Übersetzung von Dokumenten

# Überblick

Die Dokumentübersetzung 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 große Mengen komplexer Dokumente benötigen, in eine oder mehrere Sprachen übersetzt werden müssen, wobei die Struktur und das Format des Originaldokuments beibehalten werden. Ganze Dokumente werden asynchron in eine Vielzahl von Dateiformaten übersetzt, darunter **Text, HTML, Word, Excel, PowerPoint, Outlook, PDF und Markdown** in eine der **111 Sprachen und Dialekte, die**  vom Translator-Dienst unterstützt werden.

Die 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 sie separat zu übersetzen. Wenn Sätze an einem unnatürlichen Haltepunkt geteilt werden, kann der Kontext verloren gehen, was zu suboptimalen Übersetzungen führt. Nach Erhalt der Übersetzungsergebnisse muss der Kunde die übersetzten Teile in das übersetzte Dokument einfügen. Dabei geht es darum, den Überblick darüber zu behalten, welches übersetzte Stück dem entsprechenden Abschnitt im Originaldokument entspricht. Kompliziert wird das Problem, wenn Kunden komplexe Dokumente mit reichhaltigem Inhalt übersetzen möchten.



Document Translation macht es dem Kunden leicht, Folgendes zu übersetzen:

1. große Mengen an Dokumenten,
2. Dokumente in einer Vielzahl von Dateiformaten,
3. Dokumente, bei denen das ursprüngliche Layout und Format beibehalten werden müssen, und
4. Dokumente in mehrere Zielsprachen.

# Benutzererfahrung

Der Benutzer stellt eine Anforderung an den Dokumentübersetzungsdienst, in der er den Speicherort der Quell- und Zieldokumente sowie die Liste der Zielsprachen angibt. Der Dienst gibt eine Kennung zurück, mit der der Benutzer den Status der Übersetzung nachverfolgen kann. 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 unter Beibehaltung des Layouts und des Formats in den Quelldokumenten rekonstruiert und das übersetzte Dokument an einem angegebenen Speicherort gespeichert. Bei der Dokumentübersetzung wird der Status der Übersetzung entweder auf Auftrags- oder Dokumentebene aktualisiert.

Benutzer können eine benutzerdefinierte Modell-ID bereitstellen, die mit dem benutzerdefinierten Übersetzerportal, benutzerdefinierten Glossaren oder beidem als Teil der Anforderung zum Übersetzen von Dokumenten erstellt wurde. Bei der Dokumentübersetzung wird eine solche Anpassung angewendet, wobei bestimmte Terminologien beibehalten und domänenspezifische Übersetzungen in den übersetzten Dokumenten bereitgestellt werden.