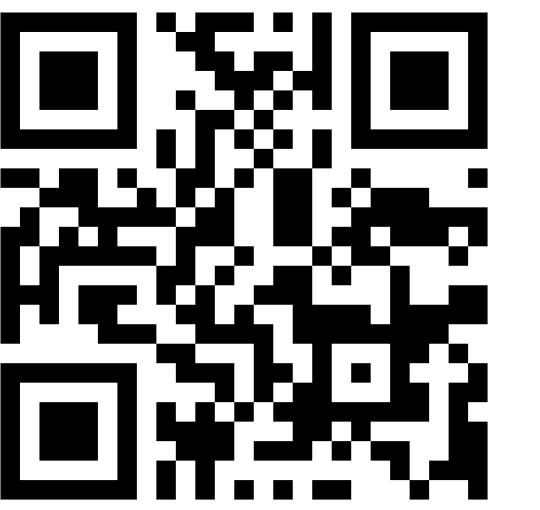


SPOT THE ODD SONG OUT: an online game for collecting music similarity data



mi.soi.city.ac.uk/camir/game/

¹Daniel Wolff , ²Guillaume Bellec and ¹Tillman Weyde

¹City University London, Music Informatics Research Group , ²École Nationale Supérieure de Techniques Avancées, Paris
daniel.wolff.1@soi.city.ac.uk, guillaume-bellec@ensta-paristech.fr, t.e.weyde@soi.city.ac.uk

Music Similarity and User Data

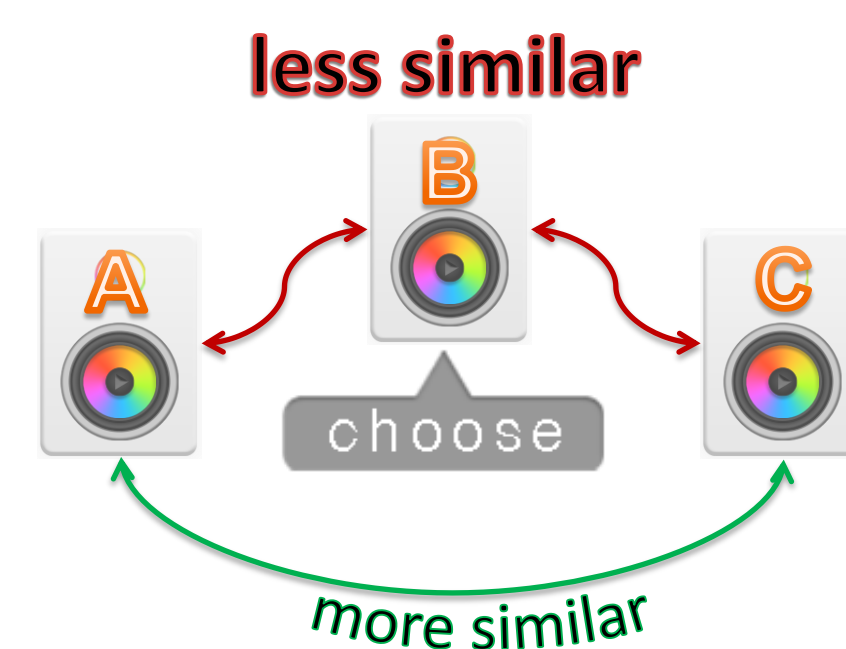
Similarity estimation is a key topic in Music Information Retrieval with many applications. In scenarios such as music recommendation, user expectations depend on music similarity. Perceived similarity is **specific to the individual user** and influenced by a number of factors such as **cultural background**, age and education.

Our goal is to adapt similarity models to similarity data from users, taking into account cultural groups of users sharing common attributes (**cultural indicators**). At this point, there are **few similarity datasets** openly available, and none contains information on user background.

Collected Data

The game **collects relative similarity judgments** of users on triplets of songs, where they are asked to **choose one song as the "odd song out"**.

- If song B is chosen, then:
 $\text{sim}(A,C) > \text{sim}(B,C)$ AND
 $\text{sim}(A,C) > \text{sim}(B,A)$



Data is annotated with anonymised **user attributes**:

Age group	Music education
Gender	Listening habits
Occupation (sector)	Favourite genres
Languages	Music activities
Current location (city)	Favourite media items
Birth location (city)	Religious affinities
Geoprofile of friends	Political affinities

Users can login with their acebook account. Given prior permission, user attributes are extracted from their profile data.

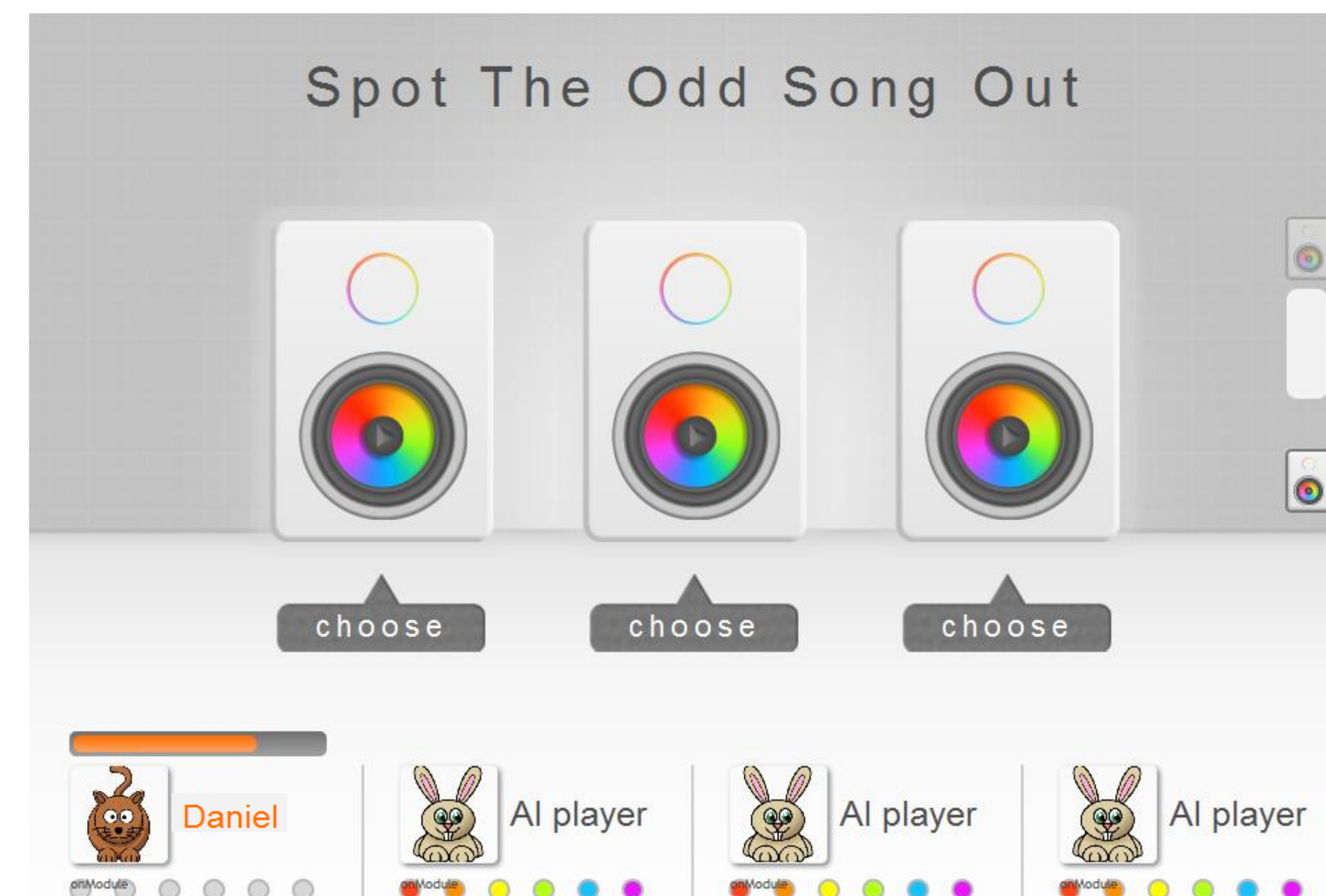
Music Database

- More than 11000 clips in total, each min. 30 seconds long
- Datasets: **1. Million Song Dataset Subset**
 - Mostly Pop/Rock music, streamed by 7digital
- 2. MagnaTagATune**
 - Only "classic" genre subset used.

The Game Interface

- HTML5 web application runs on many devices w. browser
- Game** look and feel
- Multiplayer**
- Hop-on – hop off:** enter or leave anytime

- Odd One Out** scenario
- 45 seconds **time limit**
- Progress of other players** is shown
- Decision and timings** are logged



Motivation and Rewards

- In-Match:
 - Input-agreement** of players is rewarded with points
- In-Game:
 - High-score** table motivates competition
 - Points represent **in-game currency**:
 - Buys **new** fancy **avatars**
 - Buys **music genre selection for triplet**: Classical, Jazz, Country, Electronica, Latin, Rap, R&B, ...
- Social network: acebook
 - High-scores can be shared with friends
 - Game can be advertised to friends

Result: Agreement for Each Triplet

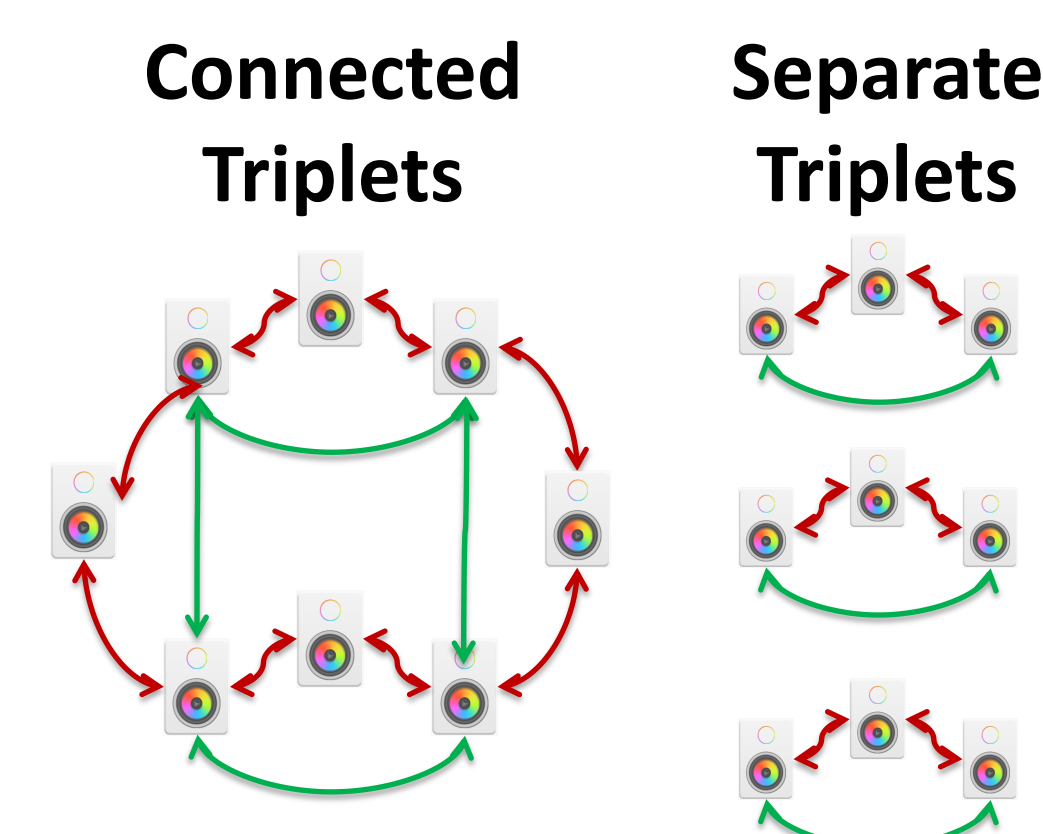
AI player	+20
AI player	+0
AI player	+20
Daniel	+20

Overall High-score

1	170
15	140
17	60
Daniel	50
19	40
18	40
20	20
22	20

Sample Selection

- With unknown participant numbers, online surveys benefit from real-time sample selection
- CASIMIR API (see right) manages an expanding music clip subset for all API clients
- For Odd One Out similarity annotation task, minimal thresholds of connections between the comparison triplets are assured, before the subset is expanded.

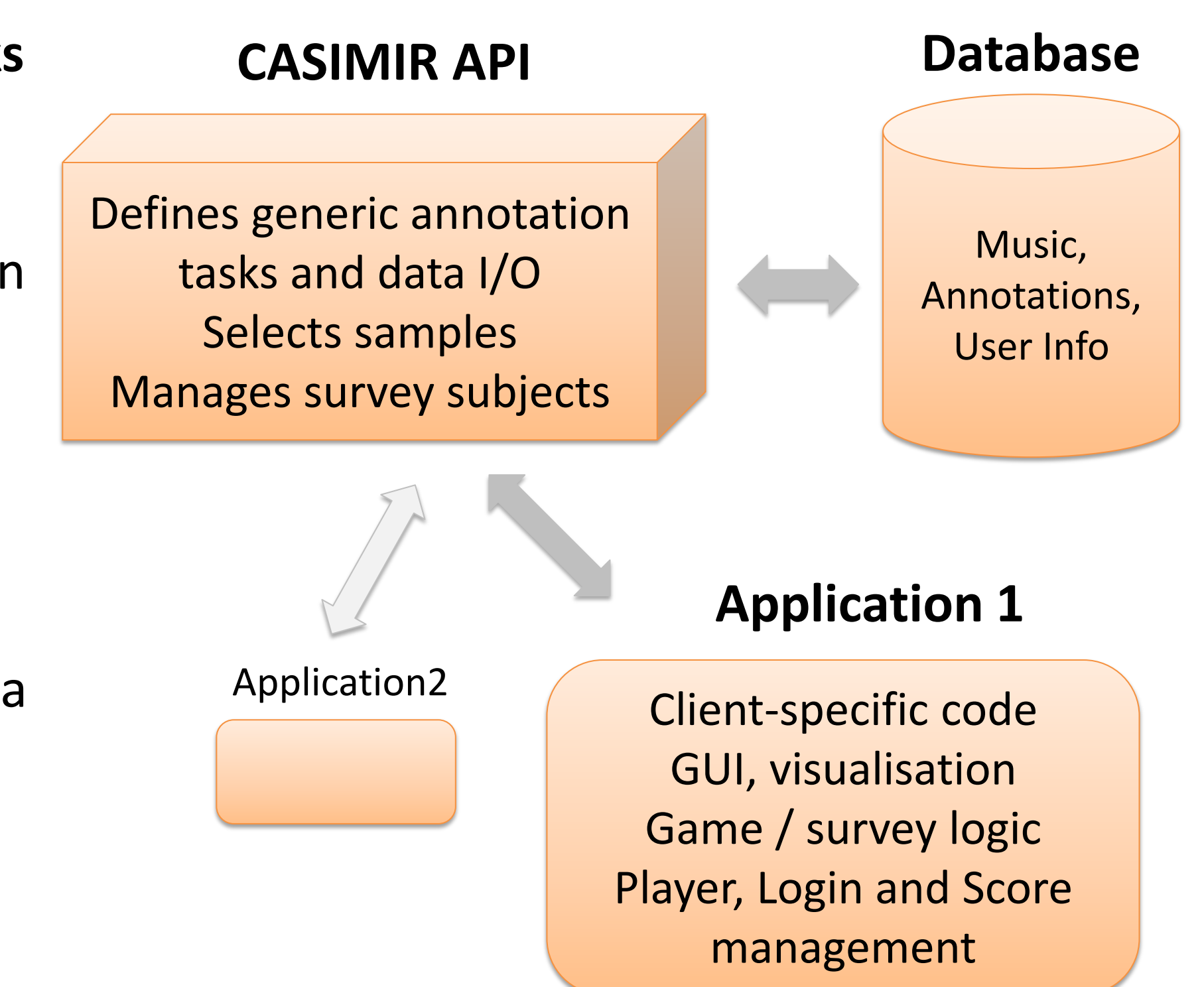


Towards a Generic Data Collection API

Being extensible to multiple question types and scenarios, the CASIMIR API manages the **storage of user input and user attributes**, as well as the music presented and the **statistical selection of samples**. **Additional data annotation tasks** such as beat annotations can be **easily integrated**.

Benefits of an Online API

- API defines annotation tasks** independent of game or survey logic.
- Generic task & I/O definition
 - Different (simultaneous) client applications
 - Reusable – Extendable
- Central collection database**
 - Real-time control on data sampling
 - Oversight of concurrent surveys
 - Possible aggregation of annotation data



Current CASIMIR API Implementation

CASIMIR API is provided by a **PHP5.x server**, via the industrial standard **SOAP** protocol. The stable **open source** implementation for Odd One Out music similarity annotations provides three API hooks:

- GetOSOTriplet(user_details,opt_genre)*: returns a triplet of clips with urls to audio.
- SendOSOVote(user_details)*: stores a vote for the Odd Song Out
- GetOSOAIVote(triplet)*: returns a AI vote for the given triplet

A hook for **beat annotations** is currently tested. Existing system and implementation allows to focus on data collection application.

- Build your own application using CASIMIR API! -



The CASIMIR HTML5+PHP Game Framework

HTML5 interface built on a **open source GWAP framework for CASIMIR**:

- Based on **LimeJS** and Google Closure - runs on any HTML5 browser
- Real-time multiplayer!** Single player and pseudo multiplayer possible
- Modular architecture**: Little implementation effort for new applications
- Create your own game** or add your module to "Spot the Odd Song Out"!

