Übersetzung von Dokumenten

# Überblick

Document Translation ist ein neues Feature im [Azure Translator-Dienst](https://azure.microsoft.com/en-us/services/cognitive-services/translator/)， mit dem Unternehmen， Übersetzungsagenturen und Verbraucher， die Volumen komplexer Dokumente benötigen， in eine oder mehrere Sprachen übersetzt werden können， wobei Struktur und Format im Originaldokument beibehalten werden.Es übersetzt asynchron ganze Dokumente in einer Vielzahl von Dateiformaten， einschließlich Text， HTML， **Word， Excel， PowerPoint， Outlook Message， PDF** usw. in einer der **90 Sprachen und Dialekte， die**  vom Übersetzerdienst unterstützt werden.

Standard-Übersetzungsangebote auf dem Markt akzeptieren nur Nur-Text oder HTML und begrenzen die Anzahl der Zeichen in einer Anfrage. Benutzer， die große Dokumente übersetzen， müssen die Dokumente analysieren， um Text zu extrahieren， sie in kleinere Abschnitte aufzuteilen und separat zu übersetzen. Wenn Sätze an einem unnatürlichen Haltepunkt aufgeteilt werden， kann dies den Kontext verlieren， was zu suboptimalen Übersetzungen führt. Nach Erhalt der Übersetzungsergebnisse muss der Kunde die übersetzten Teile in das übersetzte Dokument zusammenführen.Dabei wird nachverfolgt， welches übersetzte Stück dem entsprechenden Abschnitt im Originaldokument entspricht.Das Problem wird kompliziert， wenn Kunden komplexe Dokumente mit umfangreichen Inhalten übersetzen möchten.



Document Translation erleichtert dem Kunden die Übersetzung：

1. Mengen großer Dokumente，
2. Dokumente in verschiedenen Dateiformaten，
3. Dokumente， die die Beibehaltung des ursprünglichen Layouts und Formats erfordern， und
4. Dokumente in mehreren Zielsprachen.

# Benutzererfahrung

Der Benutzer stellt eine Anforderung an die Dokumentübersetzung， in der er den Speicherort der Quell- und Zieldokumente sowie die Liste der Zielsprachen angibt.Es gibt einen Bezeichner zurück， der es dem Benutzer ermöglicht， den Status der Übersetzung zu verfolgen.Asynchron ruft die Dokumentübersetzung jedes Dokument vom Quellspeicherort ab， erkennt das Dokumentformat， wendet die richtige Parsing-Technik an， um Textinhalte im Dokument zu extrahieren， und übersetzt den Textinhalt in Die Zielsprachen.Anschließend wird das übersetzte Dokument rekonstruiert， wobei Layout und Format wie in den Quelldokumenten vorhanden beibehalten werden， und das übersetzte Dokument an einem bestimmten Speicherort gespeichert wird.Document Translation aktualisiert den Status der Übersetzung auf Dokumentebene.

Mit der Dokumentübersetzung können Benutzer die Übersetzung von Dokumenten anpassen， indem sie benutzerdefinierte Glossare， eine benutzerdefinierte Modell-ID， die mit dem [Kundenübersetzer erstellt wurde](https://portal.customtranslator.azure.ai/)， oder beides als Teil der Anforderung bereitstellen.Bei einer solchen Anpassung werden bestimmte Terminologien beibehalten und domänenspezifische Übersetzungen in den übersetzten Dokumenten bereitgestellt.

|  |  |
| --- | --- |
| 默克瑪律[[1]](#footnote-2) | Microsoft 在線翻譯 |
| Unterstützte Dateiformate | txt， tab， tsv， html， htm， mhtml， docx， pptx， xlsx， **msg， pdf， md， xlf， doc， xls， ppt und rtf** |
| Gesamtzahl der Dateien | ≤ 1000 |
| Dokumentgröße | **≤ 40 MB** |
| Gesamtgröße des Inhalts in einem Stapel | ≤ 250 MB |
| Anzahl der Zielsprachen in einem Batch | **10. ≤** |
| Anzahl der unterstützten Sprachen | 90 |
| Übersetzungsmodelle für Kunden | 阿布格斯特 |
| Benutzerdefiniertes Glossar | Unterstützte Dateiformate – tsv， xlf |
| Größe der Glossardatei | ≤ 10 MB |
| Anzahl der Glossardatei pro Dokument | **1 pro Sprachpaar** |

1. Die Grenzwerte in dieser Tabelle würden basierend auf Benutzerfeedback während der öffentlichen Vorschau überarbeitet. [↑](#footnote-ref-2)