

पीडीएफमैथट्रांसलेट: लेआउट को संरक्षित रखते हुए वैज्ञानिक दस्तावेज़ अनुवाद

रोगशनि ओयांग^{1,†}, चांग चु^{2,*†}, झकिआंग शनि³, शियांगयाओ मा⁴ ¹नेशनल यूनिवर्सिटी ऑफ सगिापुर। ²सिंधुआ विश्वविद्यालय। ³चीनी विज्ञान अकादमी विश्वविद्यालय ⁴फनस्टोरी.एआई लमिटिड। *पत्राचार: chuc23@mails.tsinghua.edu.cn। † ये लेखक इस कार्य में समान रूप से योगदान करते हैं।

अमूर्त

वैज्ञानिक दस्तावेज़ों में भाषा संबंधी बाधाएँ विज्ञान और प्रौद्योगिकियों के प्रसार और विकास में बाधा डालती हैं। हालाँकि, ऐसे दस्तावेज़ों का अनुवाद करने के पछिले प्रयासों ने लेआउट में मौजूद जानकारी को ज्यादातर अनदेखा कर दिया। इस कमी को पूरा करने के लिए, हम PDFMathTranslate पेश करते हैं, जो लेआउट को बनाए रखते हुए वैज्ञानिक दस्तावेज़ों का अनुवाद करने वाला दुनिया का पहला ओपन-सोर्स सॉफ्टवेयर है। बड़े भाषा मॉडल और सटीक लेआउट पहचान में नवीनतम प्रगति का लाभ उठाते हुए, हम सटीकता, 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 परिचय

विज्ञान और प्रौद्योगिकियों में प्रगति विभिन्न भाषाओं में फैली और विकसित होती है (वान गज़ी-की; मोटगोमरी)। फिर भी, भाषा बाधाएँ वैज्ञानिक प्रगति के प्रसार और विकास में बाधा डालती हैं (रामírez-कास्टानेडा, 2020; हवांग, 2005; एमोन, 2012)। उदाहरण के लिए, जबकि विज्ञान में 98% प्रकाशन अंग्रेजी में लिखे जाते हैं (लियू; एमोन), इस भाषा के केवल 7.3% मूल वक्ता हैं और विश्व स्तर पर 20% से अधिक वक्ता नहीं हैं (बहाजी एट अल.)। परिणामस्वरूप, अधिकांश मानव भाषा बाधाओं के कारण विज्ञान और प्रौद्योगिकी में हुई प्रगति से बाधित होते हैं। इस बाधा को दूर 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 अद्वितीय चुनौतियों का समाधान करने में वाफ़ल रहता है। वैज्ञानिक और तकनीकी दस्तावेज़ों में, लेआउट और गैर-पाठ्य तत्वों को अनदेखा नहीं किया जा सकता है — पैराग्राफ़, गणितीय समीकरणों, तालिकाओं और आकृतियों की व्यवस्था के समृद्ध और महत्वपूर्ण अर्थ होते हैं।

वे पाठ-आधारित प्रयास लेआउट में मौजूद जानकारी को अनदेखा करते हैं और इस प्रकार वैज्ञानिक तथा तकनीकी दस्तावेज़ों 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 दस्तावेज़ों का अनुवाद करने वाला दुनिया का पहला ओपन-सोर्स टूल है। लेआउट डिटैक्शन और बड़े भाषा मॉडल (चित्र 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 डिटैक्शन, 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-वाणज्य मॉडल जो स्थायी विकास को बढ़ावा देता है।

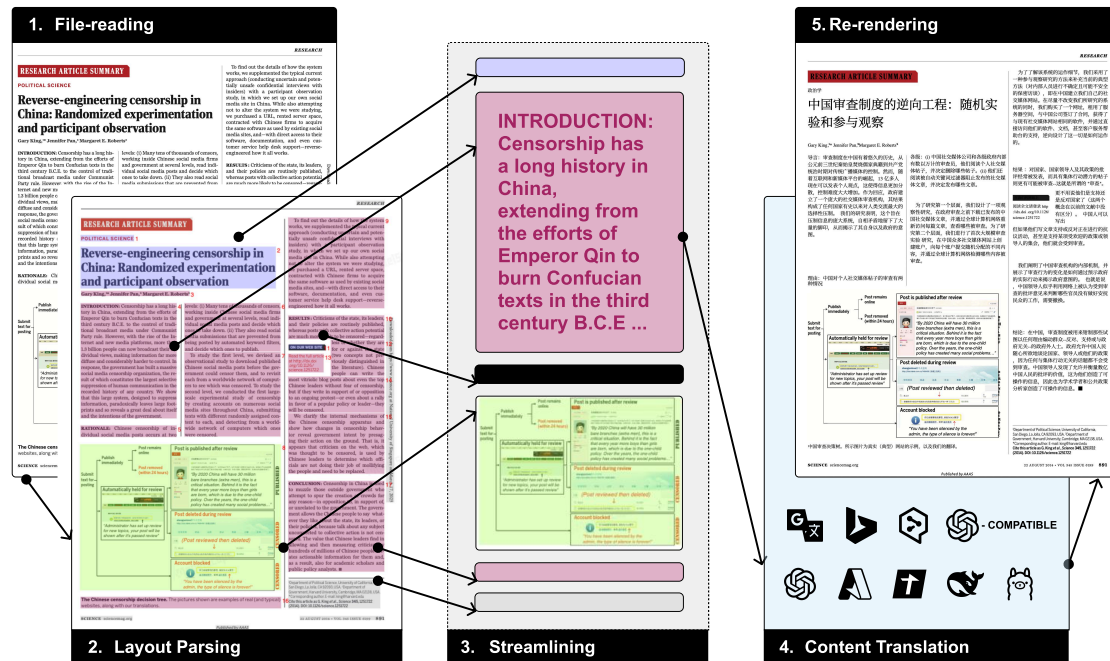
2 वास्तुकला और डिज़ाइन

कुल मिलाकर, PDFMathTranslate को दस्तावेज़ों के मूल लेआउट को बनाए रखते हुए उनका अनुवाद करने के लिए डिज़ाइन किया गया है। सबसे पहले, यह एक PDF दस्तावेज़ के साथ उपयोगकर्ता-नरिद्विष्ट पैरामीटर, जैसे कि भाषाएँ और पसंदीदा अनुवाद सेवा (परिशिष्ट A देखें) स्वीकार करता है। इसके बाद, यह दस्तावेज़ से लेआउट और पाठ्य सामग्री निकालने के लिए लेआउट का पता लगाता है। तिसरा, चयनित अनुवाद सेवा, जैसे कि GPT-4, DeepL, Google, या Ollama (परिशिष्ट A देखें) का उपयोग करके पाठों का अनुवाद किया जाता है। अंत में, अनुवादित पाठ और पहले से पता लगाए गए लेआउट को संरक्षित लेआउट वाले एक अनुवादित दस्तावेज़ के रूप में पुनः प्रस्तुत किया जाता है।

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

सटीक लेआउट पारस्पर

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



चित्र 1: पीडीएमैथट्रांसलेट की संरचना

जसिमे लेआउट पहचान, वभाजन, प्रसंस्करण और पुनरावर्तन शामिल है। शुरुआत में, हम लेआउट पहचान में हुई एक हालिया प्रगति, *DocLayout-YOLO-DocStructBench-onnx* का लाभ उठाते हैं। यह मॉडल ऑब्जेक्ट डिटिक्शन में एक SOTA समाधान, *Yolov10* (Wang et al.) का एक और संस्करण है, और वशिष्ट 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 विविध हार्डवेयर के लिए हमारी पारसिंग पाइपलाइन की अनुकूलता सुनिश्चित करने के लिए डिफॉल्ट के रूप में चुना गया था।

लचीला अनुवाद मडिलिवेयर

अनुवाद मडिलिवेयर *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 समर्थन 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 किसी भी भाषा को इनपुट या आउटपुट भाषा के रूप में स्वतंत्र रूप से चुन सकते हैं।

दूसरा, *PDFMathTranslate* अनुवाद के लिए कम से कम 23 लोकप्रिय सेवाओं का समर्थन करता है। सेवाओं की सूची में *Google Translate* और *OpenAI* जैसी प्रमुख ऑनलाइन सेवाएँ, साथ ही स्थानीय परिनियोजन भी शामिल है।

मॉडलों का, चाहे वह अनुवाद के लिए हो या संवादों के लिए (परशिष्ट A देखें)। स्पष्ट रूप से सूचीबद्ध सेवाओं के अतिरिक्त, उपयोगकर्ता *OpenAI* प्रोटोकॉल¹ का पालन करने वाली किसी भी नई सेवा को एकीकृत कर सकते हैं।

इसके अतिरिक्त, नई सेवाओं को जोड़ने को सरल बनाने के लिए, हमने एक ऐसी संरचना तैयार की है जो लेआउट प्रवाह को अनुवाद प्रवाह से अलग करती है। इस संरचना में, भाषा सेवाओं को केवल पाठ्य इनपुट प्राप्त होते हैं, जिससे हमारे प्रोजेक्ट की सामान्य प्रयोज्यता और स्थिरता बढ़ती है। परिणामस्वरूप, किसी भी भाषा सेवा को लागू करने के लिए 15 पंक्तियों से कम के एक संक्षिप्त फंक्शन की ही आवश्यकता होती है।

तीन, *PDFMathTranslate* अनुवाद में अतिरिक्त अनुकूलन क्षमता प्रदान करता है। हमने एक अनुकूलन सुविधा डिज़ाइन की है जो ऐसी रणनीतियों को प्रारूपित करती है जिससे अनुवाद की गुणवत्ता या डोमेन-वशिष्ट कार्यों में उसकी अनुकूलनशीलता संभावित रूप से बढ़ सकती है। यह सुविधा उपयोगकर्ताओं को उनमें प्रारूपित रणनीतियों का उपयोग करने की अनुमति देती है जैसे कि फ़्यू-शॉट्स (ब्राउन एट अल., बी), चैन-ऑफ-थॉट (वेई एट अल.), रोल-प्लेइंग (शानहन एट अल.), आदि।

कुशल स्ट्रीमिंग प्रवाह

वैज्ञानिक और तकनीकी दस्तावेजों का अनुवाद करने में, कुछ मामलों में, हजारों पृष्ठों के लिए एक कुशल कार्यप्रवाह की आवश्यकता होती है, जो प्रदर्शन को चुनौती देता है।

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

इस समस्या के समाधान के लिए, हमने एक स्ट्रीमिंग डिजिटल अपनाया जो एक कुशल इन-मेमोरी प्रोसेसिंग पाइपलाइन पर केंद्रित है। यह डिसिक I/O को न्यूनतम करता है जबकि दस्तावेज़ की मूलता को बनाए रखता है। विशेष रूप से, `translate_stream` फंक्शन एक PDF फाइल की बाइट स्ट्रीम को स्वीकार करता है और इसे एक परिवर्तनीय दस्तावेज़ प्रतिनिधित्व में बदल देता है। यह दृष्टिकोण सिस्टम को गतिशील रूप से संशोधन लागू करने की अनुमति देता है, जैसे आवश्यकता पड़ने पर PDF/A प्रारूप में परिवर्तित करना और लक्ष्य भाषा के आधार पर उपयुक्त फॉन्ट एम्बेड करना। अस्थायी बफर का उपयोग करके और `download_remote_fonts` जैसे फंक्शंस का लाभ उठाकर, यह डिज़ाइन संसाधनों का प्रभावी ढंग से प्रबंधन करता है और यह सुनिश्चित करता है कि प्रसंस्करण के दौरान मूल फाइल अपरिवर्तित रहे।



Translated Text: चित्र 2: आधिकारिक तौर पर समर्थित इंटरफेस

हमारा वर्कफ्लो विभिन्न अनुवाद बैक-एंड का समर्थन करता है, जिसमें अनुवादक हमारे वर्कफ्लो के भीतर प्लगइन के रूप में कॉन्फिगर किए जाते हैं। विभिन्न अनुवादक एक आधार वर्ग से इनहेरिट करके और `do_translate` विधिको लागू करके विशिष्ट अनुवाद कार्यक्षमताओं को एकीकृत करते हैं। इन अनुवादकों को विभिन्न उपयोगकर्ता इंटरफेस में सहजता से एकीकृत किया जा सकता है। समुदाय के योगदान से, हमारा वर्कफ्लो वर्तमान में 20 से अधिक अनुवादकों का समर्थन करता है। इसके अतिरिक्त, यह OpenAI-संगत API और पारंपरिक मशीन अनुवाद से वा के साथ बैकएंड LLM सेवाओं के लिए सुदृढ़ वसतिारशीलता प्रदान करता है।

प्रदर्शन में और अधिक अनुकूलन अतुल्यकालिक नष्टि पादन, कैशिंग रणनीतियों और त्रुटि-हैंडलिंग तंत्रों के संयोजन के माध्यम से प्राप्त किया जाता है। `asyncio.Event` जैसे अतुल्यकालिक निर्माणों को कार्य रद्द करने और समवर्तीता का समर्थन करने के लिए नियोजित किया जाता है, खासकर बड़े दस्तावेज़ों को संसाधित करते समय या एक साथ कई अनुवाद कार्यों को संभालते समय। इसके अतिरिक्त, डिज़ाइन में एक कैशिंग तंत्र शामिल है—जो `ignore_cache` फ्लैग के माध्यम से नियंत्रित होता है—ताकि अनावश्यक गणनाओं को रोका जा सके, LLM टोकन बचाए जा सकें और अनुवाद प्रक्रिया को गति दी जा सके। जब कौन सा सबसेट फॉन्ट एम्बेडिंग रणनीति संगतता से समझता है कि बना अंतिम फाइल आकार को कम करती है। कुल मिलाकर, यह स्ट्रीमिंग डिज़ाइन स्वचालित PDF अनुवाद वर्कफ्लो के लिए तैयार एक मजबूत, स्केलेबल और संसाधन-कुशल ढांचे को दर्शाता है।

3 उपयोग और परिनियोजन

PDFMathTranslate को विभिन्न इंटरफेस में लागू किया गया है, जिनमें एक कमांड-लाइन टूल (CLI), एक ग्राफिकल यूजर इंटरफेस (GUI), क्रॉस-प्लेटफॉर्म एप्लिकेशन (मैकओएस और वंडोज पर), और डॉकर इमेज शामिल हैं। इन आधारिक अतिरिक्त



चित्र 3: CLI के लिए प्रमुख कमांड

विशेष रूप से समर्थित इंटरफेस, उपयोगकर्ता गुआगुआस टैडअप हमारे समुदाय को एक जोड़ते प्लगइन का योगदान देता है। आम लोगों के लिए, हम मैक², वंडोज, वेब, ऑनलाइन डेमो और जोड़ते प्लगइन³ पर GUI प्रदान करते हैं। विशेष रूप से, मैक और वंडोज पर डेस्कटॉप संस्करण इंस्टॉल किए जा सकते हैं, डेमो सार्वजनिक रूप से ऑनलाइन उपलब्ध है, और `pdf2zh -i` का उपयोग करके GUI शुरू किया जा सकता है। ये ग्राफिक इंटरफेस चित्र 2 में दिखाए गए हैं।

डेवलपर्स के लिए, हम उन्नत उपयोग और सबस्ट्रीम विकास 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 व्यापक सुविधाओं का समर्थन करता है, जिनमें से प्रमुख चित्र 3 में दिखाए गए हैं और⁴ में वसतिार से प्रलेखित हैं। सर्वर पर एप्लिकेशन डिप्लॉय करने का लक्ष्य रखने वाले उपयोगकर्ताओं के लिए, हम एक डॉकर इमेज प्रदान करते हैं जिसे विशेष रूप से ऑनलाइन उत्पादन डिप्लॉयमेंट को सरल बनाने के लिए डिज़ाइन किया गया है। ये इमेज संभावित नेटवर्क समस्याओं वाले डेवलपर्स के लिए विभिन्न प्लेटफॉर्म पर बनाई और वितरित की जाती हैं।

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

हमारी तुलना दर्शाती है कि पीडीएफमैट्रान्स-लेट वैकल्पिक समाधानों से बेहतर है। सबसे पहले, *accessibility* के संबंध में, हम एक आपन-सोर्स, स्वयं-तैनात करने योग्य और मुफ्त प्रदान करके उपयोगकर्ता की लागतों को काफी हद तक कम करते हैं

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

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

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

लचीले एपीआई एकीकरण के साथ समाधान। यह लचीलापन हमारे उपकरण की व्यापक प्रयोज्यता को सक्षम बनाता है और किसी भी पसंदीदा सेवा या मॉडल से सटीक अनुवाद आउटपुट सुनिश्चित करता है। दूसरा, *read-ability* के संदर्भ में, हमारा दृष्टिकोण बुनियादी लेआउट को संरक्षित 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 कमप्यूटेशनल 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 कि यह उपकरण लेआउट में अंतरनिहित जानकारी को बनाए रखते हुए पाठ का अनुवाद कैसे करता है। पहला उपयोग मामला एक तकनीकी दस्तावेज़ है जिसमें पाठ और जटिल सूत्र दोनों हैं, और दूसरा उपयोग मामला जटिल लेआउट और चित्रों 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 अंतरनिहित महत्वपूर्ण जानकारी को भी सुरक्षित रखता है।

5 मौजूदा प्रणालियों से तुलना

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

6 टिकाऊपन

फिर भी, ओपन-सोर्स परियोजनाओं में रखरखाव चुनौतीपूर्ण है (स्टोल और अली बाबर)। सतत विकास सुनिश्चित 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 परियोजना के होमपेज पर योगदानकर्ताओं को प्रमुखता से प्रदर्शित करके और परियोजना के दस्तावेजों में हाल के योगदानों को नयिमिति 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 वैश्विक डेवलपर्स को सफलतापूर्वक प्रोत्साहित किया है, जिनोंने सामूहिक रूप से दसवीं हजार लाइनें कोड की योगदान की।

दूसरे, एक व्यावसायिक साझेदार के साथ सहयोग के माध्यम से, हम सक्रिय योगदानकर्ताओं के लिए सदस्यता वनिमिय जैसे प्रायोजित पुरस्कार प्रदान करते हैं। डेवलपर्स को भेजी गई इन 11 सदस्यताओं ने लगातार योगदानों को प्रभावी ढंग से प्रोत्साहित किया है, जिसके परिणामस्वरूप 485 से अधिक उपयोगकर्ता-रिपोर्टेड मुद्दों 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 स्थायी सहयोग मॉडल ने परियोजना को एक सप्ताह से अधिक समय तक गिटहब के वैश्विक रुझानों में शीर्ष पर पहुँचाया, जिसने 25 हजार से अधिक स्टार, 22 हजार डाउनलोड, और 49 हजार से अधिक डॉकर पुल (जून 2025 तक) प्राप्त किए। इसके अतिरिक्त, हमारी परियोजना की सफलता ने समान कार्यक्षमताओं 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, फिर भी हमारे टूल में कुछ विशिष्ट सीमाएँ हैं। पहला, अनुवाद की गुणवत्ता अंतरनिहित अनुवाद मॉडल और प्रॉम्प्ट पर अत्यधिक निर्भर करती है। दूसरा, विभिन्न दस्तावेज़ लेआउट की पहचान करने में लेआउट डिटेक्शन मॉडल की सटीकता भी उपयोग किए गए लेआउट डिटेक्शन मॉडल की गुणवत्ता पर निर्भर करती है। तीसरा, टूल का वर्तमान संस्करण अभी तक स्कैन किए गए ऐसे PDF को नहीं संभाल सकता जिनमें इष्टतम ऑप्टिकल कैरेक्टर रिकॉग्निशन (OCR) का अभाव है; हालाँकि, इस अंतिम सुविधा को लागू करने की योजना है।

8 आभार

लेखक गिटहब पर ओपन-सोर्स प्रोजेक्ट के सभी 44 योगदानकर्ताओं⁵; हमें डेमो होस्ट करने की अनुमति देने वाले प्लेटफॉर्मों (हगगि फेस, मॉडलस्कोप); डेमो में उपयोग की गई सेवाओं (गूगल ट्रांसलेट); और उपयोगकर्ताओं की सभी टिप्पणियों के लिए धन्यवाद देते हैं।

⁵They are 7shi, Byaidu, Copilot, Cycloctane, Hanaasagi, IuvenisSapiens, JEFF-dev-ui, Tql-ws1, Wybxc, YadominJinta, Zxis233, alohays, aseaday, awwaawwa, borcation, charles7668, chiu0602, czz404, damaoooo, dependabot[bot], domonnss, eltociar, hellofinch, highkay, hotwa, imClumsyPanda, kharkover, kidach1, lintian233, mydreamworldpolly, namazuchin, qqueing, reyn, tastelikefeet, timelic, treeleaves30760, tylzh97, ws051682, wx-11, xxnuo, xyzyx233, 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	Paid
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	1.88

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 दृष्टिहीन नरियातक 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 किए गए दस्तावेजों को संभाल सकता है।⁵ आईएमटी: इमर्सिव ट्रांसलेट पीडीएफ प्रो।⁶ अनुवाद के लिए मूल लाटेक्स (LaTeX) फाइल आवश्यक है।⁷ एक सामुदायिक फॉरक में उपलब्ध: PDFMathTranslate/PDFMathTranslate-next।

तालिका 1: अन्य परियोजनाओं और उत्पादनों के साथ तुलना

Original	Translated
<p>12.3. Multivariate Gaussian and Weighted Least Squares</p> <p>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</p> <p>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 ($-\infty$ 到 ∞), 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 ____。</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 \hat{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>

Original	Translated
<p>RESEARCH SOCIAL MEDIA AND ELECTIONS</p> <p>from its policy relevance and simplicity, our choice of the reverse-chronological baseline was intended to maximize the differences in user feed experiences and thus the likelihood that algorithmically driven changes in content and emphasis could shape politically relevant attitudes and behaviors. Additionally, machine-learning feed algorithms are often thought of as a "black box" because the rules they follow are not transparent. They are also optimized for certain outcomes such as user engagement, and they "learn" from user behavior and adapt over time. By contrast, a chronological feed does not exhibit any of these features because the same rule is applied to all users at all times (26).</p> <p>We focused on three primary hypotheses that represent important public concerns that have been theorized in existing research, tied to three groups of outcomes potentially affected by the altered mix of feed content. The first outcome that we studied is polarization (H1). Scholars have long been interested in the polarizing effects of social media (22). Social media feed algorithms could affect political polarization in at least two ways. First, if algorithms use past user behavior to inform output, then perspectives with which a user has engaged in the past may be prioritized in the future, potentially encouraging selective exposure to like-minded political views (2, 28–34). Repeated exposure to like-minded and reinforcing content may foster greater extremity on issue positions (37). Thus, we expect the Chronological Feed to lessen issue polarization (H1a). Second, algorithms may deprecate content from certain parts of a user's network of connections. More specifically, prior work has argued that Facebook's features may encourage partisan stereotyping (3, 35) and influence negative attitudes about outgroups (26). As a result, in the US partisan political con-</p> <p>A1</p> <p>A2</p> <p>B1</p> <p>B2</p> <p>C1</p> <p>C2</p>	<p>RESEARCH SOCIAL MEDIA AND ELECTIONS</p> <p>从其政策相关性和简单性来看, 我们选择反向编年基线是为了最大限度地扩大用户在信息源体验方面的差异, 从而使算法驱动的内容和重点变化可能塑造与政治相关的态度和行为。此外, 机器学习新闻算法通常被视为“黑箱”, 因为它们遵循的规则并不透明。它们还对特定结果 (如用户参与度) 进行优化, 并根据用户行为 (如“学习”) 随着时间的推移不断调整。相比之下, 按时间顺序排列的信息则不存在这些特点。没有相似的特征, 相同的规则都适用于所有用户 (26)。</p> <p>我们重点研究了三个主要假设, 它们代表了现有研究中的理论上的重要公共关切, 并可能与内容组合变化影响的三个结果联系在一起。我们研究的第一组结果是两极分化 (H1)。长期以来, 学者们一直对社交媒体媒体的极化效应感兴趣 (21)。社交媒体推送算法可能会以两种形式影响政治极化。首先, 如果算法使用用户过去的行为作为输出信息的依据, 那么用户过去参与过的观点可能会在未来被优先展示, 从而有可能帮助用户选择性地接触志同道合的政治观点 (1, 19, 22–24)。反复接触志同道合且具有强化作用的内容, 可能会导致用户在问题立场上更加极端 (17)。因此, 我们预计“按时间顺序推送”会减少问题的两极分化 (H1a)。其次, 算法可能会降低用户连接网络中某些内容的可见性。更具体地说, 先前研究认为, Facebook 的功能可能会鼓励刻板印象 (3, 25), 并影响对外部群体的负面态度 (26)。因此, 在美国的党派政治内容中, 我们也期望“按时间顺序推送”会减轻个人层面的情感化 (H1b) (27)。</p> <p>A1</p> <p>A2</p> <p>B1</p> <p>B2</p> <p>C1</p> <p>C2</p>

चित्र 4: उपयोग के मामले।

9 नैतिकता 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 किए जाने वाले उपकरण के रूप में, यह दस्तावेजों के कॉपीराइट के संबंध में संभावित नैतिक चर्चाएं पैदा करता है। उचित अनुमति के बिना दस्तावेजों का अनुवाद करना वैज्ञानिकों, रणनीतियों और नवाचारों के बौद्धिक संपदा अधिकारों को चुनौती दे सकता है। हम इस चर्चा को कम करने के लिए बौद्धिक संपदा विशेषज्ञों के सुझावों का स्वागत करते हैं।

A परिशिष्ट

A.1 परिशिष्ट 1. PDFMathTranslate में समर्थित भाषाएँ और सेवाएँ

संदर्भ

Ulrich Ammon. भाषाई असमानता और वैज्ञानिक विमर्श में भागीदारी तथा वैश्विक ज्ञान संचय पर इसके प्रभाव – द्वा-तीय-स्तरीय भाषा समुदायों की समस्याओं पर एक गहन दृष्टि के साथ। 3(2):333–355.

उलरिच एमोन. 2012. भाषाई असमानता और वैज्ञानिक विमर्श में भागीदारी तथा वैश्विक ज्ञान संचय पर इसके प्रभाव – द्वितीय-स्तरीय भाषा समुदायों की समस्याओं पर एक गहन दृष्टि के साथ. *Applied Linguistics Review*, 3(2):333–355.

अनीस बाजी, लौरा असयिन, ऐनी-मैरी लेस्लेट, और बायरन एडिनॉफ। वैज्ञानिक साहित्य से गैर-अंग्रेजी भाषी दुनिया का बहिष्करण: व्यसन संबंधी पत्रिकाओं और प्रकाशकों के लिए परिवर्तन हेतु सफ़ाई। 40(1):6–13.

पीटर एफ. ब्राउन, जॉन कॉक, स्टीफन ए. डेला पेट्रि, वसिंट जे. डेला पेट्रि, फ्रेडरिक जेलनिक, जॉन डी. लफ़र्टी, रॉबर्ट एल. मरसर, और पॉल एस. रूसिनी. a. मशीन अनुवाद के लिए एक सांख्यिकीय दृष्टिकोण. 16(2):79–85.

टॉम ब्राउन, बेजामिन मान, नकि राइडर, मेलानी सुबबिया, जेरेड डी कपलान, प्रफुल्ल धारीवाल, अरवि नीलकांतन, प्रणव श्याम, गरिश शास्त्री, अमांडा आस्कल, संधि अग्रवाल, एरियल हर्बर्ट-वॉस, ग्रेटचन करुगर, टॉम हेनगिन, रेवॉन चाइलड, आदित्य रमेश, डेनियल ज़गिलर, जेफरी वू, क्लेमेंस वट्टर, और 12 अन्य। b. भाषा मॉडल कुछ-शॉट सीखने वाले हैं। *Advances in Neural Information Processing Systems* में, खंड 33, पृष्ठ 1877–1901। कर्न एसोसिएट्स, इंक.

कुमजू हवांग. 2005. ज्ञान उत्पादन में नमिनतर विज्ञान और अंग्रेजी का प्रमुख उपयोग: कोरियाई विज्ञान और प्रौद्योगिकी का एक केस अध्ययन. *Science Communication*, 26(4): 390–427.

मेल्विन जॉनसन, माइक शूस्टर, क्वोक वी. ले, मैक्सिम क्रकिन, योगहुई वू, झफिंग चैन, नखिल थोरात,

फर्नेंडा वरिगास, मार्टिनि वाटेनबर्ग, और ग्रेग कोराडो। गूगल का बहुभाषी न्यूट्रल मशीन अनुवाद प्रणाली: शून्य-शॉट अनुवाद को संभव करना। 5:339–351.

वीशु लयि। 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*। शिकागो विश्वविद्यालय प्रेस।

ओएनएनएक्स. ओएनएनएक्स/ओएनएनएक्स.

वेलरिया रामरिज़-कास्टेनेडा। 2020। विज्ञान में अंग्रेजी भाषा के प्रभुत्व के कारण वैज्ञानिक पत्रों की तैयारी और प्रकाशन में कमियाँ: जैविक विज्ञान 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 बेजामिन पब्लिशिंग कंपनी.

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

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

क्लास-जान स्टोल और मुहम्मद अली बाबर। उत्पाद विकास में ओपन सोर्स सॉफ्टवेयर के उपयोग में चुनौतियाँ: साहित्य की समीक्षा। 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 विज्ञान.

आशीष वासवाणी, नोआम शजीर, नकि परमार, जेकब उश्कोराइट, एलियन जोन्स, एडन एन गोमेज़, लुकाज़ काइज़र, और इलाया पोलोसुकनि. 2017. ध्यान ही सब कुछ है. *Advances in Neural Information Processing Systems*, 30.

रेनॉल्ड वॉन गज़िकी। अंतरराष्ट्रीय वैज्ञानिक समुदाय में केंद्र और परिधि: 19वीं शताब्दी में जर्मनी, फ्रांस और ग्रेट ब्रिटेन। 11(4):474–494.

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

जेसन वेई, जुएज़ी वांग, डेल शूरमैनस, मार्टिन बॉसमा, ब्रायन डच्टर, फेई ज़िया, एड ची, क्वोक वी. ले, और डेनी झोउ। विचार-शृंखला प्रोम्प्टिंग बड़े भाषा मॉडल में तर्क को प्रेरित 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].

Wybxc. 2025. feat: onnx सपोर्ट Wybxc द्वारा · पुल रकि वेस्ट #116 · Byaidu/PDFMathTranslate. [ऑनलाइन; 18 मार्च 2025 को एक्सेस किया गया]। जयिआन झाओ, हैंग रुई कांग, बनि वांग, और कॉन्-गुई हे। डॉकलेआउट-योलओ: विविध सथिटकि डेटा और वैश्वकि-से-स्थानीय अनुकूली धारणा के माध्यम से दस्तावेज लेआउट विश्लेषण को बढ़ाना। *Preprint*, arXiv:2410.12628. जनिहुआ झू, यगिसे जयिआ, लजिन वू, डी हे, ताओ कनि, वेगांग झोउ, होउकआंग ली, और टए-यान लउ। न् यूरेल मशीन ट्रांसलेशन में BERT को शामिल करना। *Preprint*, arXiv:2002.06823.

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 XInference, 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 दिखाए गए से कहीं अधिक है। हमारा प्रोटोकॉल, जसि OpenAI जैसी सेवाओं के लिए एक सामान्य ढाँचे के रूप में डिज़ाइन किया गया है, उपयोगकर्ताओं को किसी भी अनुवाद प्रणाली को एकीकृत करने की अनुमति देता है, जसिमें स्थानीय रूप से होस्ट किए गए मॉडल या नज़ी ऑनलाइन सेवाएँ शामिल हैं, जसिसे व्यापक भाषा समर्थन सुनिश्चित होता है।

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