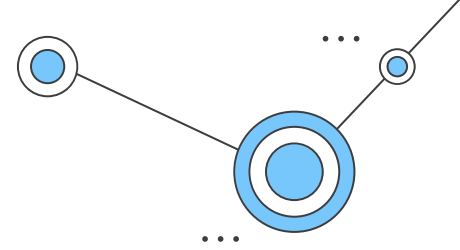


# Herramienta de clasificación y generación de textos de pacientes con enfermedades mentales mediante Deep Learning y NLP

MUCSI - 31/5/2023

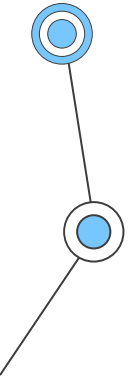
Unai Sainz, Lander Hernández e Iker Silva

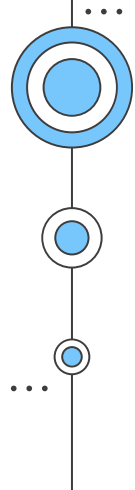
# Índice



Durante la presentación se van a tratar los siguientes temas:

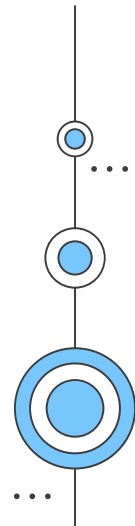
1. Introducción
2. **Arquitecturas** propuestas. **Esquema** general y componentes de aplicación
3. **Desarrollo** del proyecto
  - a. Adquisición y tratamiento inicial de **datos**
  - b. Modelos de **clasificación**
  - c. Modelos **generativos**
  - d. Integración y **despliegue** del proyecto
4. **Desviaciones** técnicas
5. Demo

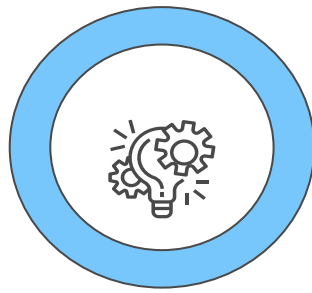




# 01

## Introducción





# Motivación

- Las enfermedades mentales afectan actualmente hasta un 19,5% de la población española.
- Se prevé que 1 de cada 4 personas tendrá un trastorno mental a lo largo de su vida.
- Todo el análisis médico se basa en entrevistas subjetivas.
- Las únicas herramientas "objetivas" que existen son unos formularios usados con poca frecuencia.

Se pretende dar respuesta a dos necesidades que actualmente no se encuentran resueltas:

- Herramientas de apoyo a la decisión del profesional basadas en textos escritos por el propio paciente.
- Herramientas sofisticadas de autoevaluación.

...

# Objetivos

- Obtener y analizar un dataset de textos escritos por pacientes con condiciones mentales definidas.
- Definir, desarrollar y entrenar una red de Deep Learning para clasificación de textos por categoría de enfermedad mental.
- Definir, desarrollar y entrenar un modelo de Deep Learning para generación de textos basados en los de los pacientes con enfermedades mentales.
- Desplegar los modelos desarrollados en forma de API REST mediante un despliegue de contenedores Docker.

...

...

...

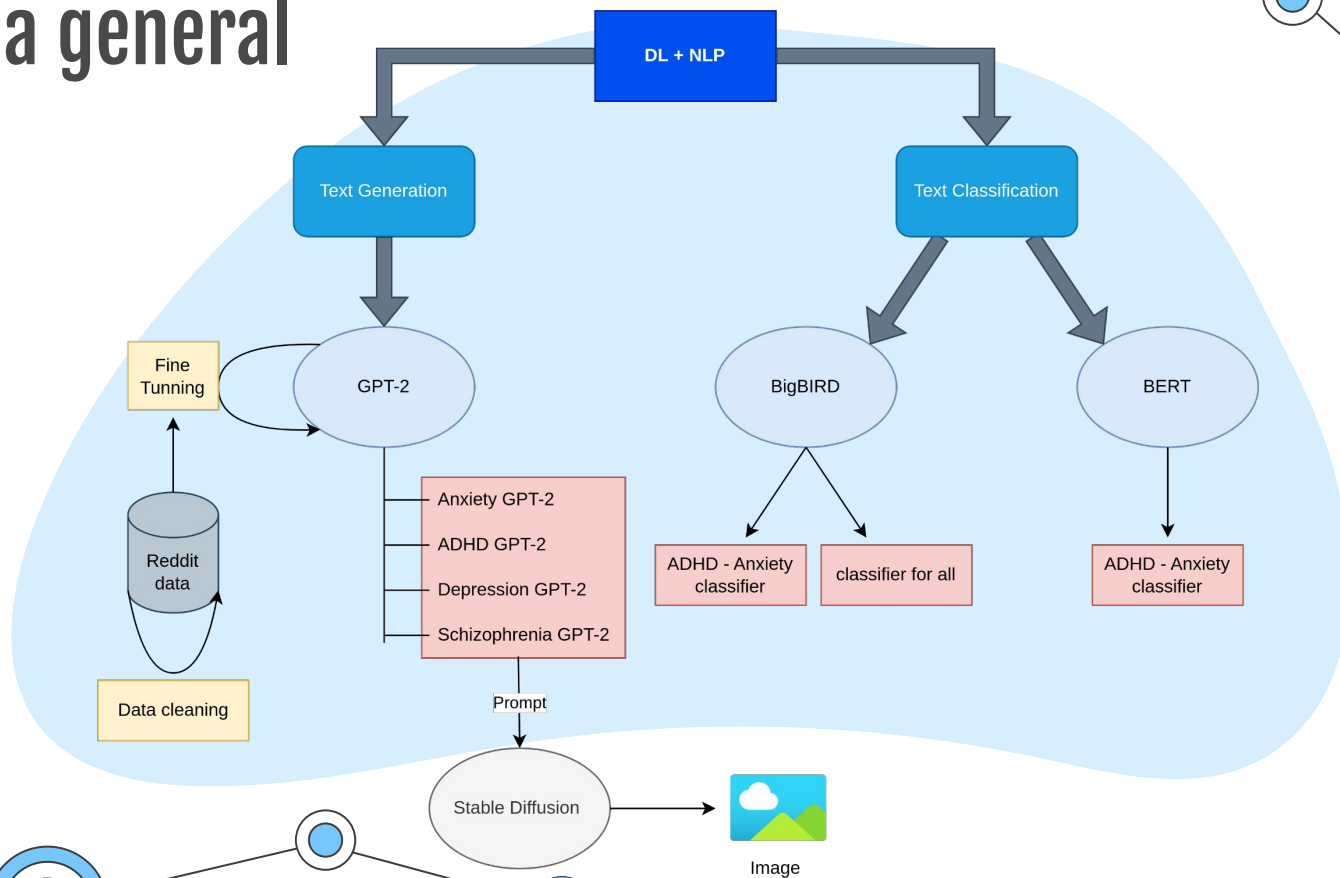


02

# Arquitectura Propuesta

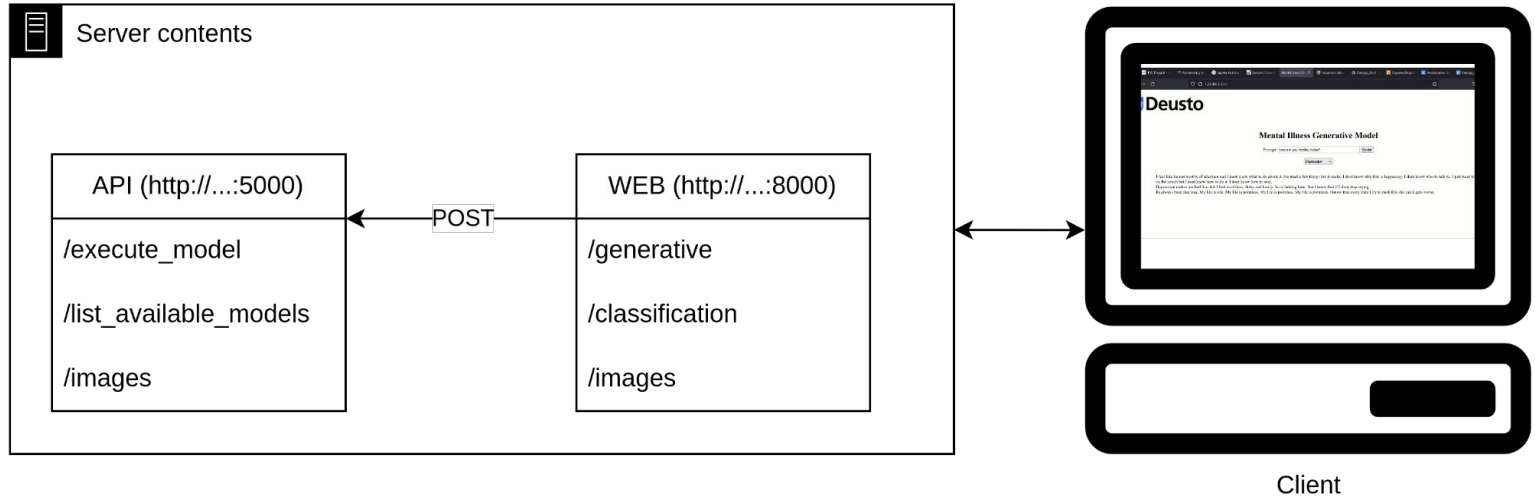


# Esquema general



# Componentes de aplicación

API → Backend  
WEB → Frontend







## Mental Illness Generative Model

Prompt:

Enviar

- ADHD
- ADHD
- anxiety
- classifier\_all
- Depression



127.0.0.1:8000



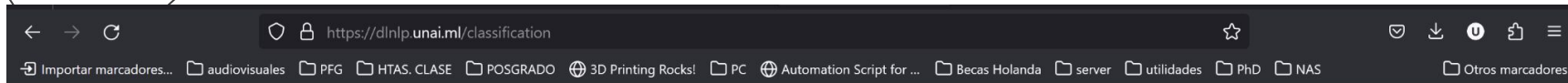
## Mental Illness Generative Model

Prompt:

Enviar

- Depression

A new girl in my life Ive been dating a girl for almost 3 months, and Ive been extremely sad for more than a year now. She always talks about how she likes me and that she could be with me or her. I really dont know what to do. I have no clue how to express how I feel about demonstrably nothing or about her. I have anxiety, depression and anxiety because of the way Im feeling. I cant express how I feel to her and I have no way of telling her that I feel this way. I feel like Im a disappointment to her and just like all of this I just think that I should just leave her out of



[Modelo Generativo](#)


## Mental Illness Classification Model

Input text:

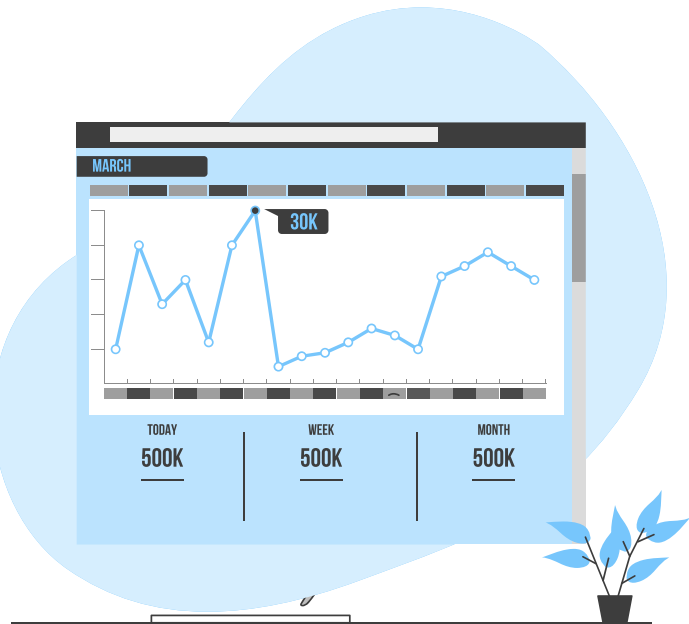


03

# Desarrollo del proyecto



# Adquisición de datos



Dataset obtenido de zenodo.

Datos de reddit para:

- Ansiedad
- ADHD
- Bipolaridad
- BPD
- Depresión
- EDA
- PTSD
- Esquizofrenia

Calidad del dataset no muy buena.

Script para combinación de todos los datos en un único dataset.

Generación de label y extracción de texto.



# Desarrollo del proyecto. Modelos de clasificación

## BERT

BERT (Bidirectional Encoder Representations from Transformers) modelo de lenguaje basado en redes neuronales de transformers.


Presentado por Google en 2018 con un rendimiento sobresaliente en:

- Reconocimiento de entidades
- Clasificación de texto
- Etiquetado de partes del discurso
- Respuesta a preguntas

## BigBird

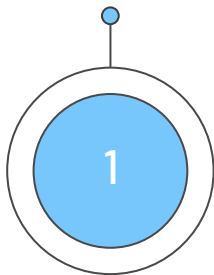
Modelo de lenguaje basado en transformers desarrollado como una extensión de BERT.

Presentado por Google en 2020 y diseñado para abordar el desafío de procesar secuencias más largas de texto de manera más eficiente.

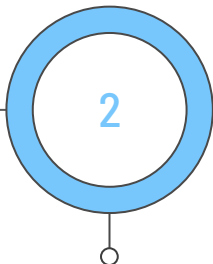


# Desarrollo del proyecto. Modelos de clasificación

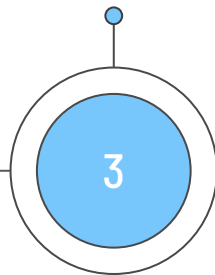
Limpieza de datos para  
clasificación



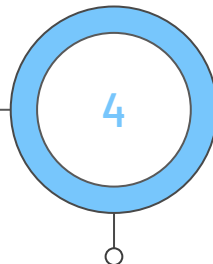
Modelo de clasificación  
binaria con BERT y  
BigBird

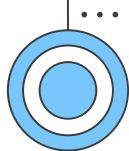


Revisión de resultados y  
Selección del modelo con  
mejores resultados (BigBird)



Modelo de clasificación  
para todas las  
condiciones con BigBird





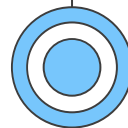
# Desarrollo del proyecto. Modelos generativos

## PROCESO DE TRABAJO

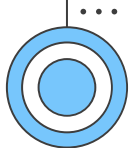
- Generación de corpus específico para cada condición mental
  - Importación del fichero combinado
  - Limpieza de textos
    - Definición de expresiones regulares para eliminación de emoticonos, símbolos, URLs y similares
    - Eliminación de ciertos términos detectados tras un primer entrenamiento (amp, TLDR, etc.)
  - Filtrado por label y generación de corpus específico
- Generación de script de entrenamiento para cada condición mental
  - Carga de corpus generado
  - Elaboración de lista con documentos divididos
  - Creación de script de dataset utilizando librería de Hugging Face
    - Generación de lista de diccionarios: “id” y “text”
    - Generación de dataset
  - Data split



...



...



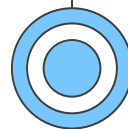
# Desarrollo del proyecto. Modelos generativos

## PROCESO DE TRABAJO

- Realización del entrenamiento del modelo de cada enfermedad
  - Empleando script “run\_clm.py”
  - Definición de parámetros (learning rate, stops, epochs, etc.)
  - Definición de rutas de output y logs
- Obtención de checkpoint
- Repetición del proceso para cada enfermedad



...



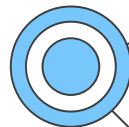
...





...

# Desarrollo del proyecto. Integración y despliegue



...

## 01

### Codificación

Realización de código limpio capaz de cargar un modelo y ejecutarlo.

## 02

### API

Creación incremental de la API Flask, añadiendo funcionalidades progresivamente y probándola con programas externos.

## 03

### WEB

Diseño del frontend utilizando Flask.

## 04

### Integración y Dockerización

Inclusión de peticiones a la API en el frontend.  
Dockerizar el proyecto. Subir proyecto a server en dominio público.



# 04

## Desviaciones





## Uso de dataset alternativo

En lugar de utilizar el dataset de la Universidad de Georgetown, se optó por un dataset de menor calidad obtenido de Zenodo. Esto se debió a problemas legales y a la falta de registros explícitos sobre la enfermedad en cuestión en el dataset original.

## Descarte de uso en sesiones de terapia

Debido a que el modelo no se ajustaba al caso de uso, se decidió descartar la opción de poder utilizarlo en sesiones de terapia reales

## Generación de textos de baja calidad

Tras entrenar el modelo generativo, se observó que los textos generados eran demasiado explícitos y de baja calidad. Esto hizo que se descartara la retroalimentación del modelo base, ya que podría sobreentrenarlo y afectar negativamente al primer propósito del proyecto.





## Descarte del módulo de speech to text

Se decidió no desarrollar el módulo de speech to text propuesto inicialmente debido a que no aportaba suficiente valor a las tareas de clasificación y generación, y por limitaciones temporales en la entrega del proyecto.

## Generación de imágenes

Se añadió una función de generación de imágenes utilizando prompts y un modelo pre-entrenado de stable-diffusion. Esto permitió ilustrar de manera gráfica conceptos relacionados con las enfermedades para las que se entrenaron los modelos de generación de texto.

## Problemas técnicos y brechas de seguridad

La integración del proyecto en una solución visual para usuarios no expertos requirió más tiempo del esperado debido a los problemas técnicos que surgieron durante el desarrollo. Esto resultó en una menor calidad en el frontend y brechas de seguridad en el servidor.



# 05

## Demo



# Demo

A continuación se va a realizar una  
demostración práctica del proyecto  
desarrollado

<https://dlnlp.unai.ml/>

## Mental Illness Generative Model

Prompt:

Depression ▾

how are you feeling today?

I've become so depressed I can't even get out of bed. I just feel like I am in a complete mental breakdown. I get upset when something bad happens and I am just angry and upset. I have no motivation to do anything. I have no energy to do anything. I don't know how to fight it. I don't think I can hold on to some of my friends anymore. I just want to get away from all this. I don't want to think about anything. I just want to stop existing.

I don't want to live. I don't want to live. I want to die. I want to fucking die.

I am depressed because I am a girl and because I am not attractive.



# Deusto

## Mental Illness Classification Model

Input text:

the insomnia, but the main issue was I was really anxious about getting my meds. I had also been diagnosed with Aspergers, and 53 mg of Concerta was the equivalent of a week of my morning dose. Ive been on Vyvanse for about a month now, and Im fairly certain that I couldnt feel anything for the first few hours of the day. I did feel something, but it was just not as intense.

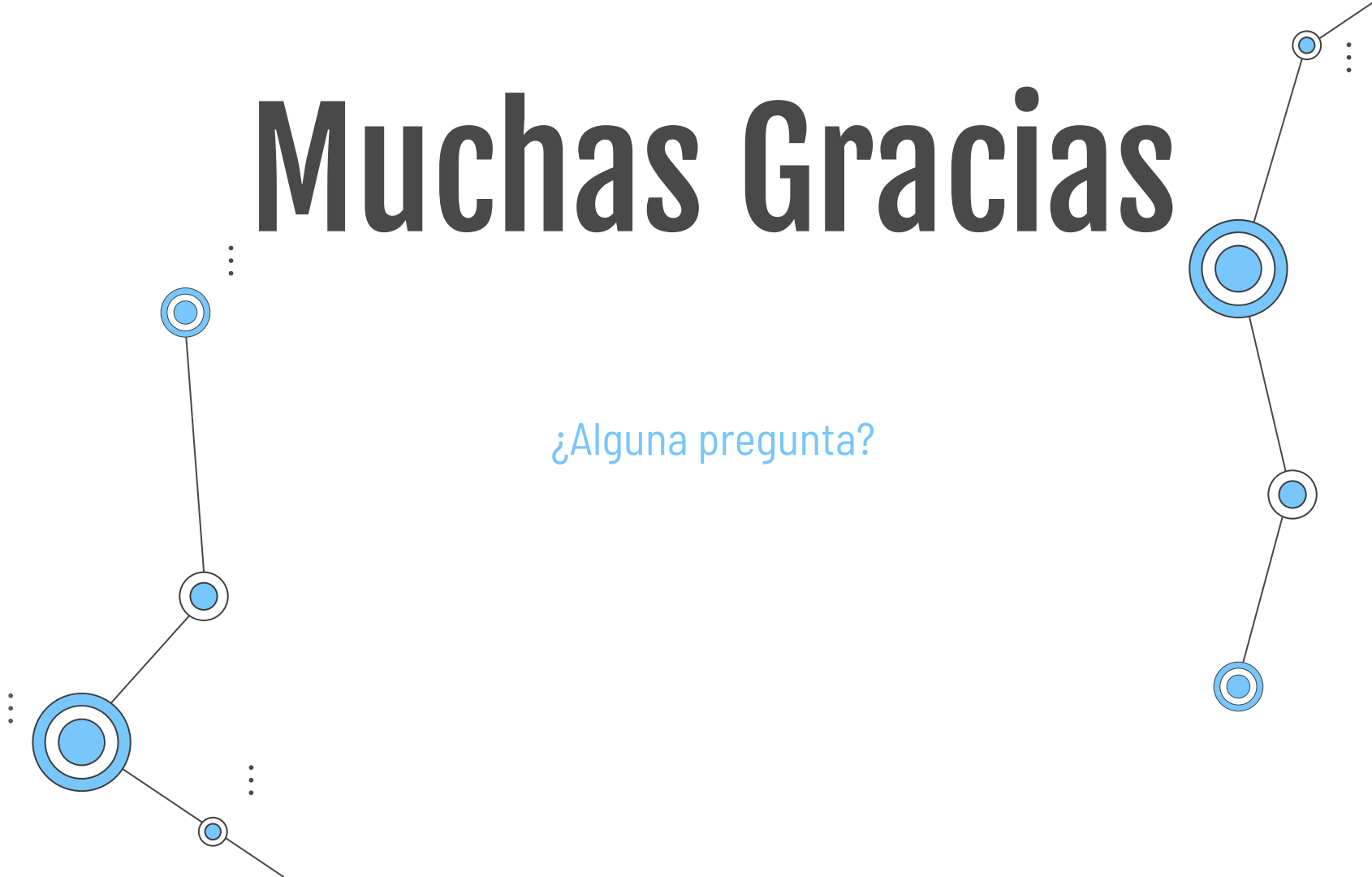
Enviar

adhd

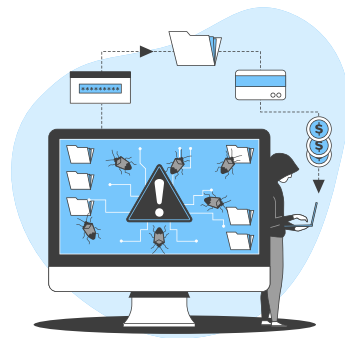


# Muchas Gracias

¿Alguna pregunta?



# Alternative Resources



Find more illustrations like these on [Stories by Freepik](#)

# Resources

## Vectors

- Programming
- Development
- Mission impossible
- Social networking
- Segmentation
- Site stats
- Responsive
- Meeting
- Analysis
- Cyber attack
- Secure server
- Control panel

## Photos

- Close-up of woman using a laptop
- Beautiful young woman holding laptop on pink backdrop

## Icons

- Project Management Icon

Find more illustrations like these on  
[Stories by Freepik](#)

# Instructions for use

In order to use this template, you must credit [Slidesgo](#) by keeping the **Thanks** slide.

## You are allowed to:

- Modify this template.
- Use it for both personal and commercial projects.

## You are not allowed to:

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>

# Fonts & colors used

This presentation has been made using the following fonts:

## **Fjalla One**

(<https://fonts.google.com/specimen/Fjalla+One>)

## **Barlow Condensed**

(<https://fonts.google.com/specimen/Barlow+Condensed>)

#494949

#477797

#77c6fc

#bbe3fe

#ddf1fe

# Storyset

Create your [Story](#) with our illustrated concepts. Choose the style you like the most, edit its colors, pick the background and layers you want to be visible and bring them to life with the animator panel. It will boost your presentation for sure! Check out [how it works](#).



Pana



Amico



Bro



Rafiki

# Use our editable graphic resources...

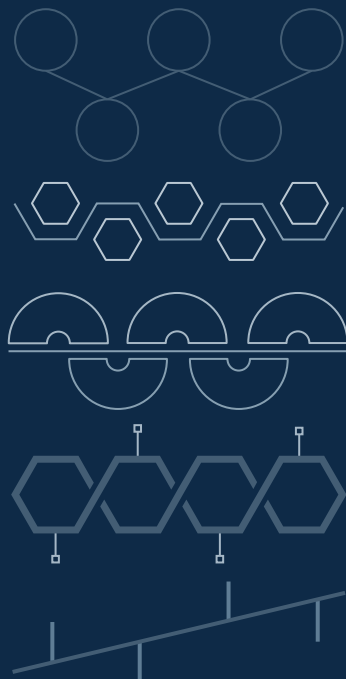
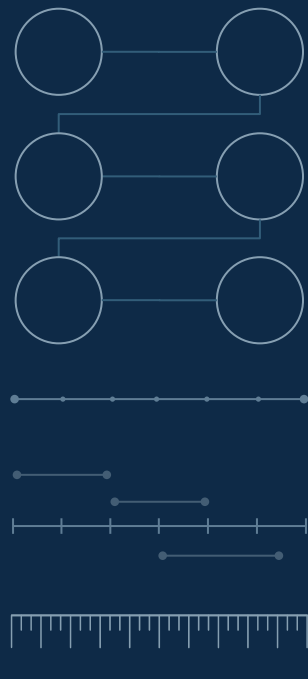
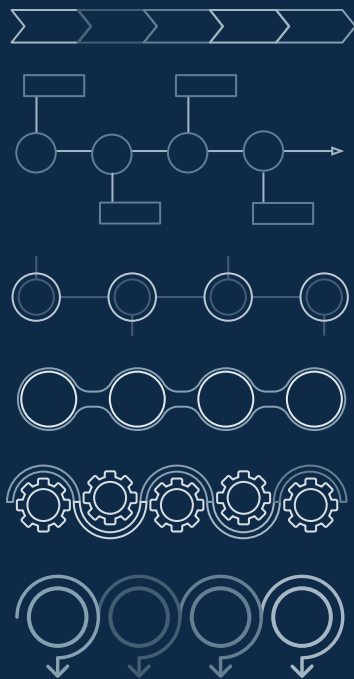
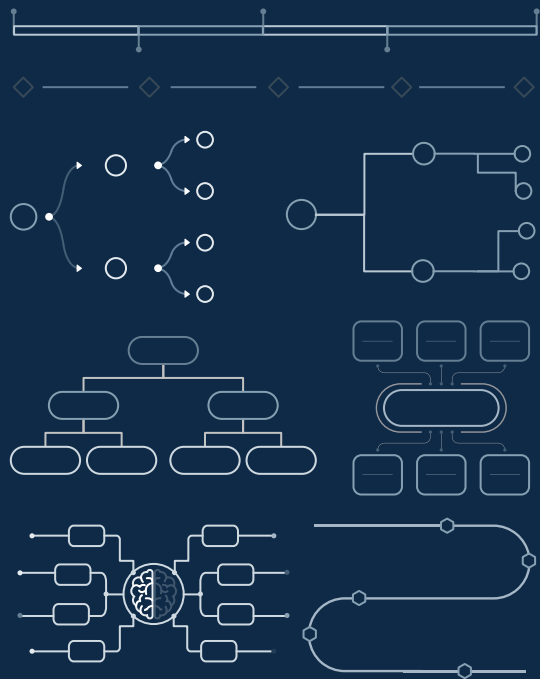
You can easily [resize](#) these resources without losing quality. To [change the color](#), just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want.

Group the resource again when you're done. You can also look for more [infographics](#) on [Slidesgo](#).

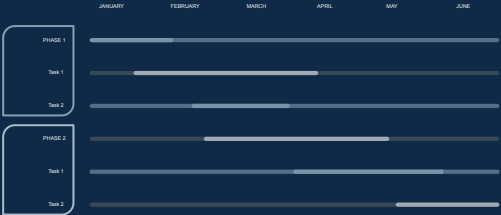


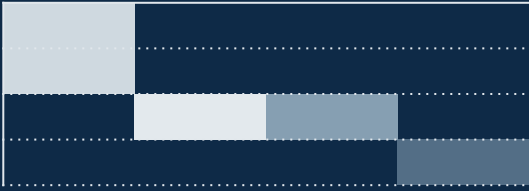


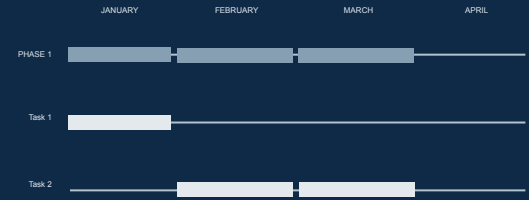




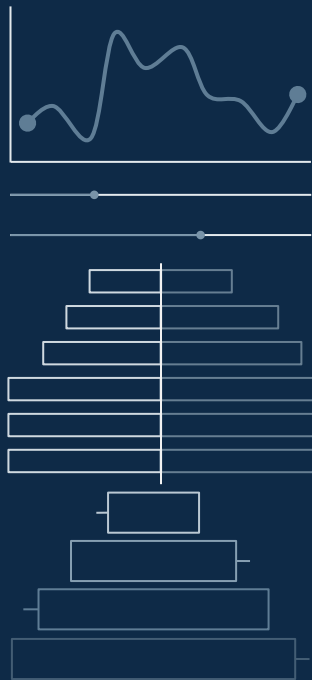
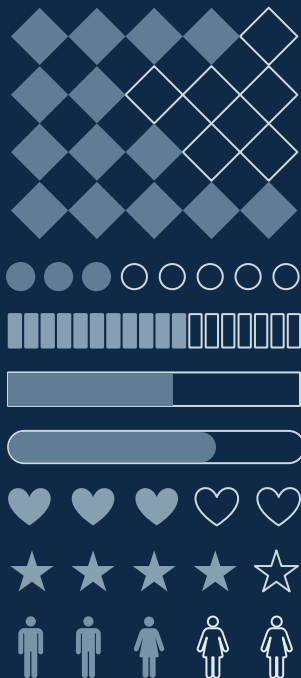
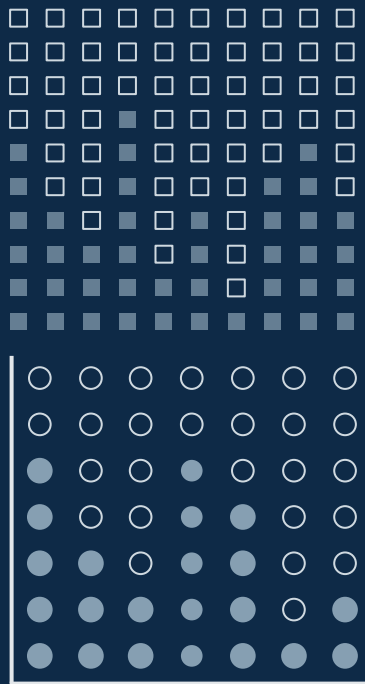






## ...and our sets of editable icons

You can resize these icons without losing quality.

You can change the stroke and fill color; just select the icon and click on the paint bucket/pen.

In Google Slides, you can also use Flaticon's extension, allowing you to customize and add even more icons.



## Educational Icons

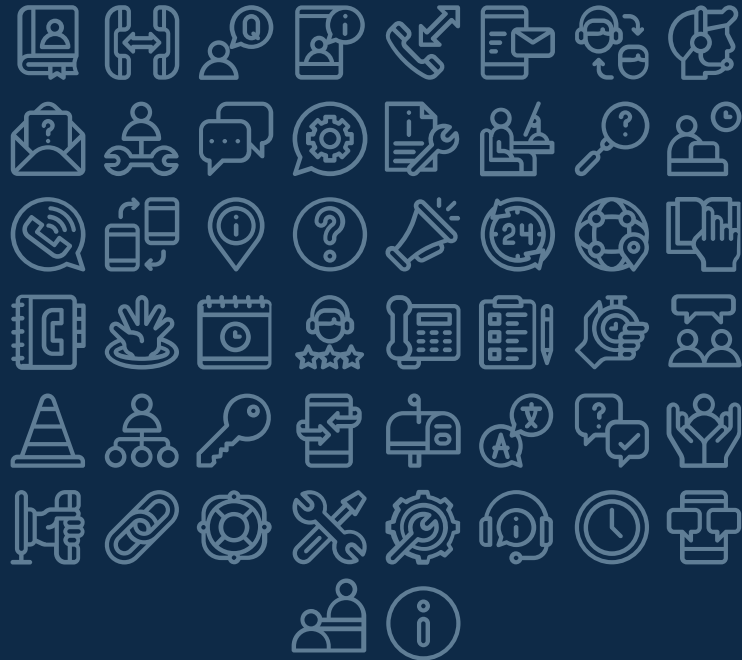


## Medical Icons





## Help & Support Icons



## Avatar Icons





## Creative Process Icons



## Performing Arts Icons



## Nature Icons



# SEO & Marketing Icons



