

# Knowledge Graph of Basque Cultural Concepts

---

Language Resources  
final project

erman ta zabal zazu

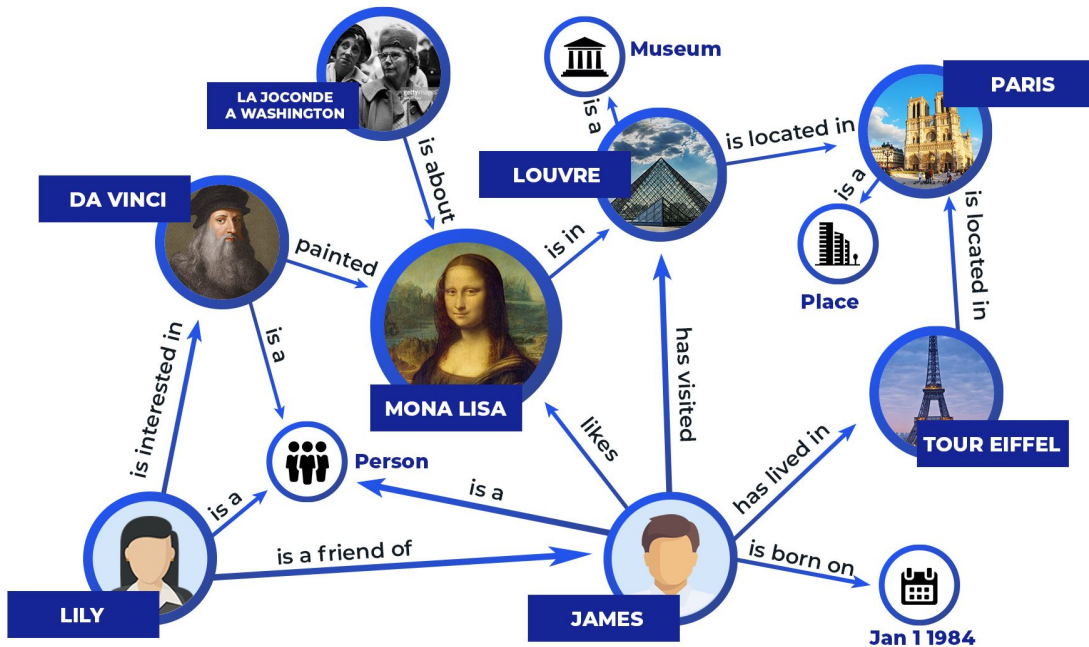


UPV EHU

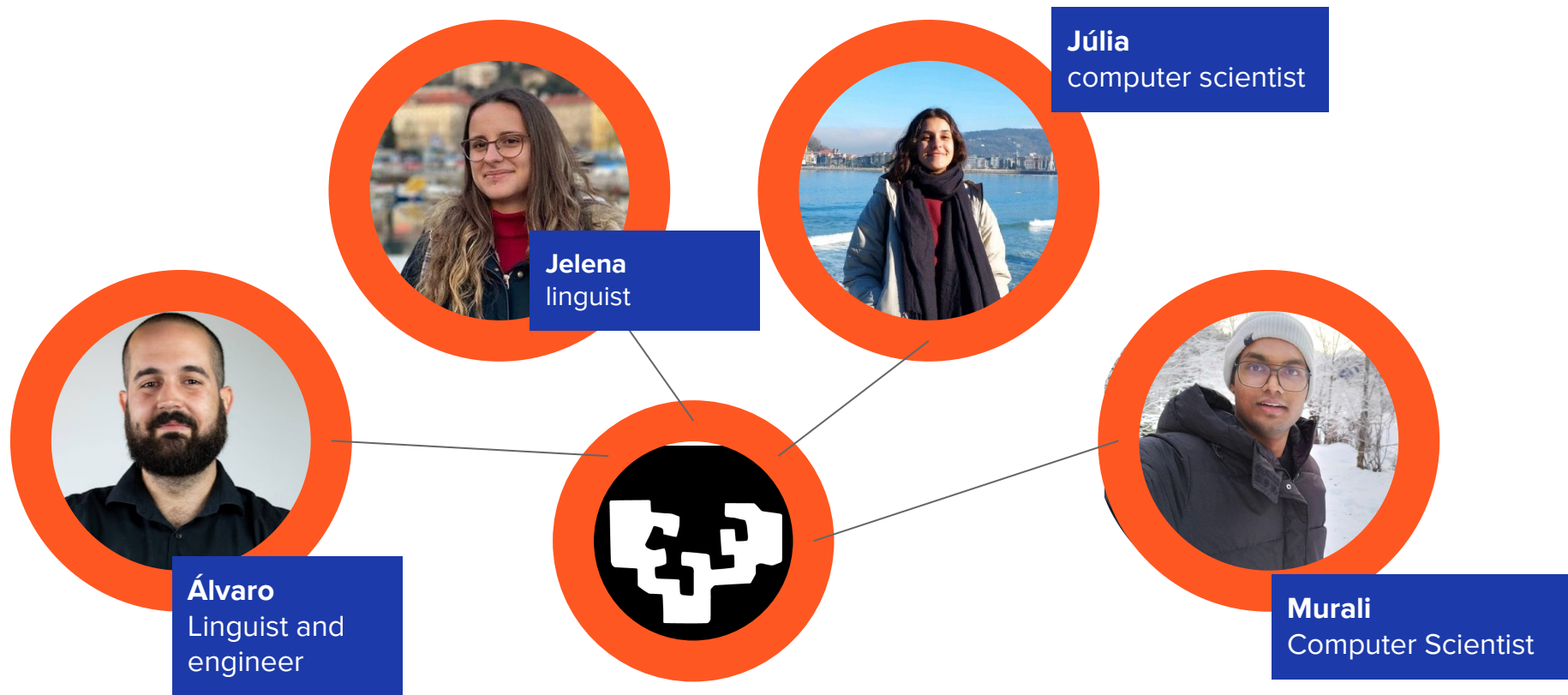
# What is a knowledge graph?

## a.k.a semantic network

A network of **real-world entities** and the **relationships** between them



# but first... our group!



**Concepts**

# Languages

CASTELLANO	EUSKERA
Playa de La Concha	Kontxa Hondartza
País Vasco	Euskadi
galerna	enbata
tarta de queso	gazta tarta
calimocho	kalimotxo



**Tortilla de patatas**



**Pintxo**



**Peine del viento**



**Tamborrada**



**Donostia**



**Eduardo Chillida**

# Categories



CATEGORY	SUB-CATEGORY 1	SUB-CATEGORY 2
Culture and the arts	Visual arts	Sculpture

# Wikipedia categories



## Culture and the arts (see in all page types)

[edit](#) · [watch](#)

*Main categories: Culture and The arts*

**Culture** and **Humanities** · Classics · Critical theory · Cultural anthropology · Folklore · Food culture · Food and drink · Language · Literature · Museology · Mythology · Philosophy · Popular culture · Science and culture · Traditions

**The arts** and **Entertainment** · Arts and crafts · Celebrity · Censorship in the arts · Festivals · Humor · Literature · Museums · Parties · Poetry

**Performing arts** · Circus · Dance · Film · Music · Opera · Storytelling · Theatre

**Visual arts** · Architecture · Comics · Crafts · Design · Drawing · Film (Animation) · New media art · Painting · Photography · Sculpture

**Games** and **Toys** · Board games · Card games · Dolls · Puppetry · Puzzles · Role-playing games · Video games

**Sports** and **Recreation** · Air sports · American football · Association football · Auto racing · Baseball · Basketball · Boating · Boxing · Canoeing · Cricket · Cycling · Exercise · Fishing · Golf · Gymnastics · Hobbies · Horse racing · Ice hockey · Lacrosse · Olympic Games · Rugby league · Rugby union · Sailing · Skiing · Swimming · Tennis · Track and field · Walking trails · Water sports · Whitewater sports

**Mass media** · Broadcasting · Film · Internet · Magazines · Newspapers · Publications · Publishing · Television · Radio



## Geography and places (see in all page types)

[edit](#) · [watch](#)

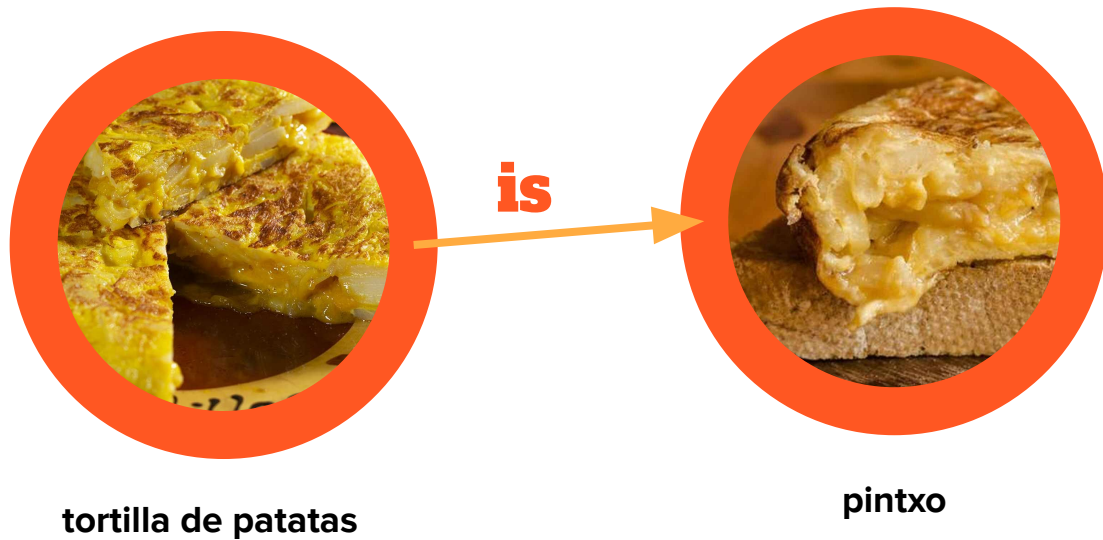
*Main categories: Geography and Places*

Earth · World · Bodies of water · Cities · Communities · Continents · Countries · Deserts · Lakes · Landforms · Mountains · Navigation · Oceans · Populated places · Protected areas · Regions · Rivers · Subterranea · Territories · Towns · Villages



# Relations

**IS**



# LOCATED\_IN/HAPPENS\_IN



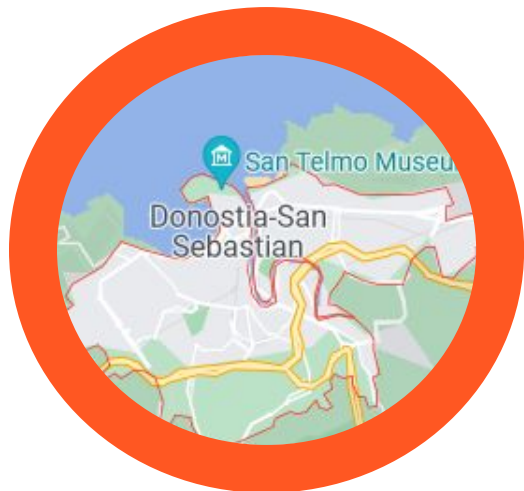
Tamborrada

**islocated/  
happens in**



Donostia

**IS\_PART\_OF**



**Donostia**

**IS\_PART\_OF**



**Basque Country**

**CREATED/MADE**



**Peine del viento**

**CREATED/  
MADE BY**



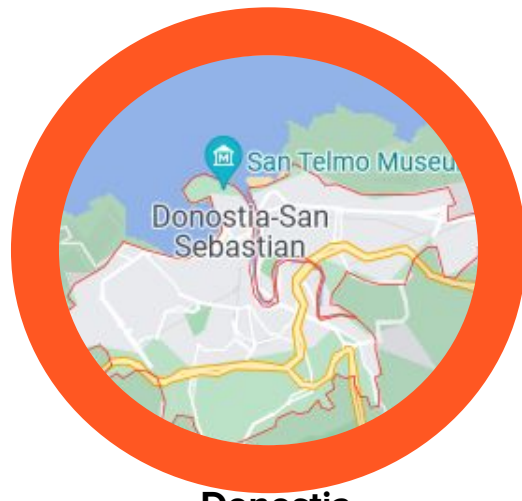
**Eduardo Chillida**

**TYPICAL\_IN**



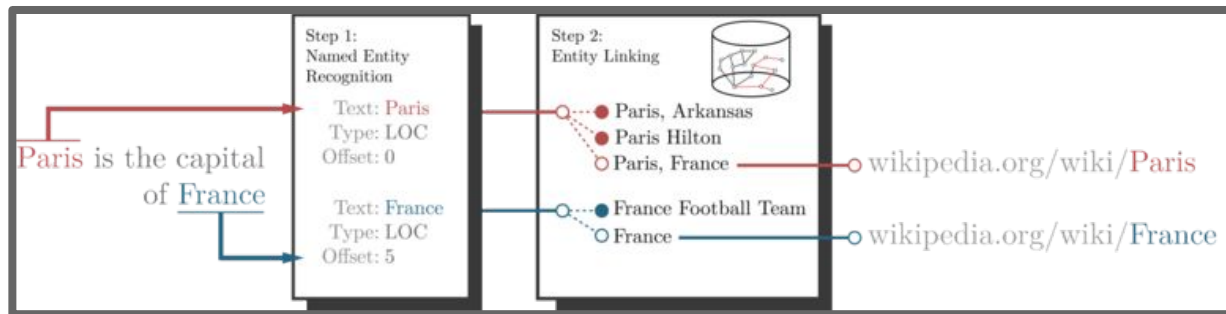
**Gilda**

**TYPICAL\_IN**



**Donostia**

# Applications



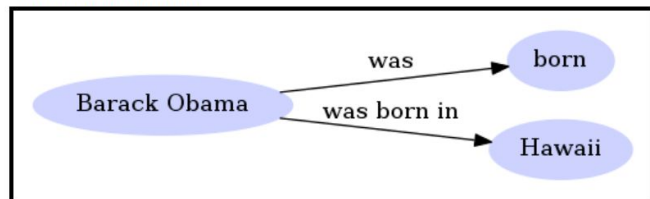
## Entity Linking

Unstructured Text

"Barack Obama was born in Hawaii."

Information Extraction

Structured Information



- Information Extraction
- Question Answering
- Common Sense Systems

# Common carbon footprint benchmarks

in lbs of CO2 equivalent

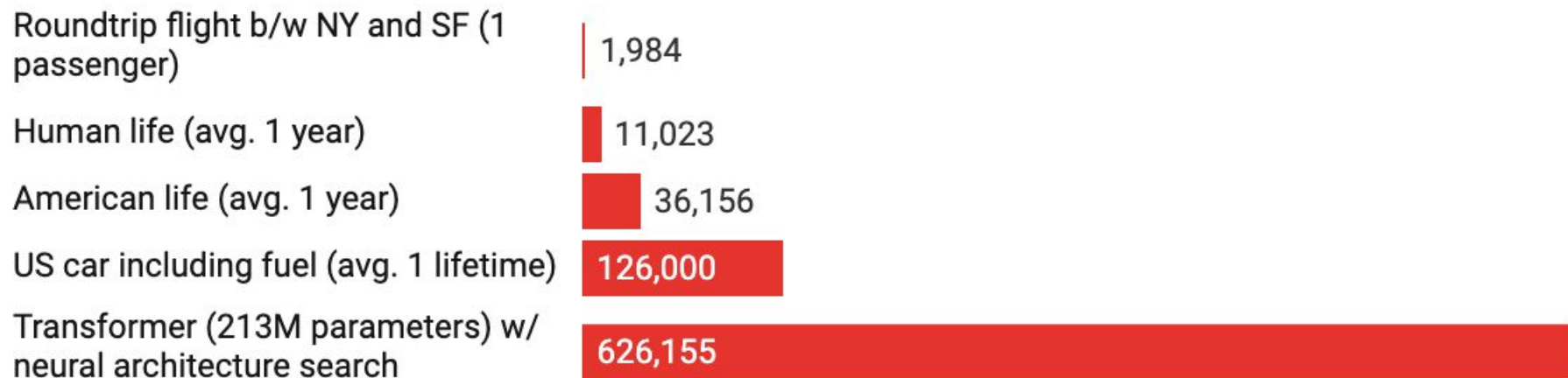


Chart: MIT Technology Review • Source: Strubell et al. • [Created with Datawrapper](#)

[Source](#)



*[Submitted on 6 Dec 2021]*

# JointLK: Joint Reasoning with Language Models and Knowledge Graphs for Commonsense Question Answering

Yueqing Sun, Qi Shi, Le Qi, Yu Zhang

[Source](#)

## KELM: Integrating Knowledge Graphs with Language Model Pre-training Corpora

Thursday, May 20, 2021

Posted by Siamak Shakeri, Staff Software Engineer and Oshin Agarwal, Research Intern, Google Research

[Reference](#)

# **Give your Text Representation Models some Love: the Case for Basque**

**Rodrigo Agerri<sup>1\*</sup>, Iñaki San Vicente<sup>2\*</sup>, Jon Ander Campos<sup>1\*</sup>, Ander Barrena<sup>1\*</sup>,  
Xabier Saralegi<sup>2</sup>, Aitor Soroa<sup>1</sup>, Eneko Agirre<sup>1</sup>**

**[IXAmBERT: Source](#)**

# Starting point

	A	B	C	D
1	SPANISH	BASQUE	DESCRIPTION	CATEGORY
2	Antiguo			Geography and places
3	Playa de La Concha	Kontxa Hondartza		
4	La Concha			Geography and places
5	Bataplán			
6	Bilbao	Bilbo		
7	Daba daba			
8	San Sebastián	Donostia		Geography and places
9	Eduardo Chillida	Eduardo Chillida		People and self
10	Egia	Egia		Geography and places
11	País Vasco	Euskadi		Geography and places
12	galerna	enbata		

Concept  
Sheet

Relation  
Sheet

CONCEPT A	RELATION	CONCEPT B
Peine del Viento	IS_LOCATED_IN / HAPPENS_IN	Ondarreta
Tamborrada	IS_LOCATED_IN / HAPPENS_IN	Donostia/San Sebastián
Gilda	TYPICAL_IN	Donostia/San Sebastián
Gilda	IS	pintxo
Gros	IS_PART_OF	Donostia/San Sebastián
pintxo	IS_PART_OF	pintxo pote
tortilla de patatas	IS	pintxo
Donostia/San Sebastián	IS_PART_OF	Euskadi

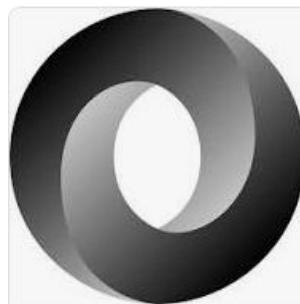
# Quantities

- Concepts:
  - Maximum of 50 various concepts related to San Sebastián or the Basque Country in general.
- Relations:
  - 5 distinct relations denoting basic entity relationships.
- Workspllit:
  - Each one have to annotate 10 concepts and its relations.
  - Code the graph

# Formats

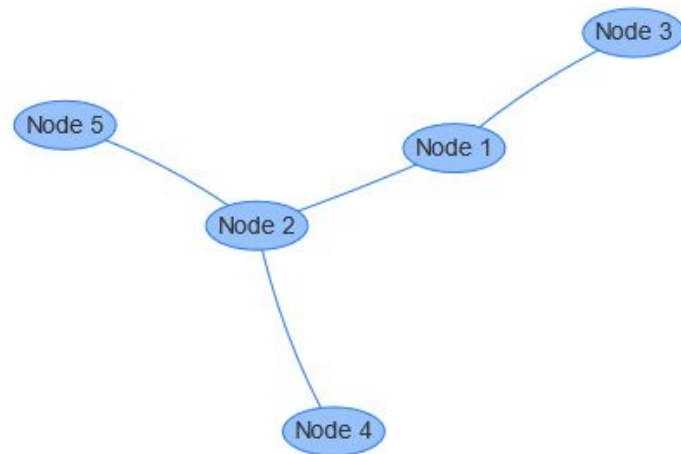


CSV



JSON

vis.js  
community edition \*



**And now...**

**Let's annotate!**

# Questions?

