



Universidad Nacional de  
Ingeniería

Maestría en Data  
Science:  
Tesis 1: Desarrollo de  
plan de tesis

# Introducción al plan de tesis

Mg. Ing. Luis Alberto Sánchez Alvarado

# ¿Qué es el plan de tesis?

El plan o proyecto de tesis es una descripción breve de lo que se pretende hacer en la investigación. Es decir, es un documento que resume los objetivos, alcance y justificación de un proyecto de investigación.



# Revisemos primero algunos puntos ...

## Definición del programa

La Maestría en Ciencia de Datos es un programa interdisciplinario diseñado para profesionales con visión estratégica y habilidades analíticas, estadísticas y computacionales. Su objetivo es formar científicos de datos especializados en el desarrollo de soluciones analíticas.

## Perfil del postulante

La Maestría en Ciencia de Datos esta diseñada para profesionales de cualquier especialidad con una visión estratégica para la toma de decisiones basada en datos, con interés de profundizar en el área de la ciencia de datos.

## Definición de tesis

La tesis en CD es una investigación aplicada que se sustenta por una revisión exhaustiva de artículos científicos. Se caracteriza por el uso intensivo de técnicas de ciencia de datos cuyo conocimiento fue adquirido por el estudiante durante sus estudios de Maestría en Ciencia de Datos. Toda investigación en CD debe seguir las siguientes etapas en el desarrollo de la solución: entendimiento y planteamiento de la problemática, análisis de la estructura de los datos y solución de las casuísticas que se presenten en la data, desarrollo de la solución vía algoritmos, validación de las métricas de performance, implementación y propuestas de estrategias.

## Procedimiento

Los temas serán propuestos por los participantes, luego deberán plantear sus proyectos de tesis y, con ello, se le asignará un docente de la especialidad como asesor. Los artículos científicos a revisar tienen que estar orientados al desarrollo de las soluciones analíticas en el end to end, con el uso de técnicas de ciencia de datos.

# Por donde empezar a buscar...



¿Qué áreas puedo encontrar?

**Big data**

**Análisis de datos**

**Machine learning**

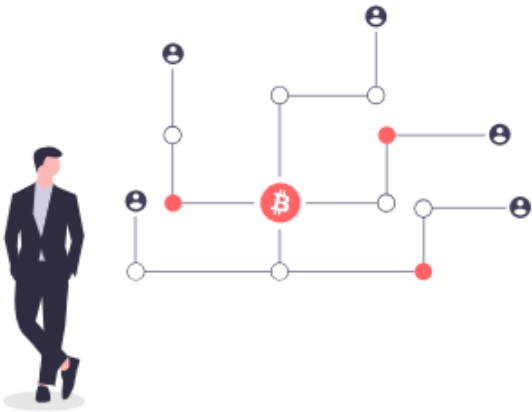
**Visualización**

**Forecasting**

# Al pensar una idea debo pensar en ...

## Idea inicial de investigación

Tengo una idea que me motiva investigar ¿**resuelve problemas** y puede ser interesante para el resto?



¿Cuál es el objetivo del proyecto?



¿Cuál es la importancia del proyecto?



# ¿Puedo comenzar con más de una idea ?

Si se puede comenzar con una lluvia de ideas, pero deben ser depuradas cuanto antes y quedarse con máximo dos

¿Cómo podemos depurar las ideas?



## Un primer filtro es revisando antecedentes

¿Ya se ha trabajado  
en este idea?

Antecedentes

¿Hay ideas similares que  
puedan mejorarse?

¿Hay ideas similares que  
puedan mejorarse?



# De los lugares de consulta los más recomendables son:



ScienceDirect

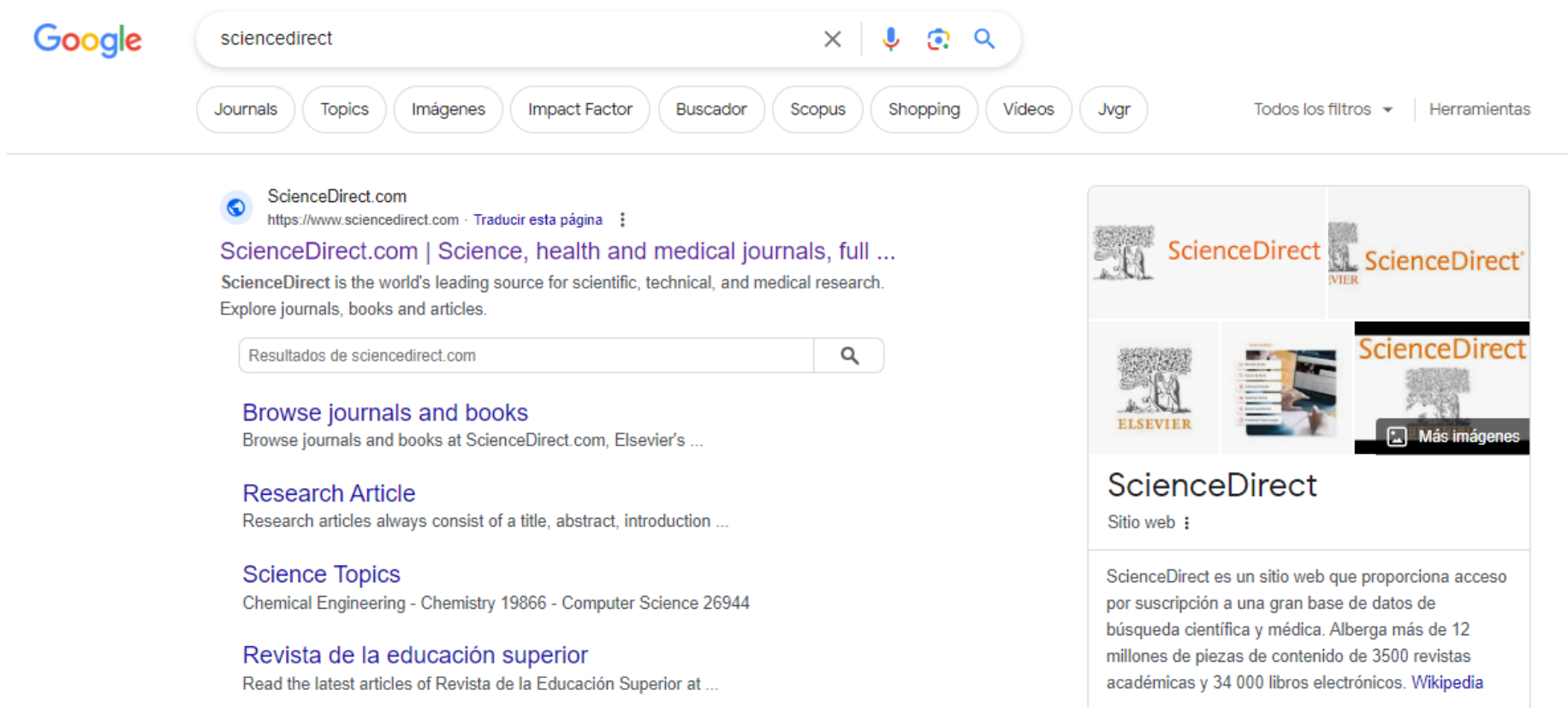
ScienceDirect es una plataforma que da acceso a las publicaciones de Elsevier, una de las principales editoriales científicas y académicas.

Su contenido es multidisciplinar: física e ingeniería, ciencias de la vida, ciencias de la salud, y ciencias sociales y humanidades.



## Paso 1:

En el buscador poner “ScienceDirect”, o ingresar al link: <https://www.sciencedirect.com/>



The image shows a Google search interface. At the top, the Google logo is on the left, and a search bar contains the text "sciencedirect". To the right of the search bar are icons for voice search, image search, and a magnifying glass. Below the search bar is a horizontal row of buttons: "Journals", "Topics", "Imágenes", "Impact Factor", "Buscador", "Scopus", "Shopping", "Vídeos", and "Jvgr". To the right of these buttons are the links "Todos los filtros" and "Herramientas".

The search results are displayed below. The first result is for "ScienceDirect.com" with the URL "https://www.sciencedirect.com" and a link to "Traducir esta página". The snippet reads: "ScienceDirect.com | Science, health and medical journals, full ... ScienceDirect is the world's leading source for scientific, technical, and medical research. Explore journals, books and articles." Below this snippet is a search bar with the text "Resultados de sciencedirect.com" and a magnifying glass icon.

Below the search bar are four links with snippets:

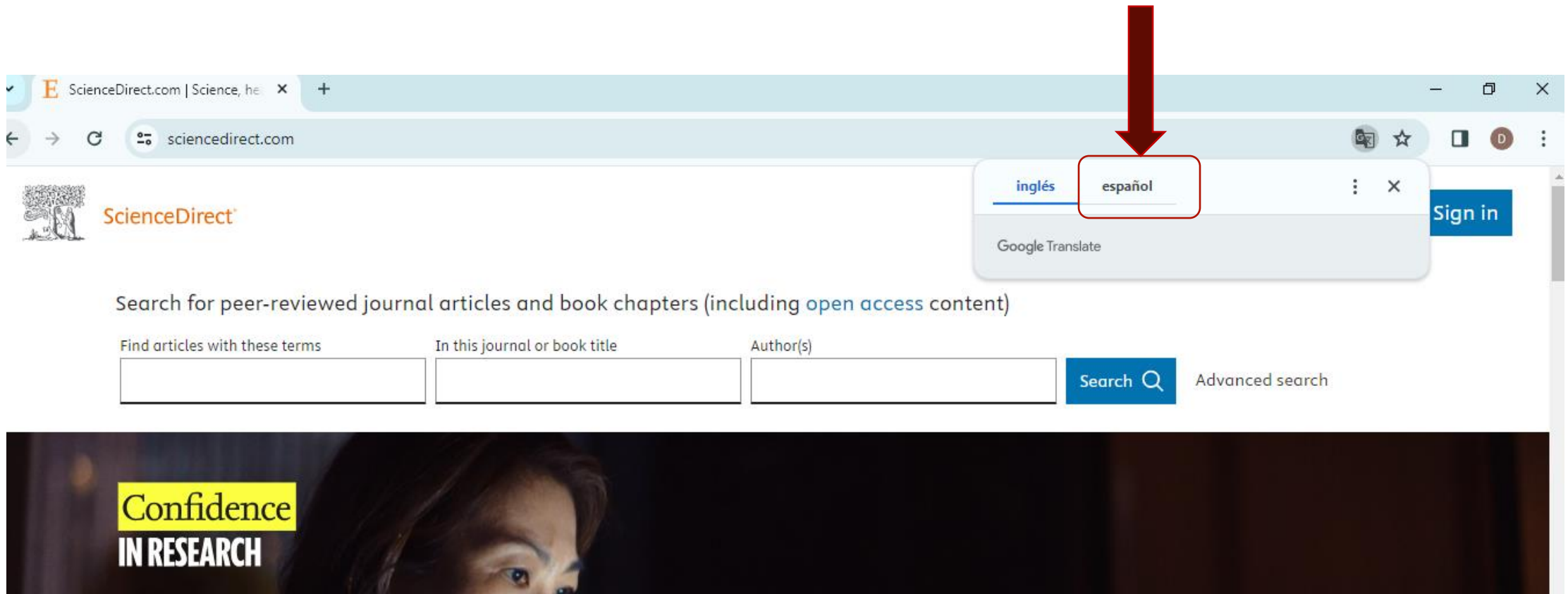
- Browse journals and books**: Browse journals and books at ScienceDirect.com, Elsevier's ...
- Research Article**: Research articles always consist of a title, abstract, introduction ...
- Science Topics**: Chemical Engineering - Chemistry 19866 - Computer Science 26944
- Revista de la educación superior**: Read the latest articles of Revista de la Educación Superior at ...

On the right side of the search results is a vertical strip of image thumbnails. The first two thumbnails show the ScienceDirect logo. The third thumbnail shows the Elsevier logo. The fourth thumbnail shows a book cover. Below the thumbnails is a link that says "Más imágenes".

Below the image thumbnails is a section titled "ScienceDirect" with the text "Sitio web :". Below this is a paragraph of text: "ScienceDirect es un sitio web que proporciona acceso por suscripción a una gran base de datos de búsqueda científica y médica. Alberga más de 12 millones de piezas de contenido de 3500 revistas académicas y 34 000 libros electrónicos. Wikipedia".

## Paso 1:

Si aparece el mensaje de traducir poner “Español”:



The screenshot shows the ScienceDirect website interface. At the top, there is a browser tab labeled "ScienceDirect.com | Science, he" and a address bar showing "sciencedirect.com". Below the browser bar, the ScienceDirect logo is visible on the left. On the right, there is a language selection dropdown menu. The menu is open, showing two options: "inglés" and "español". A red arrow points down to the "español" option, which is highlighted with a red border. To the right of the language menu is a "Sign in" button. Below the language menu, there is a search bar with the text "Search for peer-reviewed journal articles and book chapters (including open access content)". Below the search bar, there are three input fields labeled "Find articles with these terms", "In this journal or book title", and "Author(s)". To the right of these fields is a "Search" button with a magnifying glass icon and a link to "Advanced search". At the bottom of the page, there is a banner image with the text "Confidence IN RESEARCH" in white and yellow.

## Paso 1:

Hacer clic en “Registro”:



ScienceDirect

Revistas y libros



Registro

Iniciar sesión

Busque artículos de revistas y capítulos de libros revisados por pares (incluidos [acceso abierto](#) contenido)

Encuentra artículos con estos términos

En el título de esta revista o libro.

Autor(es)

Buscar 

Búsqueda Avanzada

## Paso 2:

Ingresa su correo institucional UNI



The image shows a login modal window for ScienceDirect. At the top is the Elsevier logo, which consists of a tree and a person, with the word "ELSEVIER" below it. Below the logo, the text "Bienvenido/a" is displayed. Underneath, it says "Introduzca su dirección de correo electrónico para continuar con ScienceDirect". There is a text input field labeled "Correo electrónico" with a cursor inside. Below the input field are two buttons: a grey button labeled "Continuar" and a blue-outlined button labeled "Iniciar sesión a través de su organización". The modal is set against a background of a street map.

×

  
ELSEVIER

Bienvenido/a

Introduzca su dirección de correo electrónico para continuar  
con ScienceDirect

Correo electrónico

\_\_\_\_\_

Continuar

Iniciar sesión a través de su organización

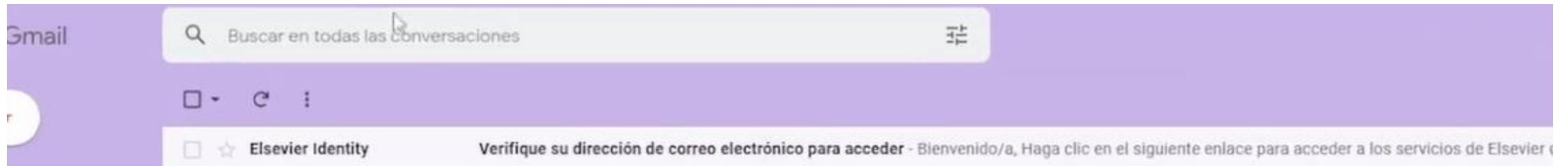
## Paso 2:

Con ello les llegará un correo de verificación:



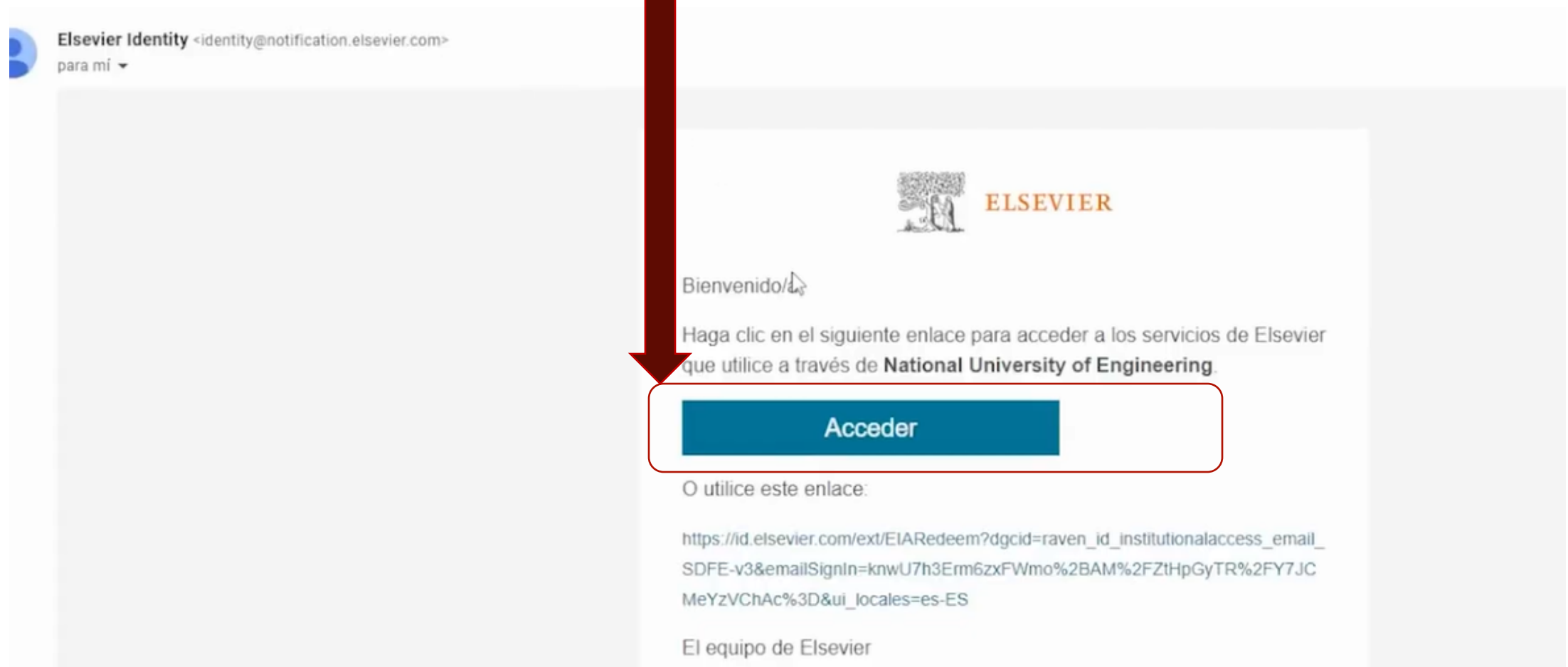
## Paso 2:

Revisar el mail:



## Paso 3:

Entrar al mail y hacer clic en “Acceder”:





## Paso 4:

Completar los datos y dar clic en “Registrarse”:



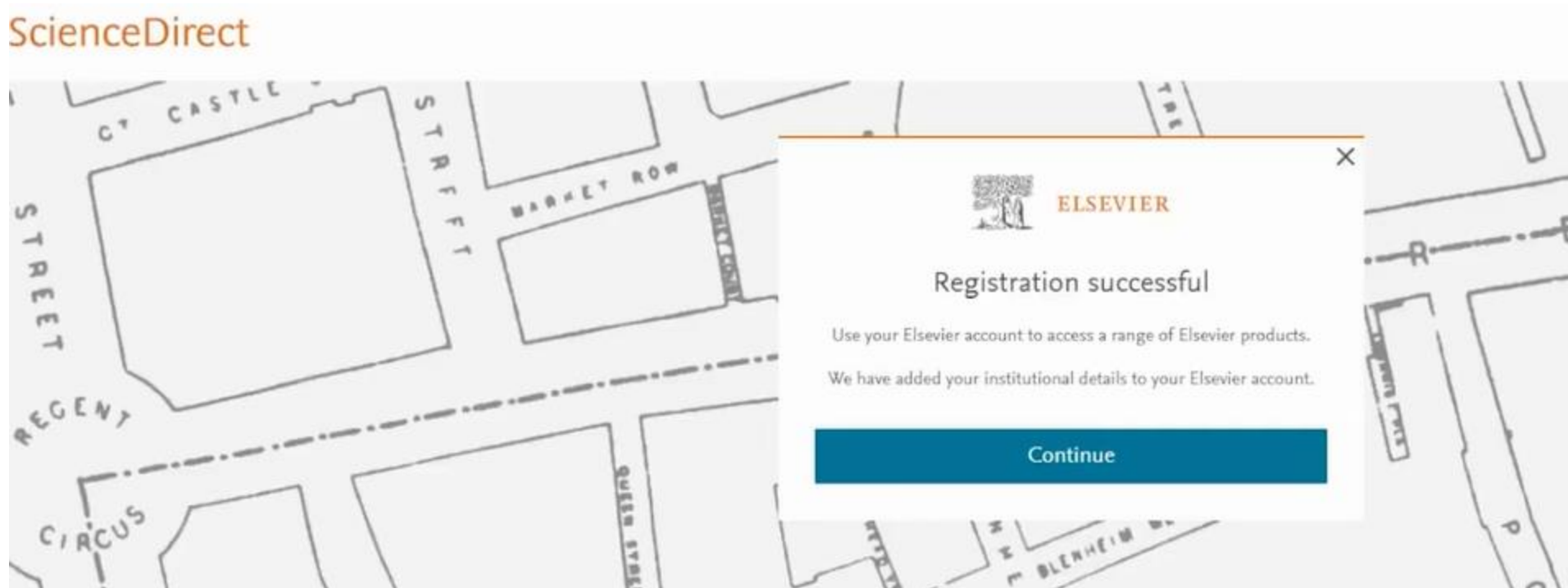
The image shows a registration form for Elsevier. At the top, there is the Elsevier logo and the word "ELSEVIER". Below this, the title "Registrarse" is centered. A subtitle reads: "Agregue sus datos y establezca su contraseña para registrarse en Elsevier,". The form contains several input fields: "Correo electrónico" with the value "lguerram@uni.edu.pe", "Institución" with the value "National University of Engineering", "Nombre de pila" with the value "L", and "Apellidos" which is empty. There is also a "Contraseña" field which is currently empty. Below the password field, there is a note: "Elija una contraseña única que contenga al menos 8 caracteres. Incluya al menos un número y un símbolo y combine minúsculas y mayúsculas para que su contraseña sea segura."



This section shows the password strength indicator and the registration buttons. The "Contraseña" field is filled with dots, and a strength indicator shows "ACEPTABLE" in a yellow box. Below this, there is a checkbox labeled "Mantener la sesión iniciada (no recomendable para dispositivos compartidos)". A paragraph of text states: "Elsevier podría enviarle comunicaciones de marketing sobre productos y eventos relevantes. Puede cancelar su suscripción en cualquier momento a través de su cuenta de Elsevier." Another paragraph says: "Al continuar, acepta nuestros Términos y condiciones y Política de privacidad." At the bottom, there are two buttons: "Registrarse" (in a blue box) and "Cancelar" (in a white box with a grey border).

## Paso 4:

Luego de ello saldrá un mensaje que el registro se hizo de manera satisfactoria:



Una vez registrados se podrá comenzar a realizar las búsquedas:



ScienceDirect®

Journals & Books



Luis Alberto Sanchez Alvarado

LA


 [You have institutional access](#)

Search for peer-reviewed journal articles and book chapters (including [open access](#) content)

Find articles with these terms

In this journal or book title

Author(s)

Search 

[Advanced search](#)

## Recomendación de mis colegas



Las exposiciones servirán para poner en vitrina nuestras investigaciones y recibir recomendaciones

## En resumen...

1

Reflexionar sobre lo que te gustaría hacer dentro de un área en particular.

2

Leer la literatura existente para identificar problemas previamente reportados.

3

Consultar a expertos, ya que pueden tener ideas sobre lo que podría valer la pena explorar dentro del sector y área.

4

Consultar a empresas y organizaciones, ya que pueden tener problemas específicos que necesitan ser investigados.

# Algunos ejemplos ...

## Caso 1:



**Título:** Análisis de la satisfacción de clientes en el centro comercial Real Plaza

¿En qué sector se desarrolla?

Retail

¿Cuál es el objetivo del proyecto?

Utilizar modelos supervisados y no supervisados para analizar la satisfacción y comportamiento del cliente del Real Plaza, con el fin de identificar patrones y tendencias que puedan mejorar la experiencia del cliente y la rentabilidad de los establecimientos.

¿Cuál es la importancia del proyecto?

Comprender la experiencia del cliente es fundamental para los centros comerciales, ya que puede influir directamente en su éxito y rentabilidad. Mejorar la satisfacción del cliente puede aumentar la fidelidad de los clientes y generar recomendaciones positivas, lo que a su vez puede aumentar las ventas y la reputación de los centros comerciales.

# Algunos ejemplos ...

## Caso 2:



**Título:** Factores Determinantes de viralidad en TikTok en Perú

¿En qué sector se desarrolla?

Redes sociales

¿Cuál es el objetivo del proyecto?

Desarrollar un modelo predictivo para analizar los factores que contribuyen a la viralidad de los videos en TikTok entre los usuarios de Perú. El objetivo es identificar patrones y características comunes en los videos virales para ayudar a los creadores de contenido a aumentar su alcance.

¿Cuál es la importancia del proyecto?

Conocer los factores que influyen en la viralidad de los videos en TikTok puede ser valioso para los creadores de contenido, las marcas y los profesionales del marketing en Perú. Esto les permitirá optimizar sus estrategias de contenido y maximizar su impacto en la plataforma, lo que puede llevar a un mayor reconocimiento de marca y oportunidades de negocio.

# Algunos ejemplos ...

## Caso 3:



**Título:** Pronósticos Deportivos en la Liga Peruana de Fútbol utilizando Machine Learning

¿En qué sector se desarrolla?

Deportes

¿Cuál es el objetivo del proyecto?

Desarrollar un modelo de machine learning para pronosticar resultados de partidos de fútbol en la Liga Peruana, utilizando datos históricos y variables relevantes como rendimiento de equipos, estado físico de jugadores, entre otros. El objetivo es proporcionar pronósticos precisos que puedan ser utilizados por aficionados, apostadores y profesionales del deporte en Perú.

¿Cuál es la importancia del proyecto?

Los pronósticos deportivos son de gran interés para los aficionados al fútbol en Lima, Perú, y pueden tener implicaciones en áreas como el entretenimiento, las apuestas y la toma de decisiones estratégicas para equipos y jugadores. Un modelo preciso de pronósticos puede proporcionar información valiosa para estas audiencias.



# Algunos ejemplos ...

## Caso 4:



**Título:** Predicción de Morosidad en Préstamos Bancarios

¿En qué sector se desarrolla?

Banca

¿Cuál es el objetivo del proyecto?

Desarrollar un modelo de machine learning para predecir la probabilidad de morosidad en préstamos bancarios. El objetivo es utilizar datos históricos de clientes, transacciones y factores económicos para identificar patrones y señales de alerta temprana que indiquen un mayor riesgo de incumplimiento de pago

¿Cuál es la importancia del proyecto?

La morosidad puede tener un impacto significativo en la estabilidad financiera de los bancos. Detectar y prevenir la morosidad de manera proactiva puede ayudar a reducir pérdidas financieras, un modelo preciso de predicción de morosidad puede mejorar la gestión del riesgo crediticio y permitir una toma de decisiones más informada por parte de las instituciones financieras.

# Estructura del plan de tesis

## Contenido

<b>RESUMEN .....</b>	<b>2</b>
<b>Capítulo I. Introducción.....</b>	<b>3</b>
<b>1.1 Situación problemática .....</b>	<b>3</b>
<b>1.2 Propósito de la investigación:.....</b>	<b>4</b>
<i>1.2.1 Objetivo principal.....</i>	<i>4</i>
<i>1.2.2 Objetivos secundarios.....</i>	<i>4</i>
<b>1.3 Preguntas de investigación .....</b>	<b>5</b>
<b>1.4 Hipótesis de Investigación .....</b>	<b>6</b>
<b>1.5 Importancia de la tesis.....</b>	<b>7</b>
<b>Capítulo II. Revisión de la literatura.....</b>	<b>8</b>
<b>Capítulo III. Metodología.....</b>	<b>9</b>
<b>Capítulo IV: Plan de trabajo o cronograma.....</b>	<b>10</b>
<b>ANEXOS.....</b>	<b>12</b>



# Actividad



Ubicarnos en la maestría de ciencia de datos y plantear una idea de tesis, siguiendo las pautas revisadas en clases, completando al final:

- Título
- Objetivo
- Importancia

Luego poner a debate las ideas propuestas

# Primer trabajo

Prepara una presentación con los siguientes puntos:

Título de investigación

Objetivo de la investigación

Importancia de la investigación

Idea de recursos que tiene o demuestre la viabilidad

Cada idea presentada será comentada por el resto de alumno de la clase, con el fin de afinar la idea de investigación, y así mejorar su proyecto propuesto.



# Ingresar los datos de sus proyectos:

## PROYECTOS DE TESIS DEL PROGRAMA EN DATA SCIENCE

### Proyecto 1

Nombre del proyecto

Datos de los alumnos:

#### Alumno 1:

Nombre:

Mail:

Celular:

#### Alumno 2:

Nombre:

Mail:

Celular: