



Artificial Intelligence to improve gender analysis

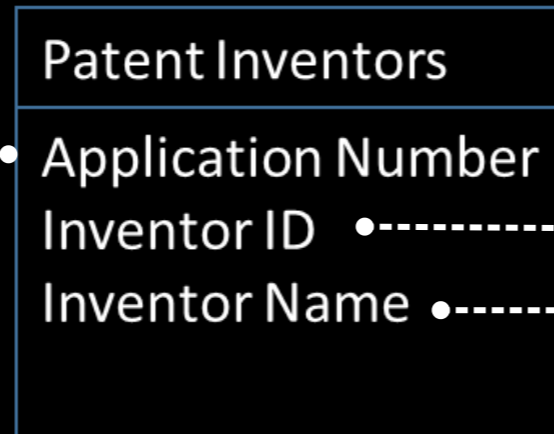
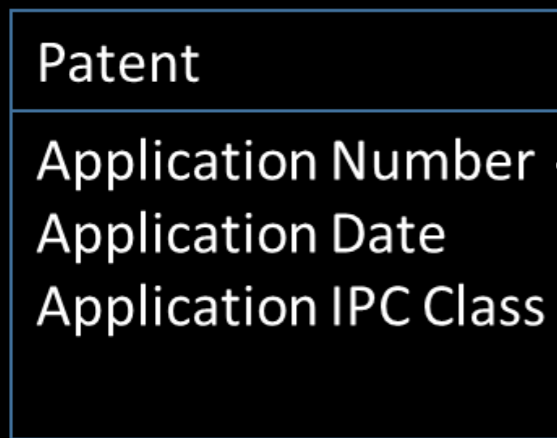
Instituto Nacional de Propiedad Industrial

INAPI

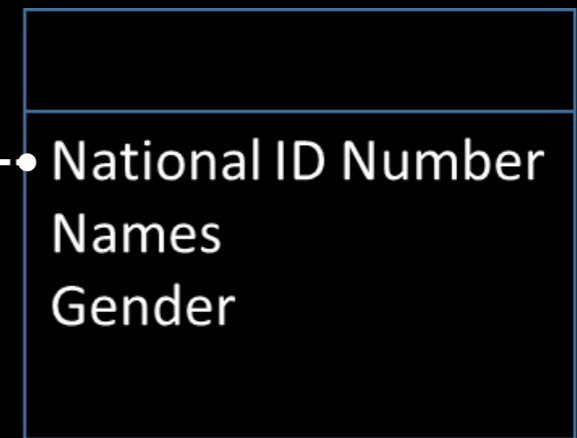
Alvaro González López
Head of Business Intelligence
March 2024

How do we obtain the gender of an inventor?

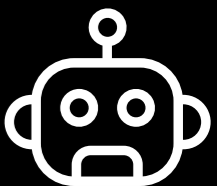
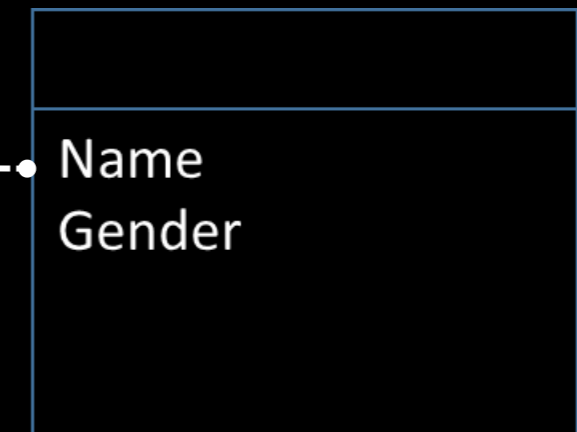
INAPI of Chile



Civil Registry of Chile



Artificial Intelligence







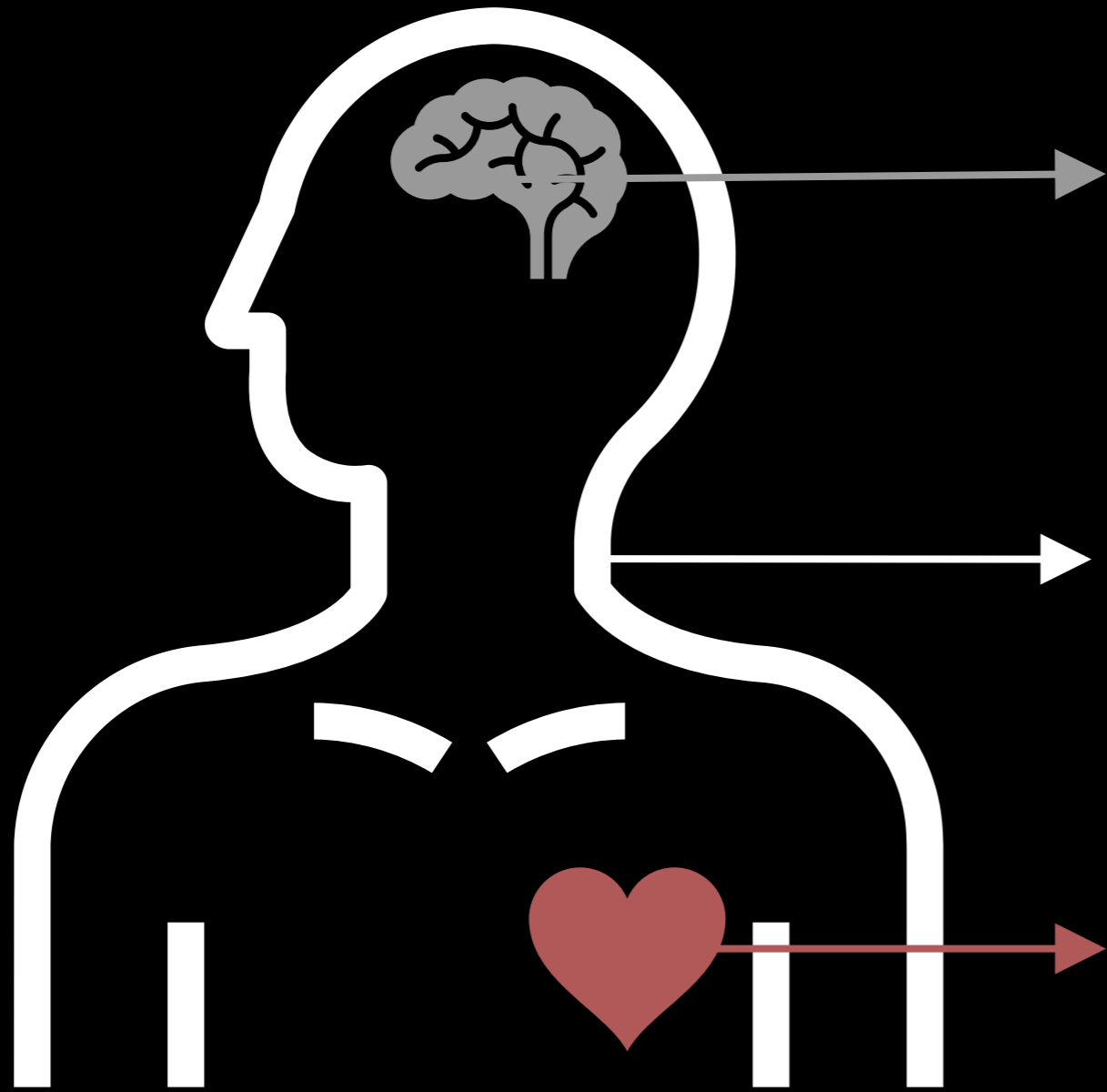
30%

70%

Typical problems to detect the gender of an inventor using the name:

- Standardized training data - BUT Non-standardized names and surnames in the INAPI's data base.
 - Names in upper or lower case.
 - Names at the beginning or surnames at the beginning.
- Ambiguous names.
 - Andreas, Alex, Ariel, etc.
- Company names instead of an inventor name.

	Generative IA provider	IA Models
Stage 1	 OPEN AI	GPT-3.5 model GPR-3.5 Turbo GPT-4
	 Anthropic	Claude Instan 1.2 Claude 2.1
	 Google	Gemini 1.0 pro
Stage 2	 Ollama	Gemma: 7 Ollama 2



Anthropic - Claude 2.1
Machine Learning + NLP
(Files)
Online database

CODE
GPT

Prompts and files with data to train the model

Results

<https://app.codegpt.co/chat/share/4f12b880-c426-41cb-a956-fa303bdcfb71/?pincode=low9zy>



Tú

JENIFFER STEPHANIA LEYTON ARAYA Andrés Cristian Macchiavello Zamora
BUSCHMANN ESPINOZA, Matías Germán María Francisca González Fernández
YERISEL CAROLAY FABIÁN MONDACA Roberto Cruz Fernández Camila Andrea
Santos Carmona



Clasificador de género

Aquí están los géneros para los nombres proporcionados:

Nombre	Género
JENIFFER STEPHANIA LEYTON ARAYA	Femenino
Andrés Cristian Macchiavello Zamora	Masculino
BUSCHMANN ESPINOZA, Matías Germán	Masculino
María Francisca González Fernández	Femenino
YERISEL CAROLAY FABIÁN MONDACA	Femenino
Roberto Cruz Fernández	Masculino
Camila Andrea Santos Carmona	Femenino

Reset Chat

Enter your message



From a sample of 1000 names, the initial findings of the pilot study were promising. The accuracy in gender prediction approached 98%.

However, inconsistencies were observed in responses for some names.

Notably, certain models demonstrated the ability to predict the owner's type with a 98% accuracy even without specific training data.

Moving forward, the next steps involve:

1. Validating the security protocols of the Ollama provider on a local server.
2. Verifying the accuracy of the model with names from diverse regions, including Arabic names, Japanese names, and others.

Gracias!

alvaro.gonzalez@inapi.cl