



# Language Understanding 시작하기

3분

Language Understanding을 사용하여 언어 이해 애플리케이션을 만드는 것은 두 가지 주요 작업으로 구성됩니다. 먼저 언어 모델을 학습시키는 데 사용할 엔터티, 의도 및 발화를 정의해야 합니다. 이 단계를 모델을 '제작'한다고 합니다. 그런 다음 클라이언트 애플리케이션이 사용자 입력을 기반으로 하여 의도 및 엔터티 '예측'에 사용할 수 있도록 모델을 게시해야 합니다.

## Language Understanding용 Azure 리소스

각각의 제작 및 예측 작업에 대해 Azure 구독의 리소스가 필요합니다. 다음과 같은 유형의 리소스를 사용할 수 있습니다.

- **Language Understanding** : 작성 또는 예측 리소스가 될 수 있는 Language Understanding 전용 리소스입니다.
- **Cognitive Services** : Language Understanding을 비롯해 다른 많은 인지 서비스를 포함하는 일반적인 인지 서비스 리소스입니다. 예측을 위해서는 이 종류의 리소스만 사용할 수 있습니다.

작성 리소스와 예측 리소스의 분리는 예측을 생성하기 위해 모델을 사용하는 클라이언트 애플리케이션과 별도로 언어 모델 학습의 리소스 사용량을 추적하려는 경우에 유용합니다.

Language Understanding 리소스를 만들도록 선택하면 작성, 예측 또는 둘 다를 선택하라는 메시지가 표시됩니다. '둘 다'를 선택하면 작성을 위한 리소스와 예측을 위한 리소스가 각각 하나씩 생성됩니다.\*\_

또는 작성을 위한 전용 Language Understanding 리소스를 사용하고, 예측을 위해서는 모델을 일반 Cognitive Services 리소스에 배포할 수도 있습니다. 클라이언트 애플리케이션이 Language Understanding 이외에 다른 인지 서비스를 사용하는 경우, 이 방법을 사용하면 단일 엔드포인트와 키를 통해 Language Understanding 예측 서비스를 포함하여 사용 중인 모든 인지 서비스에 대한 액세스를 관리할 수 있습니다.

## 작성

작성 리소스를 만들었으면 예측 모델을 학습시키는 데 사용할 수 있도록 애플리케이션이 예측할 엔터티 및 의도와 각 의도에 대한 발화를 정의하여 Language Understanding 애플리케이션을 작성하고 학습시키는 데 이 리소스를 사용할 수 있습니다.

Language Understanding은 일반적인 시나리오에 대해 미리 정의된 의도 및 엔터티를 포함하는 사전 빌드된 '도메인'의 포괄적인 컬렉션을 제공하며, 이 도메인은 모델의 시작점으로 사용할 수 있습니다.\* 사용자 고유의 엔터티 및 의도를 만들 수도 있습니다.

엔터티 및 의도를 만들 때는 어떤 순서로든 이를 수행할 수 있습니다. 의도를 만든 다음 의도에 대해 정의한 샘플 발화에서 단어를 선택하여 엔터티를 만들 수도 있고, 엔터티를 미리 만든 다음 의도를 만드는 과정에서 발화에 포함된 단어에 의도를 매핑할 수도 있습니다.

모델의 요소를 정의하기 위해 코드를 작성할 수 있지만, 대부분의 경우 Language Understanding 애플리케이션을 만들고 관리하기 위한 웹 기반 인터페이스인 Language Understanding 포털을 사용하여 모델을 작성하는 것이 가장 쉽습니다.

[팁] 모범 사례는 Language Understanding 포털을 작성에 사용하고 SDK를 런타임 예측에 사용하는 것입니다.

## 의도 만들기

사용자가 애플리케이션을 사용하여 어떤 작업을 수행할지에 따라 의도를 정의합니다. 각 의도에 대해 사용자가 의도를 어떻게 표현하면 좋은지에 대한 예제를 제공하는 다양한 발화를 포함해야 합니다.

의도를 여러 엔터티에 적용할 수 있는 경우, 각 잠재 엔터티에 대해 샘플 발화를 포함해야 하며, 각 엔터티가 발화에서 식별되어야 확인합니다.

## 엔터티 만들기

엔터티에는 네 가지 유형이 있습니다.

- **Machine-Learned** : 제공된 샘플 발화의 컨텍스트로부터 학습하는 동안 모델에 의해 학습된 엔터티입니다.
- **List** : 목록 및 하위 목록의 계층 구조로 정의된 엔터티입니다. 예를 들어 **디바이스** 목록에는 **불**과 **선풍기**의 하위 목록이 포함될 수 있습니다. 각 목록 항목에 대해 **불**의 경우 **조명**과 같은 동의어를 지정할 수 있습니다.
- **Regex** : 패턴을 설명하는 '정규식'으로 정의된 엔터티입니다. 예를 들어 \*\_555-123-4567 형식의 전화번호는 [0-9]{3}-[0-9]{3}-[0-9]{4}와 같은 패턴으로 정의할 수 있습니다.
- **\_Pattern.any\***: 샘플 발화에서 추출하기 어려울 수 있는 복잡한 엔터티를 정의하기 위해 '패턴'과 함께 사용되는 엔터티입니다.

## 모델 학습

모델의 의도와 엔터티를 정의하고 적절한 샘플 발화 세트를 포함했다면 이제 모델을 학습시킬 차례입니다. 학습이란 샘플 발화를 사용하여 사용자가 말할 수 있는 의도 및 엔터티의 자연어

표현과 일치하도록 모델을 가르치는 프로세스입니다.

모델을 학습시킨 후에는 텍스트를 제출하고 예측된 의도를 검토하여 모델을 테스트할 수 있습니다. 학습과 테스트는 반복적인 프로세스입니다. 모델을 학습시킨 후에는 샘플 발화로 모델을 테스트하여 의도와 엔터티가 제대로 인식되는지 확인합니다. 제대로 인식되지 않으면 업데이트하고 다시 학습시킨 후에 다시 테스트합니다.

## 예측

학습 및 테스트 결과에 만족하면 Language Understanding 애플리케이션이 사용될 수 있도록 예측 리소스에 게시할 수 있습니다.

클라이언트 애플리케이션은 예측 리소스의 엔드포인트에 연결하고 적절한 인증 키를 지정하고 사용자 입력을 제출하여 예측된 의도 및 엔터티를 얻음으로써 모델을 사용할 수 있습니다. 예측은 클라이언트 애플리케이션으로 반환되며, 클라이언트 애플리케이션은 예측된 의도에 따라 적절한 조치를 할 수 있습니다.

---

## 다음 단원: 연습 - Language Understanding 애플리케이션 만들기

[계속 >](#)