

# InMusic: an ontology for describing and understanding the emotional response to the listening of a musical work

Manuele Pasini

manuele.pasini@studio.unibo.it

July 23, 2021

## 1 Introduzione

“L’arte che consiste nell’ideare e nel produrre successioni strutturate di suoni semplici o complessi, che possono variare per altezza (cioè per la frequenza delle vibrazioni del corpo sonoro), per intensità (cioè per l’ampiezza delle vibrazioni) e per timbro (che dipende dal materiale del corpo sonoro), per mezzo della voce umana (m. vocale), di strumenti (m. strumentale) o della combinazione di entrambe queste fonti.”<sup>1</sup>

La musica ha accompagnato il genere umano durante tutto il corso della sua evoluzione e può essere considerata, essendo essa stessa un prodotto umano, un linguaggio attraverso il quale l’uomo si esprime, attraverso il quale riesce a descrivere, a testimoniare e cercare di comprendere la propria esistenza negli aspetti più astratti; può essere sotto certi aspetti associata alla poesia: così come quest’ultima arte non è generalmente accompagnata da suoni allo stesso modo vi sono musiche (basti pensare ad alcuni grandi capolavori di musica classica oppure ad alcune colonne sonore di film contemporanei) non accompagnate da parole che suscitano allo stesso modo una reazione nell’ascoltatore.

La filosofia della musica è una branca della filosofia che si occupa di studiare la relazione tra uomo e musica; all’interno di questa disciplina vi sono due principali correnti di pensiero: emotivismo e cognitivismo; secondo i sostenitori della teoria dell’ “emotivismo musicale” l’ascolto di un determinato prodotto musicale è in grado di suscitare una vera e

---

<sup>1</sup><https://www.treccani.it/vocabolario/musica/>

propria risposta emotiva all'interno dell'ascoltatore mentre i sostenitori della teoria del "cognitivismo musicale" sostengono che le cause di un'eventuale risposta emotiva verificatasi nell'ascoltatore sono da ricercare all'interno della struttura matematica del suono ascoltato, sostengono e ricercano l'esistenza di una "correlazione matematica".

Questo progetto segue l'ideologia emotivista ed individua nell'emozione suscitata da un ascolto musicale il concetto cardine del modello.

Consideriamo un ultimo esempio, prendendo in esame le "Instagram stories", è possibile tramite una funzionalità del social network associare ad una foto o ad un video di durata massima quindici secondi una canzone (presa dalla libreria interna di Instagram); se esaminiamo i componenti di questa storia possiamo individuare : località (determinata dalla foto), musica (inserita dall'utente su base personale), un'eventuale attività (rappresentata all'interno dell'immagine o del video) ed indicativamente un orario (derivabile dall'immagine con un certo margine di errore); sebbene questo sia un modello decisamente complesso da descrivere e di conseguenza da analizzare, in quanto l'analisi di immagini o video è materia complessa sotto tutti i punti di vista, è innegabile che questa forma di espressione ci dica qualcosa di più sull'autore rispetto a quanto non ci dica a primo impatto osservando l'immagine; se consideriamo che solamente nel 2018 la stima delle Instagram Stories è di circa 500 milioni al giorno, è logico pensare a quante informazioni potremmo estrapolare sull'aspetto emotivo di una persona semplicemente osservando ed analizzando le sue storie Instagram. Tornando all'interno del perimetro di questo progetto, se si applicasse la stessa procedura di modellazione ad esempio agli ascolti su Spotify (considerando anche il numero giornaliero di questi), ci si può rendere conto ancora di più di quanto importante potrebbe rivelarsi definire un modello di studio a riguardo. Questo progetto vuole gettare le basi per un'ontologia in grado di descrivere e modellare l'ascolto di una canzone e la reazione emotiva che suscita nell'ascoltatore.

## 2 Modellazione

Individuato e posto come concetto cardine l'emozione suscitata da un ascolto, il passo successivo è stato quello di individuare eventuali fattori che influenzano la risposta emotiva dell'ascoltatore: dopo aver osservato modelli concettuali già esistenti e disponibili in rete inerenti alla filosofia della musica ci si è accorti come una buona parte di questi tenda a scomporre un prodotto musicale nelle sue componenti sonore, relazionandole poi singolarmente all'ascoltatore; personalmente ritengo che l'ascolto di una musica sia un fenomeno olistico, che non può essere descritto semplicemente scomponendolo in sottoparti da relazionare ed analizzare poi singolarmente in quanto, soprattutto se si parla della forma di musica più diffusa oggi composta da una base sonora alla quale si sovrappone una voce, non ritengo siano tanto le singole componenti di un suono a suscitare una reazione emotiva che ritengo sia invece provocata dall'amalgamarsi di suoni che, all'interno di un determinato **contesto**, sono in grado di produrre uno stimolo emotivo. Per fare un esempio pratico, il progressive rock <sup>2</sup> è un genere musicale basato sulla continua sperimentazione e ricerca di suoni nuovi, mai sentiti, e persegue questo obbiettivo attraverso la continua fusione di suoni diversi in qualunque forma un suono possa essere riprodotto. Se prendessimo una canzone appartenente a questo genere e la scomponessimo nelle sue componenti sonore il risultato sarebbe disastroso, i suoni avrebbero ognuno la propria dimensionalità ma mancherebbe quella totalità, quella completezza di cui è dotata una composizione che la mente umana è in grado di creare. Per descrivere un prodotto musicale si è deciso di utilizzare un'ontologia denominata **Music Ontology**<sup>3</sup> che offre un modello di descrizione molto dettagliato tale da essere in grado di rappresentare qualunque forma di composizione musicale, trattando eventualmente anche la sua scomposizione in parti sonore.

La fase successiva ha previsto lo studio del contesto di un ascolto musicale: se una determinata canzone viene ascoltata in due momenti diversi, la reazione emotiva dell'ascoltatore persiste nel tempo o è influenzata da elementi esterni?

Basandosi principalmente sull'esperienza personale, si è cercato di individuare quelli che si sono ritenuti gli elementi principali che compongono il **contesto** di un ascolto: tutti quei processi fisici o emotivi che influenzano una determinata persona nel periodo di tempo in cui ascolta un prodotto musicale. Sono state definite tre entità:

---

<sup>2</sup>[https://it.wikipedia.org/wiki/Rock\\_progressivo](https://it.wikipedia.org/wiki/Rock_progressivo)

<sup>3</sup><http://musicontology.com/>

- **attività:** definisce un'attività eseguita da una persona, in particolare viene utilizzata per descrivere le azioni che una persona compie durante un ascolto musicale; questa modellazione si basa sul presupposto che una determinata composizione musicale possa essere scelta sulla base dell'attività che si sta compiendo in quel preciso momento, ad esempio sarebbe difficile trovare qualche atleta che ascolti una composizione di Fryderyk Chopin prima dell'inizio una partita di rugby; la classe "ActivityType" è stata inserita per dare una struttura logica al concetto di attività.

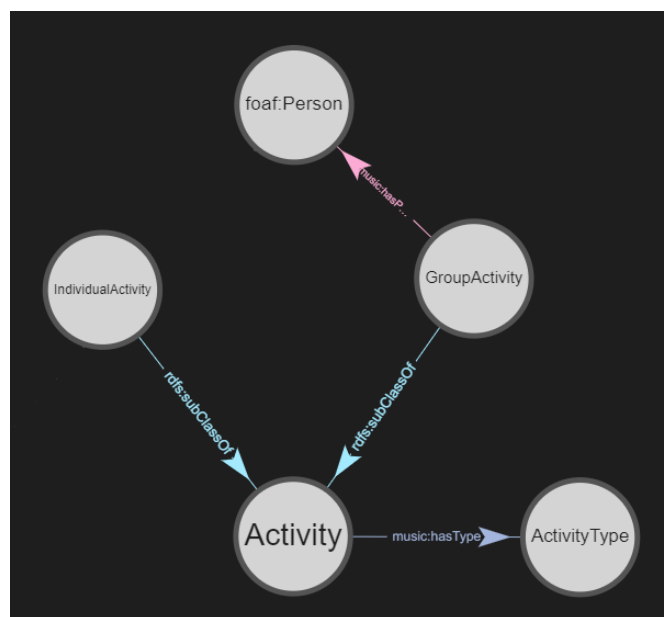


Figure 1: modellazione del concetto di attività

- **luogo:** definisce un luogo fisico, viene utilizzato per definire il luogo dell'ascolto: è uno dei fattori che si è ritenuto possano influenzare la risposta emotiva ad un ascolto musicale;
- **ricordi:** vuole modellare i ricordi collegati ad un determinato ascolto: appurato che un ascolto musicale venga effettuato in un determinato luogo mentre si svolge una determinata attività, è possibile che all'ascoltatore riemergano ricordi che, collegati direttamente o indirettamente alla canzone ascoltata, influenzino la reazione emotiva. Si fa necessariamente riferimento a ricordi appartenenti alla memoria dichiarativa (accessibile alla consapevolezza), composta da memoria semantica e memoria episodica. Si parla di memoria semantica indicando quella parte di memoria che riguarda le conoscenze generali ed in questo contesto riguarda avvenimenti non direttamente collegati e/o vissuti dall'ascoltatore; si parla invece di memoria episodica

facendo riferimento, sempre all'interno di questo contesto, a quegli avvenimenti che hanno coinvolto direttamente l'ascoltatore e dunque hanno necessariamente un valore emotivo maggiore.

Nonostante sia innegabilmente uno dei fattori principali quando si parla di emozioni o comunque di concetti astratti inerenti alla coscienza umana, non è stato possibile modellare qualunque forma di memoria implicita ovvero non accessibile alla consapevolezza; uno degli scopi di questa ontologia è anche quello di tentare di ricostruire e comprendere questa parte di contesto sulla base di ciò che invece è possibile modellare e successivamente analizzare tramite i dati.

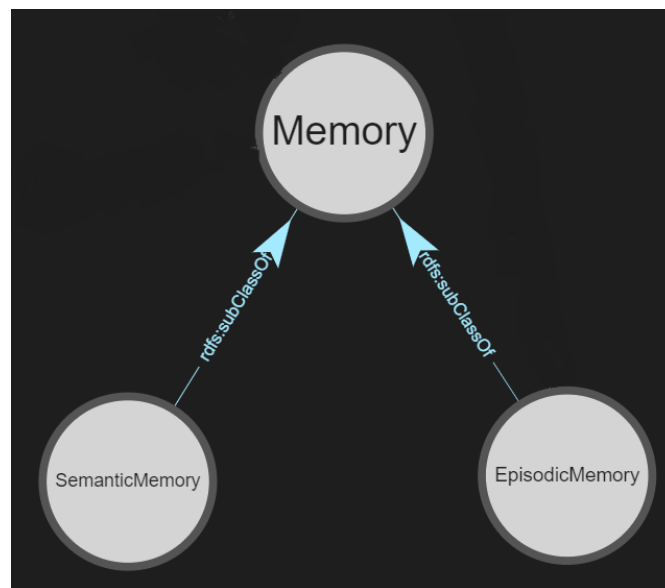


Figure 2: modellazione del concetto di ricordi

Trattandosi di un ambito in buona parte astratto, si è rivelato complesso trovare un modo per tradurlo e modellarlo in termini scientifici, ad esempio non è tanto importante il luogo all'interno del quale una persona sta ascoltando qualcosa quanto è invece molto più importante comprendere cosa quel luogo significa per quella persona, come lo percepisce e quali sono gli stimoli ad esso collegati. Per fare ciò si è associata una descrizione ad ognuna di queste entità; l'introduzione di tale descrizione personale risolve in parte il problema ed introduce complessità in termini di analisi, in quanto è difficile riuscire a comprendere ed estrapolare più informazioni possibili da un testo libero non strutturato, soprattutto se come in questo caso le informazioni da estrapolare riguardano concetti astratti e dunque potrebbero non essere direttamente evidenziati nel testo.

Definito il contesto, ci si è occupati di modellare e descrivere la reazione emotiva dell'ascoltatore,

ovvero si è cercato di rispondere alla domanda "da cosa è composta la risposta emotiva che una persona può avere al momento dell'ascolto di un prodotto musicale?". Come

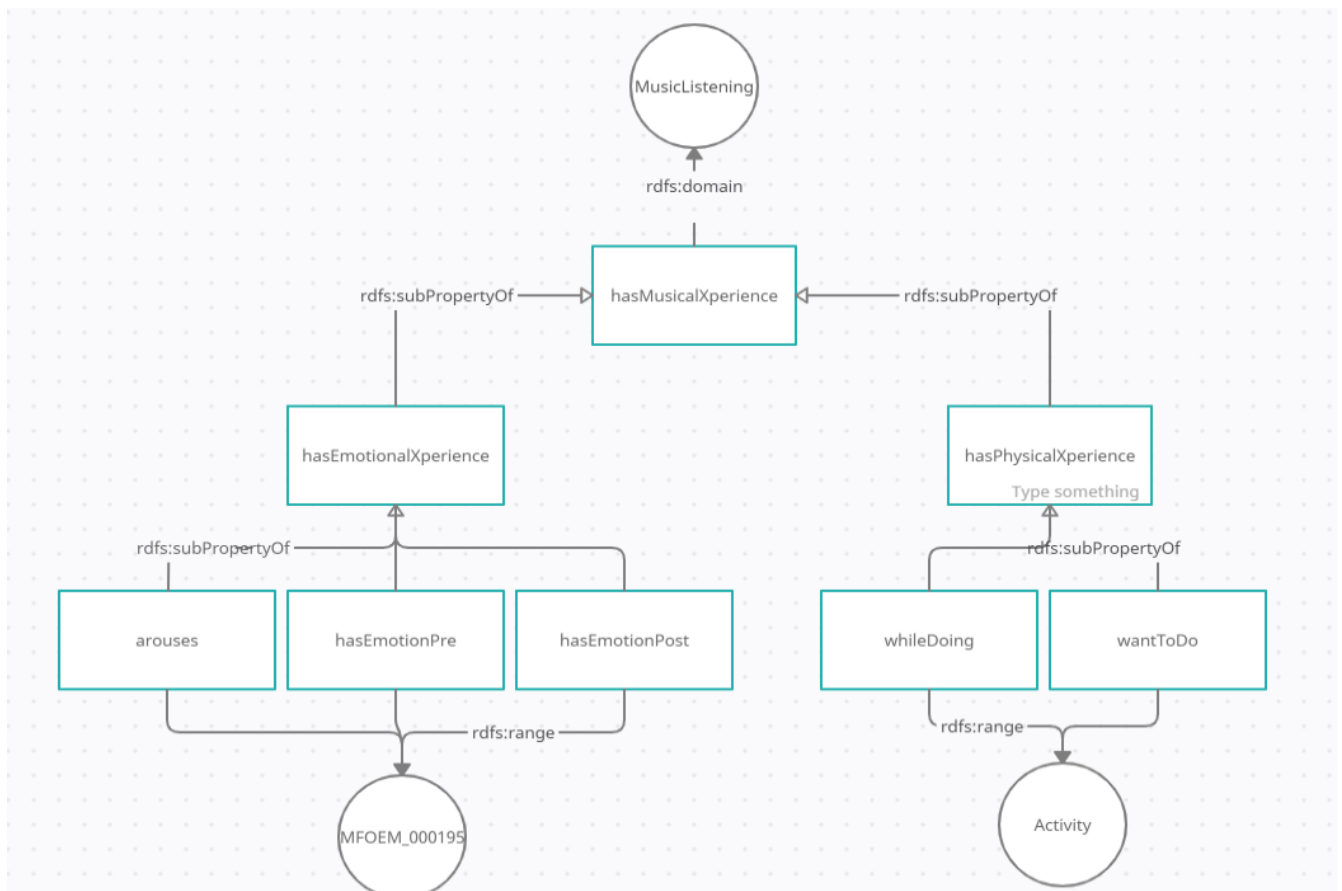


Figure 3: modellazione del concetto di esperienza musicale

mostrato in figura, si è preso il concetto di "esperienza musicale" e lo si è scomposto in sottoproprietà:

- **esperienza fisica:** con esperienza fisica si vuole descrivere come l'ascolto di una canzone possa influenzare la "produttività" di un ascoltatore relativamente all'attività che sta svolgendo. Questa proprietà è stata poi scomposta in due sottoproprietà che rappresentano l'attività che l'ascoltatore sta svolgendo al momento dell'ascolto e l'attività che la stessa persona è stimolata a compiere dall'ascolto di una canzone in un determinato contesto.
- **esperienza emotiva:** si vuole descrivere come l'ascolto di una canzone influenzi lo stato d'animo dell'ascoltatore, per cercare di modellare al meglio l'impatto che un ascolto può avere si è suddivisa questa proprietà in tre sottoproprietà, che rappresentano lo stato d'animo dell'ascoltatore precedente all'ascolto, quello successivo

all'ascolto e le emozioni evocate dall'ascolto stesso. Per rappresentare le emozioni è stata utilizzata l'ontologia "Emotion Ontology"<sup>4</sup>, in particolare viene fatto riferimento ad un'entità della suddetta ontologia denominata "**affective process**"<sup>5</sup>, che rappresenta un processo, sottoclasse di tutti i processi mentali<sup>6</sup> avente una valenza positiva o negativa; ad esempio, sono considerati "affective process" la totalità delle emozioni (l'ontologia ne definisce già una buona parte), i processi fisici (dolore, fame, sete, piacere fisico), gli stati d'animo (ansioso, irritabile, felice, disconnesso), e le sensazioni soggettive, ovvero un'emozione personale relativa al soggetto. La reazione emotiva collegata ad un ascolto musicale è stata modellata tramite inferenza: non sarebbe scientificamente corretto lasciare che sia la persona soggetto dello studio ad associare un valore alle emozioni percepite durante un ascolto in quanto introdurrebbe un bias, per questo motivo si è optato di collegare un ascolto musicale ad una coppia di immagini scelte dall'utente su un pool di sei che più raffigurano:

1. lo stato d'animo antecedente l'ascolto;
2. lo stato d'animo che più rappresenta quello percepito dalla canzone.

Ogni immagine ha associata una serie di metadati che rappresentano le emozioni collegate a quella determinata immagine, ad esempio se si prende il caso del dipinto "Campo di grano con volo di corvi" di Vincent Van Gogh, la rappresentazione in open data del quadro sarebbe la seguente:

```
<wheatfield-with-crows>
a :Image;
dc:name "Wheatfield with Crows";
dc:description "Wheatfield with Crows is one of Van Gogh's {...}
dc:creator <#van-gogh>;
:represents <http://purl.obolibrary.org/obo/MFOEM_000016>. #grief
```

Le emozioni rappresentate tramite le immagine verranno poi associate all'ascolto tramite inferenza.

---

<sup>4</sup><https://core.ac.uk/download/pdf/189205533.pdf>

<sup>5</sup>[http://www.ontobee.org/ontology/MFOEM?iri=http://purl.obolibrary.org/obo/MFOEM\\_000195](http://www.ontobee.org/ontology/MFOEM?iri=http://purl.obolibrary.org/obo/MFOEM_000195)

<sup>6</sup>[http://www.ontobee.org/ontology/MFOEM?iri=http://purl.obolibrary.org/obo/MF\\_0000020](http://www.ontobee.org/ontology/MFOEM?iri=http://purl.obolibrary.org/obo/MF_0000020)

### 3 Popolazione del database

Il database è stato popolato con un piccolo campione rappresentativo di dati; Stardog Studio<sup>7</sup> non permette una popolazione del database "user-friendly" come visto su Protégé, le modalità di inserimento dati prevedono il collegamento ad una fonte esterna dati come un database oppure l'inserimento manuale di triple RDF.

Sono state caricate all'interno del database una serie di canzoni, in particolare

- Paranoid Android - Radiohead (album);
- Tabula Rasa Elettrificata - C.S.I. (album);
- The Kids Aren't Alright - The Offspring (track);
- Hurt - Nine Inch Nails (track);
- American Jesus - Bad Religion (track);
- How To Disappear Completely - Radiohead (track);
- Heroes - David Bowie (track).

Sulla base di queste canzoni sono state costruiti gli altri dati: una serie di attività, di ricordi (sia inerenti memoria episodica che semantica) e di luoghi. L'esempio sopra mostra la modellazione di un ascolto musicale, fatto da una certa Sophia Coppola all'interno di un appartamento di New York in un momento di solitudine notturno; tale ascolto identifica un cambio di stato d'animo nell'ascoltatore (da uno stress iniziale ad una malinconia piacevole post ascolto) che, sebbene si tratti di un dato costruito ad hoc, mostra appieno come in realtà la musica possa essere un linguaggio per descrivere alcuni aspetti umani. Si è resa necessaria l'importazione dell'intera ontologia "Emotion Ontology" in quanto Stardog Studio non supporta il costrutto *owl:imports*.

#### 3.1 Applicazione Java

E' stata realizzata un'interfaccia grafica scritta in linguaggio Java che permette di interfacciarsi con il database ed inserire nuove triple: nel dettaglio l'applicazione si connette al server Stardog e permette all'utente di scegliere una canzone (tra quelle già contenute all'interno del database) per la quale descrivere un ascolto.

L'applicazione contiene al suo interno una serie di immagini che sono state parallelamente

---

<sup>7</sup><https://docs.stardog.com/>



descritte secondo questa ontologia ed inserite all'interno del database. Per ogni ascolto l'utente deve scegliere, da due pool di tre immagini, due immagini che rappresentino rispettivamente lo stato d'animo precedente l'ascolto e l'immagine che più rappresenta l'esperienza vissuta ascoltando la canzone scelta. Per interfacciarsi al database di Stardog sono state utilizzate le API proprie di Stardog e Apache Jena.

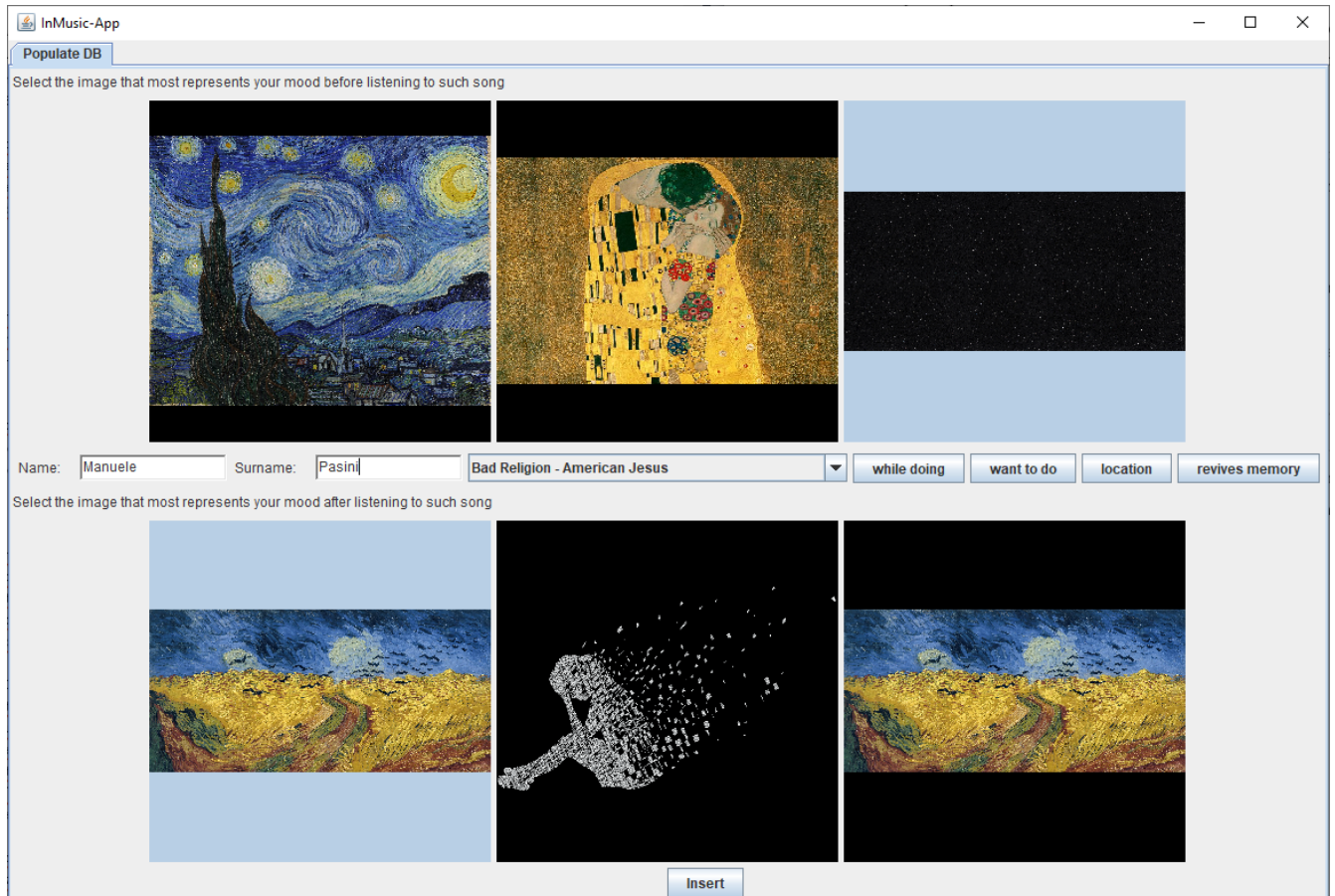


Figure 4: Interfaccia grafica applicazione

## 4 Query

Sono state realizzate cinque query, situate all'interno della cartella "query", il cui obiettivo è quello di cercare di mostrare ciò che offre questa ontologia.

- **listeners-song-aroused-activity:** mostra una panoramica del database, mostra ogni ascolto musicale presente nel database ed il suo contesto.
- **bad-pre-good-post:** elenca le canzoni che hanno indotto un cambiamento di stato d'animo nell'ascoltatore, da uno stato d'animo con una valenza negativa iniziale ad uno stato d'animo con una valenza positiva post ascolto.
- **bad-emotions-while-working:** mostra, in ordine decrescente, quali sono le canzoni ascoltate durante un'attività catalogata come lavorativa che hanno indotto un cambiamento in negativo dello stato d'animo dell'ascoltatore.
- **same-song-different-valence:** mostra i prodotti musicali che hanno indotto in almeno due ascolti differenti due emozioni contrastanti; sottolinea l'importanza del contesto
- **positive-percentage-song:** mostra una classifica delle canzoni sulla base delle emozioni suscitate direttamente ed indirettamente tramite i ricordi collegati ad un ascolto; la metrica utilizzata è il numero di emozioni positive suscitate da una canzone rapportato al numero totale di emozioni suscitate.

## 5 Deploy

Questo elaborato è stato realizzato tramite Stardog, l'esecuzione di questo IDE necessita la presenza di un server Stardog sul quale mantenere il database. Requisiti:

- Docker;
- licenza d'uso di Stardog (ottenibile qui e già inclusa all'interno della cartella "licenze").

Step per l'installazione:

1. download dell'immagine di Stardog tramite Docker Hub;

```
docker pull stardog/stardog:latest
```

2. lanciare il server Stardog:

```
docker run -it -v ~/stardog-home/:/var/opt/stardog -p  
5820:5820 stardog/stardog
```

"stardog-home" è il percorso della cartella all'interno della quale è presente solamente la licenza d'uso di Stardog ottenuta precedentemente; diventerà poi la directory contenente i file necessari all'esecuzione del server;

3. collegarsi a Stardog Studio e connettersi al server precedentemente istanziato:

- **indirizzo:** http://localhost:5820;
- **ID:** admin;
- **password:** admin.

4. spostarsi nella sezione **Databases** e creare il database "InMusic";

5. all'interno della sezione "Databases", dopo aver selezionato il database appena creato, sotto la voce **Namespaces** importare tramite la funzione "Import" in alto a sinistra il file "InMusic-namespaces.ttl";

6. spostarsi nella sezione **Models** e tramite la funzionalità "Create Model" creare un modello con le seguenti caratteristiche:

- **Model Name:** EmotionsInMusic-Ontology
- **Named Graph:** musicemotions:ontology:graph

- **Prefix:** music
7. ritornare nella sezione "Databases", selezionare il database appena creato e, tramite la funzione **Load data** importare il file **InMusic-Ontology.ttl** specificando l'opzione **Load Data to** musicemotions:ontology:graph , ovvero il grafo appena creato; a questo punto tale grafo contiene l'ontologia;
  8. per popolare il database, selezionarlo dalla sezione "Databases" e tramite la funzionalità "Load data" importare tutti i file presenti nella cartella **individuals**, lasciando l'opzione "Load Data to **Default Graph**"
  9. per aggiungere le query, spostarsi nella sezione Workspace ed effettuare l'upload dei file cliccando in alto a destra sul "+".

## 6 Conclusioni

La continua ed esponenziale evoluzione tecnologica ha portato ad un'incremento della produzione di dati che, inevitabilmente, sono diventati oggetto di studio e stanno rivoluzionando il mondo per come lo conosciamo; supportati da tecnologie che permettono l'analisi di questi in tempi relativamente brevi, i dati stanno diventando le fondamenta su cui verrà costruita la "società del futuro". La peculiarità di una società data-driven è l'inversione del ruolo del "dato", se prima era solo una conseguenza, un prodotto di un processo, in questa nuova fase diventa un punto di partenza, una causa, l'elemento sulla base del quale avviene l'innovazione, che determina i passi successivi da percorrere. Una grande mole di dati ci permette di costruire conoscenza laddove non è allo stato attuale delle cose possibile costruire un modello di rappresentazione in quanto non si ha conoscenza sufficiente per farlo; il dato diventa uno strumento di scoperta ed espansione della conoscenza.

Se è vero che la filosofia della musica è una disciplina che esiste sin dalle prime civiltà, è altrettanto vero che non si è mai arrivati a risultati concreti che andassero oltre alla teorizzazione, e credo uno dei motivi di ciò possa essere ricercabile nella complessità del fenomeno che si vuole descrivere: se prendiamo in analisi un fenomeno aleatorio il cui risultato è altamente dipendente da un numero indefinito di variabili che non sappiamo modellare (ed è il caso di qualunque fenomeno soggettivo non descrivibile matematicamente) questo diventa inevitabilmente indescribibile in quanto la sua complessità supera gli strumenti a disposizione; se tuttavia è possibile approssicare il problema da un punto di vista diverso, ad esempio adottando una soluzione euristica, allora il fenomeno diventa

“descrivibile”. Una cosa simile avviene in questo caso: descrivere una reazione emotiva all’ascolto di prodotto musicale è un fenomeno troppo complesso per la conoscenza che abbiamo a disposizione, solamente la differenza tra una persona  $a$  ed una persona  $b$  rende vano qualunque tipo di modellazione con gli strumenti attuali in quanto non sappiamo descrivere precisamente cos’è “la differenza” tra due persone e di conseguenza non siamo in grado di comprendere come quella differenza influenzi il processo di reazione emotiva. Riuscire ad attribuire a questi dati una semantica ed immagazzinarne una grande quantità potrebbe permetterci, tramite analisi approfondite, di aggiungere qualche tassello alla conoscenza di questo ambito.

Uno studio dell’università di Berkeley <sup>8</sup> il cui scopo era quello di cercare di comprendere in che modo la musica generi emozioni nell’uomo, ha sottoposto a due campioni di persone appartenenti a culture diverse (un campione costituito da abitanti degli Stati Uniti, un campione costituito da abitanti della Cina) una serie di canzoni per cercare di comprendere se l’eterogeneità culturale potesse in qualche modo influire sulla percezione di un’emozione durante un’ascolto musicale; il risultato sono tredici emozioni <sup>9</sup> che si sono scoperte presenti in entrambe le culture. Tramite un’ontologia in grado di descrivere questo ambito, l’estensione di questa ricerca avrebbe potuto essere molto maggiore e generalizzata. Per fare un altro esempio applicativo, poter comprendere in che modo l’ascolto di una canzone condiziona lo stato d’animo di una persona potrebbe permettere studi più approfonditi sulla musicoterapia, applicata soprattutto a casi di disturbi psichici.

## 6.1 Considerazioni personali

La fase di popolazione del database è quella che si è rivelata più dispendiosa in termini di tempo e tediosa in quanto sia il dover scrivere ogni tripla a mano sia il comprendere le modalità in cui Stardog permette di popolare un database sono state attività dispendiose; ciò non toglie che il toccare con mano il linguaggio Turtle ed in generale il framework RDF ha permesso di comprenderlo meglio. Questo elaborato è ovviamente una semplificazione del goal descritto nella relazione, credo comunque che questo sia un ambito in cui verranno fatti progressi ed in cui queste tecnologie e questo approccio possano portare risultati concreti, sebbene io rimanga molto timoroso di questa totale classificazione e rappresentazione in dato di ogni aspetto della vita umana.

---

<sup>8</sup><https://www.pnas.org/content/pnas/117/4/1924.full.pdf>

<sup>9</sup><https://www.ocf.berkeley.edu/~acowen/music.html>