

**Polifonia** <https://polifonia-project.eu/>

# A digital harmoniser for musical heritage knowledge

---

**Prof. Valentina Presutti**, Università di Bologna

This presentation includes contributions by Jacopo de Berardinis, Marilena Dauino, Albert Meroño Peñuela, Andrea Poltronieri, Rocco Tripodi  
The work described in this presentation is contributed by the Polifonia Consortium members



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement GA101004746.  
The communication reflects only the author's view and the Research Executive Agency is not responsible for any use that may be made of the information it contains.







# Polifonia

follow and discuss this presentation on twitter:

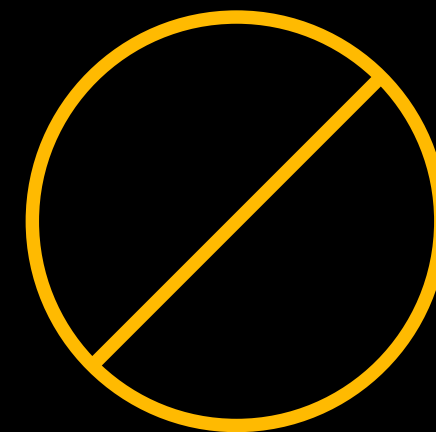
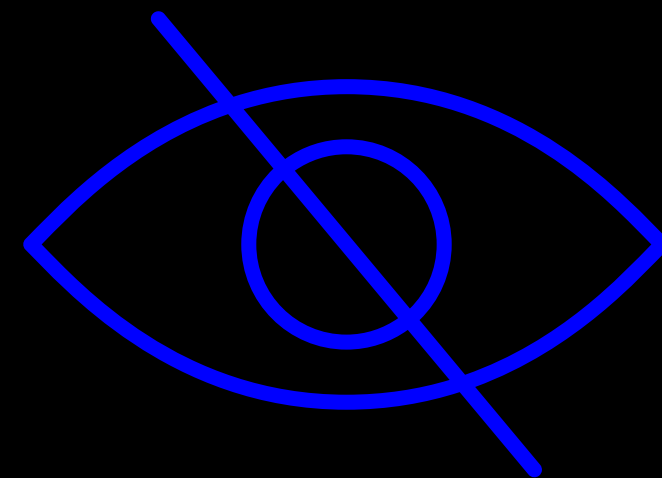
→ [#PolifoniaH2020](https://twitter.com/PolifoniaH2020)

---

**Prof. Valentina Presutti**, Università di Bologna  
(on behalf of the Polifonia Consortium)



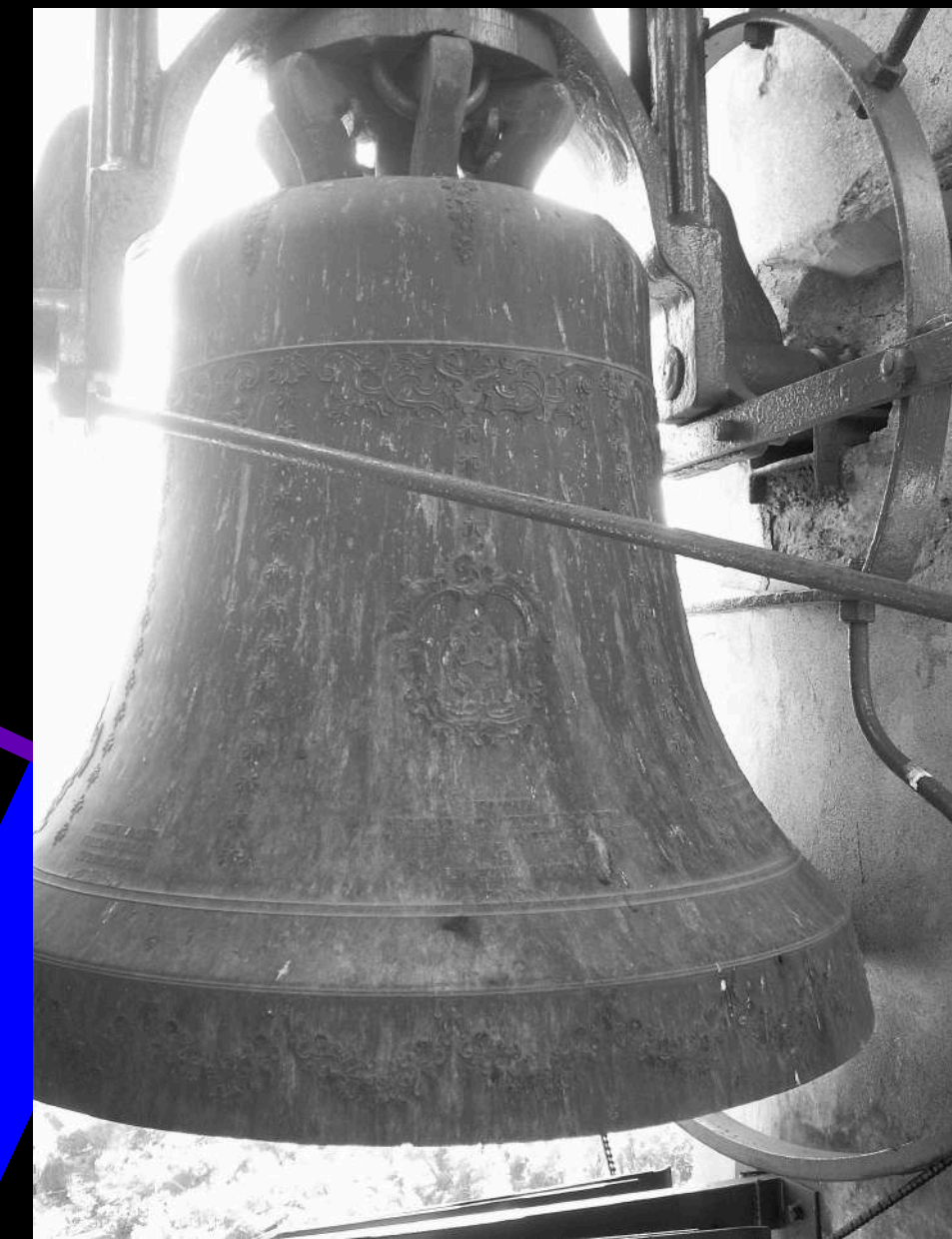
 Today I will talk about music and musical heritage 






# Did you know that...

San Pietro di Starla, Carasco  
(Genova, Liguria, Italy)



each territory has its own peculiar set of bell sounds and practices?



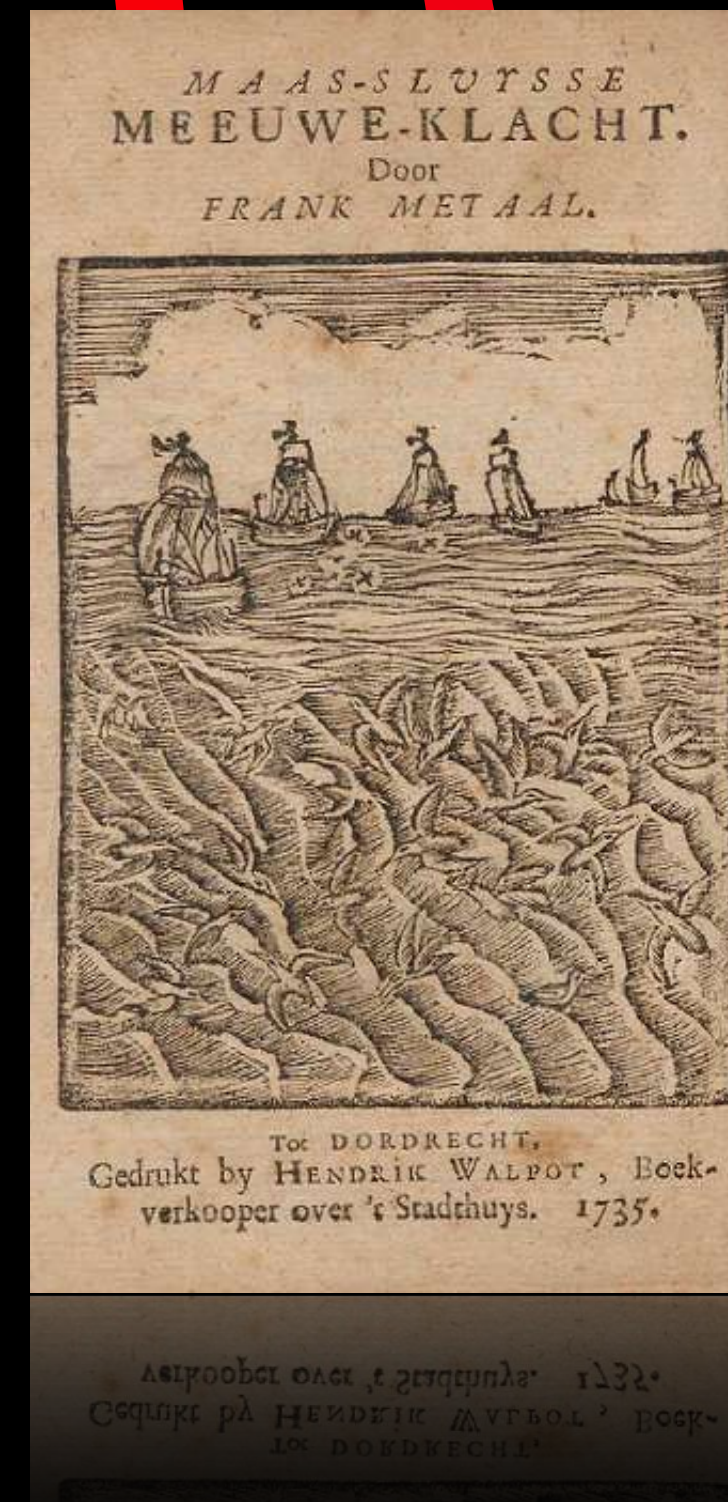


Challenge: preserving orally transmitted musical  
heritage knowledge

# Did you know that...



Jean Baptiste Lully (1632-1687): "Belle Iris"  
"Ballet de l'Impatience" (1661, LWV 14)



Dankbaarheid van Hubrecht Arentsz  
Same melody, with Dutch text about someone  
who was rescued from a shipwreck

## Other Dutch titles for the same melody

Labellerij  
Wert dan noit u straf gemoet  
t' Is goet Capucien te sijn  
Aengenaeme wildernis  
Sinte Niclaes Heyligman  
Ick segghe adieu mijn Afgodin  
Wel wat hoor ick voor gerucht  
Blaest nu op mijn brant Trompe  
Ach mijn alderliefste hert  
De haagsche kermis  
Ma Duchesse  
Van de floerse tipjes  
Van stroeyken en kooltjen vyer  
Hoog om hoog dog op myn hert  
Zielverrukkende Amaril  
'k Zag lest enen bedelaar

## many Dutch traditional songs rely on French opera music





## Prof. Louis Grijp (1955 - 2016)

Traditional Dutch music from the late sixteenth century was  
influenced by French opera from the same period



# Franco-Dutch war 1672 to 1678







How many influences or similarities like this one  
are hidden in European music?



Did you know that...



Franz Schubert and The Beatles have something in common





Challenge: to enable large scale cross-dataset analysis (distant reading) and hypotheses formulation in social science and humanities

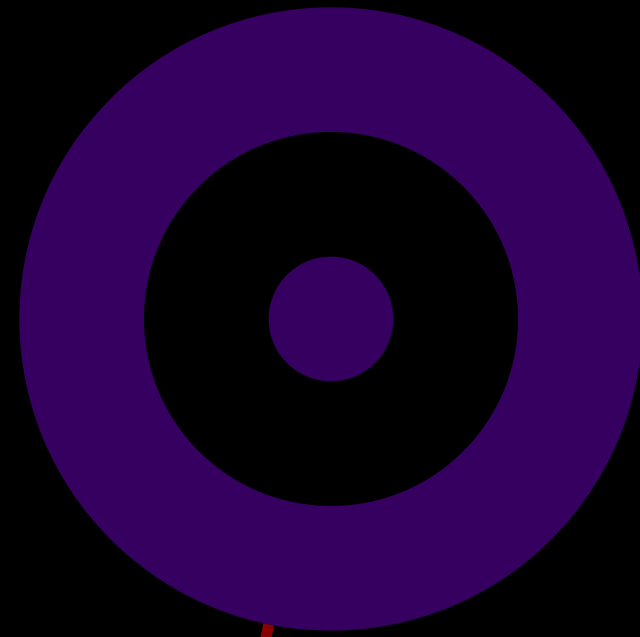


Did you know that...



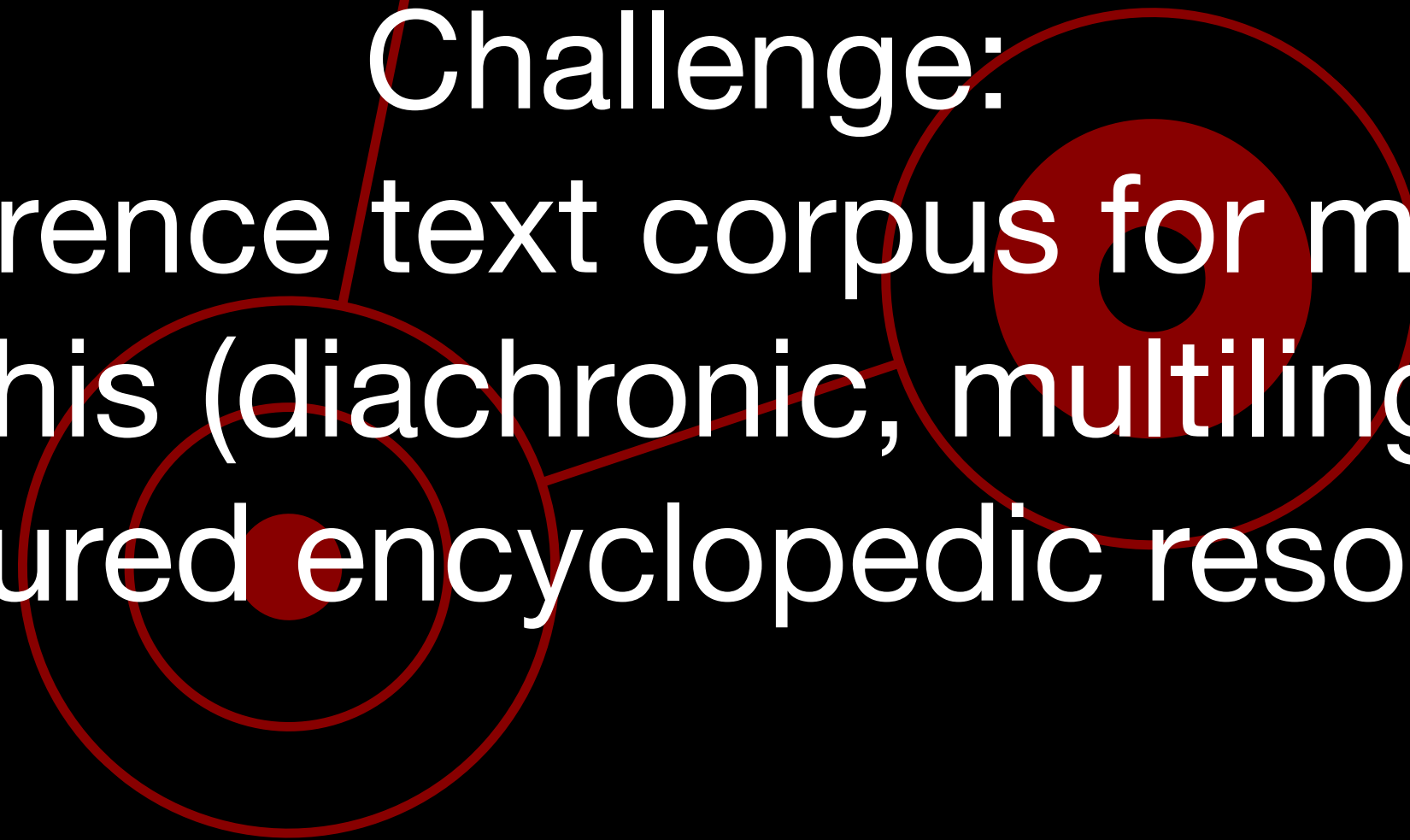
the Municipality of Bologna bestowed honorary citizenship on Wagner in 1872?






Challenge:

to create a reference text corpus for musical heritage  
to extract facts from this (diachronic, multilingual) resource to feed  
structured encyclopedic resources







Challenge: to interconnect musical heritage knowledge from heterogenous sources and create a global distributed knowledge graph publicly accessible for both humans and machines





## Project info

 **Domain:** Musical Heritage & Artificial Intelligence

 **Project lunch date:** January 1st, 2021

 **Duration:** 40 months

 **Partners in the consortium:** 10

 **Total grant:** ~3M €

## Partners



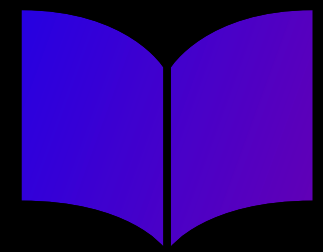
<https://github.com/polifonia-project>

Call topic

DT-TRANSFORMATIONS-12-2018-2020:

Curation of digital assets and advanced digitization Research and Innovation Action





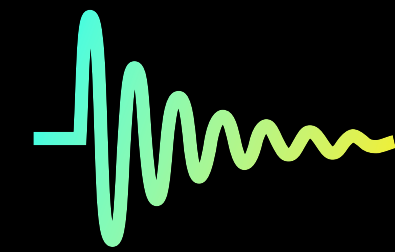
catalogues



sound



text documents



oral trasmission

---

We know there is tremendous amount of knowledge of this type and more, hidden in catalogues, texts, literature, letters, tunes, scores.

Making it explicit and interconnected for enabling processing by intelligent agents

A global interconnected knowledge base accessible everyone:  
scholars, music nerds, producers, artists, cultural institutions, creative industry, students









Type: Opera  
Title: “Tosca”  
Composer: Giacomo Puccini  
Genre: Melodramma  
First performance at: Teatro dell’Opera di Roma  
Language: Italian



Type: Representation  
Representation of: Tosca  
Date: January 14, 1900  
Venue: Teatro Costanzi  
City: Rome

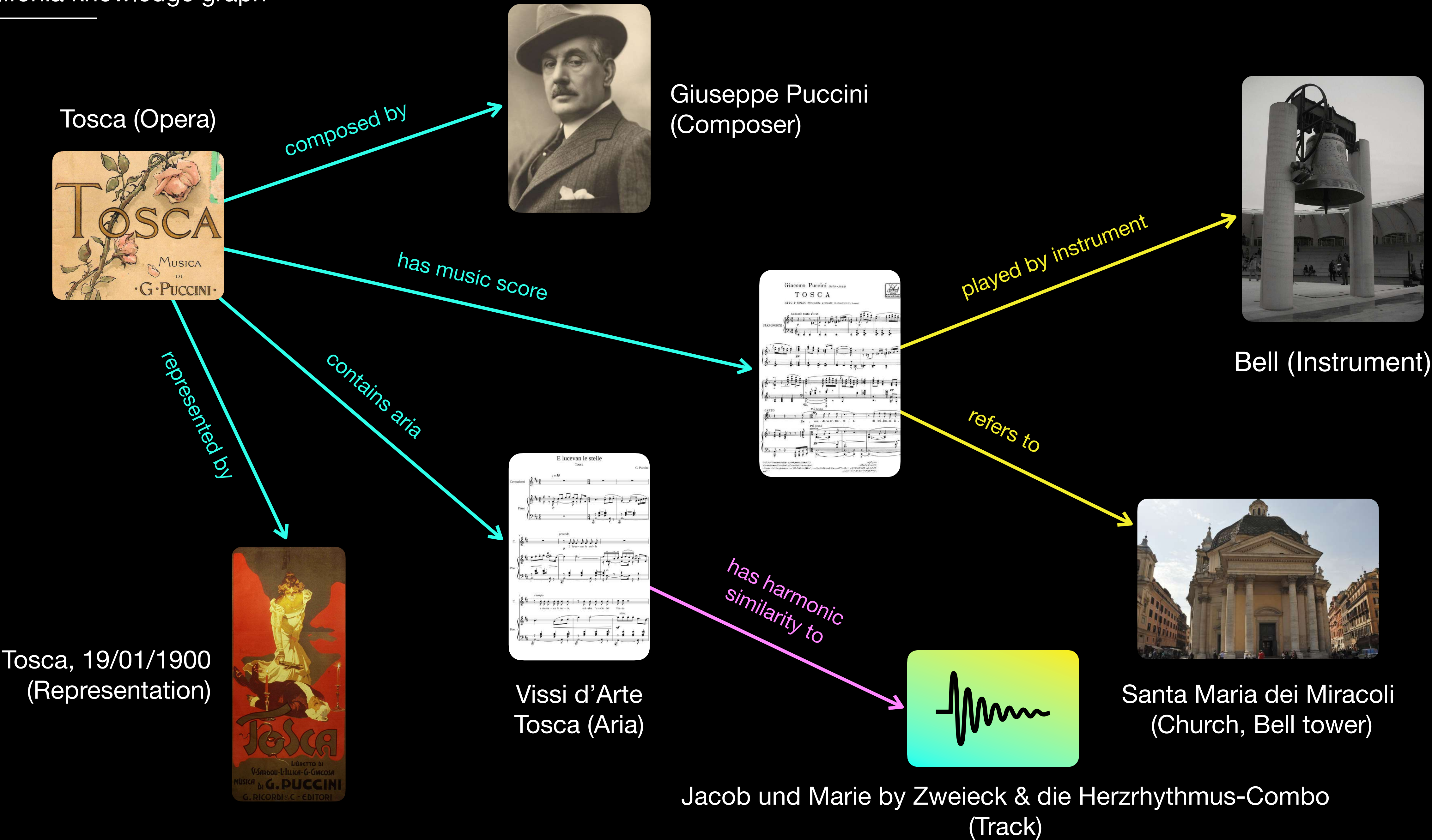
Type: Instrument, Bell  
Name: “Campana dei Caduti”  
Year: 1924  
Location: Rovereto, Italy  
Design by: Stefano Zuech  
Material: Bronze



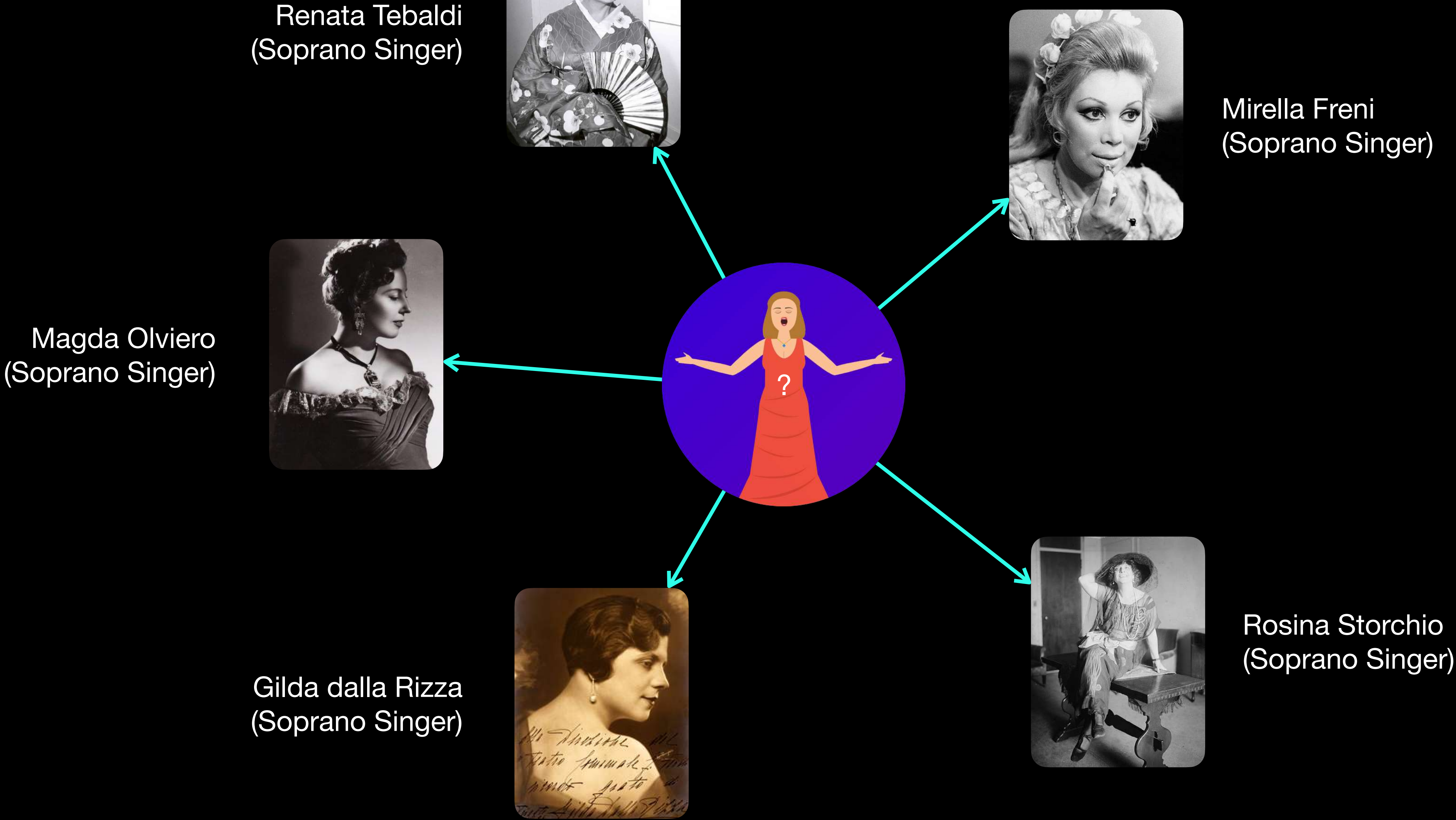
Type: Composer  
Name: “Giacomo Puccini”  
BirthDate: December 22, 1858  
DeathDate: November 29, 1924  
Place of Birth: Lucca  
Position: Senator of the Kindgdom of Italy









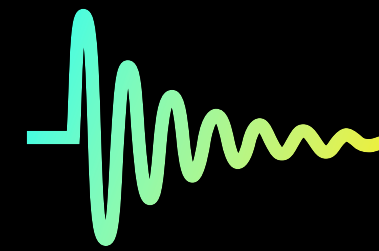
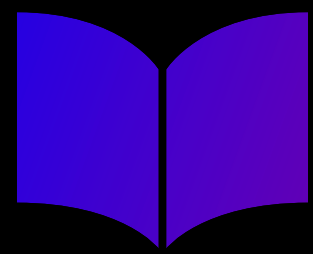




# Building the Polifonia Knowledge Graph



## Producing the Polifonia Knowledge Graph



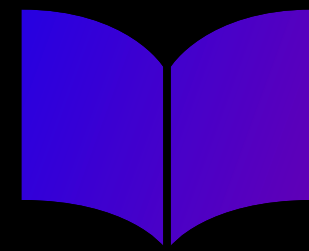


## Producing the Polifonia Knowledge Graph

---

Analysis, modelling, mapping, transformation

Semantic Web technologies



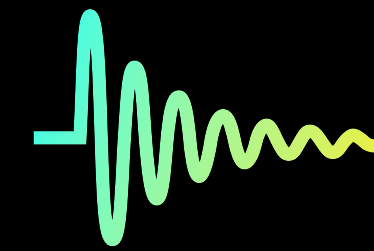
catalogues



sound



text documents



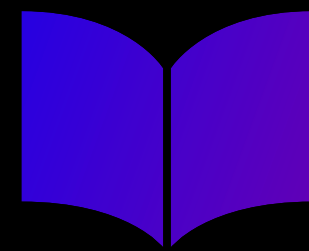
oral trasmission



## Producing the Polifonia Knowledge Graph

Analysis, modelling, mapping, transformation

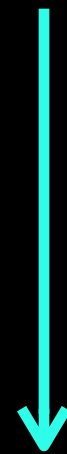
Semantic Web technologies



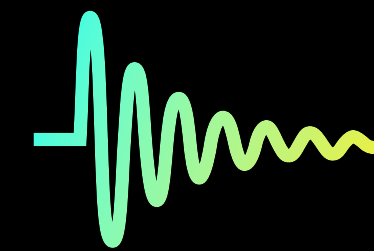
catalogues



sound



text documents

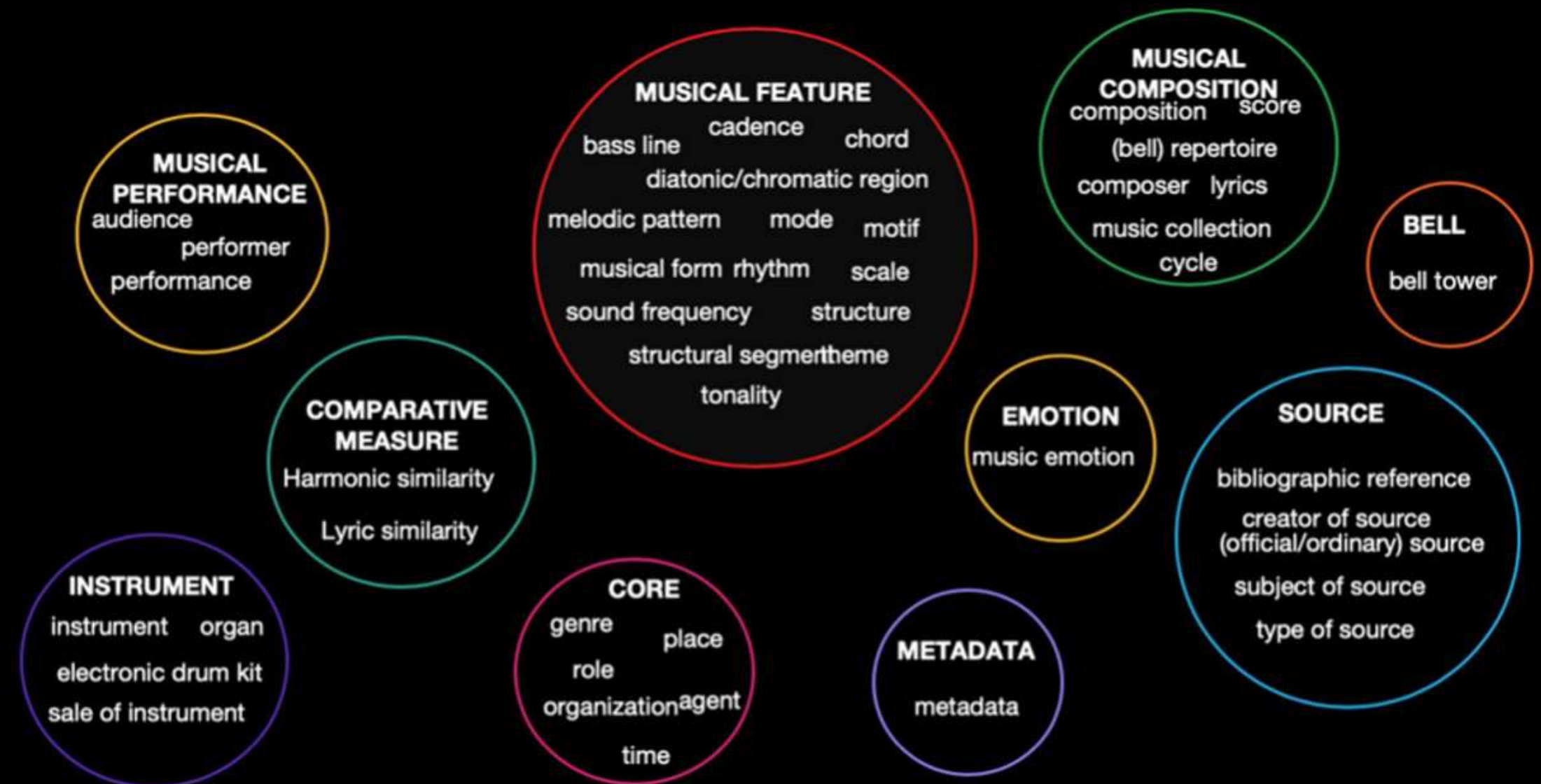
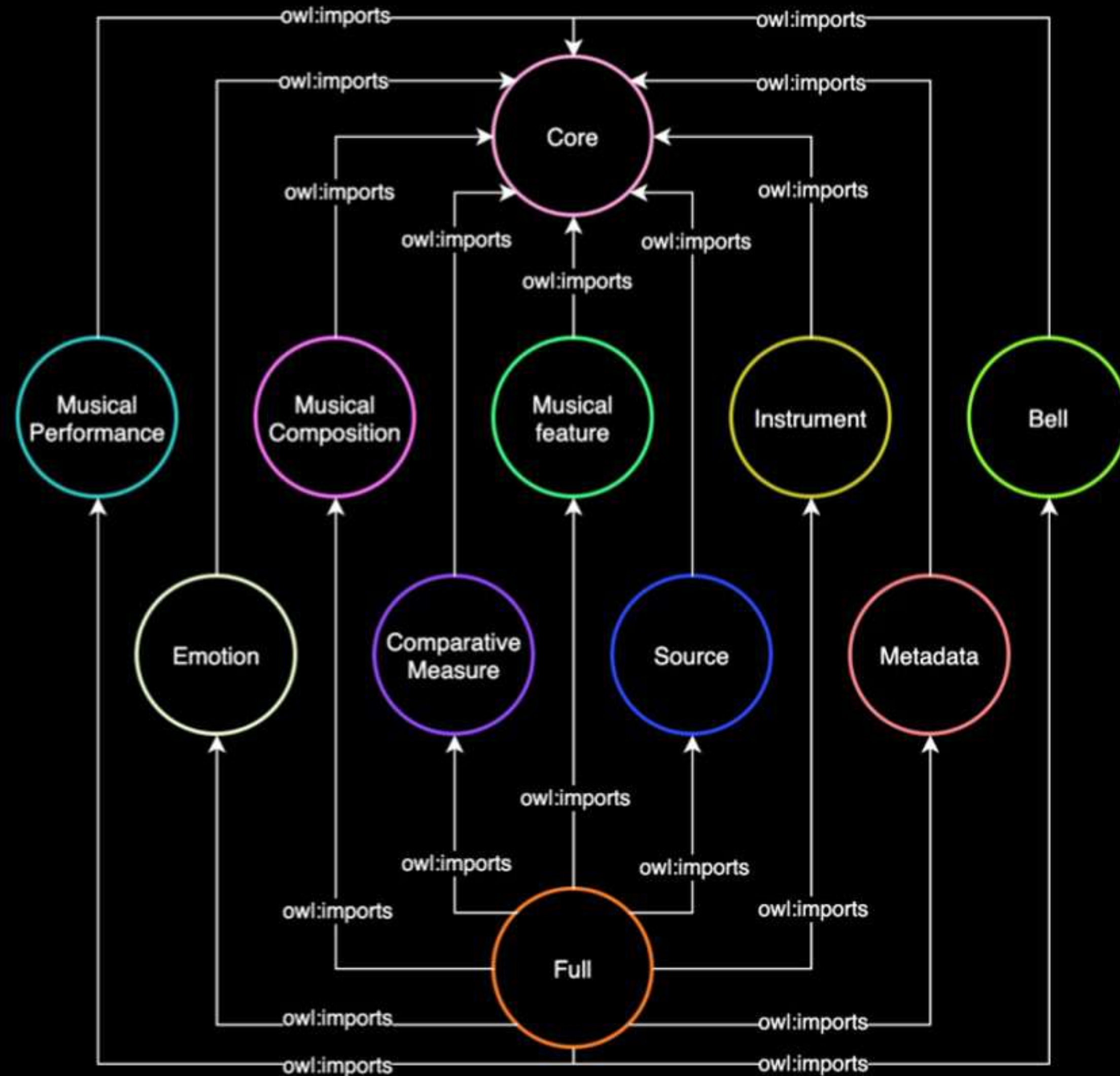


oral trasmission





Currently containing:  
**101** classes, **110** object properties  
**63** additional assertions  
(subclasses, subproperties, disjointness)  
**31** inter-module links  
and growing...



All modules browsable from  
<https://github.com/polifonia-project/ontology-network>



Challenge: preserving orally transmitted knowledge

# Bells repository and app

Filter Encounters:

Architectural complex type

COMPLEX:

☐ Church

☐ Municipality

Instrument definition

INSTRUMENT:

☐ Single Bell

☐ Bells Concert

Sound modality

SOUND:

☐ Manual

☐ Electrified

☐ Both

Bell's material

MATERIAL:

☐ Bronze

☐ Steel

Execution Technique

TECHNIQUE:

☐ A distesa

☐ A bicchiere

☐ A camp. fisse

Mounting system

MOUNTING:

☐ Batt. cadente

☐ Camp. fisse

☐ Slancio

Timespan of the construction:

1900

2015

Move on the map and discover the bells

Filter Encounters:

Architectural complex type

COMPLEX:

☐ Church

☐ Municipality

Instrument definition

INSTRUMENT:

☐ Single Bell

☒ Bells Concert

Sound modality

SOUND:

☐ Manual

☐ Electrified

☒ Both

Bell's material

MATERIAL:

☒ Bronze

☐ Steel

Execution Technique

TECHNIQUE:

☐ A distesa

☐ A bicchiere

☐ A camp. fisse

Mounting system

MOUNTING:

☒ Batt. cadente

☐ Camp. fisse

☐ Slancio

Timespan of the construction:

1900

2015

Bell Tower - San Pietro di Sturla

ARCHITECTURE:

Architectural complex type: *Church*

Complex name: *San Pietro di Sturla*

Place: *Carasco (GE)*

Bishopric: *Chiavari*

BELL TOWER:

Place of production: *Fidenza*

Date of production: *1929*

Foundry: *Fonderia Regolo Capanni*

Number of bells: *5*

Material: *Bronze*

MORE DATA ABOUT:

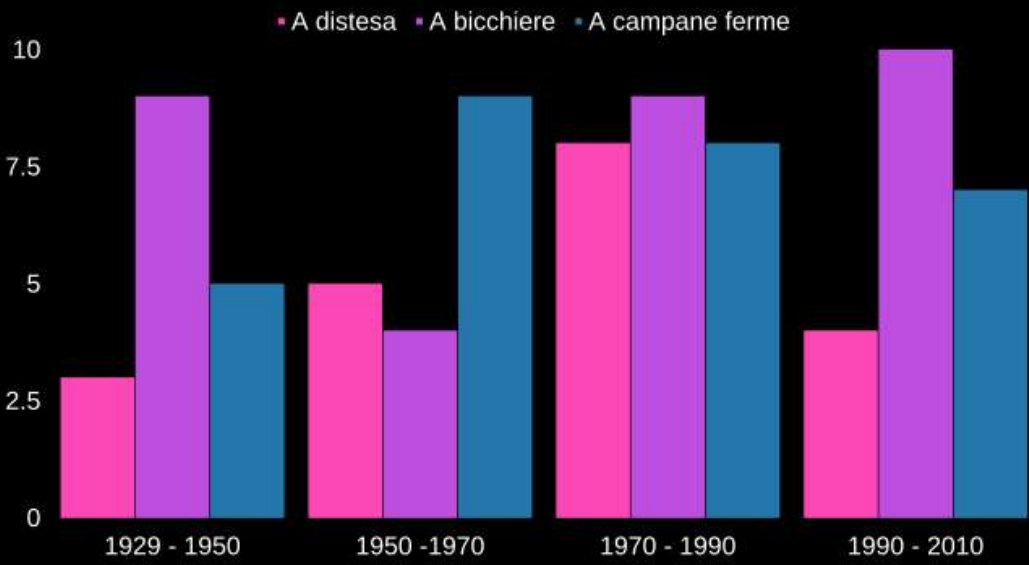
REPERTOIRE

TECHNIQUES

OCCASION

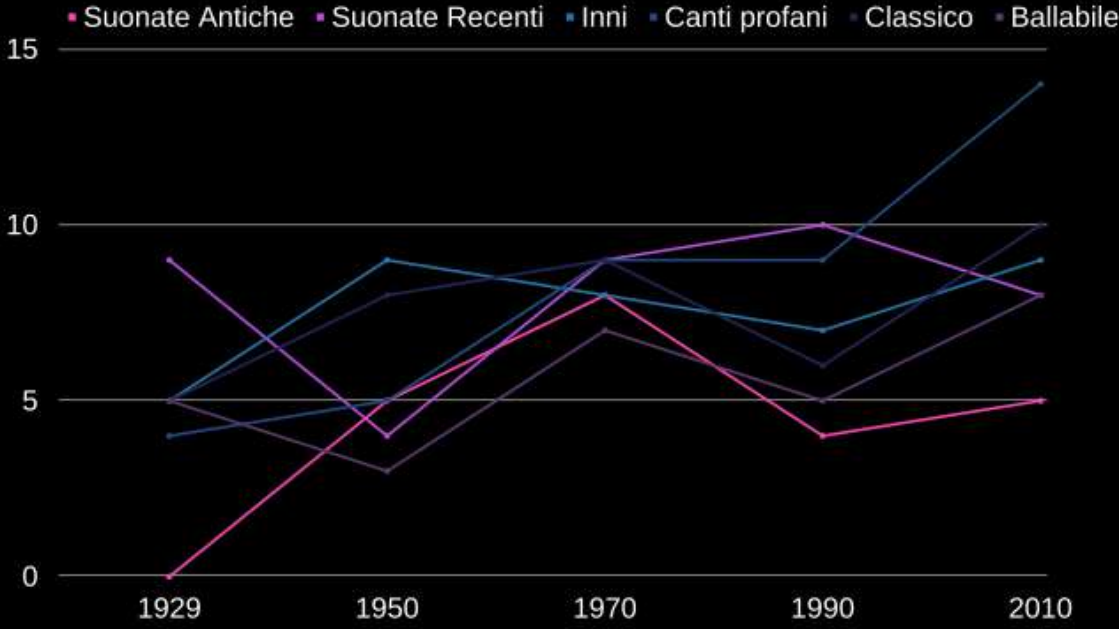
## TECHNIQUES

Execution techniques depending on the period:



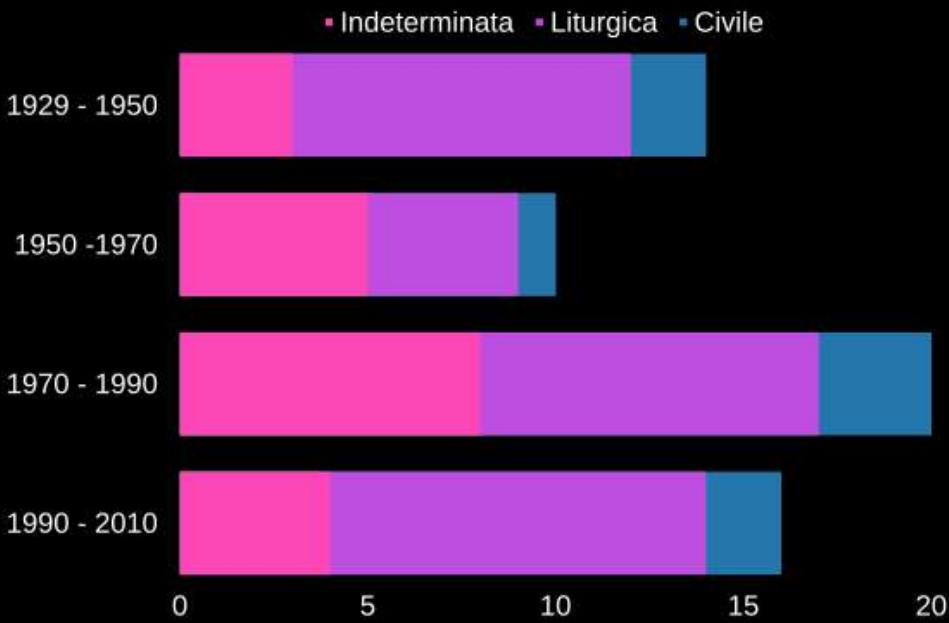
## REPERTOIRE

Evolution of the repertoire performed over time:



## OCCASION

Typical performance occasions depending on the period:





# ChoCo: Chord corpus

Partition	Type	Notation	Original format	Annotations	Genres	Ref
Isophonics	A	Harte	JAMS	300	pop, rock	[26]
JAAH	S	Harte	JSON	113	jazz	[15]
Schubert-Winterreise	A, S	Harte	csv	25 (S), 25*9 (A)	classical	[39]
Billboard	A	Harte	LAB, txt	890 (740)	pop	[5]
Chordify	A	Harte	JAMS	50*4	pop	[25]
Robbie Williams	A	Harte	LAB, txt	61	pop	[14]
The Real Book	S	Harte	LAB	2486	jazz	[27]
Uspop 2002	A	Harte	LAB	195	pop	[4]
RWC-Pop	A	Harte	LAB	100	pop	[17]
Weimar Jazz Database	S	Leadsheet	SQL	456	jazz	[30]
Wikifonia	S	Leadsheet	mxl	6500+	various	-
iReal Pro	S	Leadsheet	iReal	2000+	various	-
Band-in-a-Box	S	Leadsheet	mgu, sku	5000+	various	[10]
When in Rome	S	Roman	RomanText	450	classical	[29]
Rock Corpus	S	Roman	har	200	rock	[8]
Mozart Piano Sonata	S	Roman	DCMLab	54 (18)	classical	[21]
Jazz Corpus	S	Hybrid	txt	76	jazz	[18]
Nottingham	S	ABC	ABC	1000+	folk	[1]

**Table 1.** Overview of the 18 chord datasets currently included in ChoCo. Letters “A” and “S” are used to denote *audio* and *symbolic* annotations, respectively.



Daga, E., Asprino, L., Mulholland, P., & Gangemi, A. (2021). Facade-X: an opinionated approach to SPARQL anything. *Studies on the Semantic Web*, 53, 58–73. <https://github.com/SPARQL-Anything/sparql.anything>



J, de Berardinis, A. Meroño, A. Poltronieri, V. Presutti

Integration of eighteen datasets of chord annotations  
from scores and audio files





### Progression A

Crazy Little Thing Called Love

D, G, C, G, D, G, C, G, D, Bb, C, D, G, C, G,  
D, . . . D, Bb, C, D, Bb, C, D, Bb, C, D, Bb, C, D



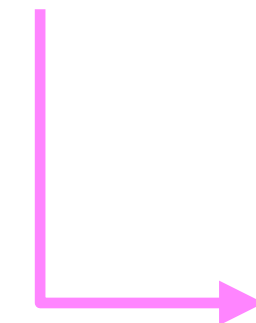
C, F, Bb, F, C, F, Bb, F, C, Ab, Bb, C, F, Bb, F, . . .  
Ab, Bb, C, Ab, Bb, C, Ab, Bb, C, Ab, Bb, C



106, 227, 96, 227, 106, 227, 96, 227, 106, 145, . . .  
145, 96, 106, 145, 96, 106, 145, 96, 106



(227, 106, 227, 96, 227, 106, 145, 96, 106, 227),  
(10, 106, 227, 96, 227) . . . (227, 106, 145, 96)



### Progression B

P.S. I Love You

G, C#:7, D, G, C#:7, D, G, C#:7, D, A, D, E:min,  
D, . . . D, A, B:min, A, Bb, C, D, Bb, C, D, Bb, C, D



F, Cb:7, C, F, Cb:7, C, F, Cb:7, C, G, C, D:min, . . .  
G, A:min, G, Ab, Bb, C, Ab, Bb, C, Ab, Bb, C



227, 172, 106, 227, 172, 106, 227, 172, 106, 10, . . .  
10, 145, 96, 106, 145, 96, 106, 145, 96, 106



(238, 106, 10, 131, 10, 145, 96, 106, 227, 106),  
. . . (145, 96, 106, 227, 106, 227, 106, 227)



## Normalisation

## Decomposition & Encoding

## Pattern EXT

### Harmonic similarity

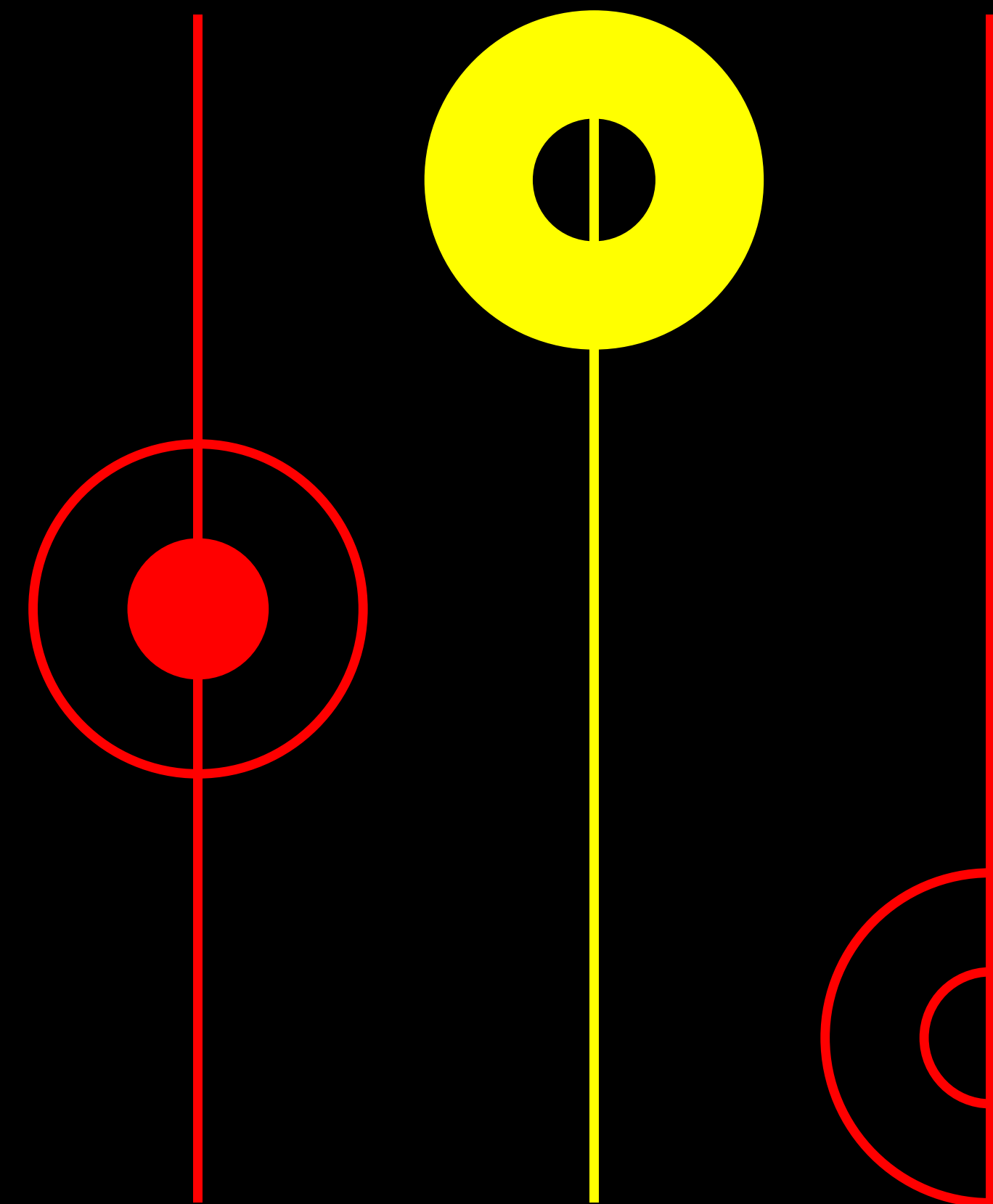
Similarity score: **0.23**  
(min 0 – max 1)

Longest pattern in common  
**Ab, Bb, C, Ab, Bb, C**



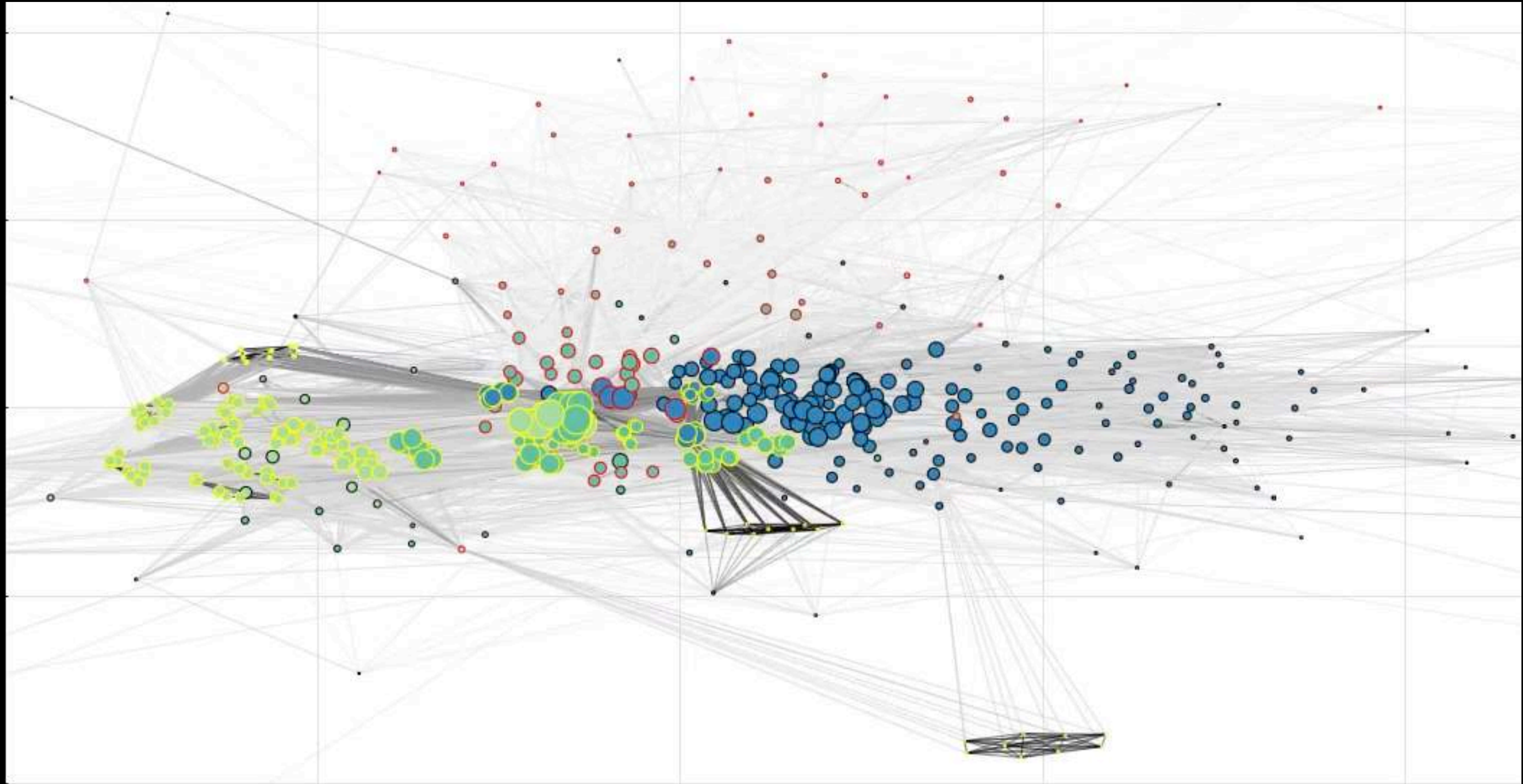
# Harmonic similarity graph

<https://polifonia-project.github.io/musilar-preview/>





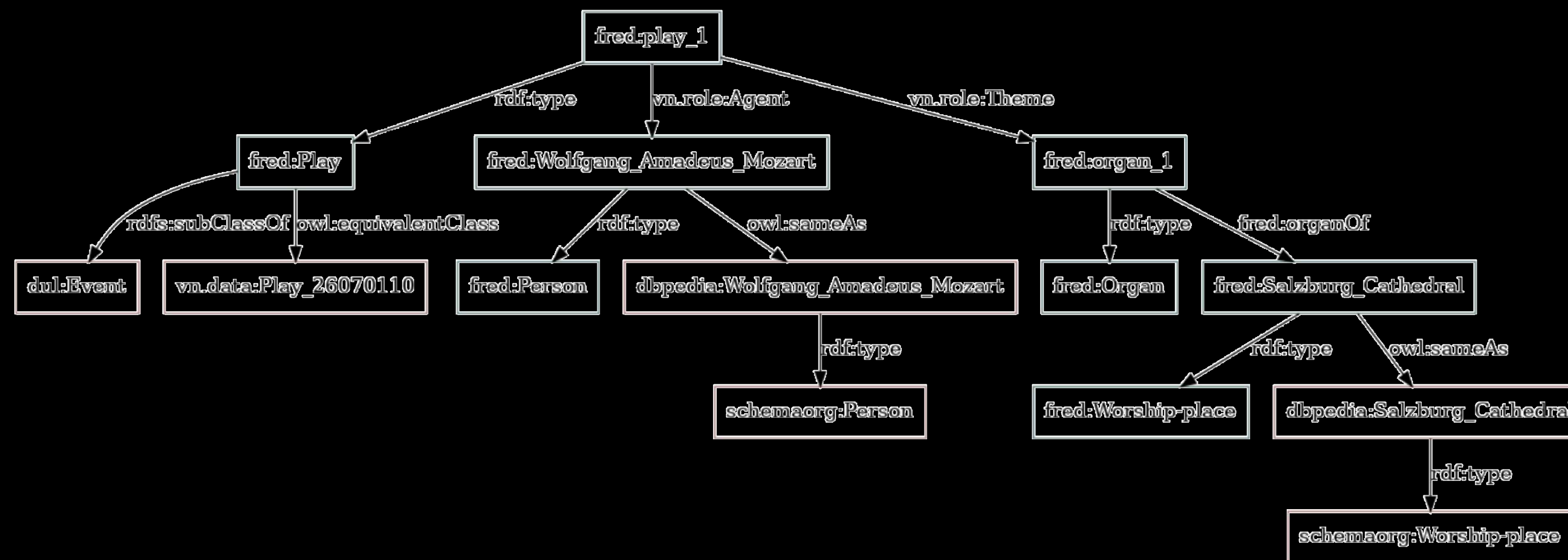
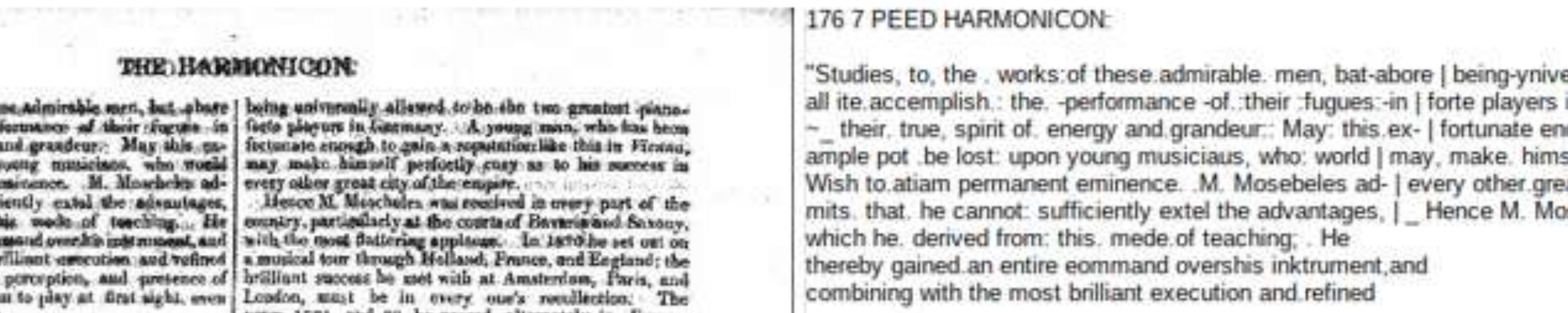
The harmonic similarity graph





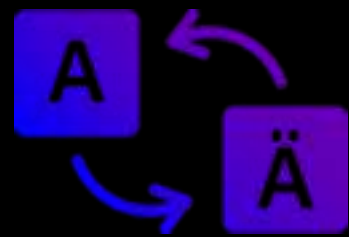
Challenge: to create a musical heritage text corpus and extract knowledge from it

# Knowledge graph extraction from text





# Polifonia text corpus and music lexicon



six languages: Dutch, English, French, German, Italian and Spanish



document types: encyclopedic articles, periodicals articles and books



from the XVIII century to nowadays

Language	Encyclopedic articles	Periodical issues	Books
Dutch	36.609	2.314	47.670
English	250.413	464	25.000
French	65.970	483	40.961
German	53.986	1.203	24.076
Italian	77.98	2.119	1.735
Spanish	57.891	924	15.894

<https://github.com/polifonia-project/Polifonia-Corpus>

<https://github.com/polifonia-project/textual-corpus-population>



An example

*Wolfgang Amadeus Mozart played the organ of the Salzburg Cathedral.*

Tokenization	POS	WSD	NER	Entity Linking
1. Wolfgang	PROPN	-	PERSON	<a href="https://en.wikipedia.org/wiki/Wolfgang_Amadeus_Mozart">https://en.wikipedia.org/wiki/Wolfgang_Amadeus_Mozart</a>
2. Amadeus	PROPN	-	PERSON	
3. Mozart	PROPN	-	PERSON	
4. played	VERB	play%2:36:01::	-	-
5. the	DET	-	-	-
6. organ	NOUN	organ%1:06:00::	-	-
7. of	ADP	-	-	-
8. the	DET	-	-	-
9. Salzburg	PROPN	-	LOC	<a href="https://en.wikipedia.org/wiki/Salzburg_Cathedral">https://en.wikipedia.org/wiki/Salzburg_Cathedral</a>
10. Cathedral	NOUN	cathedral%1:06:01::	LOC	





## From text to knowledge graphs

### Abstract Meaning Representation (AMR)

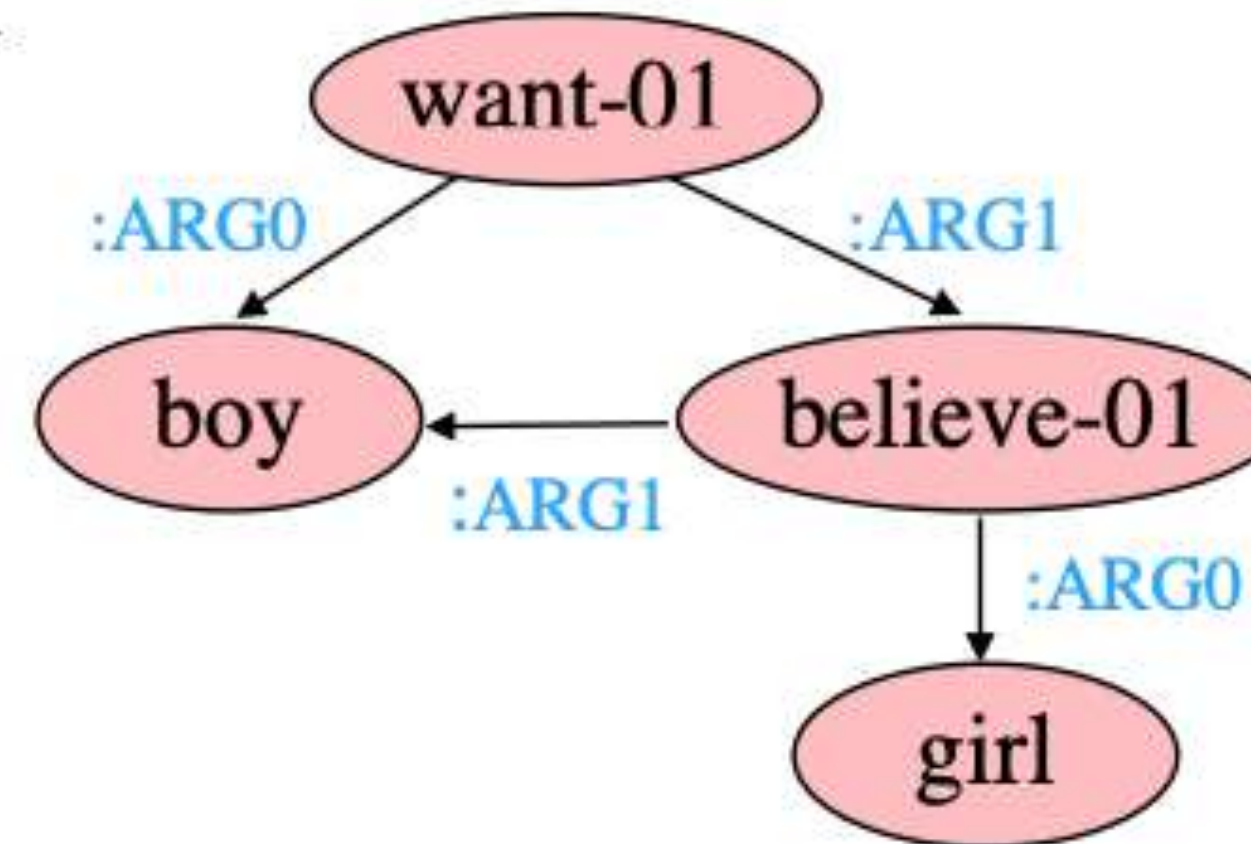
#### English

*The boy wants the girl to believe him.*

*The boy's desire is for the girl to believe him.*

*The boy wants to be believed by the girl.*

#### AMR



#### German, Spanish, Italian, and Chinese

*Der Junge möchte, dass das Mädchen ihm glaubt.*

*El chico quiere que la chica le crea.*

*Il ragazzo vuole che la ragazza gli creda.*

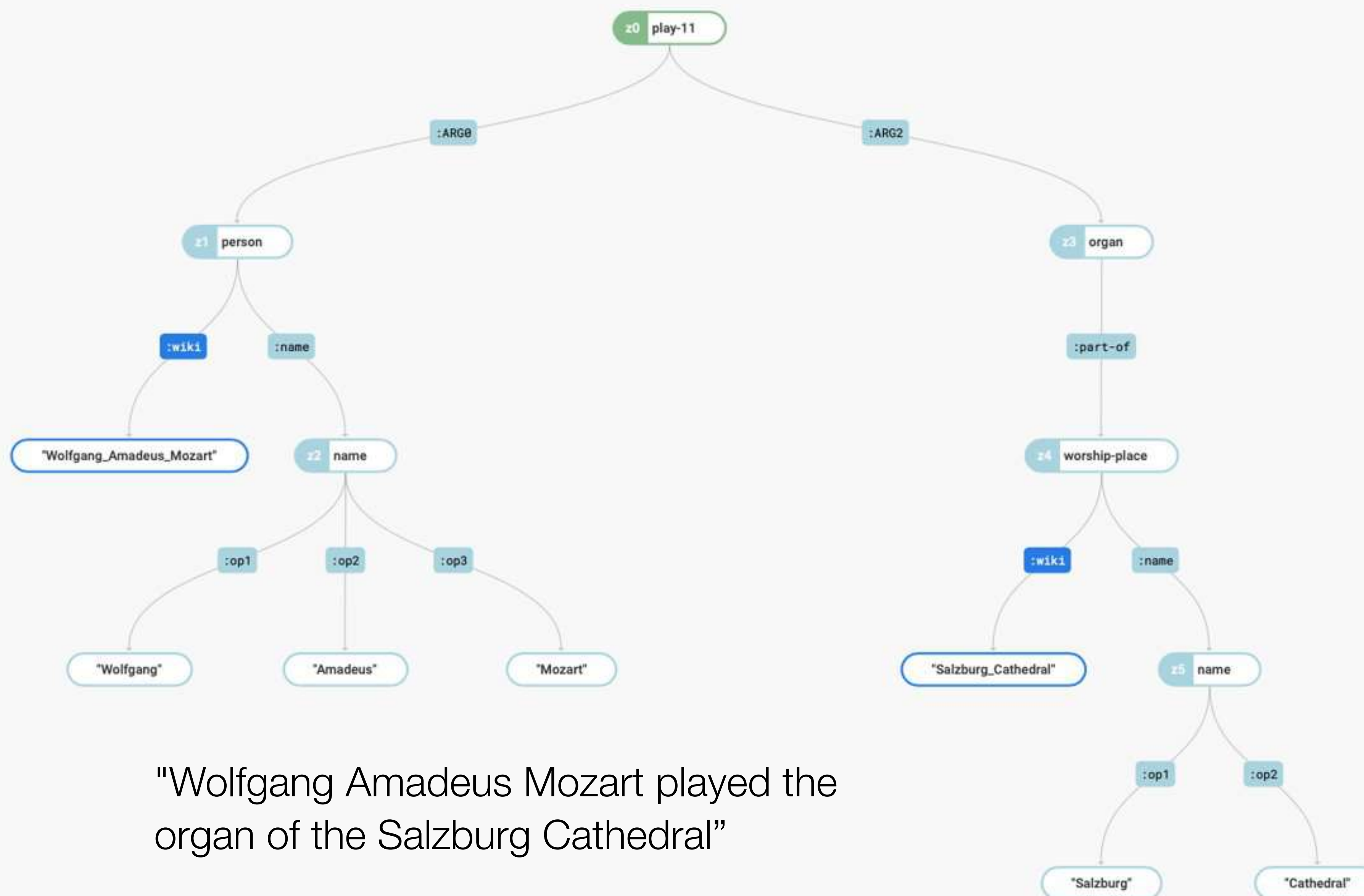
男孩要女孩相信他。

Cai et. al, 2021





# Text2AMR

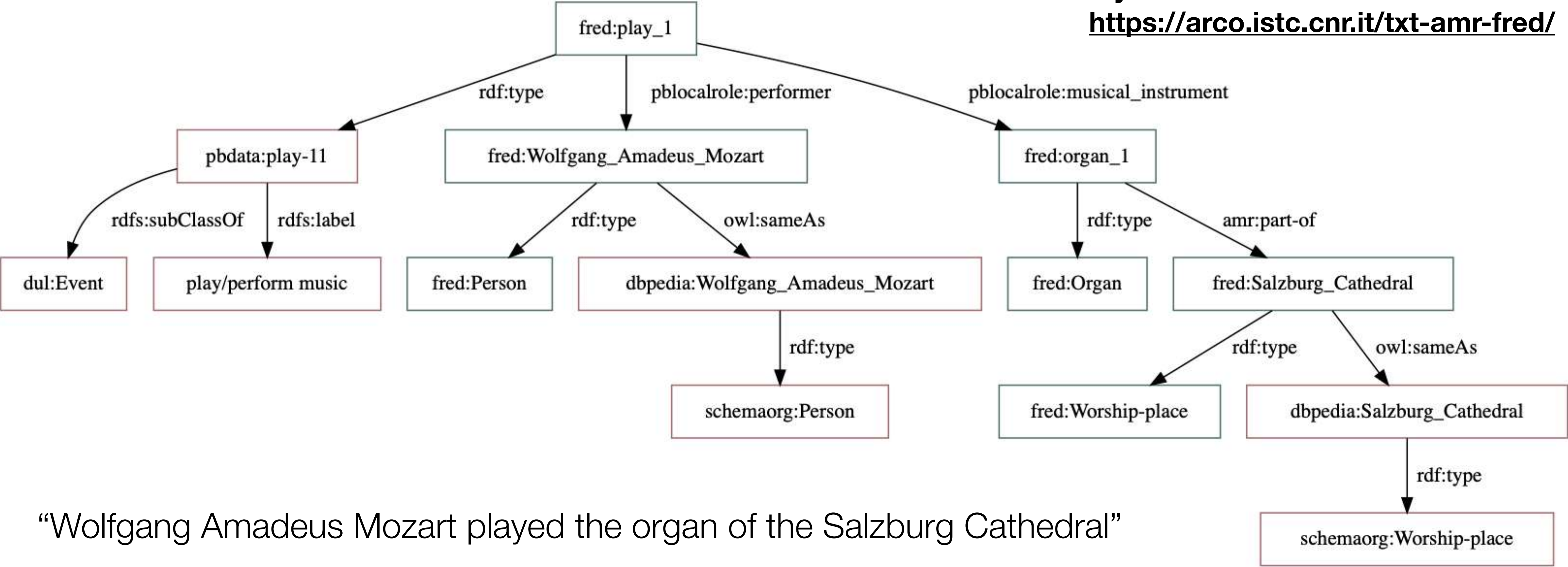


"Wolfgang Amadeus Mozart played the organ of the Salzburg Cathedral"





Try me:  
<https://arco.istc.cnr.it/txt-amr-fred/>



“Wolfgang Amadeus Mozart played the organ of the Salzburg Cathedral”



To create of a global distributed knowledge graph about musical heritage knowledge

- Interlinking data extracted from text, music scores, audio
- Integrating existing heterogeneous collections and datasets
- Encoding orally transmitted knowledge

To provoke a paradigm shift in social science and humanities research methodologies

To provide a large empirical data source for hypotheses formulation and validation

To enable countless applications in domains such as cultural tourism, culture management, music industry, scholarly studies, etc.





Adoption of methodologies and tool support for semantic integration

The diagram features three bullseye shapes. The top bullseye is purple. The bottom-left and bottom-right bullseyes are red. Thin red lines connect the center of the purple bullseye to the center of the bottom-left bullseye, and the center of the bottom-left bullseye to the center of the bottom-right bullseye.

Enabling effective question answering (querying) of KGs through any-domain ontology

End-to-end KG production from text

Harmonic similarity between music from different genres and periods



# Stay tuned and join us!

Follow us:



Web

[polifonia-project.eu](http://polifonia-project.eu)

Email

[press@polifonia-project.eu](mailto:press@polifonia-project.eu)  
[valentina.presutti@unibo.it](mailto:valentina.presutti@unibo.it)



<https://github.com/polifonia-project>



**Polifonia**  
Thanks

