

Picovoice Rhino 2.2

PICOVOICE	Speech-to-Intent engine	
	Systeemvereisten:	Cross-platform
	Ontwikkeld door:	Picovoice
Commerciële licentie	Contactpersoon:	Bert.Vanhalst@smals.be

Functionaliteiten

[Picovoice](#) is een Canadees bedrijf dat een platform biedt voor het toevoegen van voice features aan software. Het platform bestaat uit een aantal onderdelen voor onder andere transcriptie, realtime spraakherkenning en intentherkenning op basis van spraak.

Het onderdeel dat we hier bespreken is [Rhino](#), een *speech-to-intent engine*, die gebruikers toelaat om te interageren met een toepassing via spraak. Het herkennen van de intentie van een gebruiker op basis van een ingesproken vraag verloopt normaal in 2 stappen: eerst wordt de ingesproken vraag omgezet naar tekst, daarna wordt intentherkenning uitgevoerd op die tekst. Rhino daarentegen voert de intentherkenning in één stap uit, rechtstreeks van audio naar intent, en claimt daarbij hogere nauwkeurigheid en betere prestaties dan de klassieke combinatie van spraakherkenning (ASR, Automatic Speech Recognition) en natuurlijk-taalherkenning (NLU, Natural Language Understanding).

Interessant aan de oplossing is de ondersteuning voor diverse deployment modellen. Naast cloud en onpremise kan het ook lokaal op een toestel draaien, wat zowel data privacy als prestaties ten goede komt.

Er zijn heel wat SDK's beschikbaar om de oplossing te integreren in een eigen toepassing, onder andere voor Nodejs en Python.

Conclusies & Aanbevelingen

Door de grote beperkingen op vlak van intentherkenning en herkennen van parameters (slots) lijkt Picovoice Rhino in de praktijk helaas slechts bruikbaar voor heel eenvoudige scenario's met specifieke vaste spraakcommando's. Nochtans heeft de tool een interessante aanpak voor intentherkenning op basis van spraak-input. De *speech-to-intent engine* werkt heel vlot en is beschikbaar voor verschillende talen waaronder het Nederlands, Frans, Engels en Duits. De tool kan overal gedeployed worden en werkt ook offline, wat interessant is in het kader van gegevensbescherming en privacy.

Testen & Resultaten

Als test gaan we aan de slag met de gratis versie van Rhino. Als scenario kiezen we een Dimona-aangifte die één intent vertegenwoordigt en 2 slots (parameters) kan bevatten: een datum (date) en een persoonsnaam (worker). Het is meteen duidelijk dat het aantal [slot types](#) die standaard door de tool herkend kunnen worden heel beperkt is. Een datum (date) en tijd (time) wordt bijvoorbeeld niet ondersteund, iets wat toch vrij basis is. Standaard worden enkel slots ondersteund voor percentages, letters, alfanumerieke combinaties, en getallen/rangtelwoorden van één of twee cijfers. We zijn dus verplicht om voor onze case zelf slots te definiëren: een *date* met mogelijke waarden “vandaag”, “morgen” en de dagen van de week; en een *worker* met mogelijke waarden “Jan”, “Piet” en “Paul”. In het midden van de screenshot hieronder zijn de voorbeeldzinnen weergegeven waarbij slots kunnen ingevoegd worden. We kunnen het model onmiddellijk uittesten. Rechtsonder worden de herkende intent en slots weergegeven als we de zin “een aangifte voor Jan voor vandaag” uitspreken.



The screenshot displays the Picovoice Rhino configuration interface. On the left, a sidebar shows the 'Intents' list with 'dimona' selected, and 'Slots' defined as 'date' and 'worker'. The main area shows a 'New expression' input field and a list of example sentences with slots: 'ik wil een aangifte indienen', 'een aangifte voor \$worker:worker voor \$date:date', 'een aangifte voor \$worker:worker', 'een aangifte voor \$date:date', 'een aangifte', and 'een nieuwe aangifte'. A 'Click to test' button is visible. On the right, the 'Inference' panel shows the detected intent 'dimona' and slots: 'worker: "Jan"' and 'date: "vandaag"'.

Als taal voor het project kozen we Nederlands. In totaal worden een [16-tal talen](#) ondersteund, waaronder Nederlands, Frans, Engels en Duits.

Bij de testen merken we dat de intent herkenning niet gesofisticeerd is: je moet de verschillende varianten van voorbeeldzinnen expliciet opgeven voor een goede herkenning. Je moet dus werkelijk alle optionele delen van een voorbeeldzin optioneel maken. Bijvoorbeeld: op basis van de voorbeeldzin "ik wil een aangifte indienen" wordt "een aangifte indienen" niet herkend. De tool biedt wel een syntax om optionele delen aan te geven en keuzes uit een lijst, maar de intent herkenning is dus niet flexibel. De tool lijkt dan ook eerder bruikbaar voor specifieke commando's. Voor een flexibele intentherkenning moet er teruggerepen worden naar een uitgebreider NLU platform.

De performantie is dan weer wel positief, de herkenning verloopt heel snel.

Gebruiksvoorwaarden & Budget

Met Picovoice kan je gratis aan de slag, met een beperking voor Rhino van 3 actieve gebruikers per maand en ondersteuning beperkt tot de GitHub community support. Het Developer plan biedt tot 100 actieve gebruikers per maand en 6 uur technische ondersteuning, en kost \$500 per maand. Er is tenslotte ook een Enterprise plan voor “large scale deployments”. Meer informatie is te vinden op de [pricing pagina](#).