D7.4 – Data management plan v2

August 31th, 2019

Author/s: Daniel Molina (BMAT)

Contributor/s: FuturePulse Consortium

Deliverable Lead Beneficiary: BMAT

This project has been co-funded by the HORIZON 2020 Programme of the European Union. This publication reflects the views only of the author, and the Commission cannot be held responsible for any use, which may be made of the information contained therein.

Deliverable number or supporting document title	D7.4 Data management plan v2
Type	ORDP: Open Research Data Pilot
Dissemination level	Public
Publication date	31-08-2018
Author(s)	Daniel Molina (BMAT)
Contributor(s)	FuturePulse Consortium
Reviewer(s)	Stamatis Rapanakis (ATC), Rémi Mignot (IRCAM)
Keywords	DMP, Data management plan, ORDP
Website	www.futurepulse.eu

CHANGE LOG

Version	Date	Description of change	Responsible
V0.1	01/08/2019	Skeleton	Daniel Molina
V0.8	14/08/2019	Draft for peer-review	Daniel Molina
V1.0	31/08/2019	First version	Daniel Molina

Funded by the European Commission

Table of Contents

Funded by the European Commission

Funded by the
European Commission

1 Executive Summary

This document defines the data policy and data management procedures in *FuturePulse: Multimodal Predictive Analytics and Recommendation Services for the Music Industry*, Grant Agreement number 761634 ICT H2020. The purpose of this DMP is to provide an overview of the main elements of the data management policy and data sources that will be used by the Consortium.

This document is the second version of the DMP, delivered in Month 24 of the project. It corresponds to an update of the first version of the DMP (D7.2) delivered on M6, including the recommendations made by the project reviewers in the first review of FuturePulse, held in Brussels on October 31st, 2018. In any case, the DMP is not fixed, this document will evolve during the lifespan of the project (from September 1st 2017 to August 31st 2020).

This document follows the Guidelines on FAIR Data Management in Horizon 2020[1] and the Guidelines to the Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020[2]. The document follows the template provided by the European Commission in the Participant Portal and will be submitted to European Commission as D7.4 Data Management Plan.

FuturePulse participates in the Open Research Data Pilot[3]. The Open Research Data Pilot aims to make the research data generated by Horizon 2020 projects accessible with as few restrictions as possible, while at the same time protecting sensitive data from inappropriate access.

The consortium has chosen Zenodo[4], the open research repository from OpenAIRE and CERN, as the central scientific publication and data repository for the project outcomes.

Notes for version V2

This version of the DMP includes information answering recommendations received by the project reviewers during the first review process of FuturePulse, held in Brussels in October 2018. This recommendations were:

The document is written in generic terms. It needs to be revised to include more specific information on data usage policies by 3rd parties and more generally how personal data is handled (see also comments on D2.1).

According to these recommendations, a specific section of Terms of Service and strategy has been included with information and analysis on data usage policies by 3rd parties.

1 https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf
2 https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
3 https://www.openaire.eu/what-is-the-open-research-data-pilot
4 https://zenodo.org

Funded by the European Commission

2 Key terminology: Definitions and acronyms

AWS	Amazon Web Services
EC	European Commission
DMP	Data Management Plan
Data Repository	a) General term used to refer to a destination designated for data storage. b) Setup within an overall IT structure, such as a group of databases, where an enterprise or organization has chosen to keep various kinds of data (https://www.techopedia.com/definition/23341/data-repository).
DOI	Digital Object Identifier: persistent identifier or handle used to uniquely identify objects, standardised by the International Organisation for Standardisation (ISO).
FTP	File Transfer Protocol
GDPR	EU General Data Protection Regulation (https://www.eugdpr.org)
H2020	Horizon 2020 Programme
Metadata	Administrative information to help manage a resource, such as when and how it was created, file type and other technical information, and who can access it.
Zenodo	Catch-all OpenAIRE compliant repository for EC funded research, hosted by CERN since May 2013 (https://zenodo.org).

Funded by the European Commission

3 FuturePulse project

3.1 Abstract

Music is one of the fastest evolving media industries, currently undergoing a transformation at the nexus of music streaming, social media and convergence technologies. As a result, the music industry has become a mixed economy of diverse consumer channels and revenue streams, as well as disruptive innovations based on new services and content distribution models. In this setting, music companies encounter daunting challenges in dealing successfully with the transition to the new field that is shaped by streaming music, social media and media convergence. The availability of huge music catalogues and choices has rendered the problems of recommendation and discovery as key in the competition for audience, while the continuous access to multiple sources of music consumption have resulted in a dynamic audience, characterised by a highly diverse set of tastes and volatility in preferences which also depend on the context of music consumption.

To serve the increasingly complex needs of the music ecosystem, FuturePulse will develop and pilot test a novel, close to market music platform in three high-impact use cases:

- Record Labels,
- Live Music,
- Online Music Platforms.

The project will help music companies leverage a variety of music data and content, ranging from broadcasters (TV, radio) and music streaming data, to sales statistics and streams of music-focused social media discussions, interactions and content, through sophisticated analytics and predictive modelling services to make highly informed business decisions, to better understand their audience and the music trends of the future, and ultimately to make music distribution more effective and profitable. FuturePulse will offer these capabilities over a user-friendly, highly intuitive and visual web solution that will enable the immersion of music professionals in the realm of music data, and will support them to make highly informed and effective business decisions.

3.2 Project Scope and Objectives

In response to the industrial needs of the music industry the FuturePulse project has identified the following six specific technological and innovation objectives:

- Objective 1: Deliver a single tool for collecting and accessing music data from a diverse set of sources.
- Objective 2: Deliver a set of data-driven services for estimating the current and future popularity of songs, artists and genres.
- Objective 3: Deliver a set of services for enhanced audience analysis and management.
- Objective 4: Integrate music data collection, mining, and visualisation in a scalable Software-as-a-Service (SaaS) platform.
- Objective 5: Perform large-scale pilots on three clearly defined music segments.

Funded by the European Commission

- Objective 6: Develop and execute a comprehensive dissemination and exploitation plan and pave a clear path to market.

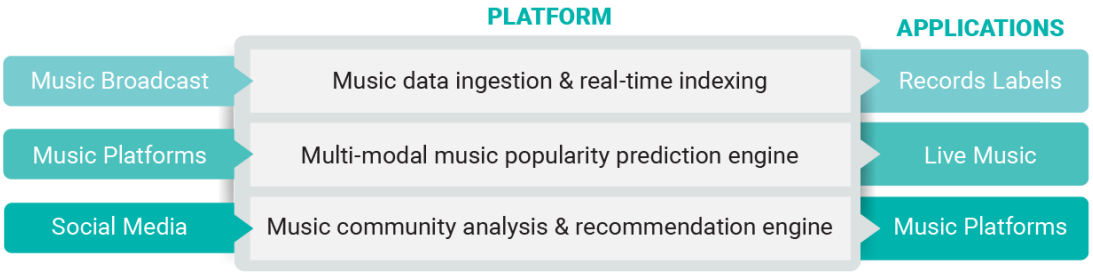

Figure 1: FuturePulse Platform and Applications

The project will result in a number of high-quality outcomes that will form the basis for the exploitation plan of the project. These are illustrated in the figure above and include the following:

- A robust and extensible multi-source music data ingestion and real-time indexing framework.
- A multi-modal music popularity prediction engine: This will produce short- and long-term predictions for popularity indices about specific artists, albums, songs, styles and genres, given a variety of incoming signals.
- An online music community analysis framework and a music recommendation engine.
- An integrated scalable cloud-based platform offering the full spectrum of FuturePulse services.
- Three market-driven applications serving the needs of record labels, event organisers and music platform.

Funded by the European Commission

3.3 Project participants

BMAT licensing S.L.	Daniel Molina dmolina@bmat.com	
Athens Technology Center S.A	Eva Jaho e.jaho@atc.gr	
Ethniko Kentro erevnas Kai Technologikis Anaptyxis	Symeon Papadopoulos papadop@iti.gr	
Musimap S.A.	Frédéric Notet frederic@musimap.com	
Institut de Recherche et de Coordination Acoustique Musique	Rémi Mignot remi.mignot@ircam.fr	
Playground Music Scandinavia AB	Anders Engström anders@playgroundmusic.com	
Soundtrack Your Brand Sweden AB	Daniel Johansson danielvinkar@gmail.com	
Advanced Music S.L.	Nacho Moya nacho.moya@sonar.es	

Table 1: Project participants

3.4 Coordinator contact

Daniel Molina	BMAT	dmolina@bmat.com

Funded by the European Commission

4 Data Summary

FuturePulse will produce several datasets during the lifetime of the project. The main purpose of these datasets is to feed the three use cases of the project: Record Labels, Live Music and Online Music Platforms[5]. The data will be both quantitative and qualitative in nature and will be analysed from a range of methodological perspectives for project development and scientific purposes.

FuturePulse consortium has performed a thorough work in order to define and update the different data sources used in the project. This work has been extensively reflected in the following deliverables:

- D2.1 – Data specifications and collection v1 (M9)
- D1.4 – FuturePulse requirements v2 (M15)
- D2.3 – Data specifications and collection v2 (M21)

This section briefly describes the data usage detailed in D2.3:

- Data requirements and abstract data model
- Data types and taxonomies used in FuturePulse and the implementation of third party data sources for data collection and processing
- Status of the implementation of all the data sourcing, collection and analysis components
- Management of the collected data

Deliverable D2.3 is an extension of Deliverable D2.1 with updated and additional information since the first definition of data specifications, and an additional overview of the current status, using Deliverable D1.4 as starting point. Thus, it includes a status-quo of which data sources are used for which of the requirements of the FuturePulse platform, and their status of implementation with respect to collection, processing and integration. It also mentions data sources which are planned to be used at a later point but not yet in place.

The list of data sources for the FuturePulse platform may still change. Furthermore, we will consider including additional data sources over the course of the project in case we find that their inclusion adds value to the platform and is feasible.

4.1 Data collection

FuturePulse partners have analysed the data needs (storage, indexing and retrieval) for the FuturePulse platform and have identified a diverse number of data sources that should be considered as inputs. These sources include both publicly available services and sites and private sources where consortium partners have privileged access and/or are in talks about partnership agreements with the providers of those sources.

The data collection activities in FuturePulse include:

1. Selected set of entities of interest for the FuturePulse pilots. This includes 2,383 artists and 41,544 tracks from PGM for the Record Label use case, 6,527 tracks from BN for the Live Music use case, and 39,467 tracks from SYB for the Background Music Provider use case.

5 http://www.futurepulse.eu/use-cases

Funded by the European Commission

2. Curated (and expanding) list of public music charts. This includes 754 charts from 84 countries containing in total 17,856,458 data points, and can contribute to track recognition and genre popularity prediction.
3. Additional data collection based on pilot-independent data sources, i.e. tracks or artists added in terms of a third party testing the FuturePulse pilots.

The main collections of data used and/or considered for the FuturePulse platform are detailed in the following subsections.

4.1.1 Core music entity metadata

It consist of artists, albums (released by artists), tracks contained in the albums, and finally the associated music genres.

We use International Standard Recording Code (ISRC)[6] for uniquely identifying sound and music video recordings, together with a FuturePulse internal identifier aimed at avoiding fragmentation merging the ISRCs corresponding to the same track.

Finally, in order to proceed with the content analysis, audio files of targeted tracks are also needed:

- PGM shared a representative set of their artists' tracks and the associated metadata. The dataset consists of artists that are or have previously been signed to PGM as well as a catalogue acquired through company mergers or acquisitions. PGM also has access to the corresponding data in platforms such as Facebook, YouTube, Spotify, Deezer and Apple Music and has granted access to these channels where applicable. The shared dataset consists of 2,383 artists, having 5,687 albums with 41,544 tracks in total.
- SYB provided a metadata set of 39,467 tracks, annotated with Song Title, Artist Name, eventual Album Name, ISRC codes, Spotify ids and URI´s as well as links to the associated YouTube videos. We were further able to annotate the tracks with additional metadata, perform acoustic content analysis and also identify the corresponding albums and artists. Furthermore, Deezer popularity ranks for those tracks have been added for this data set.
- BN delivered a set of 6,527 audio files to illustrate the specific genres of electronic music. Those are delivered as mp3, wav, m4a and mp4 files accompanied by the corresponding genre of electronic music.
- For the automatic analysis of audio recordings, collections of annotated songs have been provided by the pilots of the project: 4,835 tracks for the musical genre classification (from the SYB dataset), 4,947 tracks for the vocal gender prediction (from the SYB dataset), and additionally 4,501 tracks for the electronic genre classification (from BN). All these annotations are given in JSON files, and are used by IRCAM for the model training.

4.1.2 Broadcast data (radio, TV monitoring)

Broadcast data is one of the most representative mainstream means of consuming music since the invention of radio and TV. We currently consider a variety of the most relevant channels (most played radios in key markets), with the flexibility to add more channels based on the needs of the project.

6 http://isrc.ifpi.org/en

Funded by the European Commission

Up to date, BMAT is monitoring a total of 400 channels from different European countries. Regarding venues, BMAT has added the monitoring of 100 clubs around Europe.

4.1.3 Music streaming platform data (Spotify, Deezer, etc.)

We collect and analyse data from streaming services platforms as a key factor in the estimation of success for the entities of interest, mainly based on the number of times a track has been played and the unique number of listeners for this track in each platform. In addition to these direct measures, streaming platforms expose other attributes that may be considered as proxies of success; for example, how many users of the platform follow an artist (without necessarily implying that they listen to their tracks).

We currently support six known platforms: Spotify, Apple Music, Deezer, YouTube, Last.fm and SoundCloud.

4.1.4 Social media entities (e.g. Facebook pages, YouTube channels) and associated data (likes, comments, etc.)

We monitor profiles corresponding to artists of interest, and in messages associated with all entities of interest. The social media platforms currently used as a source by FuturePulse include Twitter and Facebook.

4.1.5 Playlists and chart data

Music charts are rankings according to specific criteria, i.e. total sales in case of physical albums, airplays in radio broadcasting, the number of downloads in case of digital music and finally the amount of streaming activity as measured by the corresponding platform. The most common period is one week, one month or even daily charts. With respect to FuturePulse requirements, charts are an important source to consider in estimating the current and future popularity and recognition of an artist, track and genre.

In addition, playlists tracking gives many signals to estimate track and artist success. By monitoring a representative set of playlists, i.e. tracking how playlists change by the inclusion and exclusion of tracks, it is possible to identify those tracks that have the potential to attract many listeners in the future.

4.1.6 Live music events and venues

We collect lists of past gigs of artists and venues where artists have performed as important indicators of popularity and success. We also consider venues alongside events as core entities in the FuturePulse platform.

4.1.7 Data derivatives (popularity, recognition, trend analysis etc.)

The data described in the previous sections are used to derive additional data:

- *Music performance metric representation*: We define a general performance tuple to represent a measure of success for an entity of interest at a certain point in time: *(Metric, Value, Timestamp, Platform, EntityID, Tags)*
- *Popularity level:* We consider popularity is a numeric score (between 0 and 100) that reflects the degree that a particular artist, album or track is currently listened or actively followed by an audience.

Funded by the European Commission

- *Recognition level:* Recognition is a numeric score (also between 0 and 100) that is associated with a track and reflects the degree that this track is recognisable by a listener.
- *Further derivatives*: Increase per day of YouTube viewers and Spotify popularity, as an absolute and as a normalized relative value, and similar such derivative data.
- *Music attributes*: FuturePulse also offers sophisticated content analysis to extract useful information from the raw audio files provided by the partners and music industry players which hold access to the audio (e.g. labels).
- *Open Data and other sources*: FuturePulse use other open and publicly available sources to provide music metadata: MusicBrainz and Wikidata.

4.2 Data sources specification

4.2.1 Genre taxonomies

About genres, since there is not a widely adopted comprehensive music genres taxonomy, in the context of FuturePulse we have defined our own set of genres based on already existing categorisations used by the three pilot partners.

Genres List (PGM / SYB joint genres)			
African	Alternative	Ambient	Americana
Bass	Blues	Breakbeat	Children's Music
Christian	Christmas	Classical	Country
Dance/EDM	Dancehall/Reggaeton	Death Metal	Disco
Doom Metal	Downbeat	Drum & Bass / Jungle	Dubstep
Electronic	Electronica	Experimental	Folk/Folklore
Funk	Garage	Hardcore	Hard Rock
Heavy Metal	Hip Hop	House	House/Techno
Indie	Industrial	Inspirational	Instrumental
Jazz	Karaoke	Latin	Lounge
Mariachi	Metal	Musical	Opera
Pop	R&B	Reggae/Dub	Rock
Rockabilly- Psychobilly	Salsa	Samba/Bossa Nova	Singer-Songwriter
Soul	Soundtrack	Spoken Word	Surf
Tango	Tech House	Thrash Metal	Trip Hop

Table 2: Genres defined by PGM and SYB

In addition, BN created an electronic music taxonomy: a first level that corresponds to a generic electronic genre and a second that indicates the subgenre inside a given first-level genre. Level 1 contains 22 terms while level 2 contains 187 subgenres.

Genre Family	Subgenres
Ambient	Drone music, Ambient dub, Dark ambient, Ambient industrial
Breakbeat	Big beat, Nu skool breaks, Florida breaks, Miami bass, Acid breaks, Broken beat, Nu-funk, Baltimore club, Jersey club

Funded by the European Commission

Disco	Space disco, Disco polo, Afro-Cosmic music, Nu-disco, Euro disco, Italo disco
Downtempo	Acid jazz, Trip hop, New age, Space music, Chill-out
Drum and bass	Hardstep - Darkstep, Funkstep, Funkstep/Funkstep soul, Neurofunk, Jump Up, Techstep, Liquid
Dub	Dub poetry, Dubtronica, Dancehall, Dub reggae
Electro	Freestyle, Electro swing
Electronica	Ethnic electronica, Funktronica, Livetronica Laptronica, Folktronica
Electronic rock	Space rock, New wave, Coldwave, Indietronica, Alternative dance, Minimal wave, Electropunk, Post-punk, Dance-punk, Ethereal wave, Dark wave, Krautrock, Electroclash, Nu-gaze, New rave, Synthwave, Synth-pop, Electronicore
Hardcore	Speedcore, Breakcore, Gabber mainstream, Industrial Hardcore - Darkcore - Terror, Frenchcore, Happy Hardcore, Digital Hardcore, Breakbeat hardcore, Hardstyle, Hardstyle/Raw, hardstyle, Hardstyle/Dubstyle
Hi-NRG	Bubblegum dance, Eurodance, Eurobeat, Italo dance
Hip Hop	Electro, Trap, G-funk, Contemporary R&B, Electro hop, Hardcore hip hop, Neo soul, Drill, New jack swing, Alternative hip hop
House	Tropical house, Hard dance, Witch house, French house, French touch, Acid house, Future house, Hard NRG, Moombahcore, Funky house, Garage house, Moombahton, Tech house, New, beat, Outsider house, Big room house, Juke, Nu NRG, Kwaito, Nu jazz Jazz house, Hard house, Subground, Progressive house, Microhouse Minimal house, Fidget house, Tribal, house, Ghettotech, Chicago house, Hardbag, Deep house, Hip house, Complextro, Kidandali, Footwork, Diva house, Ghetto house, Latin house, Ambient house, Italo house, Electro, house, Dutch house
IDM	Wonky, Glitch & Glitch hop
Industrial music	Dark electro, Industrial metal, Neue Deutsche Härte, Electro-industrial, Futurepop, Death industrial, Power electronics, Electronic Body Music, Power noise, Industrial rock, Aggrotech, Cybergrind, Japanoise
Jungle	Ragga jungle, Raggacore, Darkcore dnb
Musique électroacoustique	Musique concrète, Musique acousmatique, Musique mixte
Post-disco	Boogie, Hardvapour, Vaporwave, Dance-rock, Chillwave, Electropop, Dance-pop
Techno	Schranz, Hardtek, Minimal, Industrial techno, Dub Techno, Acid, Detroit
Trance	Dream trance, Vocal trance, Psychedelic trance - Full on - Suomisaundi, Acid trance, Euro-trance, Balearic trance, Progressive trance, Hard trance, Goa trance, Uplifting, trance Nitzhonot, Tech trance
UK garage	Future bass, 2-step garage, Dubstep, Grindie, Future garage, Drumstep, Speed garage, Reggaestep, UK Funky, Bassline, Breakstep, Brostep, Grime
Video game music	Skweee, Bitpop, Nintendocore, Chiptune

Table 3: Bass Nation electronic music genre taxonomy

Funded by the European Commission

The Deep Learning model provided by Musimap via API to predict moods from audio is currently able to predict 58 of those moods, which were selected according to market needs and AI confidence. Table 4 lists the 59 moods the FuturePulse platform is able to provide for tracks with audio (following a 3-level taxonomy as well), thanks to Musimap's analysis and API:

Level 1	Level 2	Level 3
Above (Fire)	Imagination	Dreaming
	Self-Control	Inspired
	Spirituality	Meditative
Down (Metal)	Coldness	Bitter, Depressed
	Sensibility	Delicate, Sad, Anxious, Sentimental
	Withdrawal	Discreet
In/Within (Water)	Love	Intense, Sensual, Glamorous
	Nourishment	Cool, Revitalizing, Friendly, Soothing
	Intellect	Analytical
On (Ground)	Playfulness	Performing
	Warrior	Determined, Powerful, Heroic
	Roots	Organic
Out (Wood)	Good Vibrations	Warm-hearted
	Manliness	Rebellious, Aggressive
	Extroversion	Impulsive
Up (Air)	Happiness	Innocent, Atmospheric, Free, Happy, Lively
	Dynamism	Energetic
	Temperament	Funny, Melodramatic, Humorous

Table 4: Musimap predicted moods taxonomy (available in FuturePulse)

4.2.2 Music attributes

A summary of the music attributes of a track are presented in Table 5.

Attribute	Description
Genre	Genre of the track following the different taxonomies used in the project (Section 4).
BPM	Beats Per Minute (from 10 to 360) of the track. It is supposed that the tempo does not change over the duration of the track. If it changes, it provides the median tempo value.
Fade in/fade out	Duration in seconds of the Fade-in (if existing) and Fade-out (if existing).
Major/minor mode	Mode (Major or minor) of the musical key used in the track. It is supposed that the key does not change over the track duration. If it changes, it provides the most used key.
Vocal gender/ instrumental	Gender (Male/female) of the main singer(s) in the track. If there is no singer, it returns the value "instrumental".

Funded by the European Commission

Moods	Moods of the track following the Musimap proposed taxonomy (Table 4). More than a single mood can be associated with a track.
Brand Values	Brand Values according to the SYB taxonomy (Table 12). More than one brand value can be associated with a track.
Energy level	Energy level of the track (according to a scale from 1 to 10).

Table 5: Music attributes extracted through content analysis

4.2.3 Broadcast data

The broadcast data collection provided by BMAT is based on the continuous monitoring of the selected radio channels, where BMAT identifies each use of the monitored catalogue of tracks.

Resource	Method[7]	Description
Artist	Artist Info	Get the metadata for an artist on Vericast.
Channel	Channel Info	Get the metadata for a channel on Vericast.
	Channel List	Get the list of monitored channels on Vericast.
Charts	Charts top artists	Get the top artists chart, ordered by playcount.
	Charts top channels	Get the top channels chart, ordered by playcount.
	Charts top labels	Get the top channels chart, ordered by playcount.
	Charts top tracks	Get the top tracks chart, ordered by playcount.
	Charts resolved tracks	Get the unknown tracks that have been resolved.
Label	Label info	Get the metadata for a label on Vericast.
Match	Match info	Get the metadata for a match on Vericast.
	Match list	Get all matches ordered by datetime.
Track	Track info	Get the metadata for a track on Vericast.
	Track search	Search the tracks by a query term. Searches in artist and track name.

Table 6: Vericast API methods to obtain airplay counts

4.2.4 Music Charts and Event data

FuturePulse performs the data extraction process of: (i) music charts, (ii) DJ charts and (iii) events. As a result, a unified model was created to describe these three entities. This model is depicted in the following tables:

Attribute	Type	Description
name	string	chart name
id	string	chart identification number
link	string	chart url
type	string	type of chart (album/track/artist/video)
entries	array of entry objects	music chart entries (Table 8)

Table 7: Chart's object attributes

Attribute	Type	Description
position	number	position of entity on chart

7 As Vericast API is a paid service, the exact calls are not specified in the table.

Funded by the European Commission

title	string	entity's title
artist	string	artist name
type	string	entity type (album/track/artist/video)
since	date	starting date of chart
until	date	ending date of chart
country	string	country code
source	string	url of chart
chart_id	string	chart identification number
spotify_track_id*	string	Album identification number on Spotify
spotify_album_id*	string	Track identification number on Spotify
spotify_artist_id*	array of string	A List of Artist identification numbers on Spotify related with the corresponding record
spotify_link*	string	A corresponding link to the Spotify platform
isrc*	string	The International Standard Recording Code of the corresponding record
genres*	array of string	List of genres mapped to the artists involved in the record

Table 8: Music chart entry data model

*attributes acquired from Spotify by using Music Charts Annotator component of the Data Extraction System.

Attribute	Description
title	Title of the chart
source	Link of the chart
date	Date the chart was created
artist	Artist who created the chart
artist_id	Artist identification number
genres	An array of Genres of the music covered by the chart. Each genre is described by an id, name, link (on beatport) and type.
entries	An array of DJ chart entry object (Table 10).

Table 9: DJ charts data model

Attribute	Description
id	Identification number of track
position	Position of entity on chart
title	Track title
mix_type	Track's mix type
release_date	Date of track's release
genres	An array of track's Genres. Each genre is described by an id, name, beatport link and type
artists	An array of track's Artists. Each artist is described by an id, name and beatport link.
remixers	An array of Remixers. Remixers are also artists and described by id, name, beatport link.
labels	An array of track's Labels. Each label is described by an id, name, beatport link.
link	Link of the track on beatport
release_link	Link of the release on beatport
release_id	Associated release identification number of the track

Table 10: DJ chart entry data model

Attribute	Description

Funded by the European Commission

id	Event's identification number
name	Event's name
venue	Each venue is described by an id, name and address
artists	An array of artists. Each artist is described by a name and RA link.
date	Event's date
area	Geographical area of the event. Each area is described by an id and name.
country	Country code

Table 11: Events data model

4.2.5 Brand values

With respect to the Background Music Provider's use case, 20 brand-oriented values will be used to annotate the tracks provided by SYB: acoustic, careful, discreet, down to earth, dreamy, easy going, electronic, elegant, exclusive, expressive, human, inclusive, mature, modern, provocative, rugged, serious, technological, traditional, youthful. These annotations form pairs of opposites are depicted in Table 12.

Youthful	Mature
Modern	Traditional
Inclusive	Exclusive
Elegant	Rugged
Down to earth	Dreamy
Careful	Provocative
Serious	Easy going
Discreet	Expressive
Human	Technological
Acoustic	Electronic

Table 12: Pairs of opposite brand values related to SYB's use case

4.3 Public Datasets

The following section provides an overview of the different data sets to be published by the project. At the moment of submitting the current version of the Deliverable there is one dataset published in Zenodo: T-REC Song Recognition Dataset.

Firstly, a summary of the dataset is described including the origin of data, type and format and size. Secondly, we provide information about description and purpose, utility and reuse for each dataset.

For each dataset identified, the following sections/questions are addressed:

- Data set reference and name
- Data set description
- Standards and metadata
- Data sharing
- Archiving and preservation (including storage and backup)

Dataset #1	T-REC Song Recognition Dataset
Name/ID	DOI: 10.5281/zenodo.3255311

Funded by the European Commission

Provider/s	CERTH
Description	Contains the track names, artist(s), total number of responses, measured recognition (user study), computed recognition (T-REC) and measured recognition on specific demographics i.e. male, female and age groups 18-24, 25-34, 35-44, 45-54, 55-65 for 100 music tracks.
Type/Format	CSV file (tab delimited)
Size (MB)	6.6Kb
Purpose	The conducted research and the development of the song recognition model T-REC addresses FuturePulse requirement BMP_REQ#1
Utility	Background music providers supply companies with music playlists with the purpose of optimizing the in-store experience of their customers and their brand perception. Having effective means of estimating song recognition can provide such companies with a useful tool for generating better playlists.
Scientific publications	Used in the paper "Data-driven song recognition estimation using collective memory dynamics models" accepted for publication in the ISMIR 2019 conference
Integration / Reuse	The recognition scores for the tracks contained in this dataset have been integrated in FuturePulse platform and are exposed through a corresponding API endpoint.
Data sharing	The dataset is published in Zenodo platform. Link
Archiving and preservation	The FuturePulse datasets stored in Zenodo will be preserved a minimum of 5 years according to the European Commission Data Deposit Policy. There are currently no costs for archiving data in this repository.

Table 13: T-REC Song Recognition Dataset

4.4 Terms of Service and risks

The Consortium has analysed the Terms of Service of each data source used in FuturePulse in order to guarantee that the data can be used during the project and after the project subject to win-win agreements with the parties providing the data.

This analysis was performed in a two-step process: firstly, all partners listed the currently used data sources, by adding links to Terms Of Services of the corresponding platforms. In a second step, these documents were analysed by legal counsels.

According to the results of this expertise, the Consortium has set up a plan to contact third party platforms for agreement when needed. All details about legal aspects have been provided in D2.3.

The music data presented in deliverables D2.1 and D2.3 are mainly obtained from third party services and websites. A source analysis has been started along two aspects: first a risk plan has been crafted to anticipate the impact in the case where access to one or more of these data sources stops, and second legal aspects, such as the Terms

Funded by the European Commission

of Service, have been studied in order to know the official limitations (if any) of data collection and use in the context of FuturePulse.

For the period of the project, the study of the Terms of Service revealed that only minor limitations are applied for research purposes, and they do not impact the predictive analyses of FuturePulse. However, for a commercial exploitation, the terms of use of some sources are more restrictive, and some of these data providers ask for an agreement. Consequently, an action is currently conducted to prepare the requests of authorisations for the commercial exploitation of FuturePulse.

4.4.1 Risk analysis

Different input data are used by FuturePulse for tasks, such as the prediction of genre popularity and artist recognition as described in D3.1. Given that when several data sources can be used together for a relevant prediction, the loss of one source can be largely compensated by others in that case. However, some sources are more important because the information they provide cannot be replaced by others.

In the first version of the Data Management Plan, cf. the Annex of deliverable D7.2, a contingency analysis has highlighted these important sources. All the sources have been listed task per task, then the more critical ones have been exhibited. They are: *Spotify*, *Kontor New Media*, *Beatport*, and *Resident Advisor*. The following Table presents a summary of the analysis presented in D7.2.

Note that even if *Spotify* can be partly replaced by *Deezer*, no replacement solutions have been found for these other important sources for now. In consequence, the consortium has to anticipate any possible loss, and to monitor the future creations of new services potentially usable in replacement.

Tasks	Used sources	Comments
Genre Popularity and Track Recognition	*40 charts*	Possible issues at country level.
Track Recognition	*Youtube views, Spotify popularity*	More relevant prediction if both sources are available.
Genre Popularity and Annotation of entities	***Spotify***	Spotify is used: ● for its ids to reference entities ● to annotate artists and tracks to a musical genre. Deezer can partly replace Spotify
Artist Popularity	*Spotify, Deezer, Soundcloud, Last.fm, Facebook, Youtube, and Twitter*	The loss of one source is acceptable.
	Bandsintown, Resident Advisor	The loss of one source is acceptable.
Record Labels requirements	***Kontor New Media***	Allowed access for the data of PGM.
Annotation of entities	*Wikipedia, Wikidata, MusicBrainz, Discogs*	The loss of one source is acceptable.
Emerging Artist Detection	***Beatport***	Mandatory.
	Resident Advisor	Mandatory for a requirement of the Live Music use case.

Table 14: Data providers risk analysis

Funded by the European Commission

4.4.2 Terms of Service and strategy

Even if the readings of the Terms of Service (ToS) revealed that the data use for research and development purposes are permitted, for commercial applications the policies of the data providers are different from each other. But when the restrictions impact the predictive analyses of FuturePulse, an agreement can be obtained from the third party.

Nevertheless, the obtainment of the authorisation of use for a commercial application is strongly dependent on the technical progress of the project and of the Exploitation Plan of FuturePulse. In consequence, for some sources, it is too soon to formally apply for an agreement. The strategy of the consortium is to advance separately for each source as far as possible given the progress of the project and given the respective policy of the data provider. A Memorandum of Understanding has been written by the legal team at BMAT in order to sign it with data providers when needed.

For each source, we present the conclusion on the Terms of Service, for the points of view of the project period (research and development purposes) and of the commercial application. Then when it is required, the strategy and the advance to obtain an approval are presented.

Wikidata and MusicBrainz

The data of *Wikipedia*, *Wikidata* and *MusicBrainz*, are under Creative Common license (CC-0[8]). Consequently, the use of these data are very permissive, even for commercial use.

Discogs

According to the ToS of *Discogs*[9], the commercial uses are generally permitted: "*Commercial use of Our API and the Content is generally permitted, but may not be permitted, if, in our sole discretion we determine the commercial use is prohibited*". Consequently, the use of *Discogs* as data source is not going to be problematic for the period of the project, and should not be in the case of an exploitation of FuturePulse after the project.

Facebook

For the social network *Facebook*, the restrictions of use of its service only concern the user data, and as presented below (cf. Section 7.2), FuturePulse does not collect any personal data. The ToS are visible at: https://developers.facebook.com/policy.

Twitter

This other well-known social network practices similar restrictions for the data use[10]. Since these limitations only concern user data, FuturePulse is not impacted.

Spotify

According to the Terms of Service of *Spotify*[11][12], it appears that usage of data is fully authorized in the context of the project.

8 https://creativecommons.org/publicdomain/zero/1.0
9 https://support.discogs.com/hc/en-us/articles/360009334593-API-Terms-of-Use
10 https://developer.twitter.com/en/developer-terms.html
11 https://developer.spotify.com/terms
12 https://developer.spotify.com/support

Funded by the European Commission

Moreover, even if it is asked to apply for a written approval, *Spotify* has a relatively permissive policy about *Non-Streaming* commercial applications. Note that, other commercial platforms, such as *Chartmetrics*, provide service based on data extracted from *Spotify*.

The FuturePulse consortium is currently seeking to make a first contact with the legal team of *Spotify*. Then, when the Exploitation Plan of the project is sufficiently advanced, an official application will be elaborated and submitted.

Deezer

As with *Spotify*, the use of the *Deezer API* is permitted for research and development purposes. However, according to the ToS[13], the commercial use of the data from the *Deezer API* is strictly prohibited without written approval.

A first contact has been made with a legal expert at *Deezer*, and a discussion will be conducted between *Deezer* and the FuturePulse consortium.

Kontor New Media

The access of the data of *Kontor New Media* is fully permitted through the approval of a record label. For the period of the project, FuturePulse bought legal access to the data of PGM. In consequence, the same approach can be followed for other record labels, in the context of the commercial exploitation of FuturePulse.

YouTube

Given the Terms of Service of the *YouTube API*[14], no restriction appear with the exception of user data, the use of the YouTube logo and other brand features.

However, for commercial use, a formal approval can be requested using the web form[15][16]. According to the required information on the web form, it is too early to make an application; it should be made after more progress is done during the Exploitation Plan.

Bandsintown

The readings of the ToS of *Bandsintown*[17] show that commercial uses are generally prohibited. The FuturePulse consortium got approved access to Bandsintown API for the project time period. However, for a commercial exploitation, an agreement is recommended.

Resident Advisor

Resident Advisor does not provide any API. FuturePulse employs web information extraction to obtain the data. Therefore, *Resident Advisor* does not provide API ToS, but "*General terms and Conditions*" for its web site[18].

These terms mention that "*ALL CONTENT provided by Resident Advisor is protected by copyright, ...*". Since the information provided by FuturePulse is the result of

13 https://developers.deezer.com/termsofuse
14 https://developers.google.com/youtube/terms/api-services-terms-of-service and https://developers.google.com/youtube/terms/api-services-terms-of-service-emea
15 https://services.google.com/fb/forms/ytapicommercializationapplication
16 https://developers.google.com/youtube/terms/developer-policies
17 http://corp.bandsintown.com/api-terms or https://corp.bandsintown.com/terms
18 https://www.residentadvisor.net/terms

Funded by the European Commission

aggregate predictive analysis (i.e. no original content or raw data from Resident Advisor), the consortium has the conviction that this data use is permitted by the law.

Beatport

The ToS of *Beatport* fully permit the use of its data for non-commercial applications[19]. However, they required an agreement for commercial uses.

The FuturePulse consortium originally wanted to get in touch with *Beatport* through the former partner BassNation, which should have facilitated informal discussions. Now, the next step is to find to communicate officially their legal office, then a formal application will be made to get an agreement.

Other sources

Other sources are used by the FuturePulse components, for example: SoundCloud, Last.fm, Billboard and Official Charts. Even if the data use is permitted for research and development purposes, a written approval is required for any commercial use. Using the same strategy, the consortium will look for the contacts of the legal officers and will prepare and submit an official application after some relevant progress of the project.

4.4.3 Conclusions of ToS analysis and risks

As seen above, the consortium has found no significant replacement solutions for four of the important sources (*Beatport, Resident Advisor, Kontor New Media*). The main reason is that they provide specific information that is not available by the other data providers for now. In consequence, the consortium has to anticipate any possible loss, and to monitor the future creations of new services potentially usable in replacement.

To secure the legal access of the data, the consortium has made important advances in order to get a written approval when it is required for a commercial exploitation. In several cases, we are at a step where it is too soon to formally apply for an official agreement because of the current state of the project, in terms of demonstrator and exploitation plan. Note that the legal experts of BMAT have already prepared a template of a *Memorandum of Understanding* in order to facilitate the formal discussions when they will be initiated with third parties.

Regarding the issue of local caching of data, this is only performed in those cases where absolutely needed, following the Terms of Services described. Audio data will be processed only in agreement with the license owner of the respective audio files.

Even if the consortium has a global strategy, because of the different policies of the Terms of Service and the different application ways (web form, twitter chat or direct contact with the legal team) the approach used must be adapted for each data provider.

4.5 Data Storage and backup

FTP and AWS servers from BMAT have been provided to store audio tracks and annotated data. The FuturePulse website is hosted and backed up by ATC partner.

4.6 Surveys and questionnaires

The project will perform online surveys under the activity T1.2 FuturePulse co-design process within WP1, in order to contribute to provide a concrete impression of the

19 https://support.beatport.com/hc/en-us/articles/215996708-Terms-and-Conditions

Funded by the European Commission

envisioned capabilities and services offered by the FuturePulse platform. More information about these surveys is provided in the deliverable D1.2 FuturePulse requirements v1 (M6).

Funded by the
European Commission

5 FAIR data

FuturePulse participates in Open Research Data Pilot, which requires the policy of FAIR data (findable, accessible, interoperable and re-usable research data).

FuturePulse has already setup a repository in Zenodo: https://zenodo.org/search?page=1&size=20&q=futurepulse

Currently, the Zenodo repository contains one dataset and four papers:

- **Dataset**
  - T-REC Song Recognition Dataset: described in section 4.3 of this document (Link to Zenodo)
- **Papers**
  - **"*VenueRank: Identifying Venues That Contribute To Artist Popularity"***, by Emmanouil Krasanakis, Emmanouil Schinas, Symeon Papadopoulos, Yiannis Kompatsiaris, Pericles Mitkas. Presented at **19th International Society for Music Information Retrieval Conference (ISMIR2018)**, Paris, France, 23-27 September 2018, with more than 430 participants. (Link to Zenodo)
  - "***Music retiler: Using NMF2D source separation for audio mosaicing***", by Hadrien Foroughmand Aarabi, Geoffroy Peeters. Presented at **Proceedings of the Audio Mostly 2018 on Sound in Immersion and Emotion**, Wrexham, United Kingdom, 12 – 14 September, 2018. (Link to Zenodo)
  - "***Boosted seed oversampling for local community ranking"***, by Emmanouil Krasanakis, Emmanouil Schinas, Symeon Pap (adopoulos, Yiannis Kompatsiaris, Andreas Symeonidis. **Elsevier Journal (Information Processing and Management)**, accepted 3 June 2019, https://doi.org/10.1016/j.ipm.2019.06.002 (Link to Zenodo)
  - "***Data-Driven Song Recognition Estimation Using Collective Memory Dynamics Models"***, by Christos Koutlis; Manos Schinas; Vasiliki Gkatziaki; Symeon Papadopoulos; Yiannis Kompatsiaris. To be presented at **20th conference of the International Society for Music Information Retrieval (ISMIR2019)**, Delft, Netherlands, 4-8 November 2019. (Link to Zenodo)

5.1 Making data findable, including provisions for metadata

Each dataset generated during the project will be recorded in an Excel spreadsheet with a standard format. The spreadsheet will be hosted at Zenodo. Search keywords will be provided when the dataset is uploaded to Zenodo which will optimise possibilities for re-use. Zenodo follows the minimum Data Cite metadata standards.

Funded by the European Commission

A DOI will be assigned to datasets for effective and persistent citation when it is uploaded to the Zenodo repository. This DOI can be used in any relevant publications to direct readers to the underlying dataset.

In order to clearly describe the content of the data, FuturePulse naming convention for project datasets will follow a similar approach to the deliverable naming convention described in the deliverable *D7.1 Quality Assurance and Risk Management Plan* for supporting documents:

FuturePulse_SPD_DATASET" + dataset_number + dataset_name + dataset_version

Some examples:

FuturePulse_SPD_DATASET_01_Pilot1-results_v1.xls

FuturePulse_SPD_DATASET_Annotated_audios_v1.xls

The data will be accompanied with metadata clarifying the meaning of the data and how the data has been collected. The metadata can be provided without the actual data, if requested.

Datasets metadata will follow the META-SHARE[20] schema for data sets description. META-SHARE is an open resource exchange infrastructure. The following data set description is based on the DMP template circulated by CRACKER[21].

Metadata	Description
Resource Name	Complete title of the resource
Resource Type	Conceptual resource
Media Type	The Physical Medium of the content representation, e.g. video, audio, text, numerical data, etc.
Language(s)	The language(s) of the resource content
License	The licensing terms and conditions under which the tool/service can be used
Distribution Medium	The channel used for delivery or providing access to the resource, e.g. accessible through interface, downloadable, CD/ DVD, etc.
Usage	Foreseen use of the resource for which it has been produced
Size	Size of the resource with regards to a specific size unit measurement in form of a number
Description	A brief description of the main features of the dataset

Table 15: FuturePulse Dataset metadata

The specific metadata contents, formats and schema may be further refined in the future versions of the DMP.

5.2 Making data openly accessible

The consortium has chosen Zenodo, the open research repository from OpenAIRE and CERN, as the central scientific publication and data repository for the project outcomes. The repository has been designed to help researchers based at institutions

20 http://www.meta-net.eu/meta-share

21 http://cracker-project.eu

Funded by the European Commission

of all sizes to share results in a wide variety of formats across all fields of science. Furthermore, Zenodo supports DOI versioning, allowing users to update the record's files after they have been made public and researchers to easily cite either specific versions of a record or to cite, via a top-level DOI, all the versions of a record[22].

Research data, especially those needed to validate the results of scientific publications, will be deposited in the Zenodo repository.

Zenodo enables users to:

- easily share the long tail of small data sets in a wide variety of formats, including text, spreadsheets, audio, video, and images across all fields of science
- display and curate research results, get credited by making the research results citable, and integrate them into existing reporting lines to funding agencies like the European Commission
- easily access and reuse shared research results
- define the different licenses and access levels that will be provided

Furthermore, Zenodo assigns a Digital Object Identifier (DOI) to all publicly available uploads, in order to make content easily and uniquely citable.

5.3 Making data interoperable

The data will be stored in a format readable by commonly used data management tools or office software. Direct automatic interoperability of the data with other external data sets is not sought for.

5.4 Increase data re-use (through clarifying licences)

All the research data will be of the highest quality, have long-term validity and will be well documented in order other researchers to be able to get access and understand them.

The datasets will be made available for re-use through uploads to the Zenodo community page for the project. General policies of Zenodo will apply for content, access and reuse, removal and longevity[23].

As the processed data will be public, no licences are needed for re-use of the data, as long as the data source is acknowledged.

22 http://blog.zenodo.org/2017/05/30/doi-versioning-launched

23 http://about.zenodo.org/policies

Funded by the European Commission

6 Allocation or resources

The FuturePulse datasets stored in Zenodo will be preserved a minimum of 5 years according to the European Commission Data Deposit Policy. There are currently no costs for archiving data in this repository.

Each partner has authorised a responsible of data management who takes the responsibility to control the correct storage, management, sharing and security of the dataset. FuturePulse has appointed Data Protection Officers (DPOs) for each partner in the Consortium:

Partner	DPO
BMAT	Giacomo Cortese
ATC	George Zissis
CERTH	Manos Schinas
MUSIMAP	Frédéric Notet
IRCAM	Rémi Mignot
PLAYGROUND	Anders Engström
SYB	Daniel Johansson
SONAR	Ventura Barba

Table 16: Data Protection Officers

All social media data within FuturePulse is aggregated and no personal data is stored by the platform.

Daniel Molina and Gonçal Calvo at BMAT are responsible for the creation, management and updates of the Data Management Plan in FuturePulse project.

Funded by the European Commission

7 Data security

7.1 Data storage

For the duration of the project, datasets will be stored on the responsible partner's centrally provided storage, summarised in the table below.

FuturePulse partner	Data Storage
BMAT	1. Data from Vericast (accessible through the Vericast Website): - Metadata: in redundant MySQL Database - Recordings and references: in NFS storage systems 2. Playground Music Catalog: - FTP Server with NFS storage backend
ATC	1. ATC will provide a storage mechanism to save output produced by the analysis components so that it can be used as a common facility to save generated data concerning analysis of audio tracks, predictions and Media statistics that will be gathered during the project. This will be a graph-based storage which can support partners' existing solutions to a large extent but could also be accompanied by another RDBMS or NoSQL database according to the emerging needs. 2. A temporary cloud storage will be considered in order to store the actual audio files that will be used to train the components that require training as well as for new audio files that are going to be uploaded by users. A cloud option is considered for this file storage so as to have a cost-effective and easily accessible solution to make large volumes of audio files available to many components. The provision of an API to upload new files to the storage will be also taken under consideration for the selection of the appropriate cloud service.
CERTH	CERTH is currently using four workstations for the temporary storing of data: 1. Charts collected by CERTH are stored temporarily as files in a windows workstation. 2. Predictions about artists and genres from Google trends and charts are stored temporarily as files in a windows workstation. 3. Sources from social media and streaming platforms (e.g. channels, accounts, etc) and their associations with Future Pulse targeted artists are stored temporarily in a relational DBMS in a Linux workstation. The same workstation keeps events and artists from Facebook in a graph database (Neo4j). Tweets about music genres are also stored in this workstation. Data coming from the tracking social media and streaming platforms will be also stored in the same workstation in the appropriate NoSQL solution. 4. Results from Twitter data analysis (demographics predictions) are stored temporarily in a windows workstation. All aforementioned workstations are located in CERTH premises and are part of its internal network, where standard security mechanisms and policies are applied.

Funded by the European Commission

	All the above data and any new data that will appear at later stages of FuturePulse, will eventually be migrated in a linux server, also located in CERTH. These data will be exposed by a REST API with the appropriate authentication/authorisation level.
MUSIMAP	MUSIMAP is currently using dedicated private storage hosted by AWS 1. Audio files are temporarily stored in a dedicated and secured S3 storage in order to be analysed. 2. Analysis results are also stored in a dedicated and secured S3 storage and/or secured database 3. Meta-data are stored in a dedicated Neo4j database All aforementioned servers are hosted on Amazon Web Services Cloud and are part of a MUSIMAP dedicated private network, where high security mechanisms and policies are applied. All the above data and any new data that will emerge at later stages of FuturePulse, will be exposed by a REST API with the appropriate authentication/authorisation level.
IRCAM	The web-service of IRCAM removes all uploaded audio files after analysis. It only stores the corresponding ISRCs and the predicted values (genre, tempo, etc.) into a database. These information are then retrieved using the web-service and an authentication/authorisation mechanism. IRCAM also stores audio datasets for the purpose of developing Machine-Learning algorithm, this is not accessible outside a restricted number of people at IRCAM.
Playground	Playground do not have data for this project in itself. Our data is in our IT-systems, and are exported and shared with the project as needed. In practise, all data is collected by the technical partners themselves through APIs.
BASS NATION	Bass Nation do not have data for this project in itself.
SYB	The SYB use case is temporarily stored by technical partners during the research process, e.g. IRCAM and Musimap audio analysis, CERTH charts analysis and BMAT broadcast analysis. For the pilots, sales data from SYB pilot partners will be stored locally on our own servers for the analysis.

Table 17: Data storage

After the completion of the project, the public datasets will keep stored and freely accessible in the Zenodo repository for long term preservation and curation.

7.2 Personal data

7.2.1 Social media data

FuturePulse datasets do not include personal information. All Social media information used in the project is aggregated after its collection, and as a result no personal information will be stored by the FuturePulse platform. We are only interested in data related to music usage (popularity, recognition, total plays, demographics, etc.).

We neither collect nor use (nor do we plan to) any personal data to obtain FuturePulse metrics. Social media content comprises aggregate metrics related to artists, tracks, or genres. A detailed listing and description of the collected data from different sources is

Funded by the European Commission

available in D2.1 and D2.3 (Data specifications and collection v1 and v2), where it is made clear that no personal information is among the data of interest.

Given that no personal identifier (e.g. user id) is stored by the platform, it will not be possible to extract or infer any personal information from the data that is stored in the FuturePulse servers even in the unfavourable scenario that a third party managed to gain access to it.

7.2.2 Surveys

Some surveys within the project (see Section 4.6) are using Cint[24], a survey platform that connects community owners to researchers, agencies and brands, for the sharing and accessing of consumer data. The survey participants will not include children or other groups needing a supervisor.

7.2.3 Informed consent

The consent of the survey participant will be asked in all assessment activities conducted within FuturePulse. The consent includes a description of how and why the data is to be used. The informed consent template is included as an annex in the Deliverable D8.1 – POPD – Requirement No.1_v1.1.

24 https://www.cint.com

Funded by the European Commission

8 Ethical aspects

Future WP8 Ethics requirements sets out the 'ethics requirements' that the project must comply with. There are two deliverables associated to this Work Package:

- D8.1: POPD – Requirement No. 1: A report on ethical considerations in regards to social media, addressing general ethical issues regarding social media research, ethical and legal issues in regards to informed consent by participants, ethical issues and consent in regards to children and young people.
- D8.2 H – Requirement No. 2: Providing details on the procedures and criteria used to identify/recruit research participants and informed consent procedures

FuturePulse questionnaires do not include any form of personal data. Therefore, no personal data is subject to be shared or long term preserved. We are only interested in data related to music usage (popularity, recognition, total plays, demographics, etc.).

9 Other issues

Use of national procedures for data management

IRCAM makes use of Archive ouverte HAL[25], the French national repository for scientific publications.

25 https://hal.archives-ouvertes.fr

Funded by the European Commission