

# Inteligencia Artificial con GeneXus

Más sobre el uso de servicios cognitivos

*GeneXus 16*

En la introducción a este tema vimos la información básica de qué cosas podemos hacer con GeneXus en Inteligencia Artificial.

A continuación veremos los detalles de implementación de cómo invocar los métodos y un breve ejemplo desarrollado.

## Módulo de Inteligencia Artificial

Manage Module References X

Modules: Search by pattern here...

**Chatbot (2.1.10.129299)**  
GeneXus Chatbot module is a basic set of interfaces and implementations of data structures and algorithms needed to implement a Chatbot solution.

**GeneXusAI (1.1.21.129329)** Install  
GeneXusAI contains a common set of Artificial Intelligence tasks, including audio, text and image processing, all of

**GeneXus (2.1.7.129290)**  
GeneXus Core Module is a basic set of interfaces and

**GXtest (0.4.2)**  
GXtest Module provides core functionality for creating, running and reporting tests on geneXus and over ci/cd pipelines. <https://wiki.geneXus.com/commwiki/serve!wiki?>

**Module Information:**

**GeneXusAI**

**Module is not installed**

Available Versions: 1.1.21.129329

**Author:** GeneXus S.A.  
**Owner:** GeneXus S.A.

**Description:**  
GeneXusAI contains a common set of Artificial Intelligence tasks, including audio, text and image processing, all of them provided by several Cloud Platforms (e.g. IBM Watson, Microsoft Azure Cognitive Services, SAP Leonardo)

**Platforms:**

- C# Web

**Dependencies:**

- GeneXus 1.12.13.125610

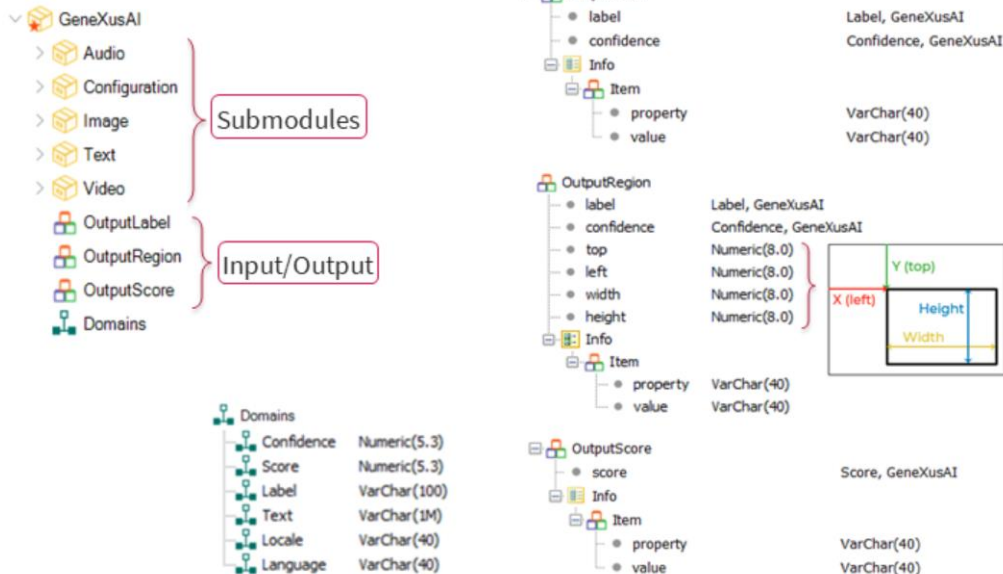
**Id:** 733e9734-9f0a-4620-b397-d84fbc2cef10

GeneXusAI

- Audio
  - SpeechToText
  - TextToSpeech
  - Domains
- Configuration
  - Provider
  - Domains
- Image
  - Classify
  - DetectFaces
  - DetectObjects
  - DetectScene
  - OCR
- Text
  - DetectLanguage
  - ExtractEntities
  - KeyPhrases
  - SentimentAnalysis
  - Translate
- Video
  - Analyze
  - OutputAnalysis
  - Process
  - Domains

La forma de utilizar los algoritmos de inteligencia artificial, es instalando el módulo GeneXusAI, que incluye 4 submódulos con funciones para trabajar con Audio, Imagen, Texto y Video y el módulo Configuration para definir el proveedor del servicio y las credenciales de acceso.

## Módulos, tipos de datos y dominios disponibles



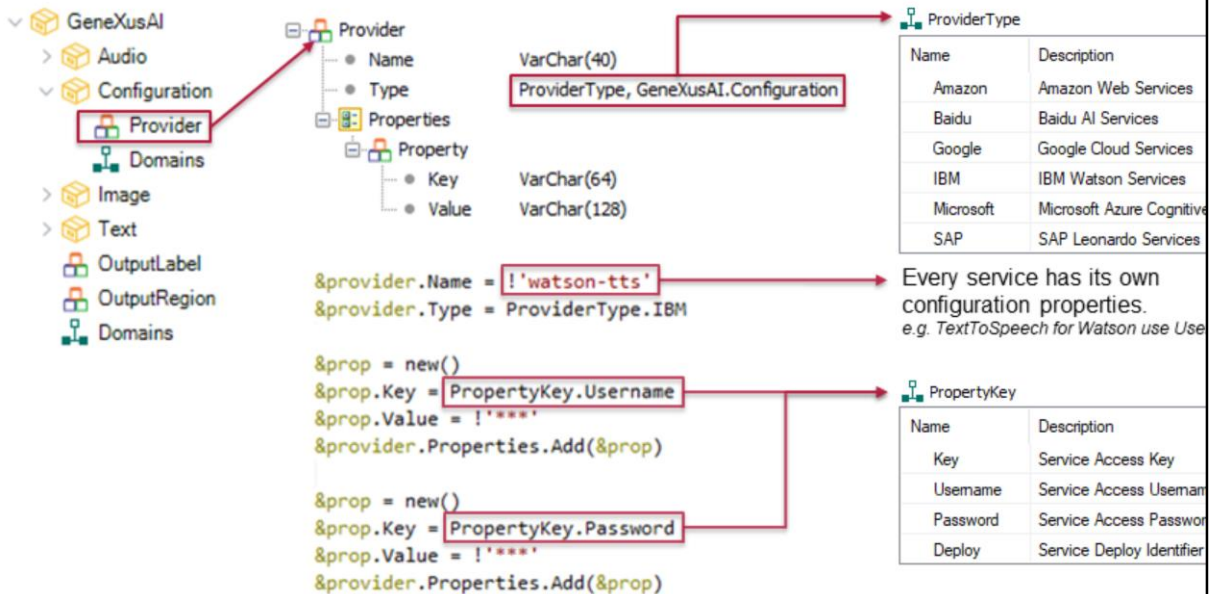
Además de los cinco sub-módulos que veremos en detalle a continuación, también disponemos de estructuras y dominios para manipular la entrada/salida de las tareas de AI, cada uno con una semántica definida.

- **Confidence** (confianza)
- **Score** (puntuación)
- **Label** (etiqueta)
- **Text** (texto)
- **Locale** (lenguaje de una región específica – por ejemplo, español de España)
- **Language** (lenguaje en el sentido genérico – por ejemplo, español, inglés)

Y por otro lado:

- **OutputLabel** nos brinda una etiqueta junto con su confianza
- **OutputRegion**, además de una etiqueta con su confianza, también nos indica una región rectangular.
  - Para visualizar cómo sería, veamos esta imagen (<X,Y> nos indica la esquina superior izquierda del rectángulo, y luego simplemente nos dice su altura y ancho).
  - Además nos brinda información adicional que nos podría devolver el proveedor.
- **OutputScore** representa un valor de score con datos adicionales

## Módulo Configuración



Comenzamos a estudiar el contenido del **módulo Configuración**.

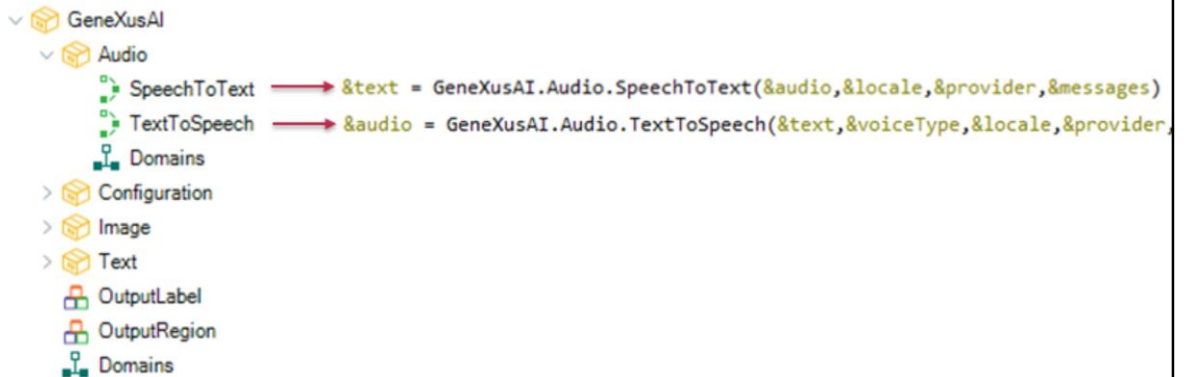
- Cada Procedure de GeneXusAI tomará como entrada la configuración de un proveedor específico.
- Para ello debemos indicar un nombre, un tipo de proveedor en el dominio ProviderType, y una serie de propiedades
  - \* El **ProviderType** contiene los posibles proveedores que podemos seleccionar.

Veamos un ejemplo:

1. Primero completamos el primer nivel del SDT indicando un nombre y un proveedor (en este caso, IBM)
  - Notar que hacemos referencia al servicio cognitivo en el nombre porque cada servicio tiene sus propias propiedades de configuración.
2. Luego completamos el segundo nivel del SDT; es decir, sus properties.
  - Primero indicamos un Username
  - Luego indicamos un Password

Esto es porque TextToSpeech para IBM necesita Username/Password para funcionar, pero otros servicios podrían requerir una API key. Para saber qué propiedades podemos configurar nos apoyamos en el dominio enumerado PropertyKey.

## Módulo Audio

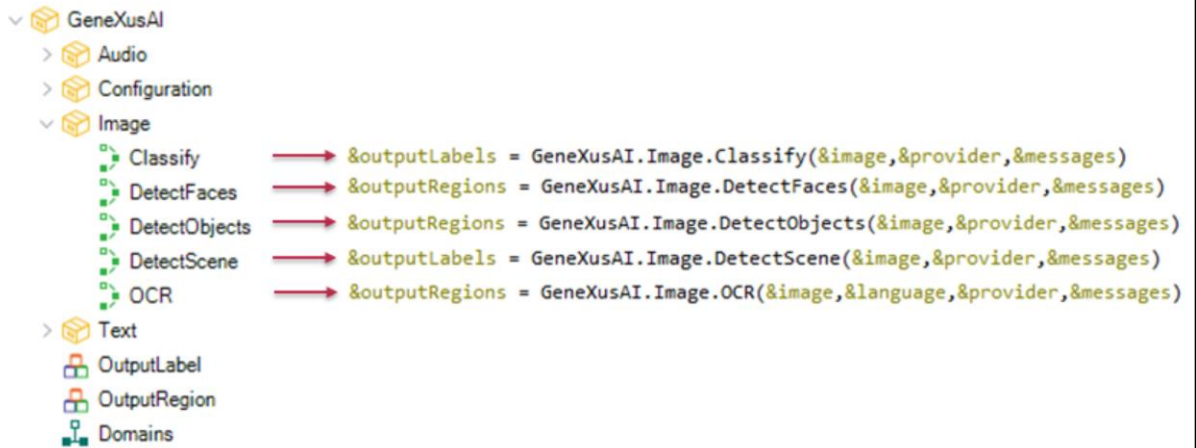


Ahora comencemos a ver el resto de los módulos.

En primer lugar, el **módulo Audio** nos permite manipular este tipo de contenido multimedia.

- **SpeechToText:** Nos permite transcribir audio a texto.
  - \* Notar cómo se recibe un audio, un lenguaje regional, la configuración para el proveedor y un messages. Con este último parámetro vamos a poder manejar errores que puedan ocurrir durante el prototipado o la ejecución – La especificación la pueden encontrar en la Wiki.
- **TextToSpeech:** Nos permite sintetizar texto en audio.
  - \* De forma análoga, recibe un texto, un tipo de voz, un lenguaje regional, etc. y devuelve un audio con el texto hablado.

## Módulo Image



En el **módulo Image** tenemos:

- **Classify**: Dada una imagen, nos devuelve una serie de categorías que describen la imagen (por eso su salida es OutputLabels – en plural).
- **DetectFaces**: Dada una imagen, reconoce caras y puede darnos información genérica sobre cada una (por ejemplo su sexo, edad estimada, etc.) Su salida es un OutputRegions (también plural) ya que nos indica la ubicación de cada rostro en un rectángulo.
- **DetectObjects**: Similar al caso anterior pero para objetos.
- **DetectScene**: Nos describe el/los escenarios de una imagen (por ejemplo si se trata de una ciudad, playa, desierto, etc.)
- **OCR**: Es una sigla para Optical Character Recognition, y nos permite extraer texto desde una imagen.

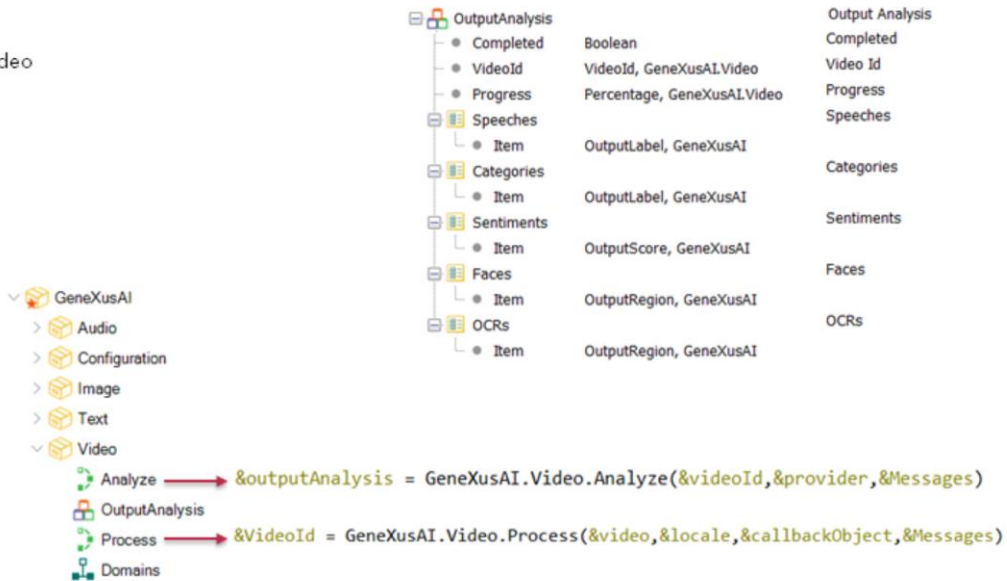
## Módulo Text



El **módulo Text** nos permite manejar aspectos del lenguaje natural, y está compuesto por:

- **DetectLanguage:** Permite detectar un idioma a partir de texto escrito. Su salida es un OutputLabel ya que no sólo nos dice el idioma, sino también un nivel de confianza de la detección.
- **ExtractEntities:** Extrae entidades a partir de texto, por ejemplo organizaciones, ciudades, nombres de personas, etc.
- **KeyPhrases:** Permite extraer las frases (o palabras) más relevantes del texto.
- **SentimentAnalysis:** Permite inferir el nivel de negatividad o positividad de un texto en 0 y 1, siendo 0.5 el sentimiento neutro.
- **Translate:** Nos permite traducir texto desde un lenguaje origen a uno destino.

## Módulo Video



El **módulo Video** nos permite procesar un video y analizarlo para detectar habla, categorías de objetos, sentimientos, rostros y hacer reconocimiento óptico de caracteres.. Está formado por:

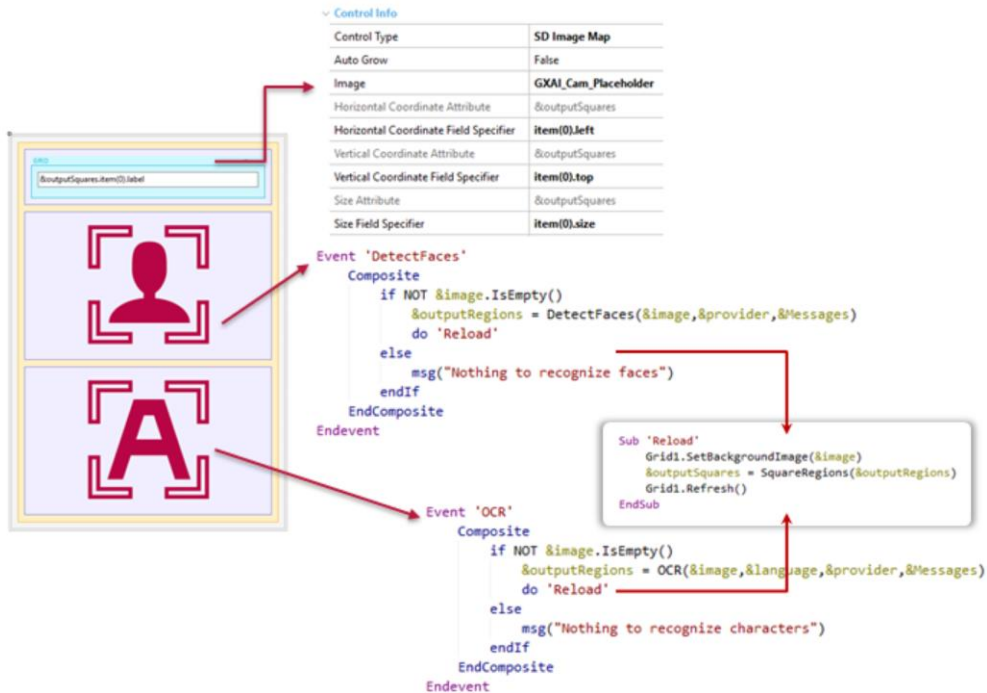
- **Analyze:** Analiza un video luego de haber sido procesado por el procedimiento Process
- **Process:** Procesa un video en forma sincrónica (en un proceso que verifica periódicamente el estado (polling) o asincrónica, definiendo un objeto de callback, es decir un objeto que es invocado cada vez que cambia algo en el proceso del video).



Detect faces and OCR example



Veamos un ejemplo sencillo para Smart Devices.



Aquí vemos cómo se implementó el ejemplo anterior.

1. Este es el layout asociado a esa pantalla. Tenemos un grid, y dos botones con las imágenes que vimos en la animación.
2. El Grid es de tipo SD Image Map para poder "dibujar los rectángulos" (en realidad, de momento, podemos dibujar cuadrados).
3. El primer botón tiene asociado un evento que verifica que la imagen no sea vacía e invoca al procedure DetectFaces de GeneXusAI.
4. Análogamente, el segundo botón invoca al procedure OCR.

La subrutina 'Reload' que vemos allí simplemente establece la imagen de fondo, obtiene los cuadrados a partir de los rectángulos del outputRegion, y finalmente realiza un "refresh" del grid.

# GeneXus™

**The power of doing.**

More videos

Documentation

Certifications

[training.genexus.com](http://training.genexus.com)

[wiki.genexus.com](http://wiki.genexus.com)

[training.genexus.com/certifications](http://training.genexus.com/certifications)