

AGREGANDO INFORMACIÓN CONTEXTUAL PARA USO EN SISTEMAS DE RECOMENDACIÓN CONTEXTUALIZADO: MUSICSENSORY

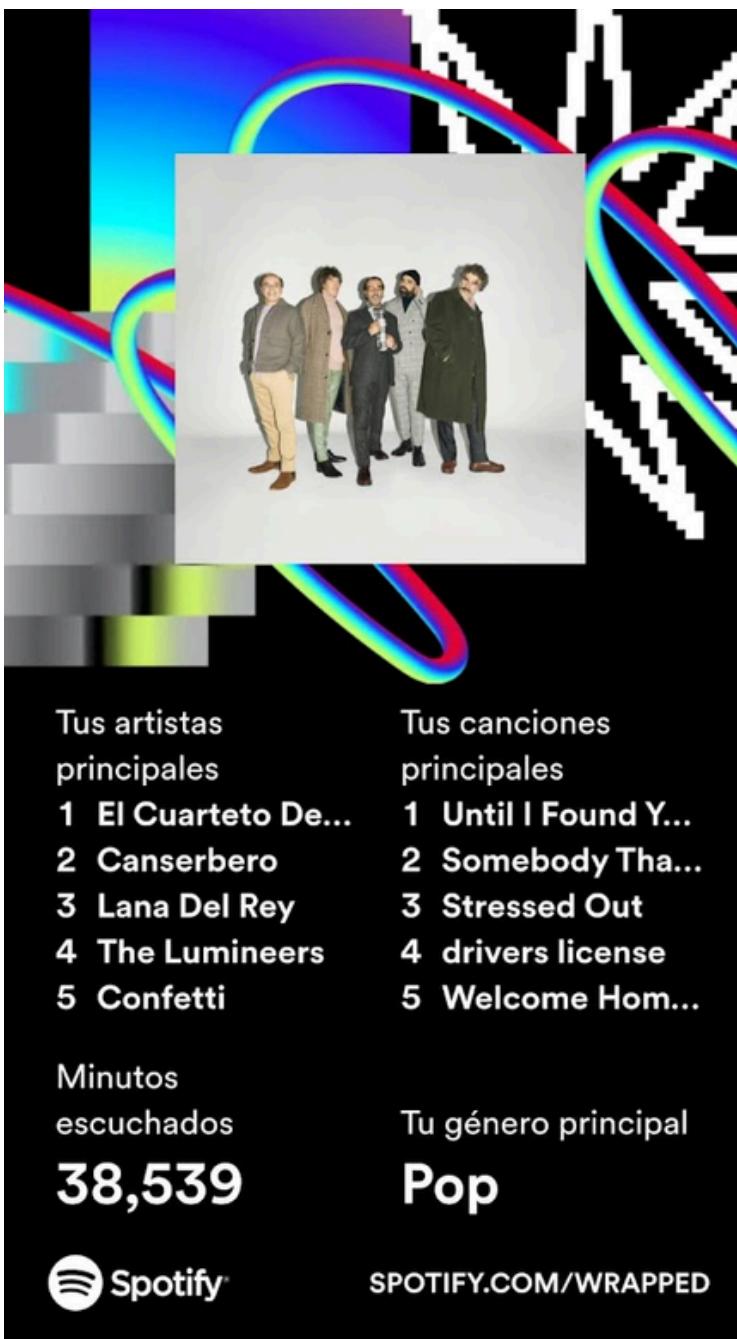
- Recomendación musical
contextualizado
- Recomendación de contenido

Victor Quicaño

CONTEXTO

- Definición: Sistemas inteligentes que sugieren ítems (productos, servicios, contactos) relevantes a usuarios según sus preferencias.
- Objetivo: Ayudar a descubrir ítems útiles e interesantes, reduciendo la sobrecarga de información.
- Funcionamiento:
 - Recopilan datos de preferencias (explícitos como calificaciones o implícitos como comportamiento).
 - Predicen calificaciones y generan rankings personalizados.

MOTIVACIÓN



JUSTIFICACIÓN INFLUENCIA DE LA MÚSICA



POR DEMOGRAFÍA



POR ACTIVIDAD



POR CONTEXTO

PROBLEMA

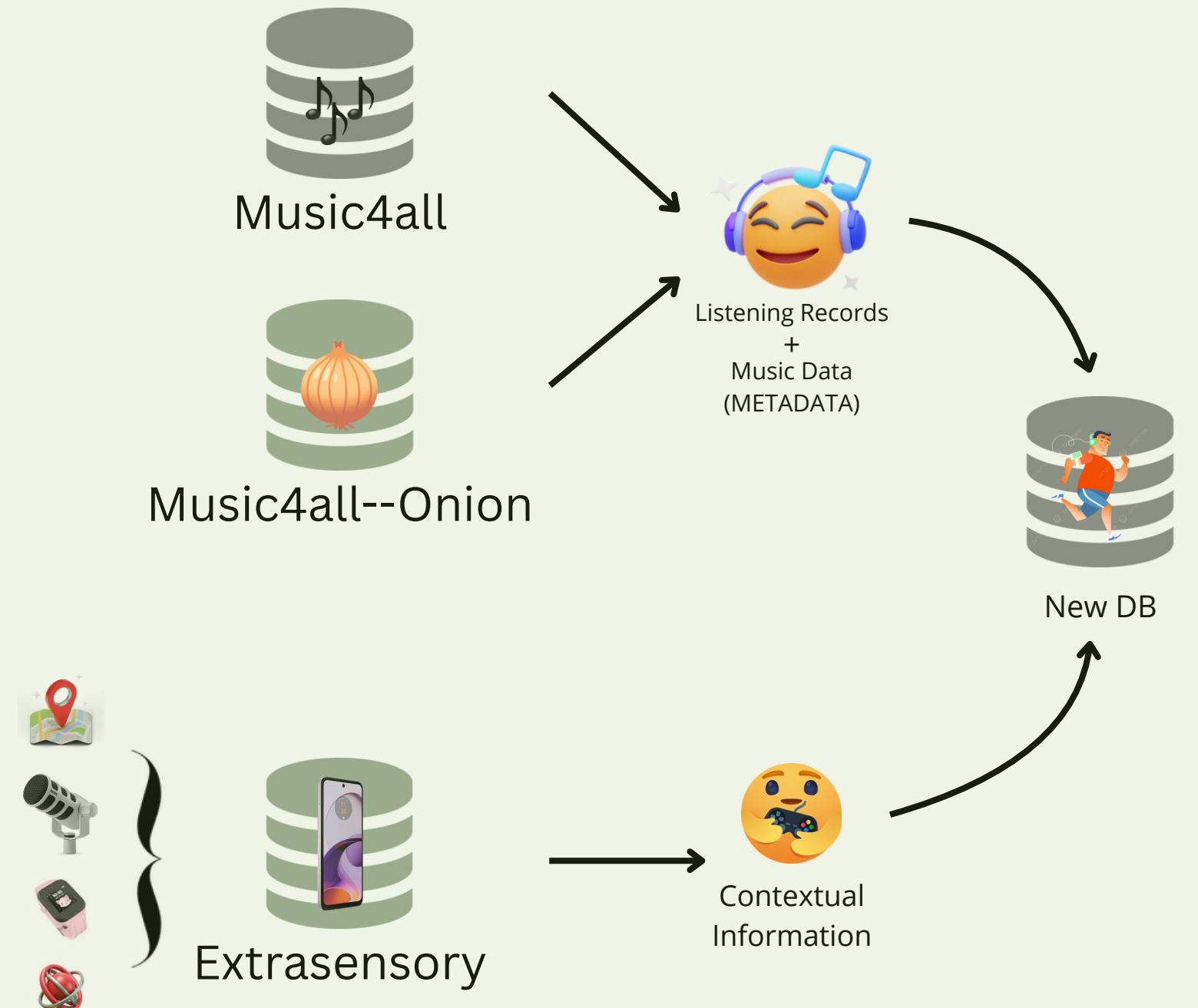
Actualmente, los sistemas de **recomendación de contenido** (basados únicamente en el historial de usuario o en similitudes de ítems) suelen ofrecer sugerencias que no consideran el contexto real del momento (por ejemplo, la ubicación, la hora del día, el dispositivo o el estado de ánimo), lo que puede derivar en recomendaciones poco relevantes o intrusivas para el usuario.

OBJETIVO

Desarrollar un dataset musical contextualizado que integre de manera multimodal variables contextuales, ubicación, emoción endebida dentro de la letra y el tiempo de escucha del usuario, abordando la heterogeneidad de datos necesaria para mejorar la calidad y relevancia de los sistemas de recomendación musical mediante la incorporación sistemática del contexto situacional, las emociones y los sentimientos del usuario.

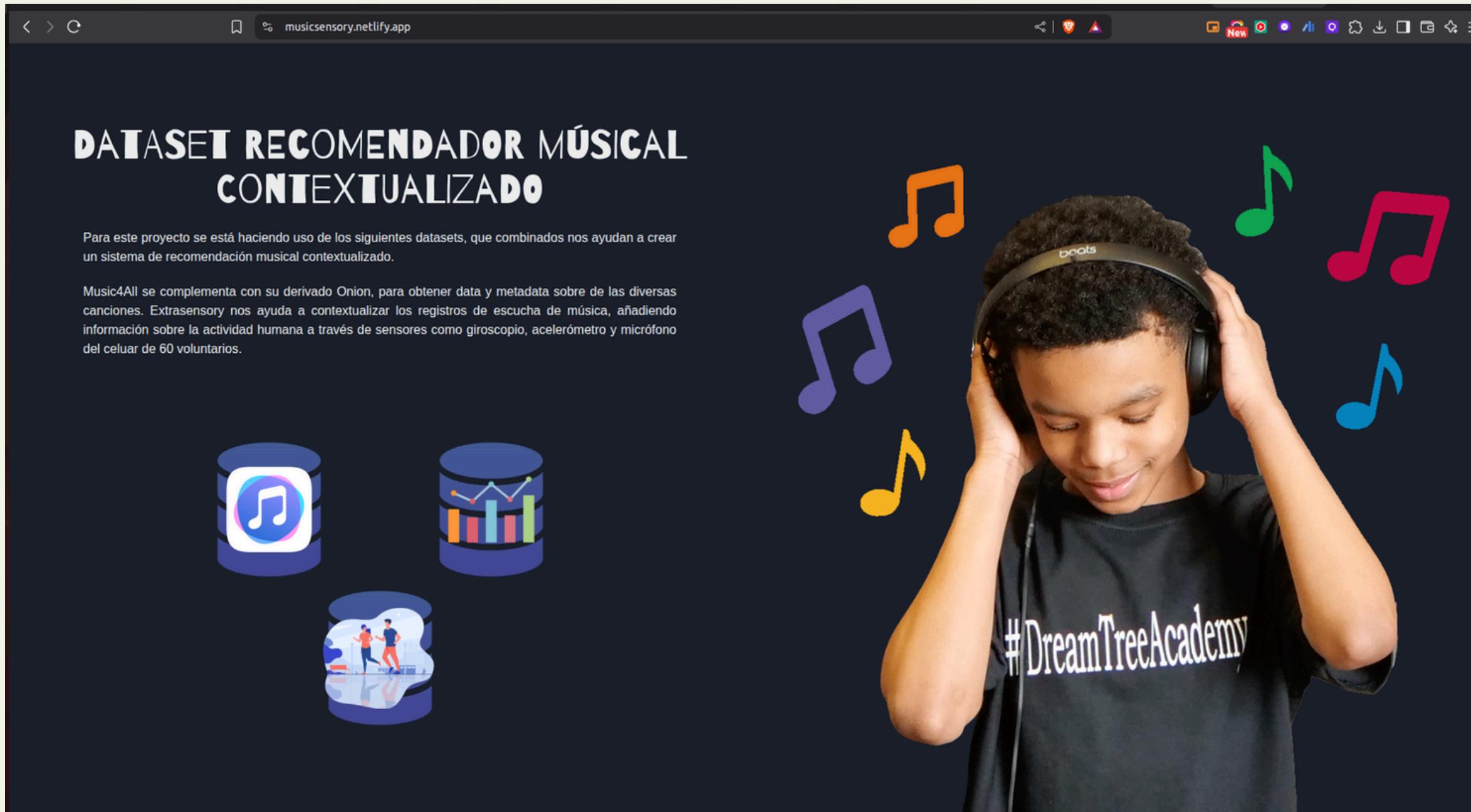
PIPELINE

Paso 1

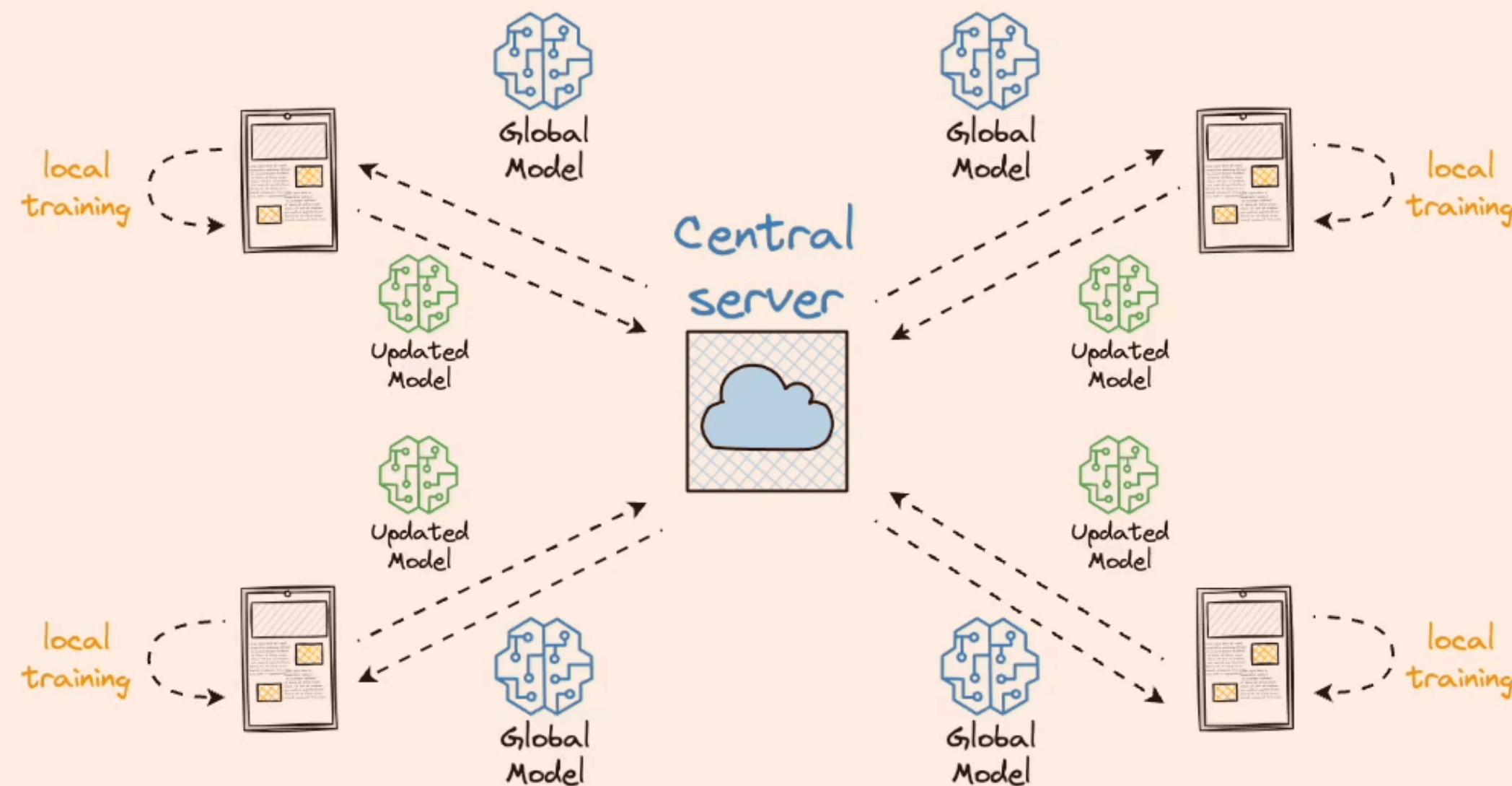


PIPELINE

Paso 2



Paso 2



DESCRIPCIÓN DEL DS

- Las bases de datos:
 - Canciones - 109270 registros [**\[Music4all\]**](#)
 - Listening events [**\[Music4all Onion\]**](#)
 - # of users: 119,140
 - # of tracks: 56,512
 - #records: 252,984,396

¿QUÉ TIPOS DE DATOS POSEEN?

Comenzando con Music4all, es un conjunto de diferentes archivos

```
music4all
└── audios/ (tiempo: 30s. formato: [ID].mp3)
└── lyrics/ (formato: [ID].txt)
└── 'id_genres.csv'
└── 'id_information.csv'
└── 'id_lang.csv'
└── 'id_metadata.csv'
└── 'id_tags.csv'
└── 'listening_history.csv'
└── 'readme.txt'
```

```
--- Información para ID: GoYo9wRiITsRDvYB ---

Archivo: id_genres.csv
id: GoYo9wRiITsRDvYB
genres: latin,reggaeton

Archivo: id_information.csv
id: GoYo9wRiITsRDvYB
artist: Bad Bunny
song: Si Estuviésemos Juntos
album_name: X 100PRE

Archivo: id_lang.csv
id: GoYo9wRiITsRDvYB
lang: es

Archivo: id_metadata.csv
id: GoYo9wRiITsRDvYB
spotify_id: 35wvL50xvKpCHEJPxLOLPI
popularity: 76.0
release: 2018
danceability: 0.672
energy: 0.594
key: 1.0
mode: 0.0
valence: 0.158
tempo: 171.854
duration_ms: 169248

Archivo: id_tags.csv
id: GoYo9wRiITsRDvYB
tags: latin,hip-hop,trap,puerto rico,puerto rican,reggaeton
```

MUSIC4ALL -- ONION

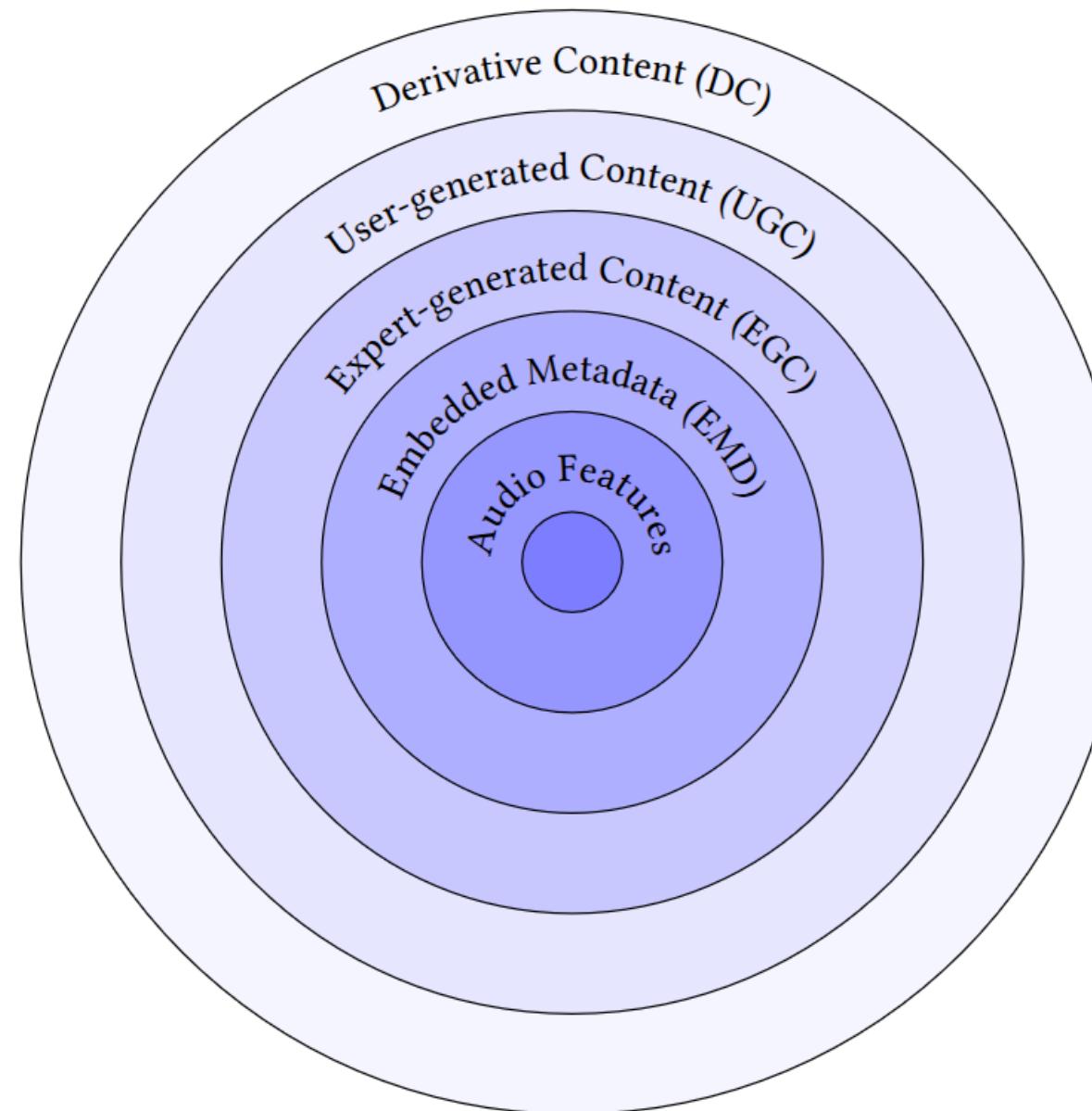
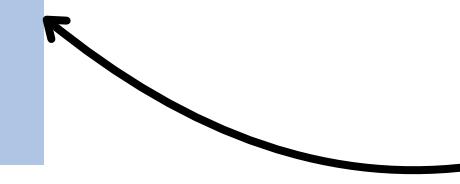


Figure 1: Onion model of music features.

- 📁 **Derivative_Content (DC)**
 - └── id_lyrics_sentiment_functionals.tsv.bz2
 - └── id_lyrics_tf-idf.tsv.bz2
 - └── id_lyrics_word2vec.tsv.bz2
- 📁 **User_Generated_Content (UGC)**
 - └── userid_trackid_count.tsv.bz2
 - └── userid_trackid_timestamp.tsv.bz2
- 📁 **Embedded_Metadata (EMD)**
 - └── id_ivec1024.tsv.bz2
 - └── id_ivec256.tsv.bz2
 - └── id_ivec512.tsv.bz2
 - └── id_resnet.tsv.bz2
 - └── id_vgg19.tsv.bz2
 - └── id_incp.tsv.bz2
 - └── id_tags_dict.tsv.bz2
- 📁 **Expert_Generated_Content (EGC)**
 - └── id_genres_tf-idf.tsv.bz2
 - └── id_tags_tf-idf.tsv.bz2
- 📁 **Audio_Features**
 - └── id_blf_correlation.tsv.bz2
 - └── id_blf_deltaspectral.tsv.bz2
 - └── id_blf_logfluc.tsv.bz2
 - └── id_blf_spectral.tsv.bz2
 - └── id_blf_spectralcontrast.tsv.bz2
 - └── id_blf_vardeltaspectral.tsv.bz2
 - └── id_chroma_bow.tsv.bz2
 - └── id_compare_audspec_stats.tsv.bz2
 - └── id_compare_f0_stats.tsv.bz2
 - └── id_compare_hnr_stats.tsv.bz2
 - └── id_compare_jitter_stats.tsv.bz2
 - └── id_compare_mfcc_stats.tsv.bz2
 - └── id_compare_pcm_stats.tsv.bz2
 - └── id_compare_shimmer_stats.tsv.bz2
 - └── id_compare_voice_stats.tsv.bz2
 - └── id_emobase_bow.tsv.bz2
 - └── id_emobase_f0_stats.tsv.bz2
 - └── id_emobase_lsp_stats.tsv.bz2
 - └── id_emobase_mfcc_stats.tsv.bz2
 - └── id_emobase_pcm_stats.tsv.bz2
 - └── id_emobase_voice_stats.tsv.bz2
 - └── id_mfcc_bow.tsv.bz2
 - └── id_mfcc_stats.tsv.bz2
 - └── id_essentia.tsv.bz2
 - └── id_vad_bow.tsv.bz2



```
User_Generated_Content (UGC)
├── userid_trackid_count.tsv.bz2
└── userid_trackid_timestamp.tsv.bz2
```



Columnas

```
User_Generated_Content (UGC)
├── userid_trackid_count.tsv.bz2
└── userid_trackid_timestamp.tsv.bz2
```

//user_id, track_id, count
//user_id, track_id, timestamp (yyyy-mm-dd hh:mm:ss)

Listening events [Music4all Onion]

of users: 119,140

of tracks: 56,512

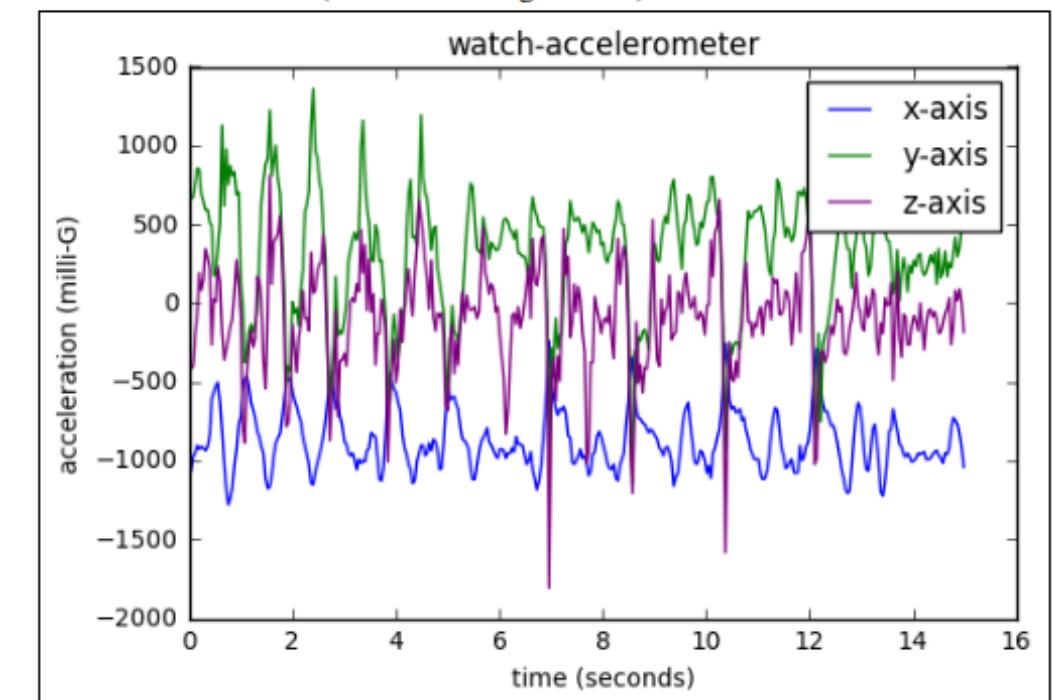
#records: 252,984,396

Tener en cuenta estos archivos

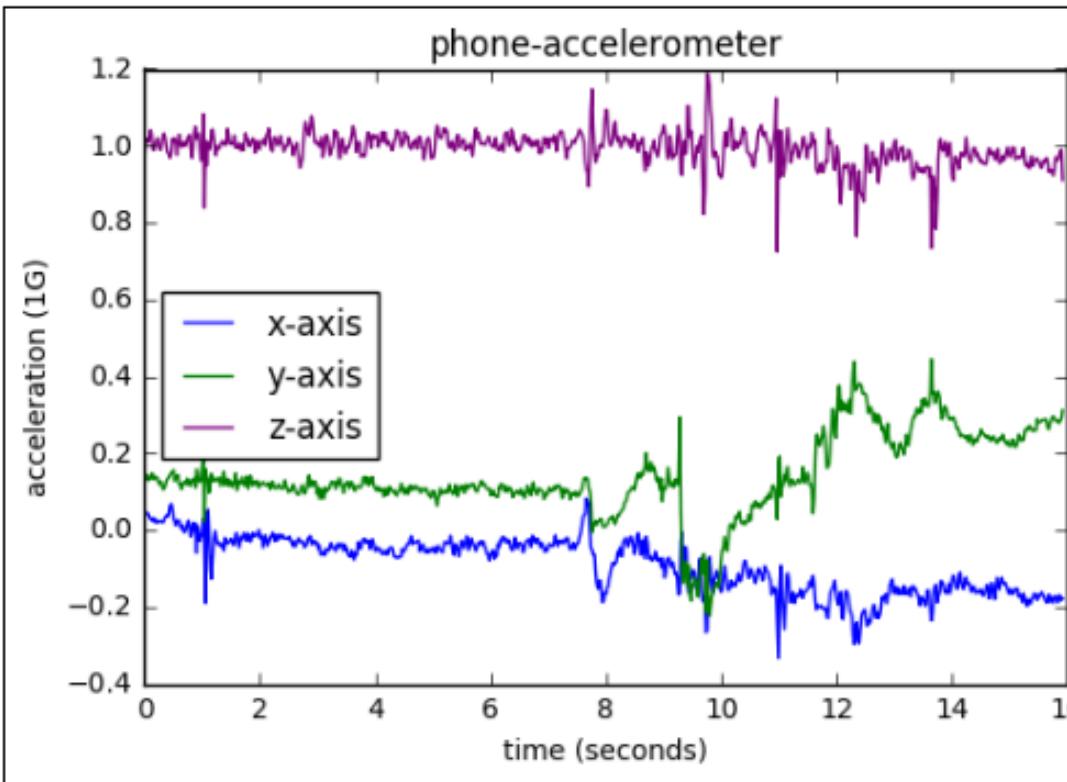
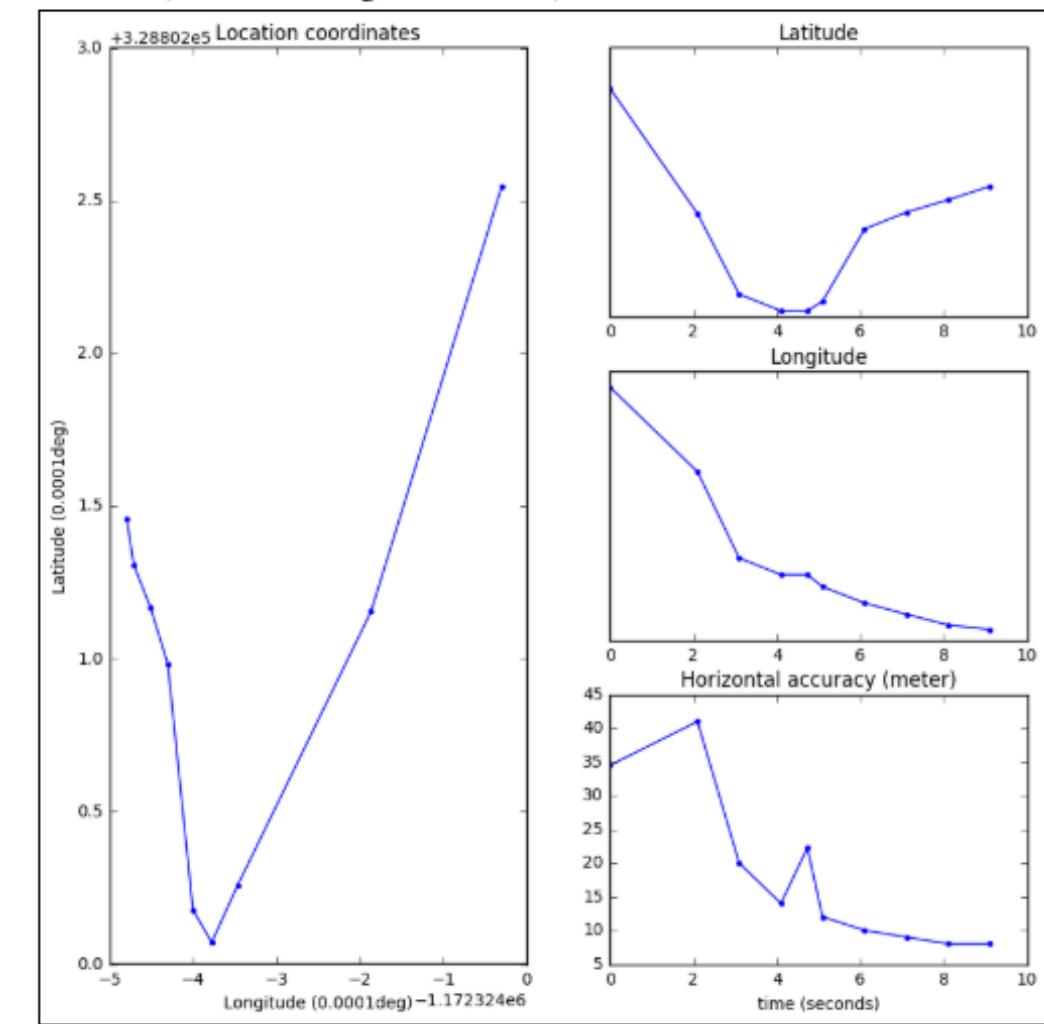
Extensivity

69 scenarios

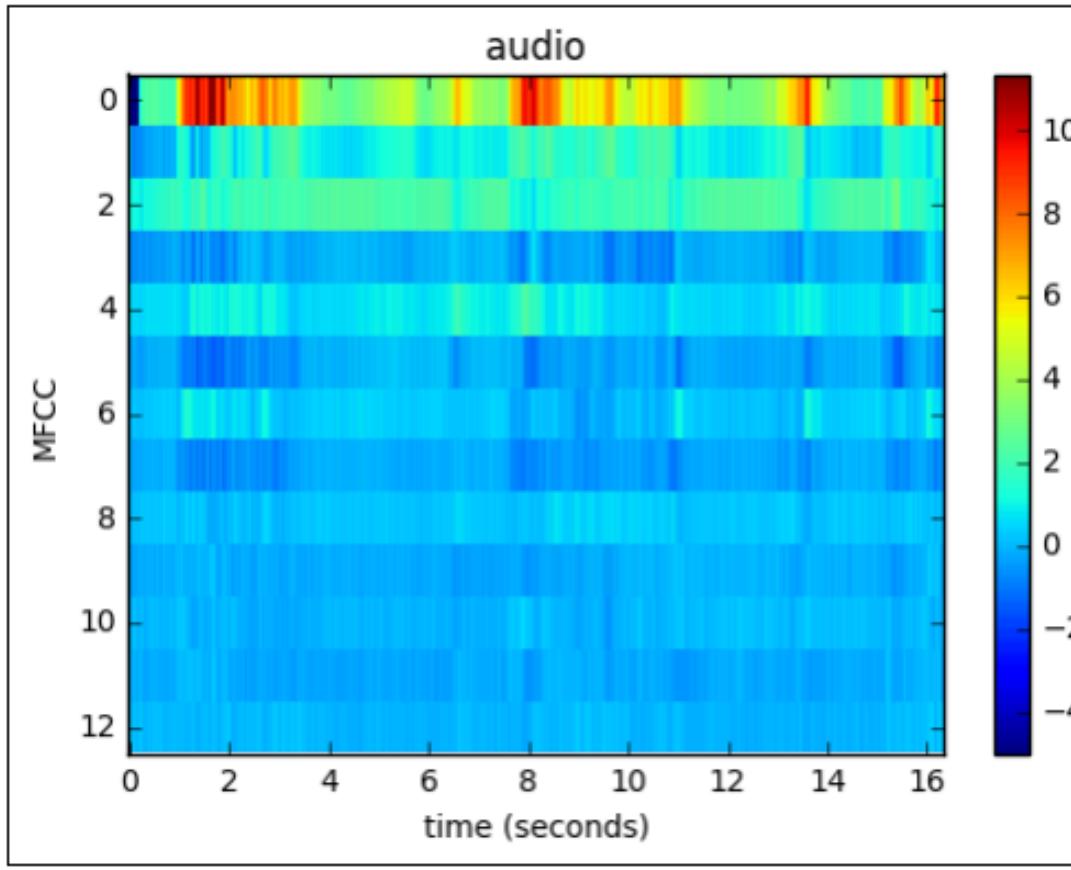
Watch-accelerometer (recorded during shower):



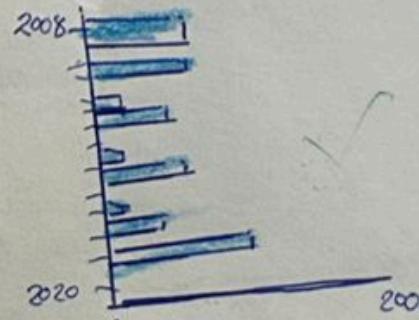
Location (recorded during drive in a car):



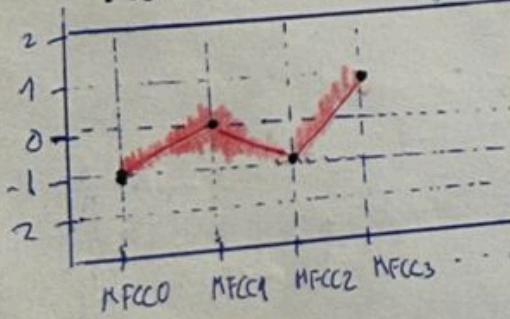
Audio (recorded while watching TV and eating at home):



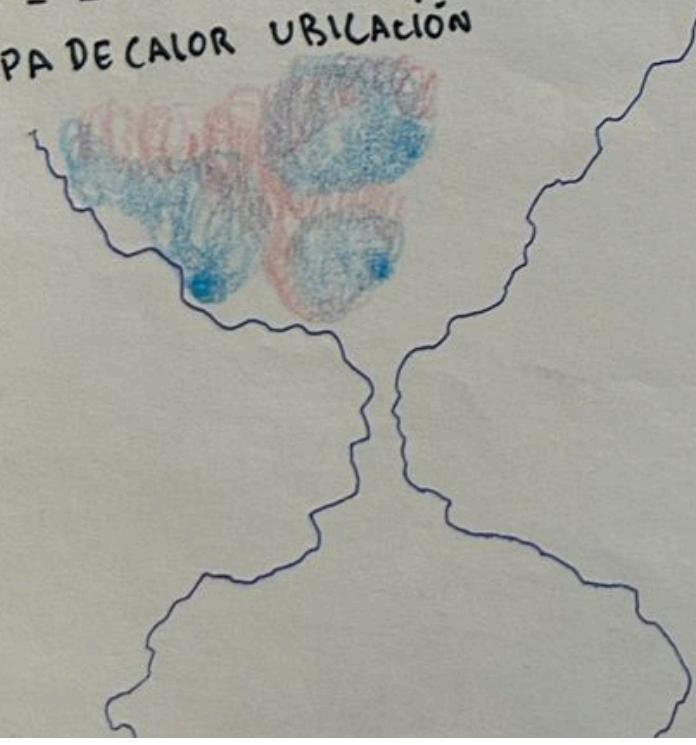
BOSQUEJO
de registros



MEDIA MFCC

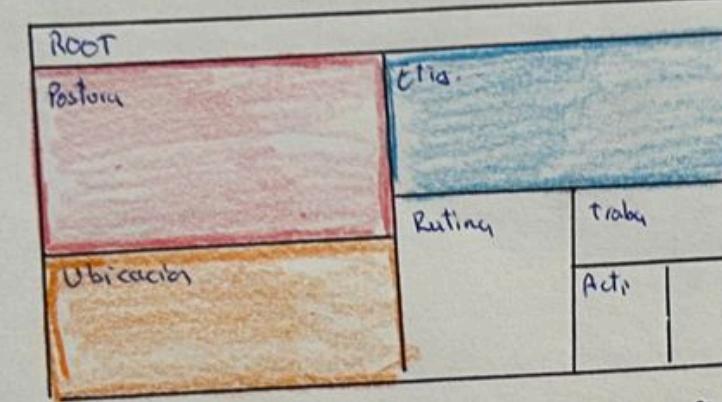


MAPA DE CALOR UBICACIÓN

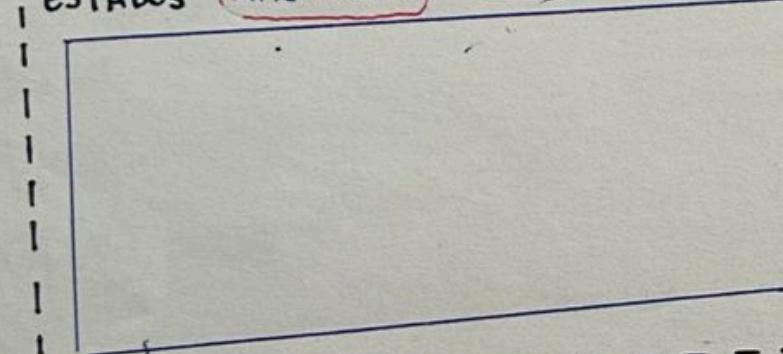


Grupos x label

(GENERIC) OR (LOC) (FIA) Erigota FIX
Correspondencia



ESTADOS (APLICACIÓN) (DISPOSITIVO)



ARTISTA más ESCUCHADO

PEPE GRILLO

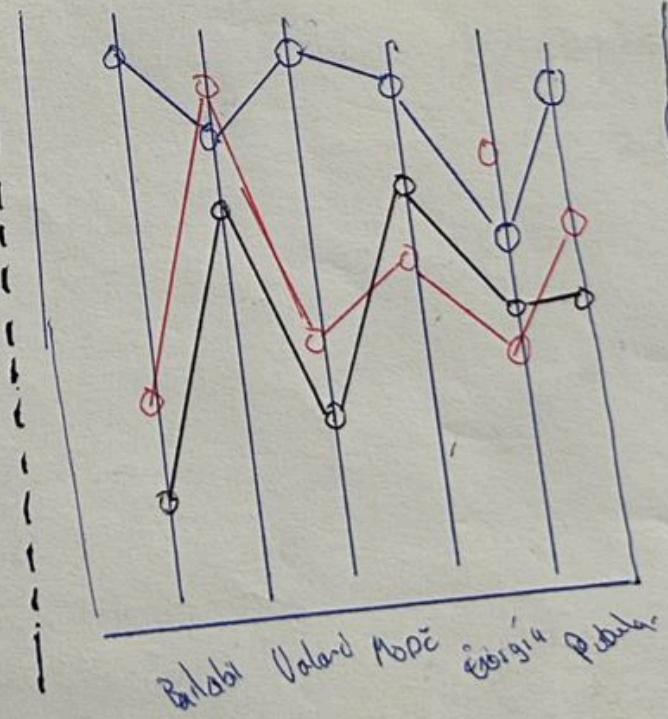
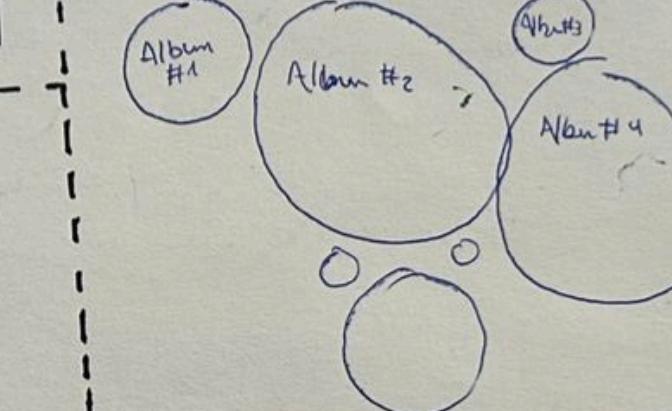


Canciones:

Álbunes

Principal lenguaje

Idioma que habla



FILTROS

ANÓ ▼

- General
- 2005
- 2006
- 2009
- :

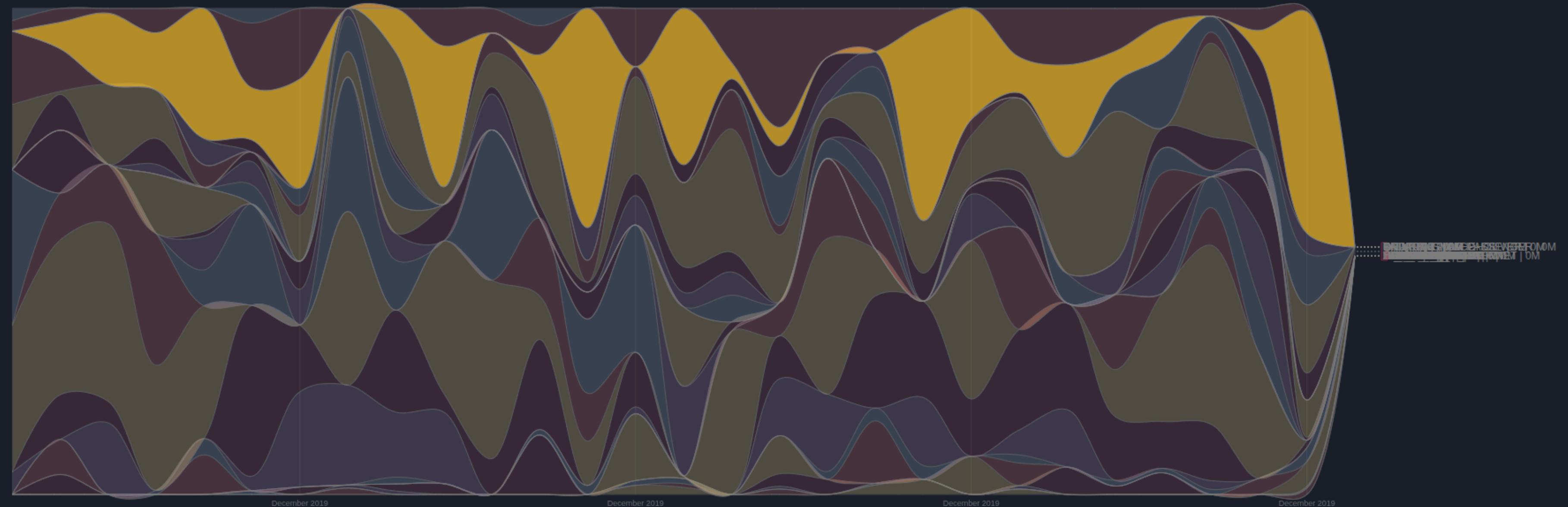
POLETIQUETAS ▼

- Actividad Física
- Transporto
- Ocio
- :

Pop Indi
ROCK
Clas'

STREAMED GRAPH

Data since December 2019



AN ACTOR-CRITIC BASED RECOMMENDER SYSTEM WITH CONTEXTAWARE USER MODELING

UN SISTEMA DE RECOMENDACIÓN BASADO EN EL ENFOQUE ACTOR-CRÍTICO CON MODELADO CONTEXTUAL DEL USUARIO

Autores: "Bukhari, Maryam and Maqsood, Muazzam and Adil, Farhan",
Artif. Intell. Rev., "Springer Science and Business Media LLC",

Feb, 2025

PIPELINE

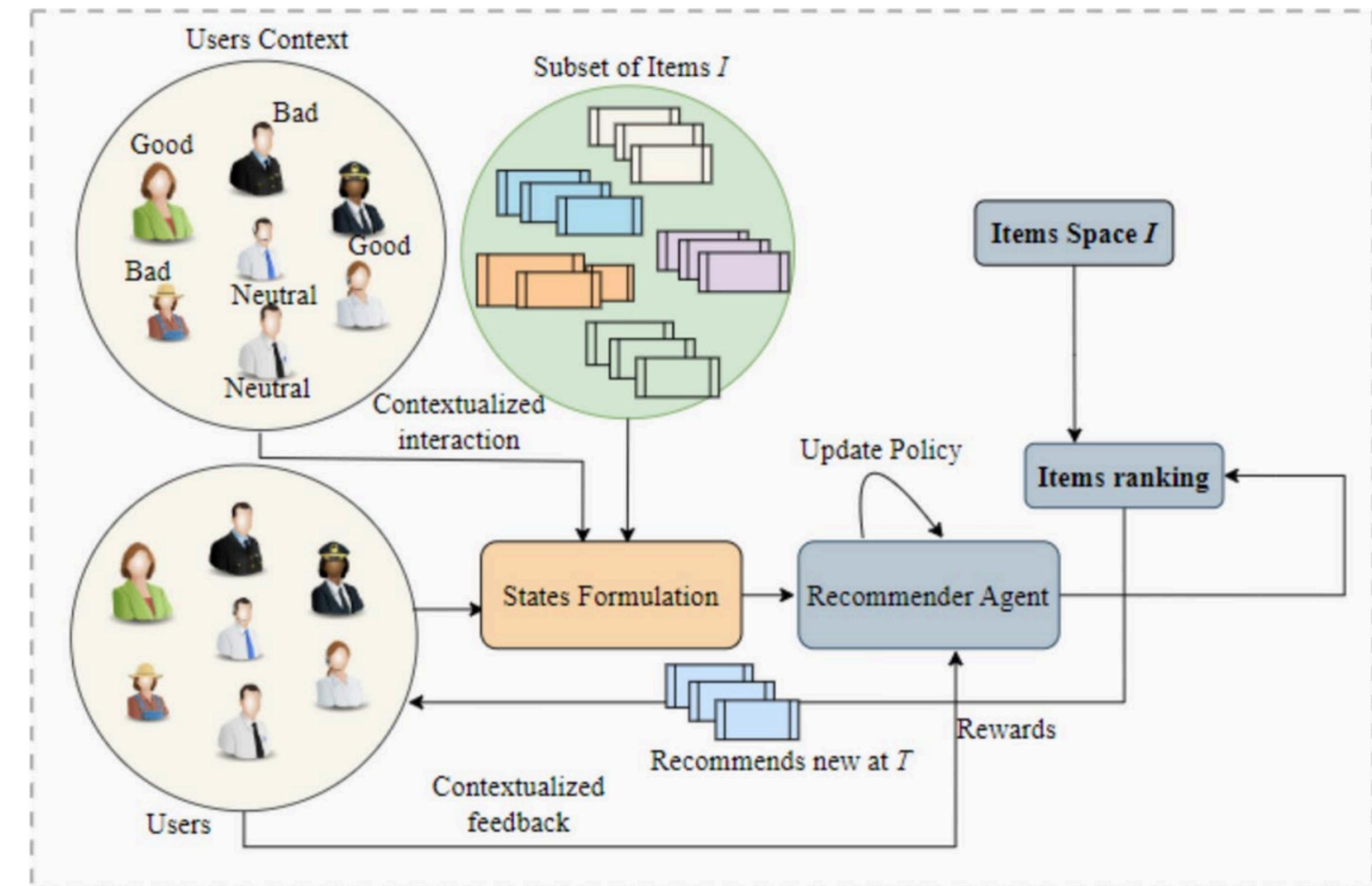


Fig. 1 Contextualized-User item interaction with the RL-agent

DATASETS

LDOS-CoMoDa Dataset

(Context-Aware Movie Recommendation Dataset)

Campos principales

- Usuario: ID, edad, sexo, ciudad, país
- Película: ID, director, país, idioma, año, géneros (3), actores (3), presupuesto
- Evaluación: Calificación (1–5)
- Contexto (12 variables):
 -  Hora del día: Mañana, Tarde, Noche
 -  Tipo de día: Laboral, Fin de semana, Festivo
 -  Clima: Soleado, Lluvioso, Tormenta, Nevado, Nublado
 -  Social: Solo, Pareja, Amigos, Familia, etc.
 -  Emociones: Al final y dominante (Feliz, Triste, Enojado, etc.)
 -  Estado físico: Saludable o Enfermo
 -  Decisión: ¿Quién eligió la película?
 -  Interacción: Primera o repetida

[\[DATASET PAGE\]](#) [\[DATASET LEGEND\]](#)

DePaulMovie Dataset

(Dataset Contextual de Recomendación de Películas)

Descripción general

Recopilado mediante encuestas a estudiantes de DePaul University, incluye 5,043 valoraciones realizadas por 97 usuarios y 79 películas, en tres dimensiones contextuales: tiempo, ubicación y compañía

Dimensiones y valores contextuales

- Tiempo (Time): Día de semana, Fin de semana
- Ubicación (Location): Hogar, Cine (u otro tipo)
- Compañía (Companion): Solo, en familia, con pareja

```
1 user_id:token,item_id:token,rating:float,time:token,location:token,companion:token,contexts:token,uc_id:token
2 1003,tt0454876,1,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
3 1003,tt0120912,4,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
4 1003,tt3793764,1,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
5 1003,tt0114148,3,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
6 1003,tt0110357,1,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
7 1003,tt0109830,5,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
1003,tt1707386,1,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
1003,tt0120338,2,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
1003,tt0088763,1,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
1003,tt0133093,3,Weekday,Cinema,Alone,Weekday_Cinema_Alone,1003_Weekday_Cinema_Alone
```



ARQUITECTURA ACTOR-CRÍTICO

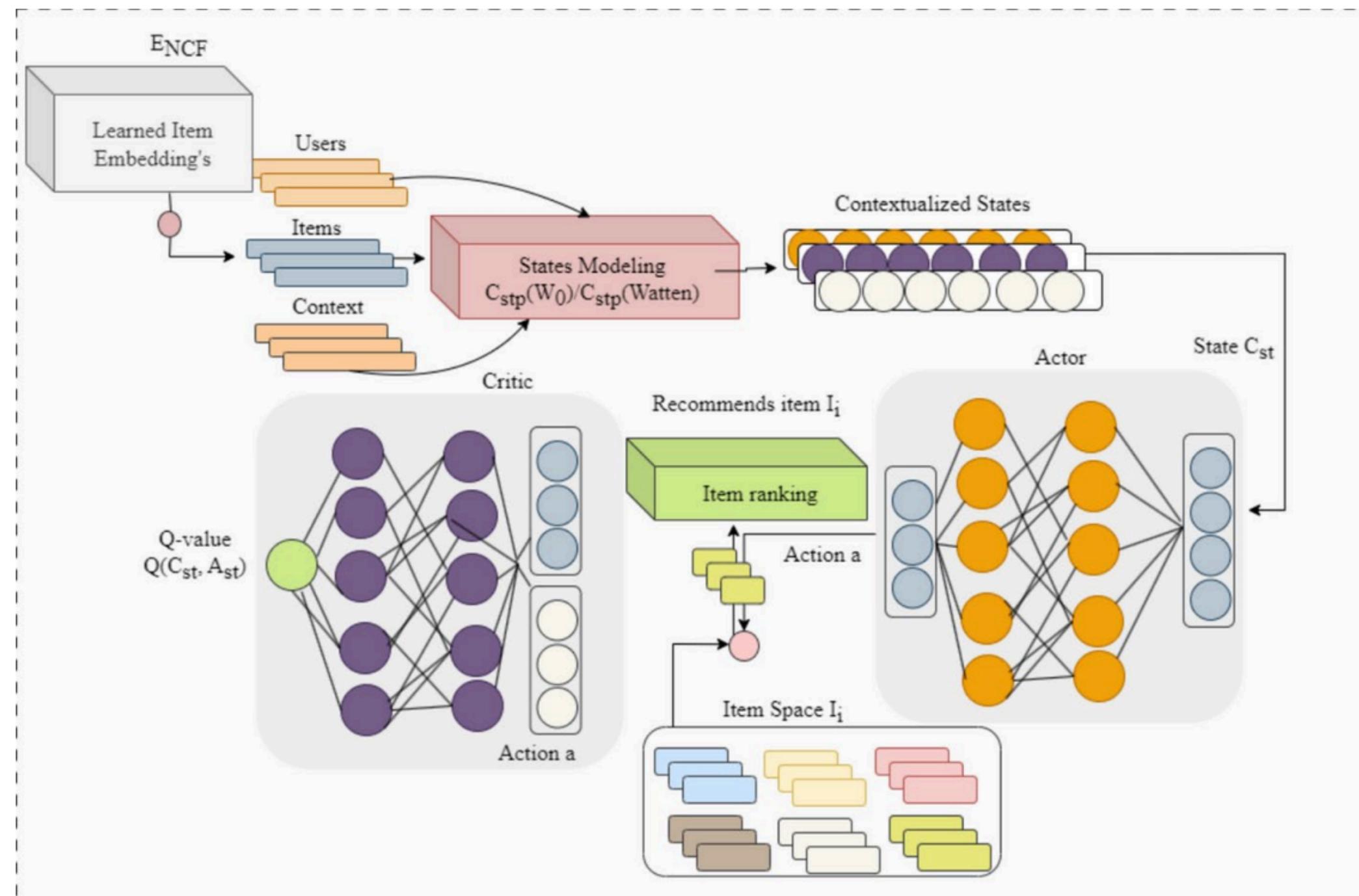


Fig. 2 A pictorial overview of the Actor-Critic model

RESULTADOS

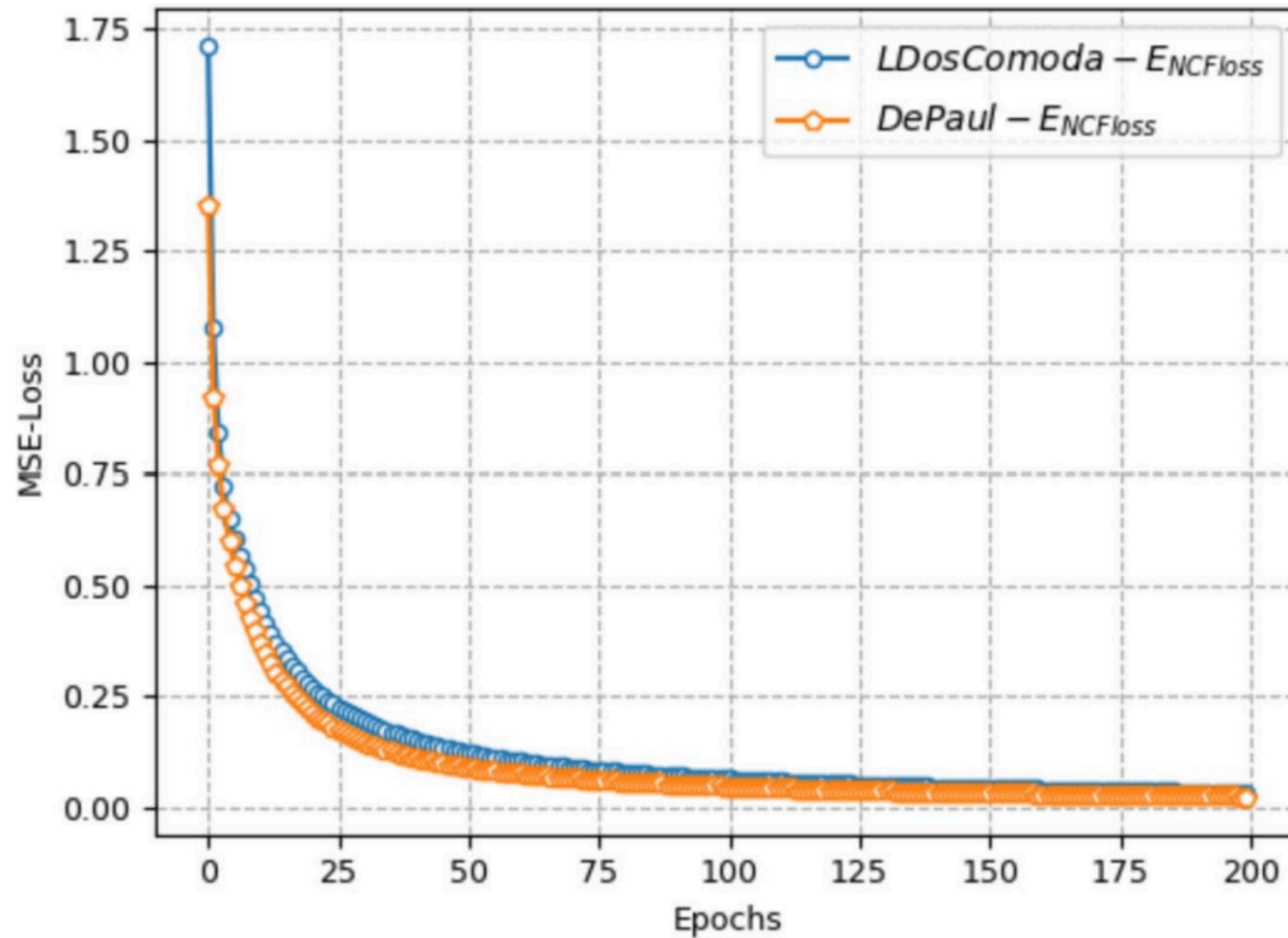
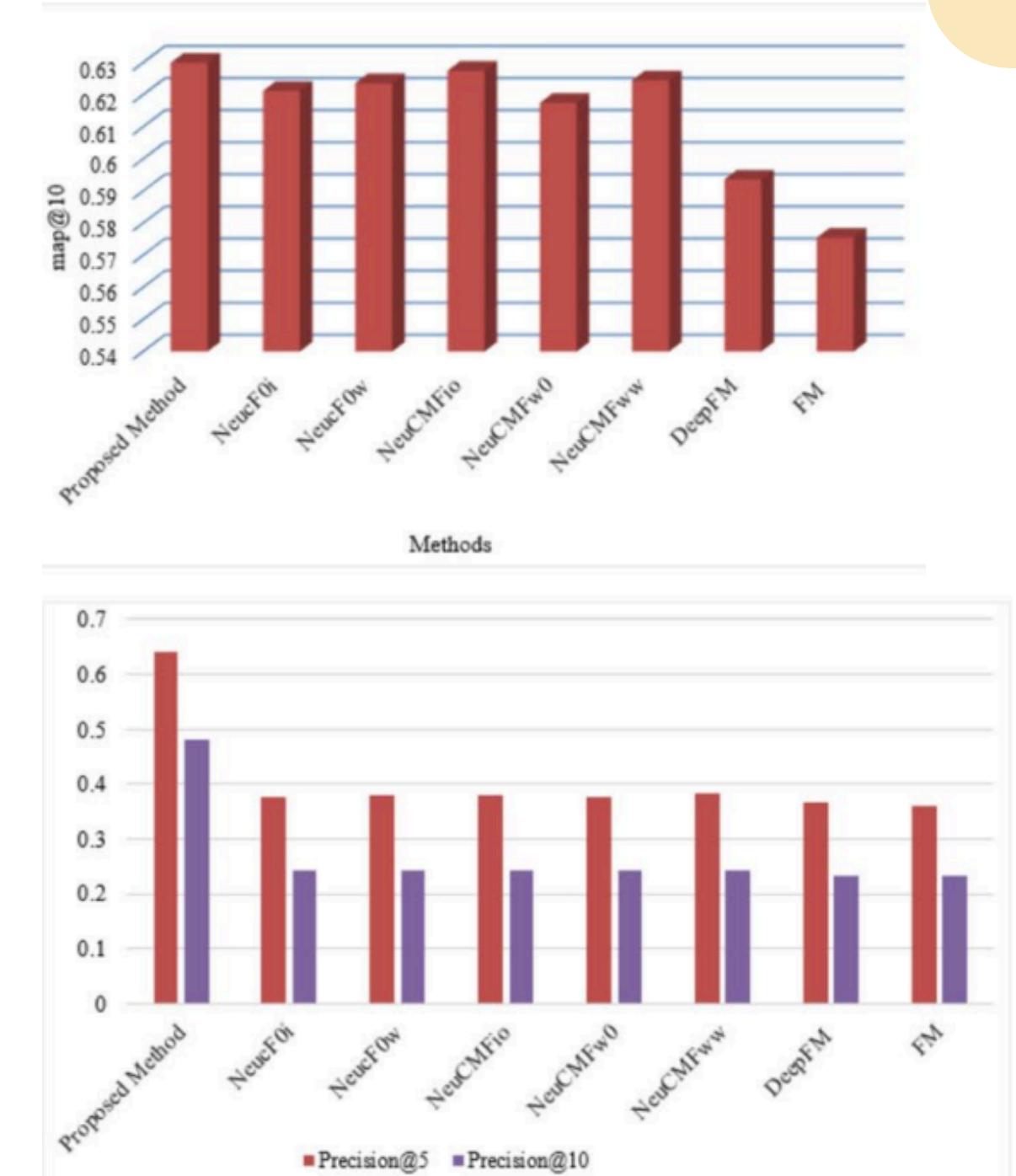


Fig. 11 Loss curves of E_{NCF} on Depaul movie and LDos-Comoda dataset



CONTEXT-AWARE SERVICE RECOMMENDATION BASED ON KNOWLEDGE GRAPH EMBEDDING

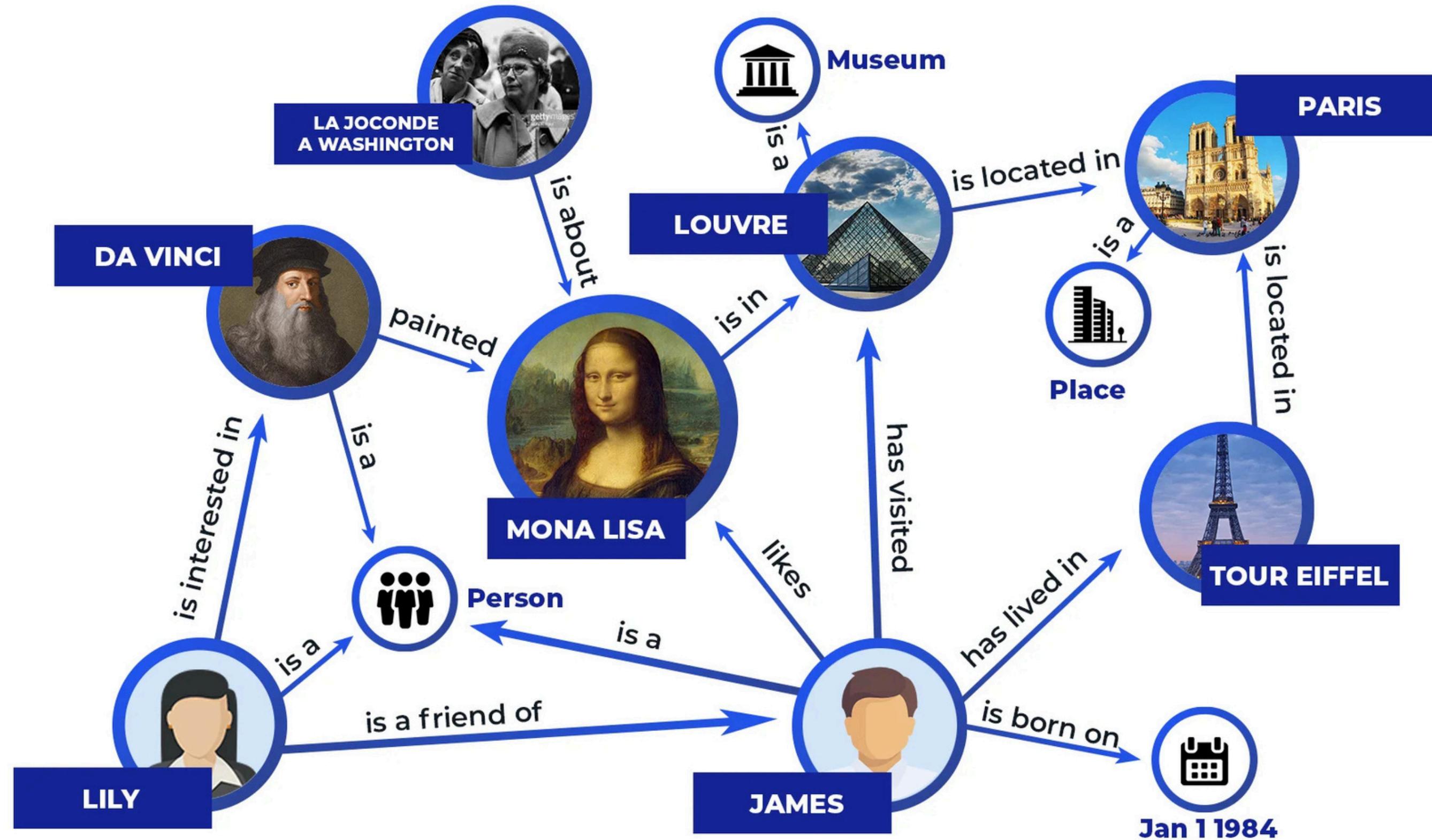
*RECOMENDACIÓN DE SERVICIOS SENSIBLE AL CONTEXTO BASADA EN
EMBEDDINGS DE GRAFOS DE CONOCIMIENTO.*

Autores: "Mezni, Haithem & Benslimane, Djamel & Bellatreche, Ladjel"
IEEE Trans. Knowl. Data Eng., Institute of Electrical and Electronics Engineers (IEEE)

Nov, 2022

PROBLEMAS EN LOS CARSS

- **Escalabilidad limitada:** Difícil de manejar grandes volúmenes de datos (QoS, contexto, reseñas, feedback).
- **Representación contextual deficiente:** Uso de matrices simples que no capturan relaciones complejas entre contextos.
- **Interacciones unidimensionales:** No consideran relaciones múltiples entre usuarios, servicios y contextos.
- **Enfoque limitado del contexto:** Se restringen a dimensiones como tiempo y ubicación, omitiendo otros factores relevantes.
- **Problemas de escasez de datos y cold-start:** Difícil hacer recomendaciones cuando hay poca información previa.
- **Falta de similitud contextual:** No se aprovecha la semejanza contextual al buscar usuarios o servicios similares.



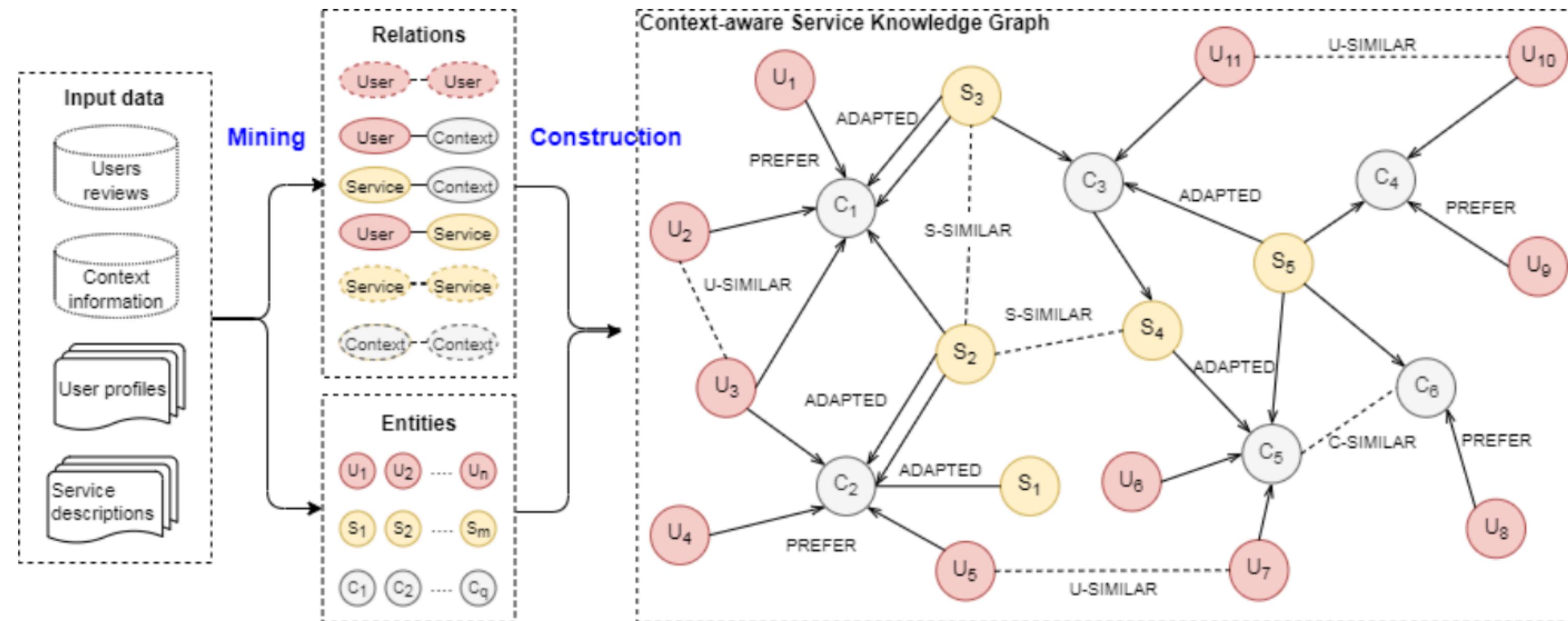


Fig. 1. Example of Contextual Service Knowledge Graph

TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING

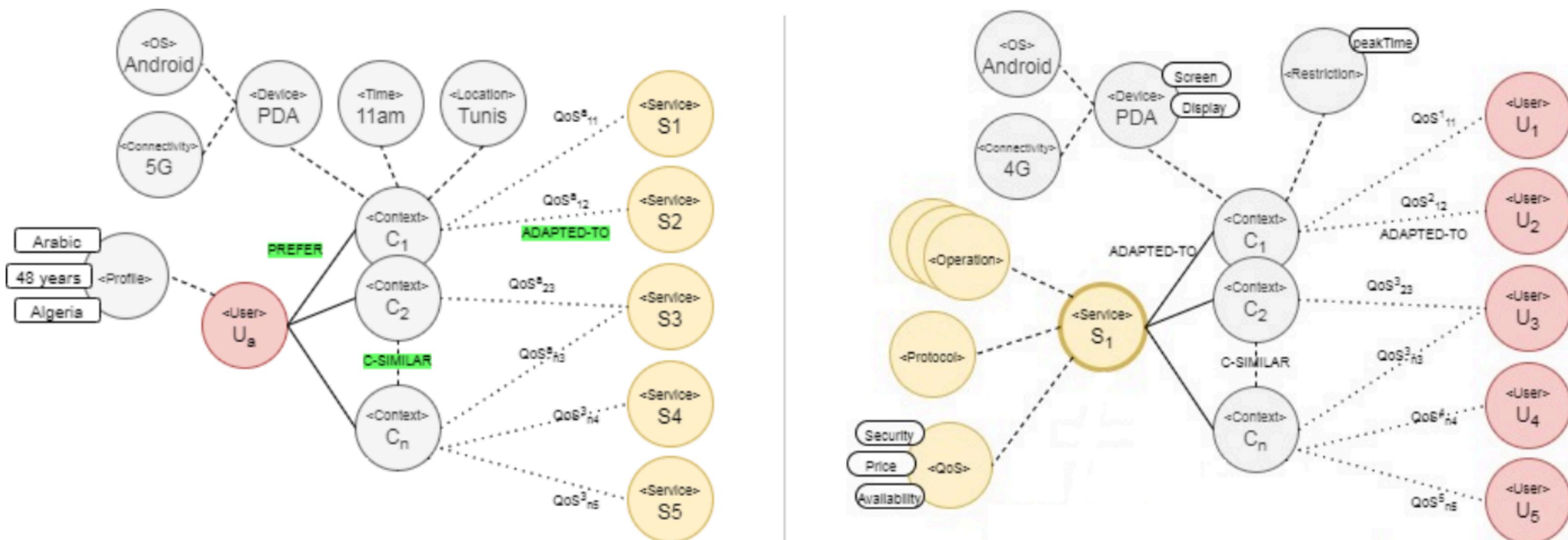


Fig. 2. (a) User's knowledge subgraph (b) Service's knowledge subgraph

PROPUESTA DEL ESTUDIO:

- Utilizar Knowledge Graph Embedding (KGE) para inferir nuevas relaciones entre usuarios, servicios y contextos.
- Superar las limitaciones de los métodos multi-vista tradicionales.
- Incorporar información contextual directamente en el proceso de embedding mediante el modelo C-SKG.

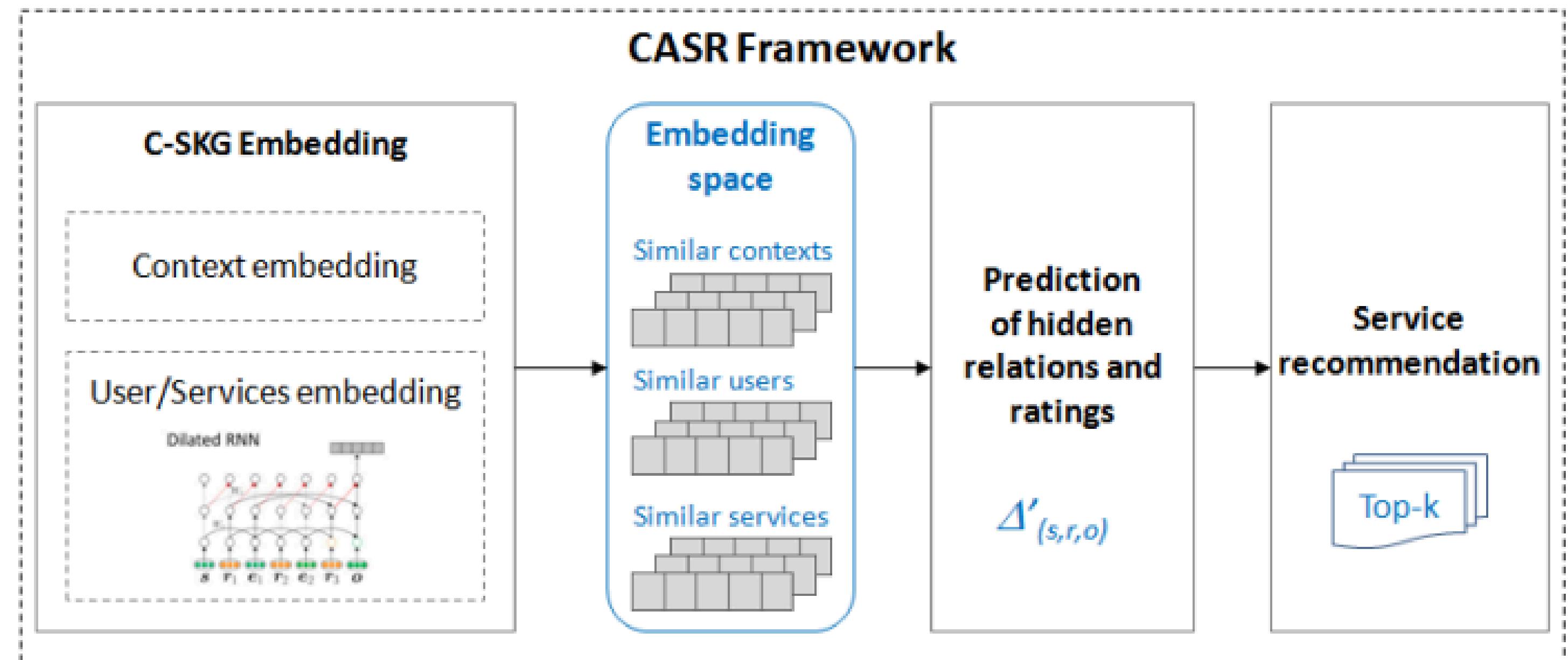


Fig. 3. Context-aware service recommendation process

VENTAJAS DEL ENFOQUE C-SKG:

- Aprendizaje conjunto de usuarios, servicios y contextos en el mismo espacio vectorial.
- Uso de medidas de proximidad de primer y segundo orden para capturar similitudes.
- Transformación del problema de recomendación en un problema de cálculo de proximidad entre nodos, facilitando la predicción de servicios con alta calificación.



Yelp Dataset – Características Principales

1. Datos de Negocios

- ID de negocio, nombre, dirección
- Categorías (ej. "Cafeterías", "Restaurantes")
- Atributos: horario, ambiente, estacionamiento, etc.
- Rating promedio y cantidad de reseñas

2. Reseñas

- Texto de la reseña
- ID de usuario y de negocio
- Calificación en estrellas (1-5)

3. Datos de Usuario

- ID y nombre de usuario
- Número de reseñas realizadas
- Promedio de calificaciones dadas
- Red de amigos

4. Check-ins

- Registros de visitas de usuarios a negocios

5. Tips

- Comentarios breves o recomendaciones sobre un negocio

6. Fotos

- Imágenes subidas por los usuarios sobre lugares o comida

DATASET