

GILkatón

Lng^{enierí}_{üística}a



Grupo de Ingeniería Lingüística



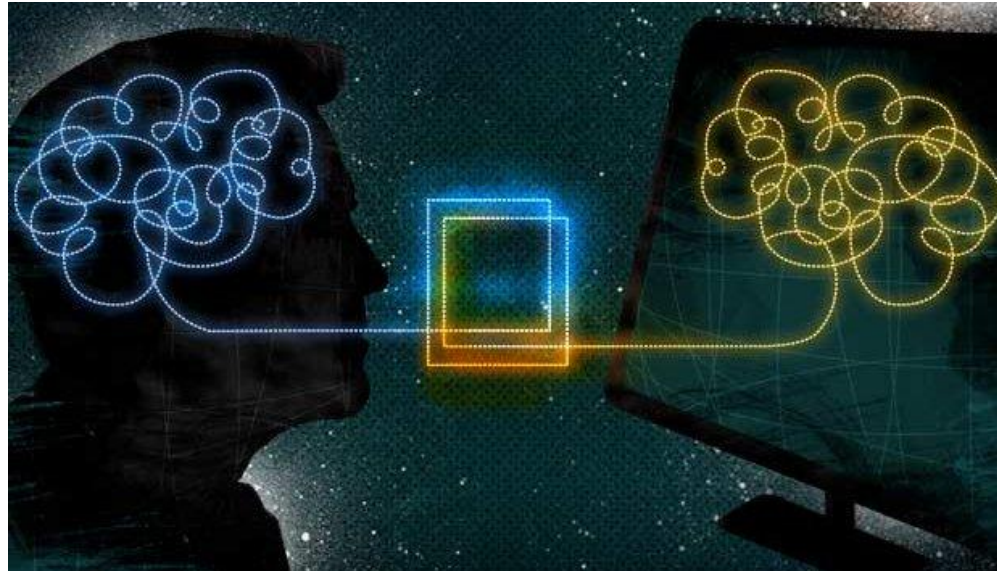
¿Qué es el GIL?



- Grupo de investigación interdisciplinario entre la lingüística e ingeniería.
- Fundado en 1999 con el fin de desarrollar formalmente el área de ingeniería lingüística en la UNAM y a nivel nacional.
- Su sede se encuentra en el Instituto de Ingeniería de la UNAM

¿Qué es la Ingeniería Lingüística?

- Aplicación de los conocimientos sobre el lenguaje para el desarrollo de sistemas informáticos que sean capaces de reconocer, comprender, interpretar y generar lenguaje humano en todas sus formas.



Integrantes

Jueces

Gemma Bel Enguix

Gerardo E. Sierra
Martínez

Ricardo Saucedo

Lucía Velázquez

Logística

Karla Jazmín Sánchez
Jerónimo

Rodrigo Mayett Guzmán

Anayeli Paulino de Jesús

Angélica Andrea Herrera
Godínez

Amparo Rocha Campos

Brayan de Jesús Rueda
González

Mentores

Octavio Sánchez

Alejandro Pimentel

Víctor Villegas

Julieta Pineda

Margarita Mota

Madai Ramírez

Lucía Velázquez

¿Cómo se formarán los grupos?

Lingüistas



Computólogos

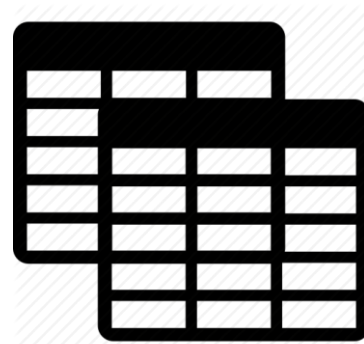


Objetivo

Estructurar datos

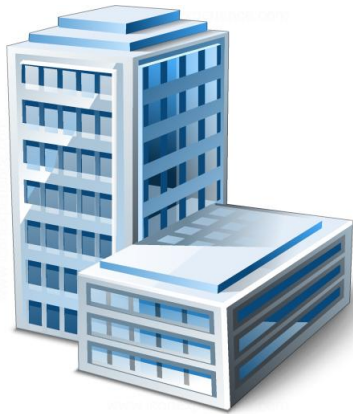
Texto

Tablas



Tres misiones

Entidades



Leyes



Sucesos



Elementos de la tabla

Entidades



- **Contexto:** El texto original que contiene los datos que se extraen.
- **Entidad:** El nombre completo de la entidad. Solo aquellos que también tengan un nombre corto.
- **Abreviatura:** El nombre corto que se le asigna a una entidad. Solo aquellos que tengan un nombre largo.

Elementos de la tabla

Entidades



Contexto	Entidad	Abreviatura
el Diario Oficial de la Federación ("DOF")	Diario Oficial de la Federación	DOF
la Solicitud de Concesión ("la Solicitud")	Solicitud de Concesión	Solicitud
Ley Federal de Telecomunicaciones y Radiodifusión (en lo sucesivo, "LFTR")	Ley Federal de Telecomunicaciones y Radiodifusión	LFTR
Estatuto Orgánico del Instituto Federal de Telecomunicaciones (el "Estatuto Orgánico")	Estatuto Orgánico del Instituto Federal de Telecomunicaciones	Estatuto Orgánico

Elementos de la tabla

Leyes



- **Contexto:** El texto original que contiene los datos que se extraen.
- **Ley:** El nombre completo de la ley o el origen de los artículos extraídos (constitución, estatuto, etc.).
- **Artículo:** El número del artículo.
- **Fracción:** El número de la fracción, si la hay.

Elementos de la tabla

Entidades



Contexto	Ley	Artículo	Fracción
los artículos 15 fracción IV y 17 fracción I de la Ley Federal de Telecomunicaciones y Radiodifusión para el otorgamiento de concesiones	Ley Federal de Telecomunicaciones y Radiodifusión	15	IV
los artículos 15 fracción IV y 17 fracción I de la Ley Federal de Telecomunicaciones y Radiodifusión para el otorgamiento de concesiones	Ley Federal de Telecomunicaciones y Radiodifusión	17	I
el artículo 174-B fracciones I y II del Estatuto Orgánico, establece la atribución	Estatuto Orgánico	174-B	I
el artículo 174-B fracciones I y II del Estatuto Orgánico, establece la atribución	Estatuto Orgánico	174-B	II

Elementos de la tabla

Sucesos



- **Contexto:** El texto original que contiene los datos que se extraen. Solo los sucesos que tengan fechas son los que deberán ser extraídos
- **Fecha:** La fecha en que ocurre el suceso. El formato deberá ser dd/mm/aaaa.
- **Lema:** El verbo principal del suceso, en infinitivo.
- **Quién**
- **Qué**
- **Dónde**
- **A quién**
- **Para qué**

Elementos de la tabla

Entidades



Contexto	Fecha	Lema	Quién	Qué	Dónde	A quién	Para qué
"Con fecha 11 de junio de 2013, se publicó en el Diario Oficial de la Federación el "Decreto por el que se reforman y adicionan diversas disposiciones de los artículos 94 y 105 de la Constitución Política de los Estados Unidos Mexicanos, en materia de telecomunicaciones" (el "Decreto de Reforma Constitucional"), mediante el cual se creó el Instituto Federal de Telecomunicaciones.	11/06/2013	publicar		Decreto por el que se reforman y adicionan diversas disposiciones de los artículos 94 y 105 de la Constitución Política de los Estados Unidos Mexicanos, en materia de telecomunicaciones	Diario Oficial de la Federación		se creó el Instituto Federal de Telecomunicaciones

Evaluación



- Deberán correr su programa con un documento de word al que tendrán acceso al final del hacking. Las tablas que obtengan se compararán con un gold standard.
- Cada una de las celdas correctas que obtengan valdrá un punto.
- Cada fila de información extra que obtengan (información no deseada) se penalizará con un punto.
- Cada una de las tres tablas tendrá el mismo valor (serán normalizadas según el número máximo de aciertos posibles).

Entrega



entidades.csv



sucesos.csv



leyes.csv



Código fuente