et al.](#)). Consequently, the majority of human beings are hindered by language barriers from the advances in science and technology. To overcome the barrier, there are attempts such as those from international organizations ([UNESCO](#)) and academic efforts in improving *Machine Translation (MT)* over text ([Brown et al., a](#); [Vaswani et al., 2017](#); [Zhu et al.](#); [Johnson et al.](#); [Sennrich et al.](#)).

However, machine translation literature fails to address the unique challenges posed by the layouts in *technical translation* ([Schubert, 2012](#)). In scientific and technical documents, layouts and non-textual elements are not ignorable — the arrangement of paragraphs, mathematical equations, tables, and figures have rich and important meanings.

Those text-based attempts ignore the information in the layout and are thus insufficient to address the barrier in scientific and technical document translation.

To fill the gap, we introduce **PDFMathTranslate**, the world’s first open-source tool to translate PDF documents with preserved layouts. By leveraging recent advances in layout detection and large language models (see [Figure 1](#)), we better address the barrier with at least five key contributions: (1) efficient workflow of layout detection, translation, and re-rendering; (2) support for multiple languages (3) support for multiple translation models and services; (4) diverse user interfaces; and (5) a community-commerce model affording sustainable developments.

2 Architecture and Design

Overall, **PDFMathTranslate** is designed to translate documents while preserving their original layouts. First, it accepts a PDF document along with user-specified parameters, such as languages and the preferred translation service (see [Appendix A](#)). Next, it detects the layouts to extract the layouts and textual contents from the document. Third, the texts are translated using the selected translation service, such as GPT-4, DeepL, Google, or Ollama (see [Appendix A](#)). Finally, the translated text and previously detected layouts are re-rendered as a translated document with preserved layouts.

Technically, we design the architecture for *precision*, *flexibility*, and *efficiency*. Specifically, we introduce three key components: a precise parser, a flexible translation middleware, and an efficient workflow, to be explained in the following three sub-sections.

Precise layout parser

Preserving layouts requires precise parsing. To precisely parse the layouts, we propose a pipeline

पीडीएफमैथट्रांसलेट: लेआउट को संरक्षित रखते हुए वैज्ञानिक दस्तावेज़ अनुवाद

रोगशनि ओयांग^{1,*†}, चांग चू^{2,*†}, झाकिउआंग शनि³, शयिंगयाओ मा⁴ ¹नेशनल यूनिवर्सिटी ऑफ सिंगापुर। ²संघीआ विश्वविद्यालय। ³चीनी वैज्ञानिक अकादमी विश्वविद्यालय ⁴फनस्टोरी.एआई लिमिटेड। *पत्राचार: chuc23@mails.tsinghua.edu.cn। [†] ये लेखक इस कार्य में समान रूप से योगदान करते हैं।

अमूरत

वैज्ञानिक दस्तावेजों में भाषा संबंधी बाधाएँ वैज्ञानिक और प्रौद्योगिकियों के प्रसार और विकास में बाधा डालती हैं। हालाँकि, ऐसे दस्तावेजों का अनुवाद करने के पछिले प्रयासों ने लेआउट में मौजूद जानकारी को ज्यादातर अनुदेखा कर दिया। इस कमी को प्राप्त करने के लिए, हम PDFMathTranslate पेश करते हैं, जो लेआउट को बनाए रखते हुए वैज्ञानिक दस्तावेजों का अनुवाद करने वाला दुनिया का पहला ओपन-सोर्स सॉफ्टवेयर है। बड़े भाषा मॉडल और स्टीक लेआउट पहचान में नवीनतम प्रगति का लाभ उठाते हुए, हम संटीकता, लचीलेपन और दक्षता में महत्वपूर्ण सुधारों के साथ समुदाय में योगदान करते हैं। इस कार्य को 22 हजार से अधिक डाउनलोड के साथ <https://github.com/byaidu/pdfmathtranslate> पर ओपन-सोर्स किया गया है।

1 परचिय

वैज्ञानिक और प्रौद्योगिकियों में प्रगतिविभिन्न भाषाओं में फैलती और विकसित होती है (वॉन गज़ी-की; मोटगोमरी)। फरि भी, भाषा बाधाएँ वैज्ञानिक प्रगतिके प्रसार और विकास में बाधा डालती हैं (रामीरेज़-कास्टानेडा, 2020; हवांग, 2005; एमोन, 2012)। उदाहरण के लिए, जबकि वैज्ञानिक प्रकाशन अंग्रेजी में लखियो जाते हैं (लयि; एमोन), इस भाषा के केवल 7.3% मल वक्ता हैं और विश्वस्तर पर 20% से अधिक वक्ता नहीं हैं (बहाजी एट अल.)। परणिमसवरप, अधिकांश मानव भाषा बाधाओं के कारण वैज्ञानिक और प्रौद्योगिकी में हुई प्रगतिसे बाधति होते हैं। इस बाधा को ढूँ करने के लिए, अंतर्राष्ट्रीय संगठनों (यूनेस्को) और टेक्स्ट पर *Machine Translation (MT)* में सुधार के अकादमिक प्रयासों जैसे प्रयास किए गए हैं (ब्राउन एट अल., ए; वासवानी एट अल., 2017; झू एट अल.; जॉनसन एट अल.; सेनरचि एट अल.)।

हालाँकि, मशीन अनुवाद साहित्य

technical translation (Schubert, 2012) में लेआउट द्वारा उत्पन्न अद्वतीय चूनौतियों का समाधान करने में वर्तफल रहता है। वैज्ञानिक और तकनीकी दस्तावेजों में, लेआउट और गैर-पाठ्य तत्वों को अनुदेखा नहीं किया जा सकता है — पैराग्राफ, गणितीय समीकरणों, तालिकाओं और आकृतियों की व्यवस्था के समृद्ध और महत्वपूर्ण अर्थ होते हैं।

वे पाठ-आधारति प्रयास लेआउट में मौजूद जानकारी को अनुदेखा करते हैं और इस प्रकार वैज्ञानिक तथा तकनीकी दस्तावेजों के अनुवाद में आने वाली बाधा को संबोधित करने के लिए अपर्याप्त हैं।

इस कमी को पूरा करने के लिए, हम PDFMathTranslate प्रस्तुत करते हैं, जो संरक्षित लेआउट के साथ PDF दस्तावेजों का अनुवाद करने वाला दुनिया का पहला ओपन-सोर्स ट्रूल है। लेआउट डिक्टिशन और बड़े भाषा मॉडल (चिर 1 देखें) में हालिया प्रगति का लाभ उठाकर, हम कम से कम पाँच प्रमुख योगदानों के साथ इस बाधा को बेहतर ढंग से संबोधित करते हैं: (1) लेआउट डिक्टिशन, अनुवाद और पुनः-रेडरिंग का कुशल कार्यप्रवाह; (2) कई भाषाओं के लिए समर्थन (3) कई अनुवाद मॉडल और सेवाओं के लिए समर्थन; (4) विविध उपयोगकर्ता इंटरफ़ेस; और (5) एक समुदाय-वाणजिय मॉडल जो स्थायी विकास को बढ़ावा देता है।

2 वास्तुकला और डिज़िन

कुल मलिकर, PDFMathTranslate को दस्तावेजों के मूल लेआउट को बनाए रखते हुए उनका अनुवाद करने के लिए डिज़िन किया गया है। सबसे पहले, यह एक PDF दस्तावेज के साथ उपयोगकर्ता-निर्दिष्ट पैरामीटर, जैसे कभी भाषाएँ और पसंदीदा अनुवाद सेवा (प्रशिष्ट A देखें) स्वीकार करता है। इसके बाद, यह दस्तावेज़ से लेआउट और पाठ्य सामग्री नकालने के लिए लेआउट का पता लगाता है। तो प्रैसरा, चयनति अनुवाद सेवा, जैसे की GPT-4, DeepL, Google, या Ollama (प्रशिष्ट A देखें) का उपयोग करके पाठों का अनुवाद किया जाता है। अंत में, अनुवादति पाठ और पहले से पता लगाए गए लेआउट को संरक्षित लेआउट वाले एक अनुवादति दस्तावेज़ के रूप में पुनः प्रस्तुत की जाता है।

तकनीकी रूप से, हम *precision, flexibility*, और *efficiency* के लिए आर्कटिक्चर डिज़िन करते हैं। विशेष रूप से, हम तीन प्रमुख घटक प्रस्तुत करते हैं: एक स्टीक पारस्पर, एक लचीला अनुवाद मॉडलवेयर, और एक कुशल कार्यप्रवाह, जिनकी व्याख्या अगले तीन उप-खंडों में की जाएगी।

स्टीक लेआउट पारस्पर

लेआउट को संरक्षित करने के लिए स्टीक पारस्परी की आवश्यकता होती है। लेआउट को स्टीक रूप से पारस्पर करने के लिए, हम एक पाइपलाइन का प्रस्तुत करते हैं।

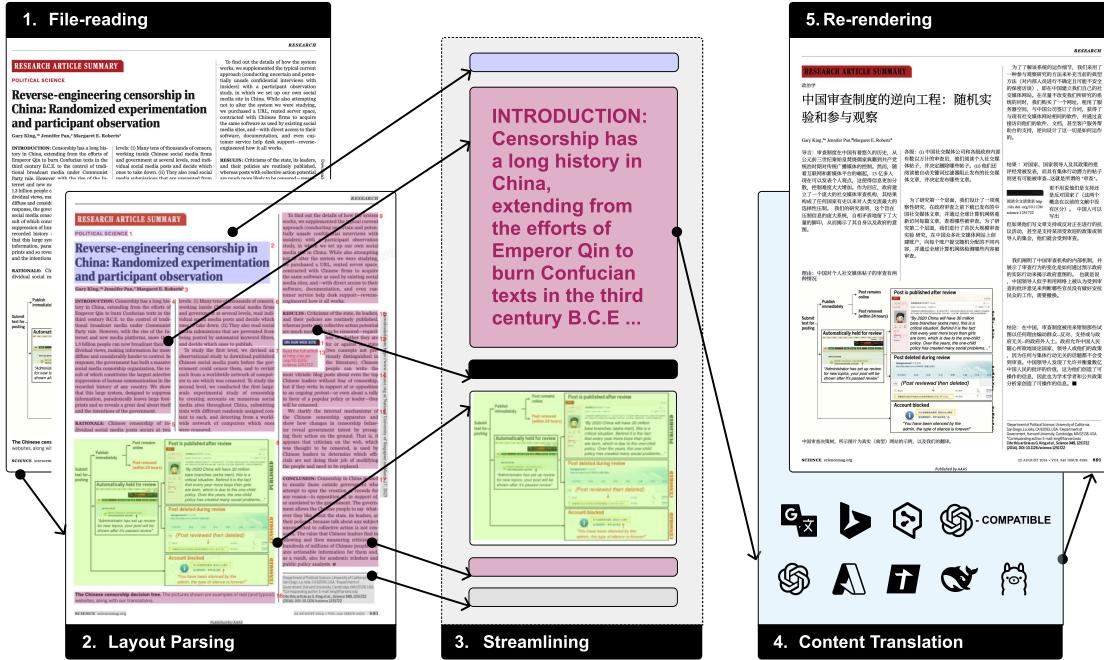


Figure 1: The architecture of PDFMathTranslate

consisting of layout detection, splitting, processing, and re-rendering. In the beginning, we exploit a recent advance in layout detection, *DocLayout-YOLO-DocStructBench-onn*. The model is another version of a SOTA solution in object detection, *Yolov10* (Wang et al.), and is fast and accurate in the specific task (Zhao et al.).

To increase the compatibility of this model, we support two model formats: the *Open Neural Network Exchange (ONNX)* (onnx) standard and a *pytorch* version. The ONNX version was chosen as default to ensure the compatibility of our parsing pipeline for diverse hardwares.

Flexible translation middleware

Translation middleware offers flexibility in terms of *language diversity*, *supported services*, and *customization capabilities*. One, **PDFMathTranslate** supports at least 56 languages (see Appendix A), including popular ones such as *English*, and some languages shared by relatively smaller communities, such as *Cherokee*. Users can freely select any of these languages as either the input or output language when translating documents.

Two, **PDFMathTranslate** supports at least 23 popular services for translation. The list of services includes prominent online services such as Google Translate and OpenAI, as well as local deployments

of models for either translation or dialogs (see Appendix A). In addition to explicitly listed services, users can integrate any new service complying with OpenAI protocols¹.

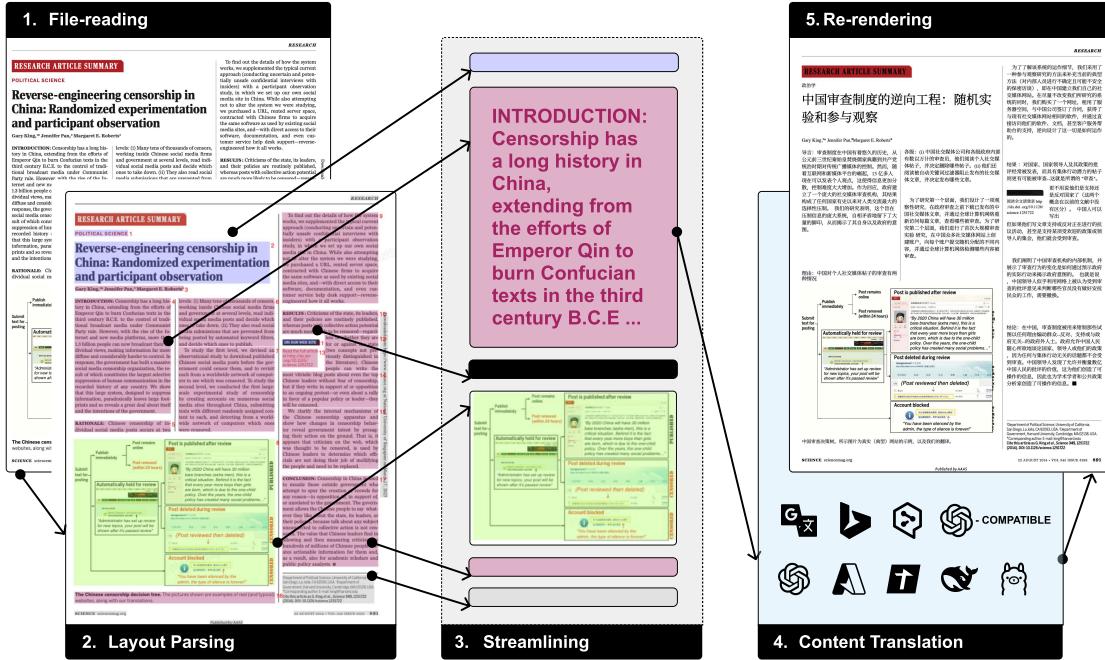
Moreover, to simplify the addition of new services, we designed a structure that separates layout flows from translation flows. In this structure, language services receive only textual inputs, enhancing the generalizability and sustainability of our project. As a result, implementing any language service requires only a concise function of fewer than 15 lines.

Three, **PDFMathTranslate** provides additional customizability in translation. We designed a customization feature prompting strategies that potentially increase the quality of translation or its adaptability in domain-specific tasks. This feature allows users to exploit advanced prompting strategies such as few-shots (Brown et al., b), Chain-of-Thought (Wei et al.), Role-playing (Shanahan et al.), etc.

Efficient streaming flow

Translating scientific and technical documents, in some cases, needs an efficient workflow for thousands of pages, challenging the performance.

¹See: <https://platform.openai.com/docs/api-reference/chat/create>



चत्तीर 1: पीडीएफमैथट्रांसलेट की संरचना

जसिमे लेआउट पहचान, वभिजन, प्रसंस्करण और पुनरा वर्तन शमलि है। शुरुआत में, हम लेआउट पहचान में हुई एक हालया प्रगति, *DocLayout-YOLO-DocStructBench-*onnx** का लाभ उठाते हैं। यह मॉडल ऑब्जेक्ट डिक्षन में एक SOTA समाधान, *Yolov10* (Wang et al.) का एक और संस्करण है, और व शिष्ट कार्य में तेज़ और सटीक है (Zhao et al.)।

इस मॉडल की अनुकूलता बढ़ाने के लिए, हम दो मॉडल प्रारूपों का समर्थन करते हैं: *Open Neural Network Exchange (ONNX)* (onnx) मानक और एक *pytorch* संस्करण। ONNX संस्करण को विधि हार्डवेयर के लिए हमारी पारसगि पाइपलाइन की अनुकूलता सुनिश्चित करने के लिए डिफिल्ट के रूप में चुना गया था।

लचीला अनुवाद मंडिलिवेयर

अनुवाद मंडिलवेयर *language diversity, supported services, और customization capabilities* के संदर्भ में लचीलापन प्रदान करता है। एक, PDFMathTranslate कम से कम 56 भाषाओं (परशिष्ट A देखें) का समरथन करता है, जिनमें English जैसी लोकप्रिय भाषाएँ, और Cherokee जैसे अपेक्षाकृत छोटे समुदायों द्वारा साझा की जाने वाली कुछ भाषाएँ शामिल हैं। दस्तावेजों का अनुवाद करते समय उपयोग करता इनमें से कर्सि भी भाषा को इनपुट या आउटपुट भाषा के रूप में सवतंतर रूप से चेन सकते हैं।

दूसरा, PDFMathTranslate अनुवाद के लिए कम से कम 23 लोकप्रयि सेवाओं का समर्थन करता है। सेवाओं की सूची में Google Translate और OpenAI जैसी परमु ख ऑनलाइन सेवाएँ, साथ ही स्थानीय परानियोजन भी शा मलिहै।

मॉडलों का, चाहे वह अनुवाद के लिए हो या संवादों के लिए (परशिष्ट A देखें)। सूपष्ट रूप से सूचीबद्ध सेवाओं के अंतरिक्त, उपयोगकर्ता OpenAI प्रौटोकाल¹ का पालन करने वाली कसी भी नई सेवा को ऐकीकृत कर सकते हैं।

इसके अतरिक्त, नई सेवाओं को जोड़ने को सरल बनाने के लए, हमने एक ऐसी संरचना तैयार की है जो लेआउट परवाह को अनुवाद परवाह से अलग करती है। इस संरचना मैं, भाषा सेवाओं को केवल पाठ्य इनपुट प्राप्त होते हैं, जो ससे हमारे प्रोजेक्ट की सामान्य परयोज्यता और स्थरित बढ़ती है। परणिमसवरप, कसी भी भाषा सेवा को लागू करने के लए 15 पंक्तियों से कम के एक संक्षिप्त फ़ंक्शन की ही आवश्यकता होती है।

तीन, PDFMathTranslate अनुवाद में अतरिक्त अनुकूलन क्षमता परदान करता है। हमने एक अनुकूलन सुवधा डिज़ाइन की है जो ऐसी रणनीतियों को प्राप्त करती है जिससे अनुवाद की गुणवत्ता या डोमेन-विशेषिक कारयों में उसकी अनुकूलनशीलता संभावित रूप से बढ़ सकती है। यह सुवधा उपयोगकरताओं को उननत प्राप्ति रणनीतियों का उपयोग करने की अनुमति देती है जैसे के फ़ियू-शॉट्स (ब्राउन एट अल., बी), चैन-ऑफ़-थॉट (वर्इ एट अल.), रोल-प्लेइंग (शानहन एट अल.), आदि

कशल सटरीमगि परवाह

वैज्ञानिकी और तकनीकी दस्तावेजों का अनुवाद करने में, कुछ मामलों में, हज़ारों पृष्ठों के लिए एक कुशल कारब्यपर वाह की आवश्यकता होती है, जो प्रदर्शन को चुनौती देता है।

¹See: <https://platform.openai.com/docs/api-reference/chat/create>

To address the issue, we landed on a streaming design that is centered around an efficient in-memory processing pipeline that minimizes disk I/O while preserving document fidelity. Specifically, the `translate_stream` function accepts a byte stream of a PDF file and transforms it into a mutable document representation. This approach allows the system to dynamically apply modifications, such as converting to PDF/A format when needed and embedding the appropriate fonts based on the target language. By employing temporary buffers and leveraging functions like `download_remote_fonts`, the design effectively manages resources and ensures that the original file remains unaltered during processing.

Our workflow supports various translation backends, with translators configured as plugins within our workflow. Different translators integrate distinct translation functionalities by inheriting from a base class and implementing the `do_translate` method. These translators can be seamlessly integrated into different user interfaces. With contributions from the community, our workflow currently supports more than 20 translators. Additionally, it provides robust extensibility for backend LLM services with the OpenAI-compatible API and traditional machine translation service.

Further optimization in performance is achieved through a combination of asynchronous execution, caching strategies, and error-handling mechanisms. Asynchronous constructs, such as `asyncio.Event`, are employed to support task cancellation and concurrency, particularly when processing large documents or handling multiple translation tasks simultaneously. In addition, the design incorporates a caching mechanism—controlled via the `ignore_cache` flag—to prevent redundant computations, save LLM tokens, and accelerate the translation process. While a subset font embedding strategy minimizes the final file size without compromising compatibility. Overall, this streaming design reflects a robust, scalable, and resource-efficient framework tailored for automated PDF translation workflows.

3 Usage and Deployments

PDFMathTranslate has been implemented in various interfaces, including a command-line tool (CLI), a graphic user interface (GUI), cross-platform applications (on MacOS and Windows), and Docker images. In addition to those offi-

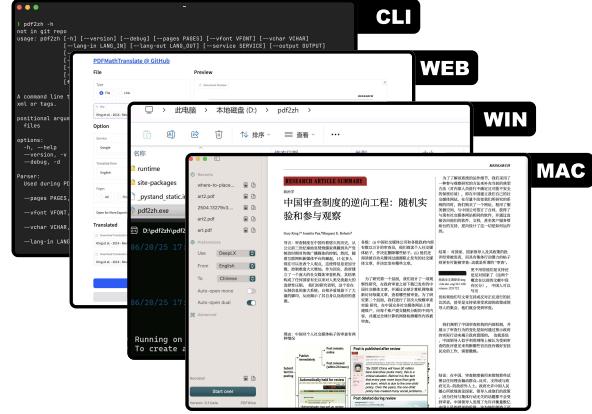


Figure 2: Officially supported interfaces



Figure 3: Major commands for CLI

cially supported interfaces, user `guaguastandup` contributes a Zotero plugin to our community.

For laymen, we offer GUI on Mac², Windows, Web, online demos, and Zotero plugin³. Specifically, desktop versions on Mac and Windows can be installed, demos are publicly accessible online, and GUI can be started using `pdf2zh -i`. Those graphic interfaces are shown in Figure 2.

For developers, we support CLI and Docker for advanced usage and substream developments. Developers who prefer command-line tools can conveniently install our Python CLI program using `pip install pdf2zh`. The command-line tool supports extensive features, with the major ones shown in Figure 3 and documented in detail⁴. For users aiming to deploy applications on servers, we offer a Docker image designed specifically to simplify online production deployment. Those images are built and distributed on different platforms for developers with potential network issues.

```
docker pull byaidu/pdf2zh
docker run -d -p 7860:7860 byaidu/
pdf2zh
```

Our comparison indicates that **PDFMathTranslate** surpasses alternative solutions. First, regarding *accessibility*, we significantly reduce user costs by providing an open-source, self-deployable, and free

²See <https://github.com/reycn/pdf2zh-mac>.

³See <https://github.com/guaguastandup/zotero-pdf2zh>; community-contributed.

⁴See <https://github.com/Byaidu/PDFMathTranslate/blob/main/docs/ADVANCED.md>

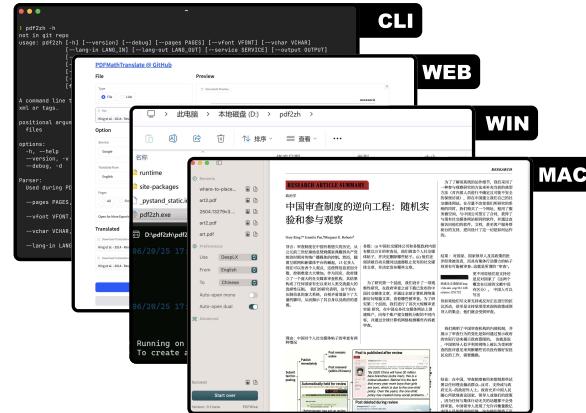
इस समस्या के समाधान के लिए, हमने एक सूटीरीमगि डज़ि इन अपनाया जो एक कुशल इन-मेमोरी परोसेसिंग पाइपला इन पर केंद्रित है। यह डिस्क I/O को न्यूनतम करता है ज बकंडिस्तावेज़ की मूलता को बनाए रखता है। वशीष रूप से , translate_stream फ़िल्टर एक PDF फ़ाइल की बाइट सूटीरीम को सवीकार करता है और इसे एक परविरतनीय द स्तावेज़ प्रतिविधित्व में बदल देता है। यह दृष्टिकोण सि स्टम को गतशील रूप से संशोधन लागू करने की अनुमतिदि ता है, जैसे आवश्यकता पड़ने पर PDF/A प्रारूप में परविर तति करना और लक्ष्य भाषा के आधार पर उपयुक्त फ़ॉन् ट एम्बेड करना। अस्थायी बफ़र का उपयोग करके और do wnload_remote_fonts जैसे फ़ंक्शंस का लाभ उठाकर, यह डज़िइन संसाधनों का प्रभावी ढंग से प्रबंधन करता है और यह सुनिश्चित करता है कि प्रिसंस्करण के दौरान मूल फ़ाइल अपरविरुद्ध रहे।

हमारा वरकफ़लो वभिन्न अनुवाद बैक-एंड का समर्थन करता है, जैसे अनुवादक हमारे वरकफ़लो के भीतर प्लग इन के रूप में कॉन्फ़िगर करि जाते हैं। वभिन्न अनुवादक एक आधार वर्ग से इनहेरिट करके और `do_translate` व धिको लागू करके वशिष्ट अनुवाद कारयकषमताओं को एकीकृत करते हैं। इन अनुवादकों को वभिन्न उपयोगकर्ता इंटरफ़ेस में सहजता से एकीकृत किया जा सकता है। समु दाय के योगदान से, हमारा वरकफ़लो वर्तमान में 20 से अ धकि अनुवादकों का समर्थन करता है। इसके अतरिक्त, यह OpenAI-संगत API और पारंपरकि मशीन अनुवाद से वा के साथ बैकएंड LLM सेवाओं के लिए सुदृढ़ वस्त्रारशी लता प्रदान करता है।

प्रदर्शन में और अधकि अनुकूलन अतुल्यकालकि नषि पादन, कैशिंग रणनीतियों और तरुट-हैडलॉगि तंतरों के संयों जन के माध्यम से प्राप्त किया जाता है। asyncio.Event जैसे अतुल्यकालकि नरिमाणों को कारय रद्द करने और स मवर्तीता का समर्थन करने के लिए नियोजिति किया जाता है, खासकर बड़े दस्तावेज़ों को संसाधनि करते समय या ए के साथ कई अनुवाद कारयों को संभालते समय। इसके अत रिक्ति, डज़िइन में एक कैशिंग तंतर शामलि है—जो ignore_cache फ़लैग के माध्यम से नर्तिरति होता है—ताकि अ नावश्यक गणनाओं को रोका जा सके, LLM टोकन बचाए जा सके और अनुवाद प्रकरया को गतिदी जा सके। जब कि एक सबसेट फ़ॉन्ट एम्बेडिंगि रणनीति संगतता से समझ तैता करि बनि अंतमि फ़ाइल आकार को कम करती है। कु ल मलिकर, यह स्ट्रीमिंग डज़िइन स्वचालिति PDF अनुव ाद वरकफ़लो के लिए तैयार एक मजबूत, स्केलेबल और सं साधन-कुशल ढांचे को दर्शाता है।

3 उपयोग और परनियोजन

PDFMathTranslate को वभिन्न इंटरफ़ेस में लागू किया गया है, जनिमे एक कमांड-लाइन टूल (CLI), एक ग्राफ़िक यूज़र इंटरफ़ेस (GUI), क्रॉस-प्लेटफ़ार्म एप्लिकेशन (मै कओएस और वडिज़ पर), और डॉकर इमेज शामलि है। इ न आधा- के अतरिक्त



Translated Text: चतिर 2: आधकारकि तौर पर समर्थति इंटरफ़ेस

Required	Optional
pdf2zh document.pdf	-s deepl -p 1-5 -li en -lo zh Using service Deep. Select pages, 1-5 From English To Chinese

चतिर 3: CLI के लिए प्रमुख कमांड

वशीष रूप से समर्थति इंटरफ़ेस, उपयोगकरता गुआगआस टैडअप हमारे समुदाय को एक ज़ोटेरो प्लगइन का योगदान देता है।

आम लोगों के लिए, हम मैक², वडिज़, वेब, ऑनलाइन डेमो और जोटेरो प्लगइन³ पर GUI प्रदान करते हैं। वशी ष रूप से, मैक और वडिज़ पर डेस्कटॉप संस्करण इस्टॉल करि जा सकते हैं, डेमो सार्वजनिक रूप से ऑनलाइन उपल बध हैं, और pdf2zh -i का उपयोग करके GUI शरू किया जा सकता है। ये ग्राफ़िक इंटरफ़ेस चतिर 2 में दर्खिए गए हैं।

डेवलपर्स के लिए, हम उन्नत उपयोग और सबसूटीम वकिस के लिए CLI और डॉकर का समर्थन करते हैं। जो डेवलपर्स कमांड-लाइन टूल पसंद करते हैं, वे `pip install pdf2zh` का उपयोग करके हमारे पायथन CLI प्रोग्राम को आसानी से इंस्टॉल कर सकते हैं। यह कमांड-लाइन टूल वयापक सुवधियों का समर्थन करता है, जनिमे से प्रमुख चतिर 3 में दर्खिए गए हैं और⁴ में वसितार से प्रलेखित हैं। सरवर पर एप्लिकेशन डप्लॉय करने का लक्ष्य रखने वा ले उपयोगकरता ओं के लिए, हम एक डॉकर इमेज प्रदान करते हैं जसि वशीष रूप से ऑनलाइन उत्पादन डप्लॉयमेट को सरल बनाने के लिए डज़िइन किया गया है। ये इमेज सं भावति नेटवर्क समस्याओं वाले डेवलपर्स के लिए वभिन्न न प्लेटफॉर्म पर बनाई और वतिरति की जाती हैं।

डॉकर पुल byaidu / pdf2zh डॉकर रन -d -p 7860:7860
byaidu / pdf2zh

हमारी तुलना दरशाती है कपीडीएफमैथट्रांस-लेट वैक लपकि समाधानों से बेहतर है। सबसे पहले, accessibility के संबंध में, हम एक ओपन-सोर्स, सवयं-तैनात करने योग य और मुफूत प्रदान करके उपयोगकरेता की लागतों को के आपी हद तक कमे करते हैं

²See <https://github.com/reycn/pdf2zh-mac>.

³See <https://github.com/guaguastandup/zotero-pdf2zh-community-contributed>.

⁴See <https://github.com/Byaidu/PDFMathTranslate/blob/main/docs/ADVANCED.md>

solution with flexible API integration. This flexibility enables broad applicability of our tool and ensures accurate translation outputs from any preferred service or model. Second, in terms of *readability*, our approach preserves basic layouts, accurately recognizes complex mathematical formulas, and supports bilingual outputs. Finally, although our solution is not the fastest—due to the computational demands of on-device inference—its *efficiency* remains comparable to other services (see 1). These advantages suggest that our tool becomes a better alternative to existing tools in the translation of scientific documents, thus potentially better reducing the language barriers.

4 Use Cases

We illustrate two typical use cases of **PDFMath-Translate** to illustrate how the tool translates text while preserving information embedded in layouts. The first use case is a technical document with both text and complex formulas, and the second use case is a research article with complex layouts and figures (see Figure 4).

Both cases show that the tool is capable of precisely translating the information within text while preserving the crucial information embedded in layouts.

5 Comparison with Existing Systems

To illustrate the advantages of our tool over other commercial and open-source alternatives in several important aspects, including *accessibility*, *readability*, and *efficiency*.

6 Sustainability

Nevertheless, maintenance is challenging in open-source projects ([Stol and Ali Babar](#)). To ensure sustainable development, we've established a *community-commerce* model, through which the incentives of developers are provided by two sources: open-source recognition and the benefits of commercial products.

First, we increase developer exposure by prominently featuring contributors on the project's homepage via third-party services and regularly highlighting recent contributions in project documentation (e.g., *1970-01-01 / Supports Google translation (by @author_handle)*). This strategy has successfully incentivized 44 global developers who collectively contributed tens of thousands of lines of code.

Second, through collaboration with a commercial partner, we provide sponsored rewards such as membership exchanges for active contributors. These 11 subscriptions sent out to developers have effectively encouraged consistent contributions, resulting in the resolution of over 485 user-reported issues (by April 2025). Such cooperation has enabled the integration of advanced academic and community-driven technologies, including ONNX model support ([Wybxc, 2025](#)) and Qwen ([ws051682, 2025](#)), thus advancing cross-linguistic scientific communication.

This sustainable collaboration model propelled the project to the top of GitHub's global trends for over a week, garnering more than 25k stars, 22k downloads, and over 49k Docker pulls (by June 2025). Additionally, our project's success has inspired a range of new commercial products with similar functionalities, significantly impacting both the scientific community and the broader public.

7 Limitations

While we have made several key improvements, our tool still has specific limitations. First, translation quality remains highly dependent on the underlying translation models and prompts. Second, the accuracy of the layout detection model in identifying various document layouts similarly depends on the quality of the layout detection models employed. Third, the current version of the tool can not yet handle scanned PDFs lacking optimal optical character recognition (OCR); however, this last feature is planned for implementation.

8 Acknowledgments

The authors thank all 44 contributors⁵ of the open-sourced project on GitHub; platforms allowing us to host demos (Hugging Face, ModelScope); services we used in the demo (Google Translate); and all comments from users.

⁵They are 7shi, Byaidu, Copilot, Cycloctane, Hanaasagi, IuvenisSapiens, JEFF-dev-ui, Tql-ws1, Wybxc, YadominJinta, Zxis233, alohays, aseaday, awwaawa, borcation, charles7668, chiu0602, czz404, damaoooo, dependabot[bot], domonnss, eltociear, helloffinch, highkay, hotwa, imClumsyPanda, kharkover, kidach1, lintian233, mydreamworldpolly, namazuchin, qqueing, reycn, tastelikefeet, timelic, treeleaves30760, tylzh97, ws051682, wx-11, xxnuo, xyzxyz233, yidasanqian, ymattw, zqqian; ordered alphabetically, retrieved through the official GitHub API by March 2025)

लचीले एपीआई एकीकरण के साथ समाधान। यह लचीलापन हमारे उपकरण की व्यापक परयोजनायता को सक्षम बनाता है और कसी भी पसेदीदा सेवा या मॉडल से सटीक अनुवाद आउटपुट सुनशिचति करता है। दूसरा, *read-ability* के संदर्भ में, हमारा दृष्टिकोण बुनियादी लेआउट को सरक्षित करता है, जटिल गणितीय सूत्रों को सटीक रूप से पहचानता है, और द्विभाषी आउटपुट का समर्थन करता है। अंत में, हालांकि हमारा समाधान सबसे तेज़ नहीं है—ऑन-डिवाइस अनुमान की कम्प्यूटेशनल मांगों के कारण—इसकी *efficiency* अन्य सेवाओं के बराबर रहती है (देखें 1)। ये फायदे बताते हैं कि वैज्ञानिक दस्तावेजों के अनुवाद में हमारा उपकरण मौजूदा उपकरणों का एक बेहतर विकल्प बन जाता है, जिससे संभावित रूप से भाषा की बाधाओं को बेहतर ढंग से कम किया जा सकता है।

4 उपयोग के मामले

हम PDFMath-Translate के दो विशिष्ट उपयोग मामलों का उदाहरण देते हैं, यह दर्शाने के लिए कियह उपकरण ले आउट में अंतर्नहिति जानकारी को बनाए रखते हुए पाठ का अनुवाद करता है। पहला उपयोग मामला एक तकनीकी दस्तावेज़ है जिसमें पाठ और जटिल सूत्र दोनों हैं, और दूसरा उपयोग मामला जटिल लेआउट और चर्टिरों वाला एक शोध लेख है (चित्र 4 देखें)।

दोनों मामले दर्शाते हैं कि यह उपकरण पाठ के भीतर की जानकारी का सटीक अनुवाद करने में सक्षम है, साथ ही लेआउट में अंतर्नहिति महत्वपूर्ण जानकारी को भी सुरक्षित रखता है।

5 मौजूदा प्रणालयों से तुलना

हमारे टूल के लाभों को अन्य वाणिज्यिक और ओपन-सोर्स विकल्पों की तुलना में कई महत्वपूर्ण पहलुओं में प्रदर्शित करने के लिए, जिनमें *accessibility*, *readability*, *efficiency*, और *shamali* हैं।

6 टकिऊपन

फरि भी, ओपन-सोर्स परयोजनाओं में रखरखाव चूनौतीपूर्ण है (स्टोल और अली बाबर)। सतत विकास सुनिश्चिति करने के लिए, हमने एक *community-commerce* मॉडल स्थापित किया है, जिसके माध्यम से डेवलपर्स को प्रोत्साहन दो स्रोतों से प्रदान किया जाता है: ओपन-सोर्स मान्यता और व्यावसायिक उत्पादों के लाभ।

सबसे पहले, हम तृतीय-पक्ष सेवाओं के माध्यम से परयोजना के होमपेज पर योगदानकर्ताओं को प्रमुखता से प्रदर्शित करके और परयोजना के दस्तावेजों में हाल के यो गदानों को नियमित रूप से उजागर करके (उदाहरण के लिए, 1970-01-01 / Supports Google translation (by @author_handle) डेवलपर्स की दृश्यता बढ़ाते हैं। इस रणनीति ने 44 वैश्वकि डेवलपर्स को सफलतापूर्वक प्रोत्साहित किया है, जिन्होंने सामूहिक रूप से दस्तावेज़ों हजार लाइनों कोड की योगदान की।

दूसरे, एक व्यावसायिक साझेदार के साथ सहयोग के माध्यम से, हम सक्रिय योगदानकर्ताओं के लिए सदस्यता विनियम जैसे परायोजिति पुरस्कार प्रदान करते हैं। डेवलपर्स को भेजी गई इन 11 सदस्यताओं ने लगातार योगदानों को प्रभावी ढंग से प्रोत्साहित किया है, जिसके परणामस् वरूप 485 से अधिक उपयोगकर्ता-रपिरेटेड मूद्दों का समाधान हुआ है (अप्रैल 2025 तक)। इस तरह के सहयोग ने ONNX मॉडल समर्थन (Wybxc, 2025) और क्यूवेन (ws051682, 2025) सहित उन्नत अकादमिक और समुदाय-प्रेरित प्रौद्योगिकियों के एकीकरण को सक्षम किया है, जिससे अंतर-भाषाई वैज्ञानिक संचार को आगे बढ़ाया गया है।

इस सथायी सहयोग मॉडल ने परयोजना को एक सप्ताह ह से अधिक समय तक गटिहब के वैश्वकि उज्ज्ञानों में शीर्ष पर पहुँचाया, जिसने 25 हजार से अधिक स्टार, 22 हजार डाउनलोड, और 49 हजार से अधिक डॉकर प्ल (जून 2025 तक) प्राप्त किए। इसके अतरिक्त, हमारी परयोजना की सफलता ने समान कार्यक्रमताओं वाले कई नए व्यावसायिक उत्पादों को प्रेरित किया है, जिसका वैज्ञानिक समुदाय और व्यापक जनसमूह दोनों पर महत्वपूर्ण प्रभाव पड़ा है।

7 सीमाएँ

हालाँकि हमने कई महत्वपूर्ण सुधार किए हैं, फरि भी हम रे टूल में कुछ विशिष्ट सीमाएँ हैं। पहला, अनुवाद की गुणवत्ता अंतर्नहिति अनुवाद मॉडल और प्रोमूट पर अत्यधिक निभ्रित करती है। दूसरा, विभिन्न दस्तावेज़ लेआउट की पहचान करने में लेआउट डिटिक्शन मॉडल की सटीकता भी उपयोग किए गए लेआउट डिटिक्शन मॉडल की गुणवत्ता पर निभ्रित करती है। तीसरा, टूल का वर्तमान संस्करण अभी तक सूकैन किए गए ऐसे PDF को नहीं संभाल सकता जिनमें इष्टटम ऑप्टिकल कैरेक्टर रिकॉर्डिंगनशिन (OCR) का अभाव है; हालाँकि, इस अंतमि सुवधि को लागू करने की योजना है।

8 आधार

लेखक गटिहब पर ओपन-सोर्स प्रोजेक्ट के सभी 44 योगदानकर्ताओं⁵; हमें डेमो होस्ट करने की अनुमति दिने वाले प्लेटफॉर्मों (हाग्गि फेस, मॉडलस्कोप); डेमो में उपयोग की गई सेवाओं (गगल टरांसलेट); और उपयोगकर्ताओं की सभी टपिपणियों के लिए धन्यवाद देते हैं।

⁵They are 7shi, Byaidu, Copilot, Cycloctane, Hanaasagi, IuvenisSapiens, JEFF-dev-ui, Tql-ws1, Wybxc, YadominJinta, Zxis233, alohays, aseaday, awwaawa, borcation, charles7668, chiu0602, czz404, damaaaa, dependabot[bot], domonness, eltociear, helloffinch, highkay, hotwa, imClumsyPanda, kharkover, kidachi1, lintian233, mydreamworldpolly, namazuchin, qqueing, reycn, tastelikefeet, timelic, treeleaves30760, tylzh97, ws051682, wx-11, xxnuo, xyzzyx233, yidasanqian, ymattw, zqqian; ordered alphabetically, retrieved through the official GitHub API by March 2025)

	OURS	IMT ⁵	Doc2X	TeX based ⁶	Google	DeepL
Accessibility	open-source	✓		✓		
	Deployment	✓		✓		
	API integration	✓	✓		✓	✓
	Price	Free	Paid	Paid	Free	Free
Readability	Layout ¹	✓	✓	✓		
	Formula ²	✓	✓	✓	✓	
	Bilingual ³	✓	Partial	Partial	✓	✓
	OCR ⁴	✓ ⁷	✓	✓	✓	✓
Efficiency	Batch tasks	✓		✓	✓	✓
	Speed (sec/page)	1.47	1.50	1.86	1.67	0.38
1 Whether the tool can preserve figures, images, and formulas (although the formulas may be incorrectly displayed).						
2 Whether the tool can translate documents containing formulas and correctly handle the positions of the formulas.						
3 Whether the tool supports bilingual exports where both the original and translated text are simultaneously readable.						
4 Whether the tool can handle scanned documents instead of digital documents.						
5 IMT: Immersive Translate PDF Pro.						
6 The original LaTeX file is required for translation.						
7 Available in a community fork: PDFMathTranslate/PDFMathTranslate-next.						

Table 1: Comparison with other projects and productions

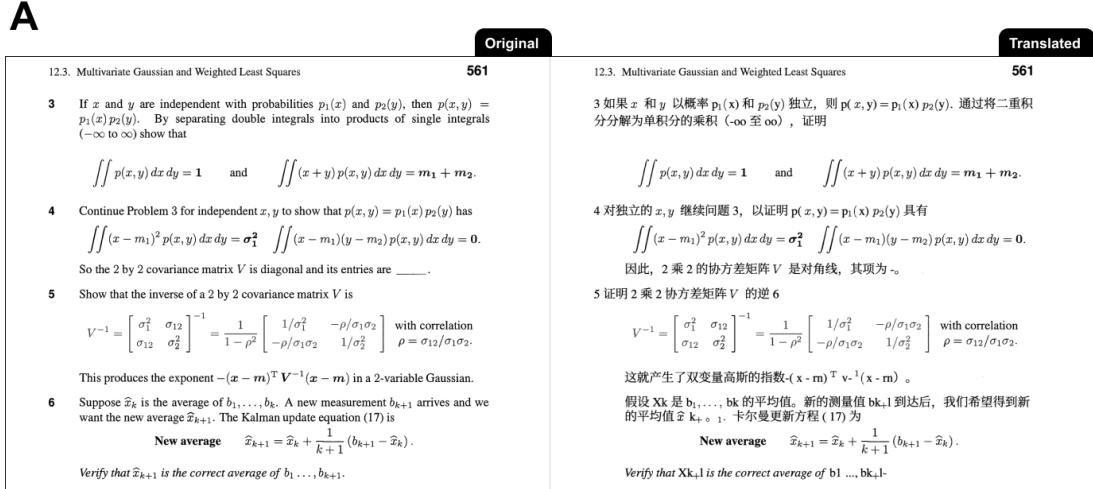


Figure 4: Use cases.

	OURS	IMT ⁵	Doc2X	TeX based ⁶	Google	DeepL
Accessibility	open-source	✓		✓		
	Deployment	✓		✓		
	API integration	✓	✓		✓	✓
	Price	Free	Paid	Paid	Free	Free
Readability	Layout ¹	✓	✓	✓		
	Formula ²	✓	✓	✓	✓	
	Bilingual ³	✓	Partial	Partial	✓	✓
	OCR ⁴	✓ ⁷	✓	✓	✓	✓
Efficiency	Batch tasks	✓		✓	✓	✓
	Speed (sec/page)	1.47	1.50	1.86	1.67	0.38
	Speed (sec/page)					1.88

2 क्या यह उपकरण आकृतियों, छवियों और सूत्रों को संरक्षित कर सकता है (भले ही सूत्र गलत प्रदर्शित हो सकते हैं)?² क्या यह उपकरण सूत्रों वाले दस्तावेजों का अनुवाद कर सकता है और सूत्रों की स्थिति को सही ढंग से संभाल सकता है?³ क्या यह उपकरण द्विभाषी नारियात का समर्थन करता है जहाँ मूल और अनुवादित दोनों पाठ एक साथ पढ़े जा सकते हैं?⁴ क्या यह उपकरण डिजिटल दस्तावेजों के बजाय स्कैन करिए गए दस्तावेजों को संभाल सकता है?⁵ आईएमटी: इमर्सिव ट्रांसलेट पीडीएफ प्रो।⁶ अनुवाद के लिए मूल लाटेक्स (LaTeX) फ़ाइल आवश्यक है।⁷ एक सामुदायिक फ़ाइल में उपलब्ध: PDFMathTranslate/PDFMathTranslate-next।

तालिका 1: अन्य परयोजनाओं और उत्पादनों के साथ तुलना

A	Original	Translated
<p>12.3. Multivariate Gaussian and Weighted Least Squares 561</p> <p>3 If x and y are independent with probabilities $p_1(x)$ and $p_2(y)$, then $p(x, y) = p_1(x)p_2(y)$. By separating double integrals into products of single integrals ($-\infty$ to ∞) show that</p> $\iint p(x, y) dx dy = 1 \quad \text{and} \quad \iint (x+y)p(x, y) dx dy = m_1 + m_2.$ <p>4 Continue Problem 3 for independent x, y to show that $p(x, y) = p_1(x)p_2(y)$ has</p> $\iint (x-m_1)^2 p(x, y) dx dy = \sigma_1^2 \quad \iint (x-m_1)(y-m_2) p(x, y) dx dy = 0.$ <p>So the 2 by 2 covariance matrix V is diagonal and its entries are _____.</p> <p>5 Show that the inverse of a 2 by 2 covariance matrix V is</p> $V^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix} \quad \text{with correlation } \rho = \sigma_{12}/\sigma_1\sigma_2.$ <p>This produces the exponent $-(x-m)^T V^{-1}(x-m)$ in a 2-variable Gaussian.</p> <p>6 Suppose \hat{x}_k is the average of b_1, \dots, b_k. A new measurement b_{k+1} arrives and we want the new average \hat{x}_{k+1}. The Kalman update equation (17) is</p> $\text{New average} \quad \hat{x}_{k+1} = \hat{x}_k + \frac{1}{k+1} (b_{k+1} - \hat{x}_k).$ <p>Verify that \hat{x}_{k+1} is the correct average of b_1, \dots, b_{k+1}.</p>	<p>12.3. Multivariate Gaussian and Weighted Least Squares 561</p> <p>3 如果 x 和 y 以概率 $p_1(x)$ 和 $p_2(y)$ 独立，则 $p(x, y) = p_1(x)p_2(y)$。通过将二重积分分解为单积分的乘积 (-oo 到 oo)，证明</p> $\iint p(x, y) dx dy = 1 \quad \text{and} \quad \iint (x+y)p(x, y) dx dy = m_1 + m_2.$ <p>4 对独立的 x, y，继续问题 3，以证明 $p(x, y) = p_1(x)p_2(y)$ 具有</p> $\iint (x-m_1)^2 p(x, y) dx dy = \sigma_1^2 \quad \iint (x-m_1)(y-m_2) p(x, y) dx dy = 0.$ <p>因此，2 乘 2 的协方差矩阵 V 是对角线，其项为 \sim。</p> <p>5 证明 2 乘 2 协方差矩阵 V 的逆 6</p> $V^{-1} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}^{-1} = \frac{1}{1-\rho^2} \begin{bmatrix} 1/\sigma_1^2 & -\rho/\sigma_1\sigma_2 \\ -\rho/\sigma_1\sigma_2 & 1/\sigma_2^2 \end{bmatrix} \quad \text{with correlation } \rho = \sigma_{12}/\sigma_1\sigma_2.$ <p>这就产生了双变量高斯的指数 $(x-m)^T V^{-1}(x-m)$。</p> <p>假设 X_k 是 b_1, \dots, b_k 的平均值。新的测量值 b_{k+1} 到达后，我们希望得到新的平均值 \hat{x}_{k+1}。卡尔曼更新方程 (17) 为</p> $\text{New average} \quad \hat{x}_{k+1} = \hat{x}_k + \frac{1}{k+1} (b_{k+1} - \hat{x}_k).$ <p>Verify that \hat{x}_{k+1} is the correct average of b_1, \dots, b_{k+1}。</p>	



चत्तर 4: उपयोग के मामले।

9 Ethics and Broader Impacts

Our work has already made a significant positive impact within both multilingual scientific and open-source communities. However, as a flexible, widely used tool, it raises potential ethical concerns regarding the copyright of documents. Translating documents without proper permissions could challenge the intellectual property rights of scientific works and innovations. We welcome suggestions from experts in intellectual property to mitigate the concern.

A Appendix

A.1 Appendix 1. Language and services supported in PDFMathTranslate

References

- Ulrich Ammon. Linguistic inequality and its effects on participation in scientific discourse and on global knowledge accumulation – With a closer look at the problems of the second-rank language communities. 3(2):333–355.
- Ulrich Ammon. 2012. Linguistic inequality and its effects on participation in scientific discourse and on global knowledge accumulation—with a closer look at the problems of the second-rank language communities. *Applied Linguistics Review*, 3(2):333–355.
- Anees Bahji, Laura Acion, Anne-Marie Laslett, and Bryon Adinoff. Exclusion of the non-English-speaking world from the scientific literature: Recommendations for change for addiction journals and publishers. 40(1):6–13.
- Peter F. Brown, John Cocke, Stephen A. Della Pietra, Vincent J. Della Pietra, Fredrick Jelinek, John D. Lafferty, Robert L. Mercer, and Paul S. Roossin. a. A Statistical Approach to Machine Translation. 16(2):79–85.
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, and 12 others. b. Language Models are Few-Shot Learners. In *Advances in Neural Information Processing Systems*, volume 33, pages 1877–1901. Curran Associates, Inc.
- Kumju Hwang. 2005. The inferior science and the dominant use of english in knowledge production: A case study of korean science and technology. *Science Communication*, 26(4):390–427.
- Melvin Johnson, Mike Schuster, Quoc V. Le, Maxim Krikun, Yonghui Wu, Zhifeng Chen, Nikhil Thorat, Fernanda Viégas, Martin Wattenberg, and Greg Corrado. Google’s multilingual neural machine translation system: Enabling zero-shot translation. 5:339–351.
- Weishu Liu. The changing role of non-English papers in scholarly communication: Evidence from Web of Science’s three journal citation indexes. 30(2):115–123.
- Scott L. Montgomery. *Does Science Need a Global Language?: English and the Future of Research*. University of Chicago Press.
- onnx. Onnx/onnx.
- Valeria Ramírez-Castañeda. 2020. Disadvantages in preparing and publishing scientific papers caused by the dominance of the english language in science: The case of colombian researchers in biological sciences. *PloS one*, 15(9):e0238372.
- Klaus Schubert. 2012. Technical translation. In *Handbook of Translation Studies: Volume 1*, pages 350–355. John Benjamins Publishing Company.
- Rico Sennrich, Barry Haddow, and Alexandra Birch. Neural Machine Translation of Rare Words with Subword Units. *Preprint*, arXiv:1508.07909.
- Murray Shanahan, Kyle McDonell, and Laria Reynolds. Role play with large language models. 623(7987):493–498.
- Klaas-Jan Stol and Muhammad Ali Babar. Challenges in using open source software in product development: A review of the literature. In *Proceedings of the 3rd International Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Development*, pages 17–22. ACM.
- UNESCO. Open science.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need. *Advances in Neural Information Processing Systems*, 30.
- Rainald Von Gizycki. Centre and Periphery in the International Scientific Community: Germany, France and Great Britain in the 19th Century. 11(4):474–494.
- Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, Jungong Han, and Guiguang Ding. YOLOv10: Real-Time End-to-End Object Detection. *Preprint*, arXiv:2405.14458.
- Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Brian Ichter, Fei Xia, Ed Chi, Quoc V. Le, and Denny Zhou. Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. 35:24824–24837.
- ws051682. 2025. feat(translator): Add Ali Qwen MT translator by ws051682 · Pull Request #585 · Byaidu/PDFMathTranslate. [Online; accessed 18. Mar. 2025].

9 नैतकिता और व्यापक प्रभाव

हमारे काम ने पहले ही बहुभाषी वैज्ञानिक और ओपन-सोर्स दोनों समुदायों में एक महत्वपूर्ण सकारात्मक प्रभाव डाला है। हालांकि, एक लचीले, व्यापक रूप से उपयोग कए जाने वाले उपकरण के रूप में, यह दस्तावेजों के कॉपीराइट के संबंध में संभावित नैतिक चतिआएं पैदा करता है। उचिति अनुमति के बिना दस्तावेजों का अनुवाद करना वैज्ञानिक कार्यों और नवाचारों के बौद्धिक संपदा अधिकारों को चुनौती देसकता है। हम इस चतिआ को कम करने के लिए बौद्धिक संपदा विशेषज्ञों के सुझावों का स्वागत करते हैं।

A परशिष्ट

A.1 परशिष्ट 1. PDFMathTranslate में समर्थन भाषाएँ और सेवाएँ

संदर्भ

Ulrich Ammon. भाषाई असमानता और वैज्ञानिक विभिन्न में भागीदारी तथा वैश्वकि ज्ञान संचय पर इसके प्रभाव – दृष्टीय-सूत्रीय भाषा समुदायों की समस्याओं पर एक गहन दृष्टिकोण के साथ। 3(2):333–355.

उलरचि एमोन. 2012. भाषाई असमानता और वैज्ञानिक विभिन्न श में भागीदारी तथा वैश्वकि ज्ञान संचय पर इसके प्रभाव – दृष्टीय-सूत्रीय भाषा समुदायों की समस्याओं पर एक गहन दृष्टिकोण के साथ। *Applied Linguistics Review*, 3(2):333–355.

अनीस बाजी, लौरा असयिन, ऐनी-मैरी लेसलेट, और बायरन एडनॉफ। वैज्ञानिक साहित्य से गैर-अंग्रेजी भाषी दुनिया का बहुषक्तरण: व्यसन संबंधी पत्रकाओं और प्रकाशकों के लिए पर्यावरण हेतु सफारियाँ। 40(1):6–13.

पीटर एफ. बराउन, जॉन कॉक, सटीफन ए. डेला परितरा, वर्सिटी जे. डेला परितरा, फरेडरकि जेलनिक, जॉन डी. लफर्टी, रॉबर्ट एल. मरसर, और पॉल एस. रूसनि. a. मशीन अनुवाद के लिए एक सांख्यिकीय दृष्टिकोण। 16(2):79–85.

टॉम ब्राउन, बेजामनि मान, नकि राइडर, मेलानी सुब्बया, जेरेड डी कपलान, परफूल धारीवाल, अरवदि नीलकांतन, परिणव शयाम, गरीश शास्त्री, अमांडा आसकेल, संधनिअग्रवाल, एरवि ल हरबरट-वॉस, गेरेटचन कुगुर, टॉम हेनगिन, रेवॉन चाइल्ड, आदतिय रमेश, डेनयिल जुगिलर, जेफरी वू, कूलेमेस वटिर, और 12 अन्य। b. भाषा मॉडल कुछ-शॉट सीखने वाले हैं।

Advances in Neural Information Processing Systems में, खंड 33, पृष्ठ 1877-1901। करून एसोसिएट्स, इंक.

कुमज हवांग. 2005. ज्ञान उत्पादन में नमिनतर वैज्ञानिक और अंग्रेजी का पूरमुख उपयोग: कौरियाई वैज्ञानिक और प्रौद्योगिकी की का एक क्षेत्र अध्ययन। *Science Communication*, 26(4): 390–427.

मेलवनि जॉनसन, माइक शूस्टर, क्रोक वी. ले, मैक्समि क्रकिन, योगद्वार्ड वू, झर्फिंग चेन, नेखिलि थोरात,

फरनेंडा विएगास, मार्टनि वाटेनबर्ग, और ग्रेग कोराडो। गूगल का बहुभाषी न्यूरूल मशीन अनुवाद प्रणाली: शून्य-शॉट अनुवाद को सक्षम करना। 5:339–351.

वीशु लियू। अकादमिक संचार में गैर-अंग्रेजी शोधपत्रों की बदलती भूमिका: वेब ऑफ साइंस के तीन जर्नल उद्धरण सूचकांकों से प्रमाण। 30(2):115–123.

स्कॉट एल. मॉन्टगोमरी। *Does Science Need a Global Language?: English and the Future of Research*। शक्तिगोवि श्ववदियालय प्रेस।

ओएनएनएक्स. ओएनएनएक्स/ओएनएनएक्स।

वेलरया रामरिज़-कास्टेनेडा। 2020। वैज्ञान में अंग्रेजी भाषा के प्रभुत्व के कारण वैज्ञानिक पत्रों की तैयारी और प्रकाशन में कमयाँ: जैवकि वैज्ञान में कोलंबियाई शोधकर्ताओं का मामला। *PloS one*, 15(9):e0238372।

Klaus Schubert. 2012. तकनीकी अनुवाद. में *Handbook of Translation Studies: Volume 1*, पृष्ठ 350–355. जॉन बेजामनि पब्लिशिंग कंपनी।

Rico Sennrich, Barry Haddow, और Alexandra Birch। दुखल शब्दों का उप-शब्द इकाइयों के साथ तंत्रकिय मशीन अनुवाद। *Preprint*, arXiv:1508.07909।

मरे शनाहन, काइल मैकडोनेल, और लारया रेनॉल्ड्स। बड़े भाषा मॉडलों के साथ भूमिका नभिना। 623(7987):493–498।

कूलास-जान स्टोल और मुहममद अली बाबर। उत्पाद विकास में औपन सोर्स सॉफ्टवेयर के उपयोग में चुनौतयाँ: साहित्य की समीकृता। में *Proceedings of the 3rd International Workshop on Emerging Trends in Free/Libre/Open Source Software Research and Development*, पृष्ठ 17–22। ACM।

यूनेस्को। मुक्त वैज्ञानिक।

आशीष वासवाणी, नोआम शज़ीर, नकी परमार, जेकब उश्कोरा इत, एलियन जोन्स, एडन एन गोमेज़, लुकाज़ काइज़र, और इलीया पोलोसुकनि। 2017। ध्यान ही सब कुछ है। *Advances in Neural Information Processing Systems*, 30।

रेनाल्ड वॉन गजिकी। अंतरराष्ट्रीय वैज्ञानिक समुदाय में केंद्र और पराधि। 19वीं शताब्दी में जर्मनी, फ्रांस और ग्रेट ब्रिटेन। 11(4):474–494।

Ao Wang, Hui Chen, Lihao Liu, Kai Chen, Zijia Lin, J. ungong Han, and Guiguang Ding. YOLOv10: Real-Time End-to-End Object Detection. *Preprint*, arXiv:2405.14458।

जेसन वेई, ज्येष्ठी वांग, डेल शर्मैनस, मार्टेन बॉसमा, बरायन इच्टर, फर्ड ज्याया, एड ची, क्र्वॉक वी. ले, और डेनी झोउ। विचार-शुरुखला प्रौम्पटगि बड़े भाषा मॉडल में तरक को प्रेरति करता है। 35:24824–24837।

ws051682. 2025. feat(translator): अली क्वेन एमटी अनुवादक ws051682 द्वारा जोड़े। पुल अनुरोध #585। Byaidu/PDFMathTranslate. [ऑनलाइन; 18 मार्च 2025 को एक्सेस किया गया]।

Wyxyc. 2025. feat: onnx support by Wyxbc · Pull Request #116 · Byaidu/PDFMathTranslate. [Online; accessed 18. Mar. 2025].

Zhiyuan Zhao, Hengrui Kang, Bin Wang, and Conghui He. DocLayout-YOLO: Enhancing Document Layout Analysis through Diverse Synthetic Data and Global-to-Local Adaptive Perception. *Preprint*, arXiv:2410.12628.

Jinhua Zhu, Yingce Xia, Lijun Wu, Di He, Tao Qin, Wengang Zhou, Houqiang Li, and Tie-Yan Liu. Incorporating BERT into Neural Machine Translation. *Preprint*, arXiv:2002.06823.

Wybxc. 2025. feat: onnx सपोर्ट Wybxc द्वारा · पुल रक्त
वेस्ट #116 · Byaidu/PDFMathTranslate. [ऑनलाइन; 18
मार्च 2025 को एक्सेस किया गया]। जियुआन ज्ञाओ, हेंग रुई
कांग, बनि वांग, और कॉन-गुई है। डॉकलेआउट-योलओ: वर्विधि
सशीटकि डेटा और वैश्वकि-से-सथानीय अनुकूली धारणा के मा
ध्यम से दस्तावेज लेआउट वश्लेषण को बढ़ाना। *Preprint*, ar
Xiv:2410.12628. जनिहुआ झु, यगिसे जयिं, लजिन वू, डी है,
ताओ कनि, वेगांग झोउ, होउकाँआंग ली, और टए-यान लिं। न्
यूरल मशीन ट्रांसलेशन में BERT को शामिल करना। *Preprint*,
arXiv:2002.06823.

Table 2: Supported languages and services for translation in PDFMathTranslate

Category	Details	Total
Service or models	Google (<i>default</i>), Bing, DeepL, DeepLX, Ollama, Ali Qwen Translation, Ollama X inference, Gemma, OpenAI, OpenAI-like ^a , DeepSeek, AzureOpenAI, Zhipu, ModelScope, Silicon, Gemini, Azure, Tencent, Dify, AnythingLLM, Argos Translate, Grok, Groq	22 ^a
Input and output languages	Amharic, Arabic, Basque, Bengali, English (UK), Portuguese (Brazil), Bulgarian, Catalan, Cherokee, Croatian, Czech, Danish, Dutch, English (US), Estonian, Filipino, Finnish, French, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Korean, Latvian, Lithuanian, Malay, Malayalam, Marathi, Norwegian, Polish, Portuguese (Portugal), Romanian, Russian, Serbian, Chinese (Simplified), Slovak, Slovenian, Spanish, Swahili, Swedish, Tamil, Telugu, Thai, Chinese (Traditional), Turkish, Urdu, Ukrainian, Vietnamese, Welsh	56 ^b

Note. ^a The number of supported languages far exceeds what is shown in the table. Our protocol, designed as a general framework for OpenAI-like services, allows users to integrate any translation system, including locally hosted models or private online services, ensuring extensive language support.

^b Based on the default service (Google Translate) and interface (CLI). However, the actual number may be significantly higher due to the extensibility of services and models.

तालिका 2: PDFMathTranslate में अनुवाद के लिए समर्थित भाषाएँ और सेवाएँ

Category	Details	Total
Service or models	Google (<i>default</i>), Bing, DeepL, DeepLX, Ollama, Ali Qwen Translation, Ollama X inference, Gemma, OpenAI, OpenAI-like ^a , DeepSeek, AzureOpenAI, Zhipu, ModelScope, Silicon, Gemini, Azure, Tencent, Dify, AnythingLLM, Argos Translate, Grok, Groq	22 ^a
Input and output languages	Amharic, Arabic, Basque, Bengali, English (UK), Portuguese (Brazil), Bulgarian, Catalan, Cherokee, Croatian, Czech, Danish, Dutch, English (US), Estonian, Filipino, Finnish, French, German, Greek, Gujarati, Hebrew, Hindi, Hungarian, Icelandic, Indonesian, Italian, Japanese, Kannada, Korean, Latvian, Lithuanian, Malay, Malayalam, Marathi, Norwegian, Polish, Portuguese (Portugal), Romanian, Russian, Serbian, Chinese (Simplified), Slovak, Slovenian, Spanish, Swahili, Swedish, Tamil, Telugu, Thai, Chinese (Traditional), Turkish, Urdu, Ukrainian, Vietnamese, Welsh	56 ^b

Note. ^a समर्थित भाषाओं की संख्या तालिका में दर्खिए गए से कहीं अधिक है। हमारा प्रोटोकॉल, जैसे OpenAI जैसी सेवाओं के लिए एक सामान्य ढाँचे के रूप में डिज़िलन किया गया है, उपयोगकर्ताओं को कसी भी अनुवाद प्रणाली को एकीकृत करने की अनुमति दीता है, जिसमें स्थानीय रूप से होस्ट किए गए मॉडल या नज़ी ऑनलाइन सेवाएँ शामिल हैं, जिससे व्यापक भाषा समर्थन सुनेशिच्चित होता है।

^b डफिल्ट सेवा (गूगल ट्रांसलेट) और इंटरफेस (सीएलआई) के आधार पर। हालांकि, सेवाओं और मॉडलों की वसितारणीयता के कारण वास्तवकि संख्या काफी अधिक हो सकती है।