

# Enhancing Entity Alignment Between Wikidata and ArtGraph using LLMs

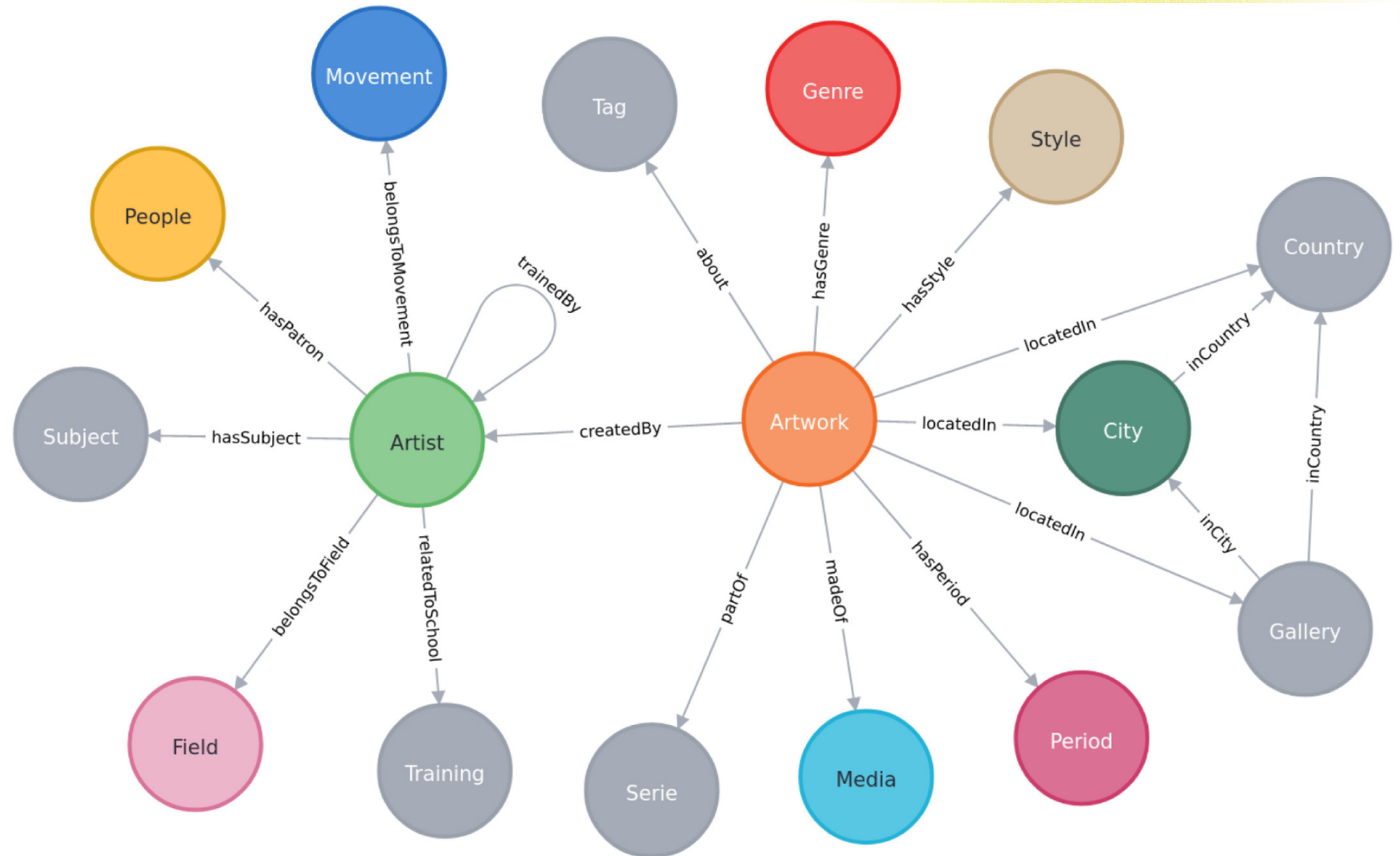
Anna Sofia Lippolis, Antonis Klironomos, Daniela F. Milon-Flores,  
Heng Zheng, Alexane Jouglar, Ebrahim Norouzi and Aidan Hogan





# Art Graph

- 135,038 resources
- Based on WikiArt





# Problem

# Analysis of data resources

- Discrepancy between the published KG and the ontology
- Missing information

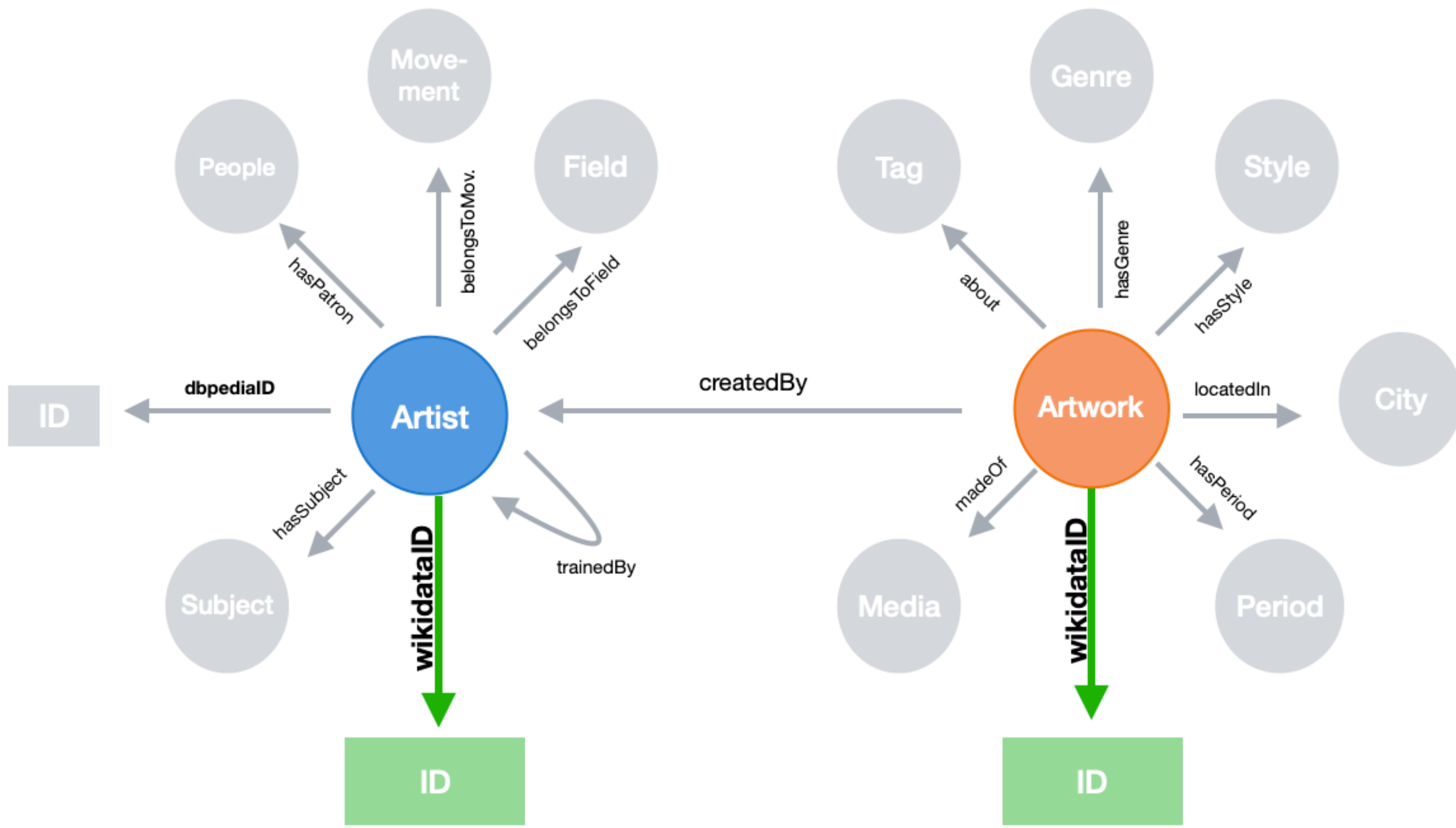
Several information gaps were discovered for artists: gender, birth and death date, school and movement the artist belongs to...

# 1,917

Missing values in belongsToField

Regarding artworks, similar gaps in information were identified, for example location of artworks, missing for 96,943 pieces.







# Methodology



# WES method (artworks)



Entity retrieval and linking  
to Wikidata for artworks

Aligning the artwork's title and  
creator with the corresponding  
info in Wikidata using SPARQL  
queries

Results

Over 14,000 artworks  
found matches



# WES method (artists)



Entity retrieval and linking  
to Wikidata for artists

Extracting Wikipedia URLs to  
pinpoint the precise artist page  
title corresponding on Wikidata

Spacy OpenTapioca on the  
remaining

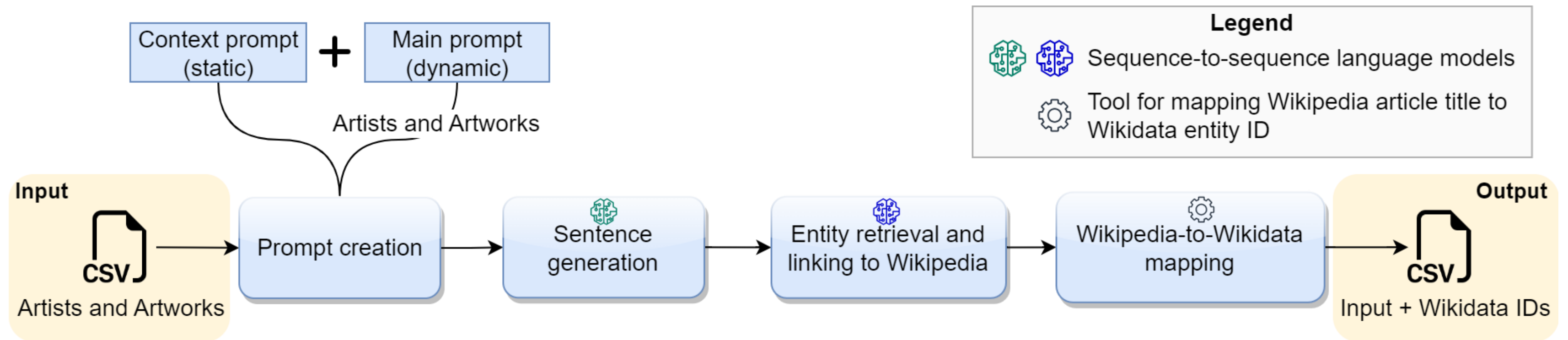
If no match was found, the  
method tried again using artists'  
names with Spacy OpenTapioca,  
good for NER

Results

423 of the 2,452 total  
artists were not linked:  
adding a pArtLink method

# pArtLink method (artists)

SWODCH ISWC 2023



pArtLink architecture diagram for artists



# pArtLink method (prompt creation)

“You are a helpful AI assistant for finding information about artists. Your answer must contain only one line. Your answer must follow this template: “**artist\_name: info\_about\_artist**”. Your answer must not include any other text than the answer itself. Write a short sentence about the artist {**artist**} who created {**artwork**}.”

# Results and Evaluation



<b>Entities</b>	<b>Methods</b>	<b>Total</b>	<b>Linked</b>	<b>Evaluated</b>	<b>Correctly linked</b>	<b>Incorrectly linked</b>
Artworks	WES	116 475	14 982	100	<b>99</b>	1
	pArtLink	-	-	-	-	-
Artists	WES	2452	<b>2029</b>	100	<b>100</b>	0
	pArtLink	700	332	332	252	76



# Evaluation agreement

We performed a manual evaluation to **measure the level of agreement** between reviewers.



A sample of 100 random artwork/artist entities was created and **evaluated by 5 different reviewers.**

## Sample

A	C	D	E
Entity ID	Creator Label	Input Text	correct or not
<a href="http://www.wikidata.org/entity/Q2203">http://www.wikidata.org/entity/Q2203</a>	Eleanor Fortescue-Brickdale	An Outcast	Y
<a href="http://www.wikidata.org/entity/Q1106">http://www.wikidata.org/entity/Q1106</a>	John Duncan	Phlegethon	Y
<a href="http://www.wikidata.org/entity/Q2044">http://www.wikidata.org/entity/Q2044</a>	Grace Cossington Smith	Quaker girl	Y
<a href="http://www.wikidata.org/entity/Q1026">http://www.wikidata.org/entity/Q1026</a>	Josefa de Óbidos	Cordeiro Pascal	Y
<a href="http://www.wikidata.org/entity/Q2018">http://www.wikidata.org/entity/Q2018</a>	Pierre Puvis de Chavannes	The Rest	Y
<a href="http://www.wikidata.org/entity/Q2965">http://www.wikidata.org/entity/Q2965</a>	Gabriel Metsu	The Poultry Seller	Y
<a href="http://www.wikidata.org/entity/Q3311">http://www.wikidata.org/entity/Q3311</a>	Hans Gude	Landskap Med Buskap Og Foss	Y
<a href="http://www.wikidata.org/entity/Q1120">http://www.wikidata.org/entity/Q1120</a>	Pietro Longhi	The Family Concert	Y
<a href="http://www.wikidata.org/entity/Q1734">http://www.wikidata.org/entity/Q1734</a>	Gabriel Metsu	Woman Eating	Y
<a href="http://www.wikidata.org/entity/Q8539">http://www.wikidata.org/entity/Q8539</a>	Eduardo Paolozzi	I Was a Rich Man's Plaything	Y
<a href="http://www.wikidata.org/entity/Q2044">http://www.wikidata.org/entity/Q2044</a>	Tom Roberts	Harper's Weekly	Y
<a href="http://www.wikidata.org/entity/Q1990">http://www.wikidata.org/entity/Q1990</a>	Claude Monet	The Bodmer Oak, Fontainebleau	Y
<a href="http://www.wikidata.org/entity/Q1021">http://www.wikidata.org/entity/Q1021</a>	Salvador Dali	Dionysus Spitting the Complete Image of Cad	Y
<a href="http://www.wikidata.org/entity/Q4324">http://www.wikidata.org/entity/Q4324</a>	Ivan Shishkin	Mountain path. Crimea	Y
<a href="http://www.wikidata.org/entity/Q3944">http://www.wikidata.org/entity/Q3944</a>	Agnolo Bronzino	The Panciatici Holy Family	Y
<a href="http://www.wikidata.org/entity/Q2393">http://www.wikidata.org/entity/Q2393</a>	Joan Brull i Vinyoles	Paisatge Amb Figura	Y
<a href="http://www.wikidata.org/entity/Q2495">http://www.wikidata.org/entity/Q2495</a>	Ernst Ludwig Kirchner	Self-Portrait as a Soldier	Y
<a href="http://www.wikidata.org/entity/Q6460">http://www.wikidata.org/entity/Q6460</a>	Martin Schongauer	Bust of a Monk Assisting at Communion	Y
<a href="http://www.wikidata.org/entity/Q2053">http://www.wikidata.org/entity/Q2053</a>	Camille Pissarro	Poultry Market at Gisors	Y
<a href="http://www.wikidata.org/entity/Q5082">http://www.wikidata.org/entity/Q5082</a>	Abraham Storck	The Four Days' Battle	Y



# Evaluation agreement

Three scores were used for the evaluation: **Fleiss Kappa, Percentage of agreement, and Rodolph Kappa.**

	Fleiss Kappa	Percentage of Agreement	Rodolph Kappa
Artwork	0.018	95%	0.9
Artist	0.88	95%	0.89



## Discussion :

- Previous criteria definition is necessary between the reviewers.
- Example of Disagreement issues: The language of the artwork title in the sample (e.g., **La Pergola**) diverged from that of the Wikidata page (e.g., **After lunch**).
- Fleiss' Kappa measure is affected by label imbalance.
- Complementary measures are needed to get a more complete picture of matching.
- For both artist and artwork, entities were usually correctly associated with their Wikidata ID.



# Limitations and future work



## Limitations

1. False positives
2. Quality of the sentences generated by the LLM
3. Measurement of agreement between reviewers

## Future work

1. Potential collaboration with Wikidata
2. Test with other CH-related knowledge bases
3. Expand the method to languages other than English





# Resources

## Contact us

[annasofia.lippolis2@unibo.it](mailto:annasofia.lippolis2@unibo.it)

[antonios.klironomos@students.uni-mannheim.de](mailto:antonios.klironomos@students.uni-mannheim.de)

[daniela.milon-flores@univ-grenoble-alpes.fr](mailto:daniela.milon-flores@univ-grenoble-alpes.fr)

[alexane.jouglar@unamur.be](mailto:alexane.jouglar@unamur.be)

[ebrahim.norouzi@fiz-karlsruhe.de](mailto:ebrahim.norouzi@fiz-karlsruhe.de)

