

Context-Aware Recommender Systems in the Music Domain: A Systematic Literature Review

Álvaro Lozano Murciego, Diego M. Jiménez-Bravo, Adrián Valera Román, Juan F. De Paz Santana and María N. Moreno-García *

Department of Computer Science and Automation, Science Faculty, University of Salamanca, Plaza de los Caídos s/n, 37008 Salamanca, Spain; loza@usal.es (Á.L.M.); dmjimenez@usal.es (D.M.J.-B.); adrianvalrom.usal@usal.es (A.V.R.); fcofds@usal.es (J.F.D.P.S.)

* Correspondence: mmg@usal.es

Abstract: The design of recommendation algorithms aware of the user's context has been the subject of great interest in the scientific community, especially in the music domain where contextual factors have a significant impact on the recommendations. In this type of system, the user's contextual information can come from different sources such as the specific time of day, the user's physical activity, and geolocation, among many others. This context information is generally obtained by electronic devices used by the user to listen to music such as smartphones and other secondary devices such as wearables and Internet of Things (IoT) devices. The objective of this paper is to present a systematic literature review to analyze recent work to date in the field of context-aware recommender systems and specifically in the domain of music recommendation. This paper aims to analyze and classify the type of contextual information, the electronic devices used to collect it, the main outstanding challenges and the possible opportunities for future research directions.

Keywords: context-aware recommenders; cars; music recommendation; internet of things; slr; systematic literature review

Citation: Lozano Murciego, Á.; Jiménez-Bravo, D.M.; Valera Román, A.; De Paz Santana, J.F.; Moreno-García, M.N. Context-Aware Recommender Systems in the Music Domain: A Systematic Literature Review. *Electronics* **2021**, *10*, 1555. <https://doi.org/10.3390/electronics10131555>

Academic Editors: Marco Lombardi, Domenico Santaniello, Zilu Liang and Muhammad Khan

Received: 1 June 2021
Accepted: 24 June 2021
Published: 27 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the area of data mining and big data technologies, recommender systems (RS) [1] are a widely used collection of algorithms aimed to link the items offered by any service or business with its consumers through personalization techniques. The importance of these systems is such that it is noticeable how they have attracted the interest of the scientific literature and large technology companies over the last few years. In some companies, in their early days they were a central part of their business models, as in the cases of Amazon or Netflix.

Therefore, the main objective of recommender systems is to offer the most relevant items to users based on their profile, history, their explicit preferences such as ratings or opinions of other items or implicit preferences related to their interactions in the system with other items or even their relationships with other customers.

Furthermore, the rise of mobile technologies and ubiquitous computing employing Internet of Things (IoT) devices has allowed these systems to include information about the user's context when making their personalized recommendations. These types of recommender system are called context-aware recommender systems (CARS) [2] that seek to improve the relevance of the recommendations provided to the user by taking advantage of contextual information such as the user's location, time of day, physical activity, emotional state and other aspects of their environment. These contextual characteristics may influence the rating that the user would give an item depending on the specific

situation in which the user is expecting the recommendation, and this issue has been extensively studied and confirmed in the scientific community [3] and by companies such as Netflix [4].

This type of recommender system has started to be used in very diverse domains such as cinema [5], news [6], points of interest (POIs) [7], hotels [8] and tourism [9], e-learning [10], e-retailing [11], e-commerce [12], food [13], restaurants [14], books [15] and music among many others.

One of the application domains that has attracted the interest of researchers and large companies is the music domain. This interest has been motivated by the rise of online music shops and music streaming services that have extensive music catalogues to offer their customers. In this domain it is possible to find what are known as music recommender systems and companies such as Pandora, Last.fm or Spotify (among others) together with the research community are looking for new ways to include context to improve the recommendations provided in this domain.

The problem of music recommendation [16] has a special set of characteristics when it is compared to other domains such as, for instance, travel, books or films. There is a significant difference in the time needed to consume the item in each domain: days or weeks in the case of a book, hours or less in the case of a film, and a few minutes in the case of a single song (although this could increase if it is a music playlist). As a consequence, the time it takes for the user to shape an opinion about the item being consumed is much shorter than in the case of books, and the likelihood of discarding the item (skipping the song) is much higher than in other domains.

Music is also characterized by the fact that an item can be consumed repeatedly even on the same day, which is less frequent in domains such as cinema or books. This makes the user not only tolerate recommendations of items they have already consumed, but also appreciate it and see it as a positive factor in the recommendation, something that in domains such as films or travel is highly unlikely to be the case. Regarding user penalization for a bad recommendation, the user is more tolerant due to the short time it may take to consume the item and the possibility of skipping the recommendation if it is necessary. Linked to the aforementioned, in the music recommendation it is possible to find a recommendation at different levels of abstraction, being able to recommend a single song, a playlist, an album or an artist.

From the point of view of the ratings that can be found, this is a domain in which explicit ratings are uncommon and, if they are available, they tend to be sparse due to the magnitude of the existing song catalogues. Therefore, it is a domain where implicit information is often used and where content-based approaches have often been used in the literature. The latter approaches are linked to another field called music information retrieval (MIR) and are still very popular today, extracting semantic information from songs based on audio features, artists, lyrics, cover art, etc.

From a user's point of view, who consumes music on a day-to-day basis, the content-based approach may be decisive, however, several studies in the field of music psychology [17] show how short-term preferences are influenced by their emotional, physical, or social context. This context can be determinant as it is evident that certain types of music can be linked to different emotional states as well as other music can be tied to certain locations, user activities, moments of the day, etc. For this reason, incorporating contextual information into recommender systems and exploiting it effectively has become one of the challenges in this field.

However, before tackling how new contextual recommenders in the domain of music can be effectively designed, it is necessary to understand and explore existing research lines and solutions regarding this domain. The analysis and study of case studies from the literature and the identification of the most relevant contextual information is vital for the creation of new systems that build on the strengths and eliminate the weaknesses of existing work. In order to achieve this, a systematic literature review about contextual recommender systems in the domain of music has been carried out.

This paper follows the following structure: Section 2 describes the methodology chosen for the systematic literature review and details each of the steps and protocols established for conducting the review. Section 3 shows the selection process of the research articles following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) approach, later, in Section 4, the results obtained from the analysis of the selected publications and the answers to the research questions initially posed are presented. After the analysis, Section 5 discusses the results presented and finally Section 6 presents the conclusions drawn from the study and future lines of research.

2. Methodology and Review Planning

We have developed this systematic literature review following the guidelines of Kitchenham and Charters in [18]. Prior to this systematic literature review, a preliminary revision has been carried out to verify that a state-of-the-art review of context-aware recommender systems in the music domain has not been carried out recently. This review has been carried out in different scientific databases such as IEEE Xplore, Springer, Web of Science (WoS) or Scopus using search terms such as “slr” “state of the art”, “survey” together with terms related to the object of this review such as “context-aware”, “music”, “recommender systems”, etc. The results of these searches confirmed that there were no reviews such as that discussed in this work, justifying its development.

The following describes the review process and their foundations in order to achieve the maximum degree of reproducibility. The protocol followed, together with the methodological steps of the review, are defined below.

2.1. Research Questions

We defined the following research questions (RQ) that would allow us to understand the current state of the art of CARS in the domain of music recommendation. These research questions are listed below:

- **RQ1:** What are the kinds of contextual factors more commonly employed in the music recommendation domain?
- **RQ2:** What are the devices or technologies employed for extracting this contextual information?
- **RQ3:** How is context information exploited along the recommendation process and which algorithms are employed?
- **RQ4:** What evaluation metrics and methods have been employed to validate the effectiveness of these music RS?
- **RQ5:** Which publicly available datasets are mainly used?
- **RQ6:** What are the main open lines of research in this domain?

The review scope was established employing the PICOC method by Petticrew and Roberts [19]:

- **Population (P):** Recommender systems.
- **Intervention (I):** Context aware recommender systems in the music domain.
- **Comparison (C):** there is no comparison intervention in this work due to the main objective of this work is review all the available context aware approaches in the literature and explore new lines of research in this domain.
- **Outcomes (O):** works about CARS in the music domain.
- **Context (C):** works related to music recommendation in the context of industry or academia.

2.2. Inclusion and Exclusion Criteria

Once the scope of the systematic review has been established, following the methodology of [20], it is necessary to indicate the inclusion (IC) and exclusion (EC) criteria in order to select only those papers that are relevant to answer the research questions posed above. The inclusion criteria established are as follows:

- **IC 1:** The paper covers a recommender system.
- **IC 2:** The recommender system is applied in the music domain.
- **IC 3:** The recommender system is aware of the user's contextual information.
- **IC 4:** The paper is written in English.
- **IC 5:** The work is published in peer review conferences, books or articles.
- **IC 6:** The work was published after 2010.

Most of the exclusion criteria are related to the inclusion criteria previously shown:

- **EC 1:** The paper is not concerned with a recommender system.
- **EC 2:** The recommender system is applied to a domain other than music.
- **EC 3:** The recommender system does not use context information from the user, even though it is a music recommender system.
- **EC 4:** The work is not written in English.
- **EC 5:** The work is not published in peer review conferences, books or articles.
- **EC 6:** The work was published prior to 2010.

2.3. Search Strategy

Upon establishing the inclusion and exclusion criteria, it is necessary to find the most relevant databases in which to search for research papers that meet the established criteria. At the same time, appropriate selection of the concepts is required in order to subsequently perform the queries in these databases.

The following databases have been selected for this work: IEEE Xplore, Web of Science, SpringerLink and Scopus. The main reasons for this decision are that they are databases in the scope of the research of this work and allow searches using search queries in a similar way to each other. The search concepts used were as follows:

- The term "context" or "aware" to search for context-aware recommendation algorithms.
- The term "recommender", "recsys", to search for papers presenting recommendation algorithms.
- Terms such as "song", "music", "playlist", "album" and others related to the music recommendation domain.
- Other terms such as "social", "location", "time", "emotion", "activity" that refer to the context of the user.

In the final search, some terms such as "recsys" have been omitted as wildcard searches such as "recommend*" are flexible enough and some searchers have a limited number of wildcard terms in the search query.

2.4. Query Strings

The query strings for the search in each of the databases have been elaborated using the terms derived from the PICOC method [21] and with the help of the Parsifal software [22]. Parsifal is a tool designed to help researchers plan and conduct systematic reviews of the scientific literature. It allows the collaboration of several researchers in the review and filtering of articles in a distributed way.

The terms are connected using logical operators such as AND/OR/NEAR and special characters such as * (wildcard). The queries performed in each database are presented below (Table 1).

Table 1. Databases and query strings.

Database	Query String
IEEE Xplore	("recommend*" NEAR/10 ("*aware" OR "context*" OR "time" OR "social" OR "emotion*" OR "activity")) AND (("music*" OR ("song*" OR ("album" OR ("playlist*"))
Web of Science (WoS)	TS = (("recommend*" NEAR/10 ("*aware" OR "context*")) NEAR/10 (("music*" OR ("song*" OR ("album" OR ("playlist*")))) [23–49]
SpringerLink	("recommend*" NEAR/10 ("*aware" OR "context*")) NEAR/10 (("music*" OR ("song*" OR ("album" OR ("playlist*"))
Scopus	TITLE-ABS-KEY (("recommend*" W/10 ("*aware" OR "context*")) W/10 (("music*" OR ("song*" OR ("album" OR ("playlist*")))) AND (LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2013) OR LIMIT-TO (PUBYEAR, 2012) OR LIMIT-TO (PUBYEAR, 2011) OR LIMIT-TO (PUBYEAR, 2010)) AND (LIMIT-TO (LANGUAGE, "English")) AND (EXCLUDE (DOCTYPE, "cr") OR EXCLUDE (DOCTYPE, "re"))

2.5. Quality Criteria

Although the inclusion and exclusion criteria are a good tool for the selection of papers, the overall quality of the papers obtained after querying the various established databases is not taken into account. For this reason, it is necessary to establish a quality criterion for the articles before deciding whether to add them to the study. Using the Parsifal tool, three different ratings have been established to score each of the criteria:

- Score 1: The article meets the established quality criterion.
- Score 0.5: The article partially meets the quality criterion.
- Score 0: the article does not meet the quality criterion.

The quality criteria are as follows:

1. The purpose of the work is the development of a music recommender system.
2. The paper applies the recommendation algorithm to datasets and presents a case study.
3. The work includes the user's contextual information when making recommendations.
4. The paper describes how contextual user information is obtained and the technology used.
5. The paper presents used evaluation metrics for the proposed algorithm
6. The work uses publicly available datasets.

After applying the quality criteria to the previously filtered papers, the papers that may be taken into account for the systematic review are selected. This allows the selection of those papers that can best help answer the research questions initially formulated.

3. Review Process

This section describes each of the steps followed in the process of extracting and reviewing papers from the literature using the process described by the PRISMA flow diagram shown in Figure 1.

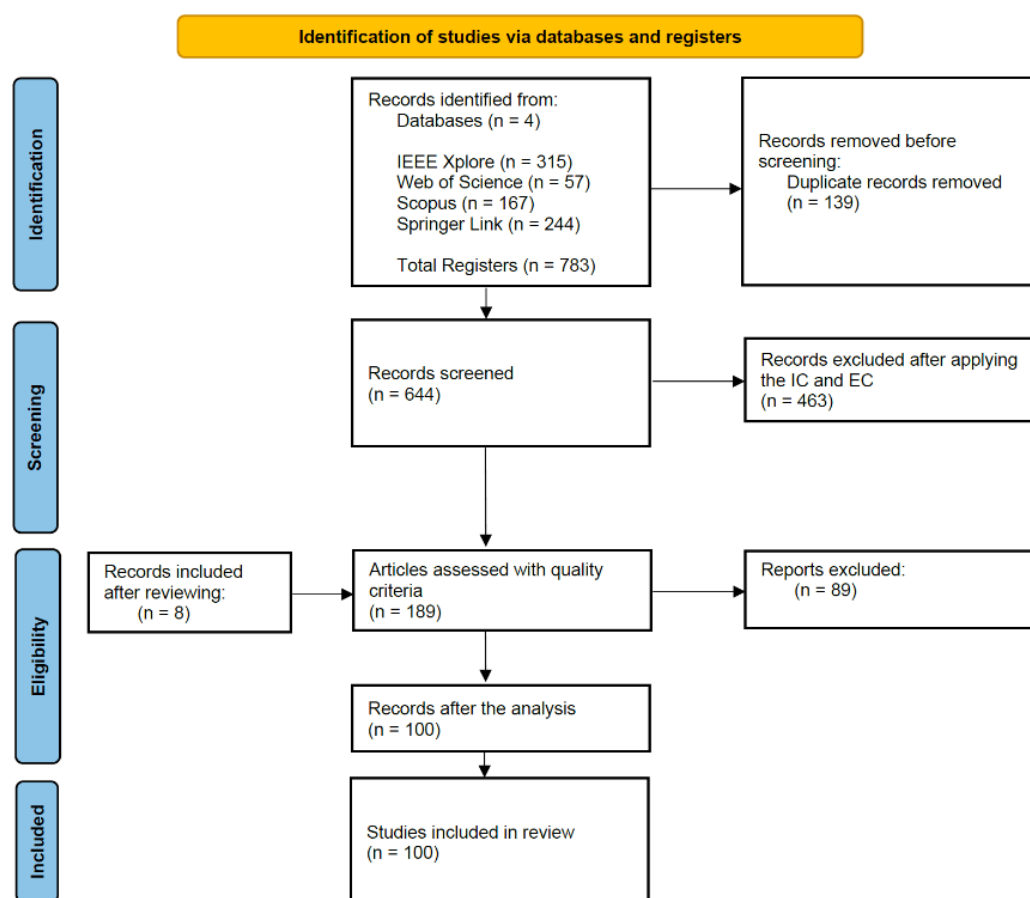


Figure 1. Review process using the PRISMA 2020 flow diagram.

1. The results after querying the selected databases (identification phase) report the following results: 783 papers collected: 57 (7.28%) from Web of Science, 167 (21.33%) from SCOPUS, 315 (40.23%) from IEEE Xplore and 244 from Springer Link. After removing 139 duplicates (17.75%) the total results of unique registers are 644 (82.25%). All these results are included in a Google Sheets [50] and the BibTeX files are included in a git repository [51].
2. We have imported the BibTeX files to Parsifal tool in order to conduct the selection of papers which meet IC (screening phase) after reading title and abstract. In this step we have accepted a total number of 181 papers (23.12%) and we have rejected 463 (59.13%). After this step we have continued the systematic review directly in the Google Sheet mentioned previously.
3. The next step involves analyzing the articles according to the quality criteria. We have set the minimum threshold score for this point at 5.0. In this phase we have included eight new papers discovered during the review. After reviewing the quality criteria (eligibility phase) we have filtered a total of 100 papers out of 189.
4. Papers selected after reading the full text (phase of final papers inclusion): 100 (12.77% of the total papers considered, 52.91% of the papers read). The analysis performed to achieve this number of papers was only based on their content, and without concerning bibliometric measurements (journal source, number of citations) or other aspects.

After reviewing and analyzing the articles finally selected, it is possible to begin to answer the research questions initially posed.

4. Results

4.1. What Are the Kinds of Contextual Factors More Commonly Employed in the Music Recommendation Domain?

In previous reviews of the state of the art such as [16,52,53] or articles such as [54] oriented to CARS in a non-specific domain, different categories of context are identified at a more general level: individual context of the user, time, activity carried out by the user and relational context. In the chapter of the book [16] dedicated to music recommender systems the context in the music domain is classified into two major groups: environment-related context and user-related context. In this paper we have taken the two main categories above and added further sub-categories of context in this domain as follows:

Environment-related context: this group covers the context of the users' physical environment and the types of devices they use. We could include in this category the following sub-types of contexts:

- **Physical context:** factors related to the user's physical environment at the time the recommendation is made or when the item is consumed. This category includes factors such as the following:
 - Location: certain places are associated with a particular type of music and it is possible to take advantage of this contextual information for recommendation.
 - Weather: the weather conditions of the day (cloudy, rainy, sunny, etc.) can influence the user's preference for different types of music, differing greatly from a sunny day in summer to a snowy day in winter.
 - Time: Time slot (morning, afternoon, evening), time of day, type of day of the week (work or holiday), are very influential factors in the domain of music recommendation because music preference may vary between leisure and work periods.
 - Other environmental factors: other external factors such as noise level, current traffic or ambient light may have an influence on the music preference for a specific situation.
- **Interaction media context:** these types of element can be treated as a type of context and are in many cases a proxy to identify a certain context of user activity. The following contexts of interaction can be identified:
 - Mobile devices: whether the user is using mobile devices such as a smartphone may imply a specific context of interaction that may influence the type of music the user prefers to listen to as doing so on a mobile device or a desktop device may be related to different user activity contexts.
 - Desktop devices: activity performed from a desktop device may involve a specific user context linked to a music preference.
 - Vehicles: it is possible to extract context information if interacting with an in-vehicle device while driving such as vehicles with Android Auto or other software.
 - Wearables: smartwatches, headphones and other wearable devices can indicate a specific interaction context by providing contextual information about the user, such as their physical activity.
 - Virtual assistants and conversational agents: the use of this type of interaction context is currently growing and attracting the attention of large companies for recommendations through conversational assistants that capture the user's context through dialogue. It is possible to extract emotions from the user's voice implicitly or explicitly through questions to the user.
 - Session context: in several works, the authors refer to session context relating to songs that are often found together in the same playlist. In this case the context of the item during playback.

User-related context: all the factors most closely linked to the user, such as their current activity, their background in terms of knowledge or their social context (reflected in social networks or other sources of information). We classify this context into the following types:

- **Social context:** the presence of or relationship with other people that may influence the user's musical tastes. Usually information extracted from social networks or systems where it is possible to identify the user together with a group of users and provide recommendations based on their environment (this could be related to another broad field such as recommendation systems for groups [55]). Within social networks the following factors can be taken into account:
 - People with whom the user meets at a given time.
 - People with whom the user has connected or interacted.
 - Roles of people in the user's environment and relationships of trust or reputation.
- **Modal context:** related to the state of mind of the user, the user's goals, mood, experience, and cognitive capabilities. We could highlight:
 - Emotional state of the user: musical tastes can change a lot depending on the emotional state of the user. Many studies focus on detecting this emotional context either directly or indirectly through the gestures on the user's face or the comments they have recently posted on their social networks.
 - User experience or skills: different music recommendations based on the user's experience of a certain genre or type of music.
 - Cultural background: this type of information is about the context of the user's origin.
- **User activity:** this contextual factor refers to the action or task performed by the user while listening to music. We can distinguish between:
 - Task or physical activity performed among a possible set of actions: resting, swimming, running, cooking, studying, etc.
 - Physiological state of the user: this physical state can be defined by biometric variables obtained by sensors such as: heart rate, oxygen saturation, brain wave activity, etc. This may be more related to identifying a type of music suitable for a particular physiological state.

Furthermore, based on the knowledge [54] that the recommender system has about the context, it can be classified into:

- Fully observable: contextual factors are explicitly known. For example, the time when a song was played, the day of the week and whether it was a bank holiday or not.
- Partially observable: only some context information is known to the recommender system, but it is not complete.
- Unobservable: there is no explicit contextual information available, but the recommender system could infer it implicitly using latent knowledge.

Moreover, given how contextual factors can change over time [54], it is possible to distinguish between two types of contextual factors:

- Static: contextual factors are stable over time (they do not change too often). In the case of music, for example, a context of study activity is very often related to a type of music and this type remains stable over time regarding this contextual factor.
- Dynamic: contextual factors change over time in some way. In the music domain this could be related to the user's changing preferences over time in the same context.

Regarding how this contextual information can be acquired by the recommender system, we can distinguish between different types:

- Explicit: when users explicitly state in the recommender system their contextual information of any kind: mood, location, etc. This can be done through different interfaces such as a text box or a question in a dialogue.
- Implicit: when the recommender system observes the user's activity and context directly, such as their geo-positioning via the Global Positioning System (GPS) or another sensor on their smartphone, or the time at which a certain playlist or song was played.
- Inferred: this approach is widely used as contextual information is often difficult to obtain. By using data mining or machine learning techniques, it is possible to use the user's interaction with the system or their song playing patterns to predict their contextual state.

In the papers reviewed in this study, we have found a large number of papers that acquire information implicitly as shown in Figure 2. In the category “inferred” we have included those who expressed those contextual factors such as emotions were obtained from implicit contextual data as well as those who used other contextual information to obtain, for example, session contexts.

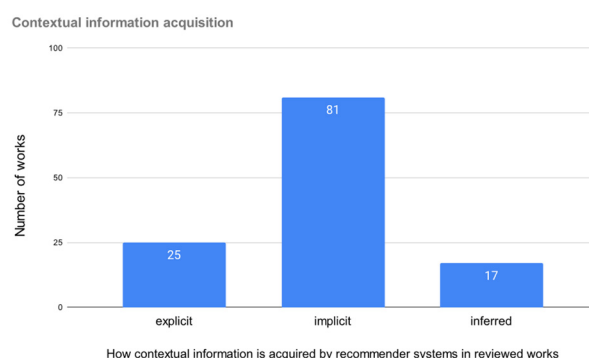


Figure 2. How contextual information is acquired by recommender systems in reviewed works.

Table 2 shows the classification of the articles included in this work and the different contexts according to the above classification.

Table 2. Context classification of included works.

Context		Works
Environment-related context	Location	[40,56–86]
	Weather	[32,40,56,58,59,63–65,71,74–77,79,83–85,87]
	Time	[23–31,33,34,36–42,44,46,47,49,56,58–60,63–65,67–71,73–77,79,83–85,88–101]
	Other	[40,56,64,65,74–77,79,83–86,102–104]
	Mobile devices	[42,43,46,57,60,64,66,70–72,80,81,83,88,105,106]
	Desktop devices	[42,44,46,86]
	Vehicles	[42,46,56,65,74–77,79,84]
	Wearables	[102,104]
	Virtual assistants and conversational agents	[42,46,49,105]
	Session context	[23–31,33,34,36,38,40–42,44,46,47,49,83,86,88,90–92,94,95,97,98,100,101,103,107,108]
User related context	Social context	[25,26,35,37,39,42,46,67–69,85,91–94,100,105,109–119]

Modal context	[25,26,32,35,37,39,42,44–47,49,56,58,59,61,63–65,68,70,74–77,79,80,84,87–89,93,94,99,100,105,112,117,120,121]
User activity	[40,43,58,59,62,64–66,70,71,74–76,79,85,106]

Figure 3 shows several charts illustrating the number of CARS works in the music domain covered in this study and the different types of contexts they use. It is worth pointing out that the above contexts can be used at the same time in the same work.

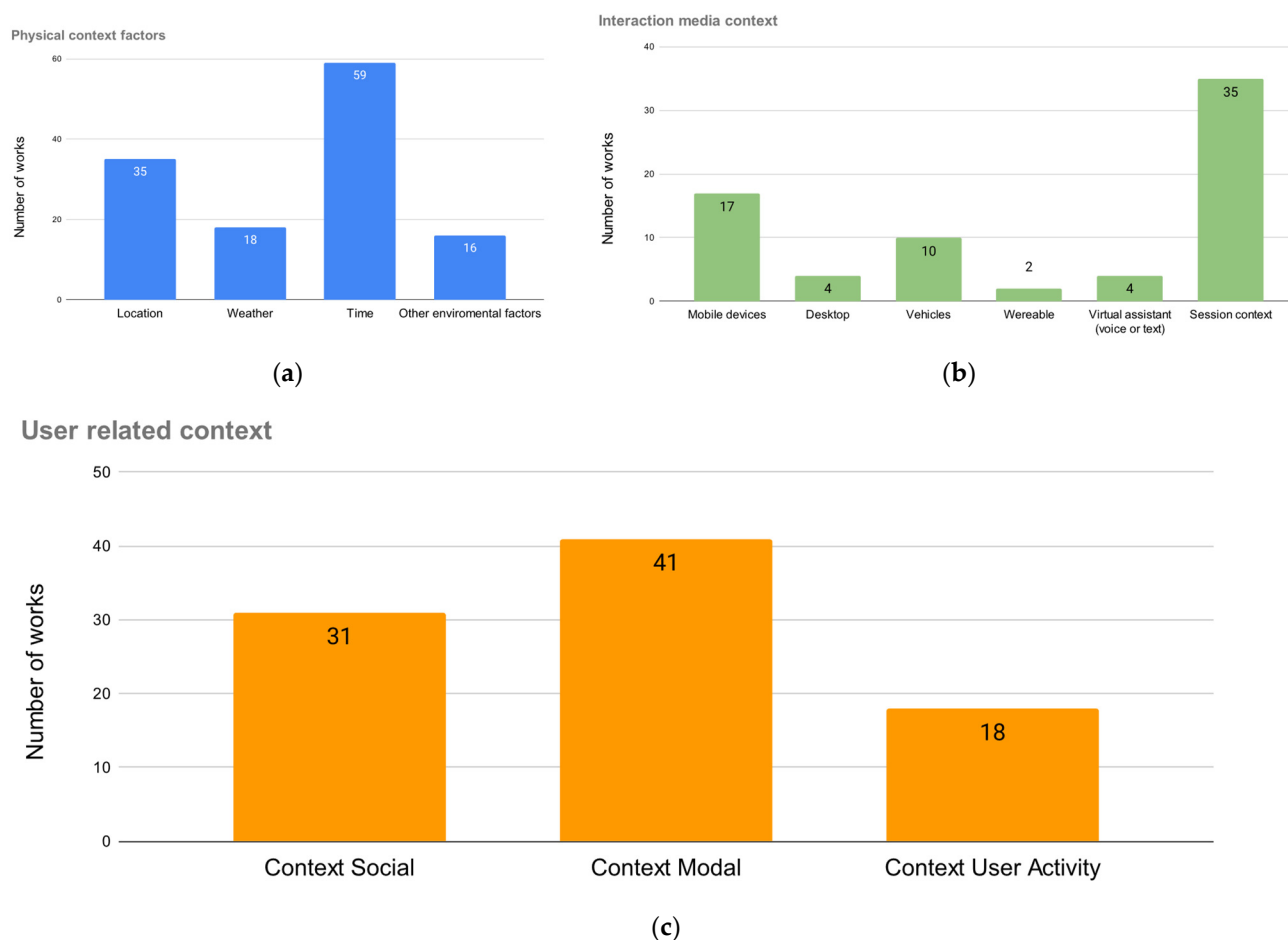


Figure 3. (a) Works employing physical context factors, (b) works employing different interaction media context (c) works involving user related context.

Figure 3a shows the number of papers using the different physical contexts in recommender systems. It is possible to highlight how the contexts of time and location are those mostly used in the works covered. In the case of time, it is a type of context that can be acquired in a relatively simple way, since it represents the time at which the song was listened to, and from this other context attributes can be derived, such as the time of day, the day of the week, whether it is a weekend or a public holiday. Moreover, geographical location has been widely used in the reviewed works, highlighting the importance of the effect that this context has on users' music preferences. Similarly, weather and other environmental factors have been used for music recommendations, but to a lesser extent than in the previous contexts.

Figure 3b shows the interaction media contexts that have been used in the reviewed papers. Note the large number in the context session category and that this is linked to the temporal context, as it exploits this information. Furthermore, there is a large amount of

research that uses mobile devices to obtain context, as many physical contexts are obtained through the sensors of these devices. Likewise, mention should be made of the high number of works related to context in vehicles, which is due to the large number of works using the InCarMusic [65] dataset which will be referred to in later sections. Finally, new forms of interaction mentioned in recent works such as conversational agents (virtual voice assistants or conversational text bots) and the attraction that this type of context is currently having.

Figure 3c shows the works using the user-related context. The modal context, and within this context especially the emotional state of the user, is likely to be the factor mainly exploited in the works reviewed. Mood or emotional state is mostly obtained through sensors or photographs from facial gestures or inferred through social media elements such as recently taken texts or photographs. Furthermore, there is a large body of work that exploits the user's social context and friendship relationships in order to find more relevant recommendations. Many of the studies analyzed exploit social networks not only to find relationships but also to obtain latent context by using social tagging information from songs, posts and photographs published by users on these networks.

4.2. What Are the Devices or Technologies Employed for Extracting This Contextual Information?

Once the different types of context have been identified, it is necessary to know which technologies, devices or sensors have been used in the scientific literature to extract these contextual factors that are later incorporated into recommender systems.

To classify the different sensors that can be used for this task, we have used the taxonomy proposed by [122] and [123]. Therefore, sensors can be classified into the following categories:

- **Physical or hardware sensors:** sensors that provide a raw measurement of the environment. Within this category we can find sensors, devices and technologies such as:
 - GPS: employed to extract information about the geographic position of the user.
 - Accelerometer and Gyroscope: used to detect the user's movement in order to later be able to infer the user's physical activity.
 - WIFI, Bluetooth: to extract information about the user's position or the presence of nearby devices.
 - Camera: used to infer the user's state of mind by recognizing the user's facial expressions or to detect the geographical position of the user based on images of the environment.
 - Microphone: used for the recognition of ambient noise or to obtain information from the user in a spoken conversation.
 - Biosensors: such as EEG which is an electrophysiological monitoring method to record electrical activity on the scalp and is used in some works to detect the state of the user in order to recommend a certain type of music. Other sensors such as heart rate monitors have been used for this purpose.
- **Virtual or software sensors:** provide measurements of the user's context, but at a higher level of abstraction and combining measurements from different sensors. Here it is possible to find several examples such as the use of external services and their APIs, for instance, geolocation based on an IP or information fusion techniques through different physical sensors to obtain a more precise measurement of a contextual factor.
- **Social sensors:** this category refers to all the information which is possible to be extracted from the content posted on social networks and from users' interactions on these networks. For example, the metadata in a given photo, the detection of emotions based on the recognition of expressions on people's faces in photos or the sentiment analysis of the text published in recent posts.

- **Human sensors:** this category is related to how contextual information is explicitly elicited from the user, such as the textual description of a playlist or explicitly expressing a mood the task being performed at that moment.

In the music domain, many of these devices and technologies have been used to obtain valuable contextual information. Table 3 shows the selected works and the types of sensors they employ according to the above categorization.

Table 3. Categories of sensors employed in included articles.

Type of Sensor	Works
Physical or hardware sensors	[43,44,60,62,64,66,70,72,80,81,83,88,102,104–106,124–126]
Virtual or software sensors	[23–49]
Social sensor	[25,26,35,37,39,42,45–49,67–69,85,87,91–94,98–100,103,105,108–120]
Human	[42,44,46,58,59,65,70,74–77,79,80,85,89,91,94,103,105,110,111,121]

In Figure 4a it is possible to see the number of papers in each of these categories. It is necessary to highlight the large presence of works within the “Social Sensor” category, since much of the contextual information (of different types) is extracted thanks to social networks. The second category with more works is the one corresponding to “Virtual Sensor”, this category is very present since much of the implicit information derives from this type of sensors. Likewise, the category human sensor, which refers to jobs in which information is explicitly requested from the user, such as their emotional state, has a large presence. Finally, physical sensors is the category with the least amount of work, but it is still a relevant number.

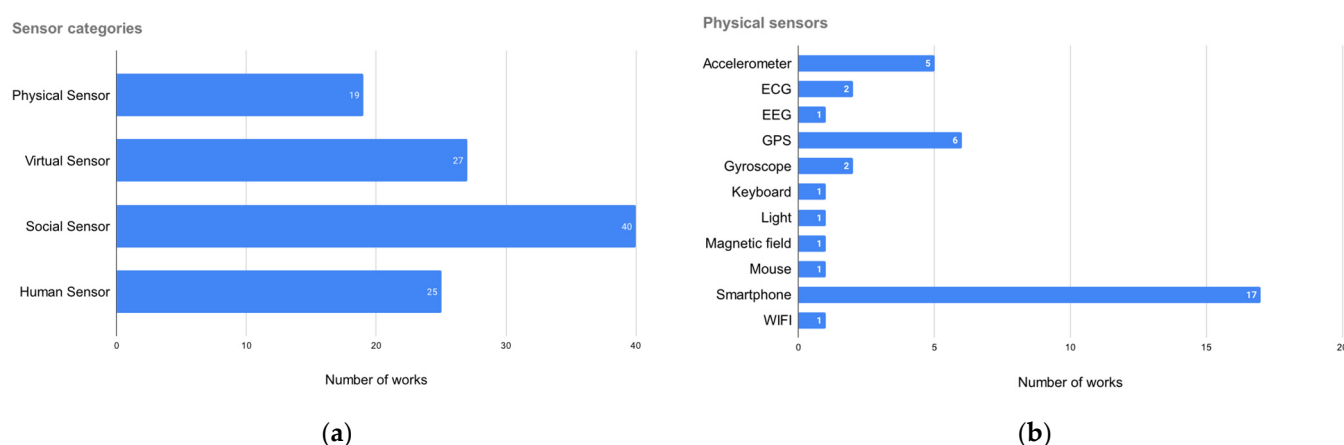


Figure 4. (a) Works employing different sensor categories (b) works employing different physical sensors.

In Figure 4b shows the physical sensors in the latter category, and there is a general category that we have called “smartphone” in order to include those works that only refer to the smartphone and not to the sensors it includes. Among the sensors shown, it is possible to observe works in which explicit reference is made to signals from the accelerometer, gyroscope, ambient light, WiFi signal, magnetic field and others linked to bio signals such as ECG and EEG.

4.3. How Is Context Information Exploited along the Recommendation Process and Which Algorithms Are Employed?

According to Adomavicius et al. [127] the different paradigms in which information can be added into a CARS throughout the recommendation process are:

- **2D Methods:** these methods are used by the vast majority of recommender systems. They work on a two-dimensional space; therefore, the rating function would correspond to (Equation (1)):

$$f_{Rating} : User \times Item \rightarrow Rating \quad (1)$$

However, contextual attributes that can be taken into account at different points in the recommendation process are also considered, giving rise to the paradigms of:

- **Contextual prefiltering:** in this paradigm, contextual attributes are taken into account to filter the data before applying traditional algorithms used in recommender systems. The main advantage of this approach is that all classical recommendation algorithms can be applied; however, if the initial filtering reduces the available data too much then the model may not have enough information to generate relevant recommendations. To mitigate such scenarios, it is possible to use generalization techniques to obtain less specific contexts and to group contextual attributes into hierarchies or to use latent factor models and dimensionality reduction models. In this approach we can also highlight approaches such as item splitting, user splitting and user-item splitting aimed at splitting the profile of an item or a user if the ratings are very different depending on the context, creating new entities of items or users linked to a given context.
- **Contextual postfiltering:** in this paradigm contextual attributes are ignored in the initial part of the process and only considered in the last phase of the recommendation process; therefore, this contextual filtering is applied on recommendations obtained with classical recommendation methods. The recommendations obtained are contextualized according to the contextual attributes in different ways:
 - Filtering or selection: recommendations that are irrelevant to a given context are discarded.
 - Ranking adjustment: the ranking is modified according to a given context.

Furthermore, these techniques can be classified into different approaches:

- Heuristic: seeks the common characteristics of an item for a specific user in the given context.
- Model-based: predictive models are built to calculate the probability with which the user chooses a certain type of item in a given context.

As with the prefiltering approach, this type of approach allows the use of traditional recommendation techniques.

- **Contextual modeling:** in this approach, contextual information is added directly into the recommendation algorithm, so that the working space is a multidimensional space (Equation (2)):

$$f_{Rating} : User \times Item \times Context \rightarrow Rating \quad (2)$$

This context can be made up of multiple dimensions, incorporating this information into the recommendation model, with these contextual dimensions acting as predictors of the user's rating of the item. Two types of approach can be found:

- Model-based: in this approach the contextual dimensions are added directly to the recommendation space, and it is possible to employ a variety of machine learning techniques such as classification and regression like decision trees, SVM (Support Vector Machines), probabilistic models, etc. It is also possible to extend collaborative filtering based on matrix factorization using prominent approaches in this category such as tensor factorization (TF), factorization machines (FM) and context aware matrix factorization (CAMF).
- Heuristic: such approaches employ an extension of k-nearest neighbor (kNN) techniques.

It is worth mentioning that in the literature it is possible to find works that combine these paradigms or carry out contextual modelling following their own type of approach. In this paper we have included many of these papers in the category of contextual modelling as many of them have characteristics of this paradigm.

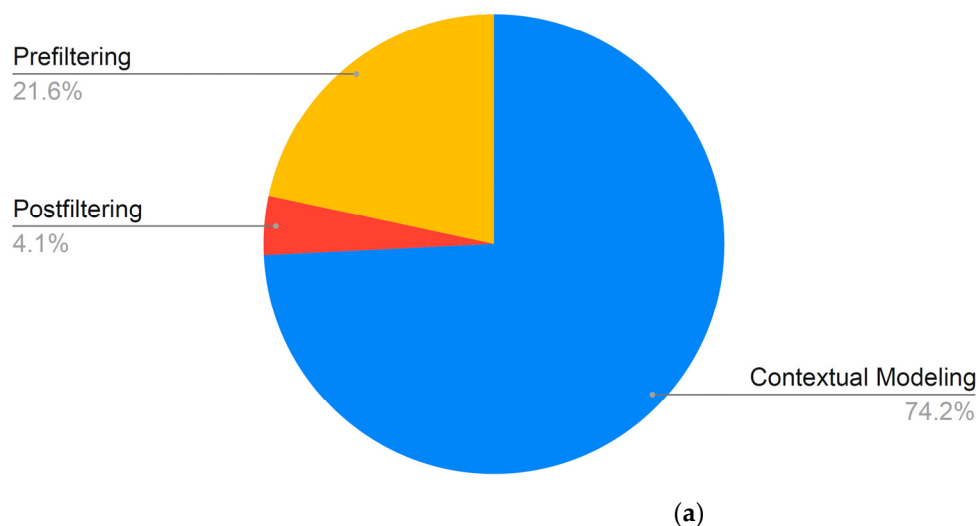
The following Table 4 shows the papers analyzed in this study according to their focus.

Table 4. Contextual paradigms for using context in recommendation.

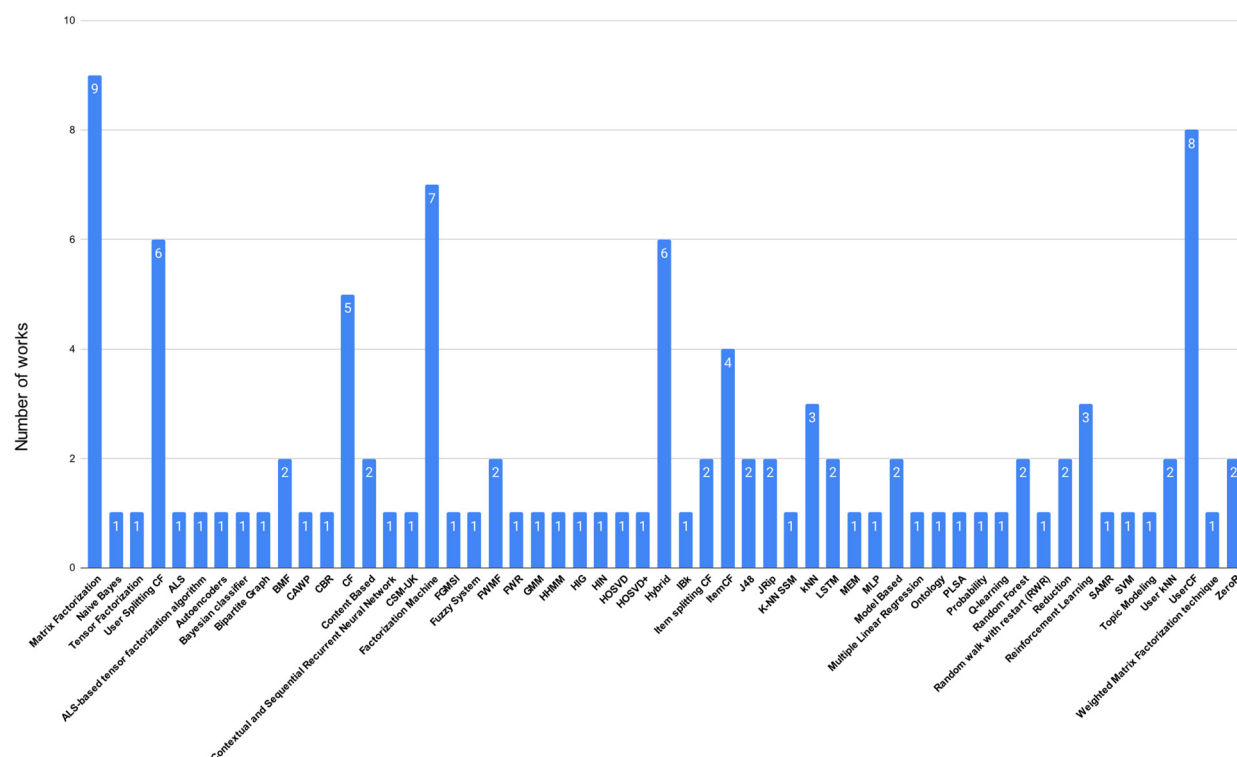
Paradigm	Works
Prefiltering	[24,27,30,31,34,36,38,40,43,56,66,72,74,76–78,81,82,84,98,99,101,109,118,124,128]
Postfiltering	[45,69,80,115]
Contextual Modeling	[23,25,26,28,29,33,35,37,39,41,42,44,46–49,56–65,67,73,75,79,80,83,85–93,96,97,100,102–108,110–117,120,121,125,129–138]

Figure 5a shows the percentage presence of each of the paradigms described above in the papers analyzed. It is worth mentioning that the contextual modelling approach is the most widely used due to the multitude of works that directly include context as predictors of the songs to be recommended. This is followed by the pre-filtering approach that is widely used as it has the advantage of being able to use traditional recommendation methods and includes the context at an earlier stage; and finally the post-filtering approach, which is the least used in the works reviewed.

Contextual paradigms in included works



Algorithms employed in the included works



(b)

Figure 5. (a) Percentages of works employing contextual paradigms (b) algorithms found in the included works.

Figure 5b shows the large number of different algorithms used in the reviewed works. It is possible to highlight in this graph, as mentioned above, the presence of approaches based on matrix factorization including the context such as factorization machines and matrix factorization. It is possible to highlight many works employing User splitting and Item splitting techniques together with classical collaborative filtering techniques in the prefiltering approach and the appearance of new approaches based on Reinforcement Learning in several works. As it is possible to see in this graph there is a great variety of different algorithms in this domain.

4.4. What Evaluation Metrics and Methods Have Been Employed to Validate the Effectiveness of These Music Recommender Systems (RS)?

When assessing the performance of a CARS, different evaluation criteria have been previously categorized in the literature [52,139] and in which we can identify three main groups:

- **User studies:** this type of study is conducted on a set of test subjects who are asked to interact with the recommender system. While the study subjects perform the tasks, their actions are observed, and quantitative data are collected about their interaction. In addition, satisfaction surveys are conducted before and after the experiment to qualitatively measure satisfaction with the system and to check whether the recommendations were relevant.
- **Offline evaluation:** this type of evaluation is carried out when a dataset is available to design the system. The system uses this dataset to predict user ratings on certain items. The performance of the system is, therefore, measured in terms of its ability to correctly predict the rating of users on the items in the dataset. Within this group it

is possible to find the following categories of performance evaluation metrics [140], divided into:

- **Rating prediction metrics:** these measure the accuracy of recommendations in terms of their error. This type of metrics has been broadly used in traditional recommender systems. Among them we can highlight:
 - Mean Absolute Error (MAE): mean absolute error between prediction and rating.
 - Root Mean Squared Error (RMSE): similar to the previous one but this error penalizes major errors more heavily.

A lower value for these metrics indicates a higher predictive power of the models used.

- **Usage prediction metrics:** this type of metric is based on ratios between recommended and consumed items. In this approach, having a dataset with items that the user has consumed, a part is hidden, and the system is asked to recommend a number of items. From the results obtained after the recommendation it is possible to evaluate the number of true positives (relevant and recommended), false positives (recommended and irrelevant), true negatives (not recommended and irrelevant) and false negatives (not recommended and relevant). The following metrics are derived from the relationships between these results:
 - Precision: proportion of items recommended to users that are relevant to them (true positives) over the total number of items recommended (true positives and false positives). If only a certain number of items are recommended it is possible to find the metrics Precision at K or Precision@K.
 - Recall (sensitivity or true positive rate): this metric measures the proportion of consumed items (true positives) that were correctly recommended. This metric is usually displayed together with Precision. As before, it is possible to find the Recall at k or Recall@K metric.
 - Specifity (true negative rate): measures the proportion of non-recommended items that are irrelevant to users.
 - F-measure (F1): this metric combines the above metrics into a single metric to be able to compare different recommender systems.
 - Area under the curve (AUC): the receiver operating characteristics (ROC) curve plots the recall against fallout (false positive rate) and the area under the curve is an indicator of the overall quality of the recommender system.
 - Hit rate: number of items in the test set that were also present in the system's recommendation list for each user is the hit number. The hit rate is the number of hits over the total number of users in the system. Its version at k can be found as Hit@K.
- **Ranking metrics:** these metrics measure the recommender system's ability to predict the correct order of items with respect to the user's preferences, known as rank correlation. Therefore, this type of metric is used when the user is presented with a list of ordered items, as it only takes into account the relative order of these items in the list and not their exact ratings. The most commonly used of this type of metric are:
 - Discounted cumulative gain (DCG) and normalized discounted cumulative gain (NDCG): is a measure based on the assumption that the lower the position in a list of an item, the less useful it is to the user. The relevance value of items in the lowest positions decreases by an amount logarithmically proportional to the position they are in.
 - Mean reciprocal rank (MRR) and cumulative reciprocal rank (CRR): these metrics measure the ranking level at which the system returned the first relevant recommendation.

- Mean average precision (MAP or MAP@K): It is the average of the precision value obtained over all users and items in the top-N lists, where the precision for a position K (prec@K) is the proportion of recommended items in the top-K set that are relevant.
- **Online evaluation:** this evaluation occurs when the experiment is conducted in real time and evaluates user feedback and measures the acceptance of the recommendation in the real world. The techniques to be highlighted in this type of evaluation include:
 - Click-through rate (CTR): this is the count of recommendations that are clicked by the user.
 - Bounce rate: percentage of users who have seen the system's recommendation lists, but instead of consulting those recommendations choose to exit the recommender system.

When evaluating the performance of a recommender system, there are other properties that need to be taken into account [141] beyond the above metrics [131] concerning the predictive power of the models to know whether the recommendations were satisfactory, useful or effective to the users. Among them, special mention may be made of the following:

- Coverage: is defined as the proportion of items on which the system is able to generate recommendations. This metric is important in the music domain because much of the music catalogue is often not recommended due to cold start problems and item popularity bias.
- Novelty: measures the ability of a recommender system to recommend new items that the user was not previously aware of. This property is also relevant in the music domain, in order to discover new artists, although in other domains such as cinema it becomes more important as consumed items are less frequently re-consumed.
- Serendipity: tries to measure relevant and surprising recommendations. There is consensus on the need to increase this property, although the definition of the measure of serendipity is a matter of controversy in the literature.
- Diversity: is another property that measures how different the recommended items are from each other, taking into account several aspects such as musical style, artist, lyrics, instrumentation, etc. Analogous to the previous property, it is possible to find different definitions in the literature. The preferred one is to measure the distance between pairs of items according to an established distance measure and then to sum these distances or average them. In the specific case of music, it is common that importance is given to this aspect since the user's preference may be linked to the fact that there is a variety of styles or artists in the list of recommendations.
- Sequence-aware evaluation measures: in the case of evaluating a recommendation of a music playlist, it is necessary to have other properties such as the transition from one song to another between different styles. Although both songs may be rated very positively by the user, the transition from one style to another such as rock to classical may affect the user's preference. In this case, given a song playing at a specific moment and the presence of other songs that are also a good option to be played right after it, it is necessary to take into account the sequence of items [142,143] and a set of multi metrics (such as intern coherence and diversity) that take into account the list as a whole.

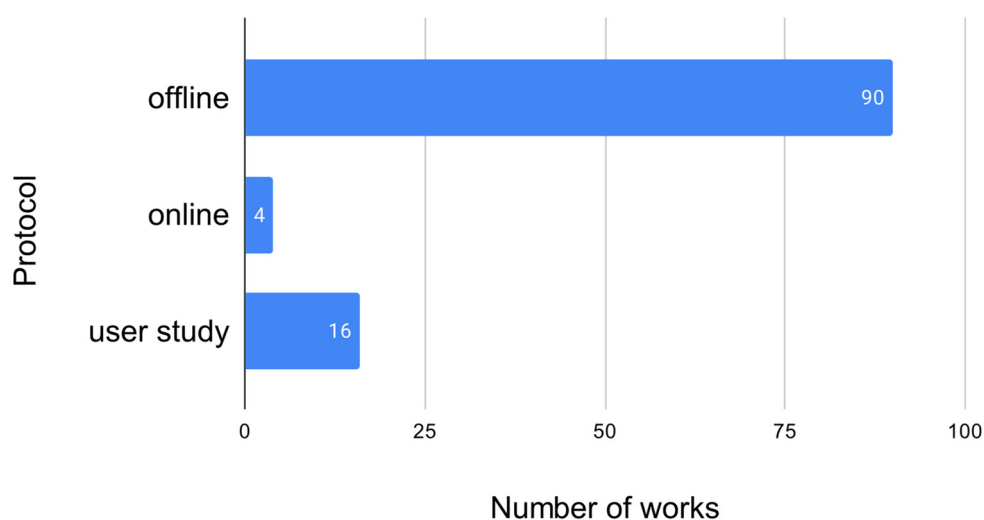
A summary of the evaluation metrics used in the papers analyzed in this study is presented in Table 5 below.

Table 5. Evaluation protocols and metrics employed in the included works.

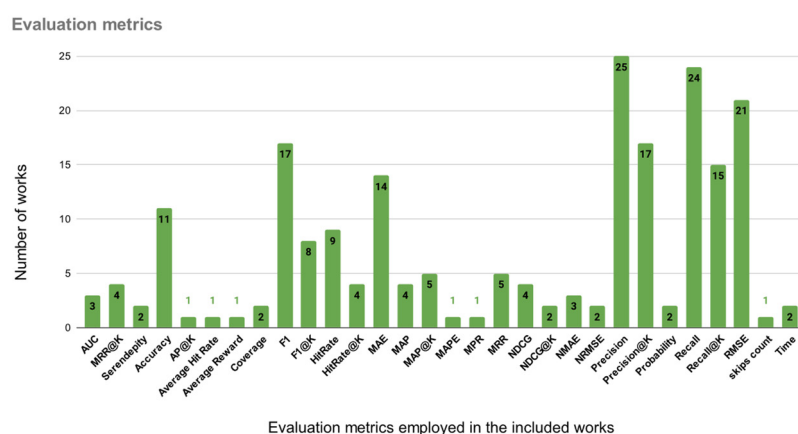
Evaluation		Metrics	Works
Offline	Rating prediction metrics	MAE, RMSE	[27,30,34,39,44,56,59,62,65,74–77,84,94,98,106,108,111,116,118,119]
		Precision	[24,25,29,31,41,44,58,60,61,63,64,67–71,77,79,84,86,87,89,92,95–97,99–102,104,106–108,112,114,115,120,133]
		Recall	[23–25,28,29,31,33,37,39,41,47,48,58,63,64,67–69,71,77,84,86,87,89,92,95–99,102,104,106,108,112,113,115,120]
	Usage prediction metrics		[24–
		F1	26,31,41,58,63,64,69,76,77,84,86,87,89,92,94,95,99,108,112,120]
		AUC	[101,129]
		Hit Rate	[32,36,38,42,46,49,67,87,93,95,96,99,115]
	Ranking metrics	NDCG, MAP, MRR, CRR	[32,36,38,42,46,49,67,87,93,95,96,99,115]
	Online		[48,82,116]
User study		[43,44,57,66,70–73,81,84,85,88,102,104,117,121]	

Similarly, Figure 6a shows the number of papers included in this study that employ each of the evaluation approaches. Notably, a large majority of the papers employ offline evaluation, although we should mention that one of the quality criteria applied is the use of publicly available datasets, so we are aware that this work is biased towards the presence of such studies. Nevertheless, this type of offline evaluation is still the most common in the literature on recommender systems.

Evaluation protocols



(a)



(b)

Figure 6. (a) Works employing different evaluation protocols (b) works employing different evaluation metrics.

In Figure 6b it is possible to highlight the wide use of metrics such as Recall and Precision in their classic versions and top N. Also note the scarce presence of metrics beyond offline evaluation, being present only serendipity and coverage and the absence of metrics such as diversity, making evident the need to work on the use of these alternative metrics that in the domain of music have a special importance.

4.5. Which Publicly Available Datasets Are Mainly Used?

After reviewing the works included in this study, the main datasets used in CARS applied to the music domain have been identified. Among all the works analyzed, it is possible to find some of them that use non-publicly available datasets that have been collected by the authors themselves through users of applications in their studio or through the use of APIs in websites such as Spotify and Last.fm among others.

Table 6 presents the publicly available dataset that are currently used in the music domain, showing their main information, where they can be found and the works in this review that use them.

Table 6. Public datasets for context-aware recommender systems (CARS) in the music domain.

Dataset	Source	Contextual Information Used	Works
InCarMusic	[65]	Modal, Weather, Driving Style, Road Type, Traffic conditions	[56,58,59,65,74–77,79,84]
Hetrec2011-LastFM2K	[144,145]	Social, Timestamp, Tags	[112,114,118]
LastFM1K	[146]	Timestamp	[23,24,28–30,33,38,40,41,98,110,111]
LastFM360K	[147]	Timestamp	[27,34]
LFM-1b	[148,149]	Timestamp	[91]
Million Song Dataset (MSD)	[150]	Music information, Tags	[107]
AotM-2011	[151]	Timestamp	[67,107,115]
#nowplaying-RS	[152]	Hashtags, Sentiment Score, Timestamp	[91]
#nowplaying	[153]	Hashtags, Timestamp	[26,45]
Spotify Playlists Dataset	[154]	Playlist Name	[26]
Million Musical Tweets Dataset (MMTD)	[155,156]	Location, Timestamp	[131]
Yahoo KDD2011	[157,158]	Session Context, Timestamp	[101]
LJ2M	[159]	Mood, Social	[35]
Douban	[160]	Mood	[136]

Figure 7 below shows the datasets used in the works analyzed in this study. It is possible to highlight how the InCarMusic dataset has a strong presence in CARS oriented to the music domain as it was one of the first datasets to include multiple contextual information. Similarly, datasets with information from the company Last.fm [161] with social tags and timestamps have become reference datasets in this type of context.

Datasets employed in the included works

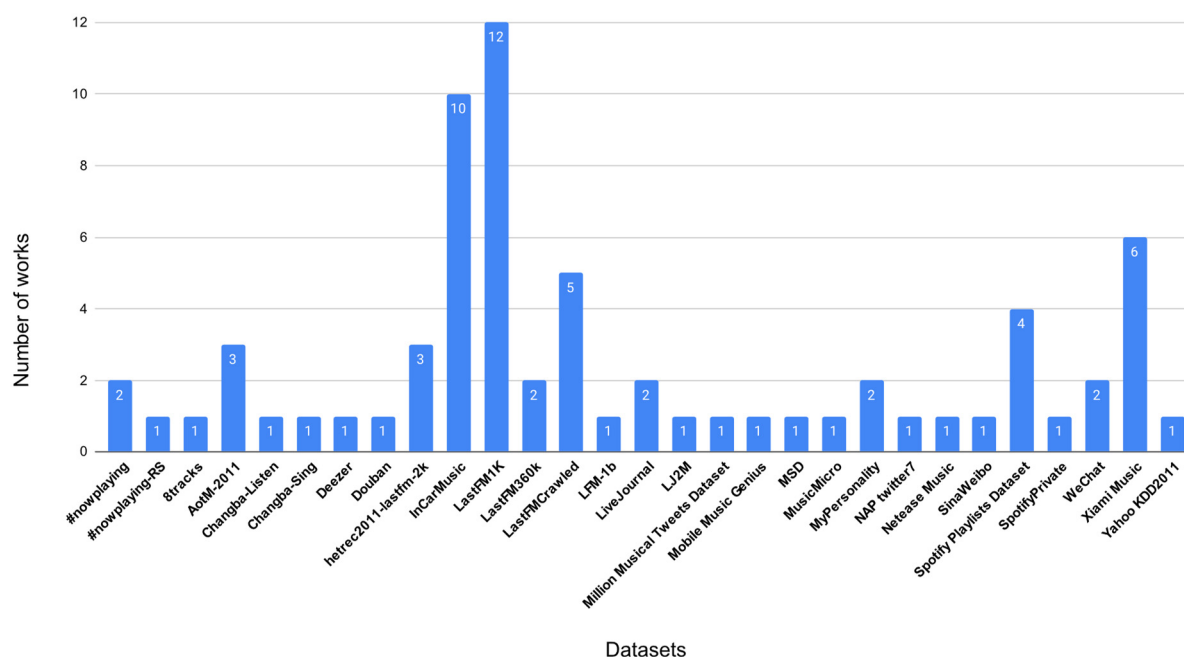


Figure 7. Amount of works which employ each dataset.

Moreover, the aspect of social context has gained much relevance, with the appearance of datasets such as #nowplaying-RS (among many others) which, from information published on social networks not directly dedicated to music, it is possible to infer moods, social contexts and events of consumption of musical items, suitable elements to exploit them in a CARS oriented to the domain of music.

It is common among the works reviewed in this study that the authors make the most of the information available in the datasets published by large companies such as Spotify, in some cases by extracting contextual information simply from the name of the music playlists.

The use of different sources of contextual information and the need to create datasets with all this contextual information that serve as a benchmark is one of the challenges in this domain.

4.6. What Are the Main Open Lines of Research in This Domain?

The work of Schedl et al. [162] provides an extensive review of the current challenges in the domain of music recommender systems; however, it is necessary to point out also the existing challenges in CARS oriented to the music domain:

- Cold start problem: the cold start problem is one of the most common problems in recommender systems and could be aggravated in CARS by the difficulty of obtaining contextual information or because the latent representation of the context is biased [163]. This problem occurs when a new music item is added to the catalogue or

when a user registers and no previous information is available. Linked to this problem is the level of sparsity [164], as the number of ratings is always much lower than the number of possible ratings, a fact that happens when there are platforms with many users and items, as is the case in the music domain. These problems can be mitigated by using contextual information such as social tags, temporal information linked to the plays or by using information from the content of the songs or by exploiting implicit information resulting from the user's interaction with the recommender system.

- Popularity bias: or long tail problem [165] is another problem linked to the previous sparsity challenge. The recommendation system tends to recommend popular items and items with which users have no interactions are not recommended. Kowald et al. [166] develop an extensive work studying this problem in one of the previously mentioned datasets (LFM-1b) [148]. In this work they form three groups of users according to their level of *mainstreamness* score regarding all other users of the dataset. They conclude that both users of this dataset who have limited interest in popular items and users who are interested in unpopular items receive worse recommendations. This type of problem is being addressed in the literature with works such as [111,167,168], but there is still a need to investigate how context can help [69] to mitigate the bias produced by the popularity of songs.
- Dimensionality: using all available contextual information is not the best option as the number of new contextual variables included in these systems leads to the manifestation of "the curse of dimensionality". It is necessary to balance this dimensionality and manage the level of importance of each contextual dimension.
- Evaluation of CARS applied to music domain: there is a need for more contextual datasets that serve as a reference for the community and with aggregated data from different sources such as song content, social tags, friendship relations between network users, social media posts, etc. There is a lot of separate information that could be aggregated to create CARS datasets applied to the music domain that could become benchmarks like MovieLens datasets [169] in the movie domain.
- Enhancing novelty and diversity: suggesting more novel, or less known, products along with diversity within recommendations. Recommendation systems generally provide recommendations of popular items, but it is necessary to introduce mechanisms to improve metrics such as diversity or novelty. In the studies analyzed in this review, only a few of these metrics are included and it is important to take them into account in the design of the CARS.

Furthermore, some of the open lines of research in this field are proposed below:

- Better emotion-aware recommender systems: the emotional state of the user has a great impact on their musical preferences, and it is essential to be able to correctly identify this state at the same time as the recommendations are made. In this sense, the recent social information of the user can be exploited, such as texts with the latest natural language processing (NLP) techniques, photos with artificial vision techniques or even through the tone of voice in interfaces such as voice assistants.
- Situation and intention-aware: accurately inferring the activity that the user is performing more precisely by fusing available information together with identifying the intention with which the user listens to music are contextual factors that can be determinant in the relevance of recommendations.
- Voice-driven interaction with recommender systems in virtual assistants: the rise of devices that include a virtual voice assistant has led to the development of conversational recommender systems that allow recommendations to be made through a dialogue with the user. These devices are bringing the use of streaming services to new users who need to be engaged in dialogues that allow them to extract contextual information, thus providing more relevant recommendations.

- Explainability: another line of work is the generation of reasons using NLP techniques for music recommendations provided [170]. While it is possible to find such explanations in terms of users who listened to a certain group, research is needed on how to find explanations at the context level, such as “users who do sports in the morning also listened to ...” [171].
- Biological data and wearables: the use of devices that monitor vital signs such as smart bands or wearables has not been extensively studied in this domain, or at least there are no large datasets with sufficient contextual information. Exploiting this information and creating a reference dataset could be a future line of research in this area.

5. Discussion

After answering the research questions initially posed in this study and presenting the results of the review, we are now able to discuss the results in each of the questions.

The first question asked was what kind of contextual information is mainly exploited in the music domain. The CARS in the music domain uses a wide variety of types of contextual information. At this point it is possible to highlight how a large number of the reviewed works extensively exploit the following three contexts: (1) temporal and session context of songs, mainly because this is implicit information that can be easily acquired and can provide other derived information from which to extract many contextual factors; (2) emotional context of the user, as moods directly impact the user’s preferences, a large body of work seeks to incorporate these contextual factors when providing recommendations; and (3) contextual factors concerning the user’s physical context and also the context of the user’s activity, with music preferences being closely linked to the task being performed, since music is an item that can be consumed in the background. From the results extracted after reviewing the works in this study, it is possible to observe that mobile devices and music session contexts predominate in the Interaction media contexts. This highlights that studies increasingly tend to try to infer the user’s context based on the music session (in their recent interaction with the system) or simply on the information they have available. This is mainly because contextual information is scarce, especially in the case of the user’s mood. In this case, authors tend to infer this information through patterns in our playback behavior or from external information from social networks. There is a clear trend in the works to use contextual information inferred using machine learning techniques and especially with embeddings that capture such contextual information from different sources.

Secondly, regarding the devices used to obtain contextual information, since most of the works obtain it implicitly, it is logical to observe how the social sphere is exploited as a source, as well as the information derived from the user’s interaction with the system (virtual sensors). However, the number of works that use contextual information from physical devices is not very abundant, since this type of context is much more costly to obtain in real time. Moreover, not many reference datasets with this type of information are publicly available, highlighting the need for the creation of datasets with contextual information obtained from sensors in order to be able to evaluate different approaches.

As wearables, sports wristbands or simply the continuous use of smartphones proliferate, obtaining contextual information about the user’s physical activity from music playback applications (such as Spotify, Google Music, etc.) on mobile devices (Android [172] or iOS [173]) has become trivial. However, there are currently no large public reference datasets in the literature that provide this contextual information. The release of datasets by these large companies could help research into new ways of obtaining context, such as using part of the context when available and using pre-trained models in large datasets to complete partial contextual information.

Thirdly, in terms of how to incorporate contextual information into recommender systems, there is a clear focus on models that use context as a predictor of song ratings, including this information in the model itself under the paradigm of contextual modelling.

Pre-filtering techniques are highly relevant in the literature and post-filtering techniques have also been used, although with a much smaller representation in this study. The myriad of techniques in the contextual modeling approach, seeking to make the most of context as a predictor of musical preferences, is evident.

There is a clear emphasis in the literature on the use of deep learning and embeddings to condense all available information, both contextual, from song content (lyrics, rhythm, valence, audio signal characteristics) and social user information (posts, images, etc.) with the aim of incorporating this information to predict the rating that a user is likely to provide. In most of the works on context-aware recommender systems in which deep learning techniques are applied, contextual states are obtained from topic modeling or image processing, although the literature on the application of these methods in the field of music is much more limited.

Fourthly, the evaluation protocols have been analyzed, highlighting the offline protocol over the online and user studies. There are very few studies that present both an offline evaluation and a user study, and only three studies address other types of metric such as serendipity or coverage. In the music domain, it is worth noting that users can search for new content and for this it is essential to enhance approaches that allow the latter metrics to be enhanced. Another aspect that is currently being explored in the literature is the bias introduced by the long tail effect, in which recommender systems are only able to recommend to their users the most popular items in the catalogue. In this point, several of the recent works analyzed focus on how to address this issue and how context can help in identifying new music beyond the most popular songs. Furthermore, in the specific case of playlist recommendations, there is a need to ensure that the items are not only suitable for the user, but that it is possible to control the diversity in the playlist.

With these results, it becomes quite evident that there is a need to work on the evaluation of metrics such as diversity in this domain.

Fifthly, the datasets used by the works included in this study have been analyzed and the publicly available datasets have been detailed. At this point, it is possible to highlight how most of the works use datasets based on Last.fm and how in recent years new datasets have emerged that incorporate social information from social networks such as Twitter. However, the information that is available in these datasets is partial and often needs to be enriched by researchers by inferring certain contextual factors from the external information available, which in many cases is incomplete (location of the tweet if it is geolocated, language of the tweet, time of publication, etc.). The information fusion from different sources and the increasing size of published datasets becomes a challenge for researchers, increasingly requiring big data processing approaches.

At the same time, new datasets need to be published to serve as a reference for the reproducibility of the results of the different approaches proposed in the analyzed studies. As shown in Section 4.5, some of the reference datasets are publicly available but a large number of research papers compile their own datasets and do not release them to the rest of the scientific community, hindering reproducibility and the creation of new recommendation approaches with such data.

Finally, some of the current challenges in CARS applied to the domain of music have been described and some of the future lines of work in the field proposed in the literature and others raised by this paper after the review of the selected studies have been pointed out.

6. Conclusions and Future Lines of Work

A systematic literature review has been carried out to analyze previous CARS work in the specific domain of music recommendation. This systematic literature review addresses different important topics such as the main types of contexts used in recent works in the literature and pointing out those that have not been explored deeply enough.

In addition, the main devices and technologies used in the literature to extract contextual information have been identified. It has been shown in what proportion they have been used in this specific domain, showing clear trends in the exploitation of social context.

Subsequently, the different paradigms for incorporating this contextual information into the recommendation process were analyzed and the alternatives taken by the authors of the works analyzed, showing a wide variety of approaches, mostly under the paradigm of contextual modelling.

Afterwards, the evaluation protocols and metrics commonly used in each of them in the music domain have been identified.

The protocols and metrics used in each of the papers have been analyzed and, despite the fact that this study may be biased as the quality criteria for the inclusion of papers valued the use of public datasets, we are able to conclude a wide use of offline protocols and metrics from Precision and Recall and their @K versions as the most widely used to evaluate the performance of CARS in the music domain.

Next, an extensive review of the datasets used in the works included in the study has been carried out, showing a selection of the mainly used and publicly available datasets. This is a very useful resource for researchers new to this domain and encourages the use of reference datasets in the field in order to encourage reproducibility of research and comparison of different approaches. In addition, the datasets that have been widely used by the included works and other recent datasets with contextual information have been presented. The latter will allow new advances in this domain, exploiting information in social networks from different types of data such as text or images.

Finally, some of the challenges in this domain have been identified, such as the use of metrics that favor diversity and novelty or how to deal with the problem of excessive dimensionality introduced by contextual variables, among others. Likewise, potential new lines of research in this field have been presented, such as the development of systems capable of explaining recommendations or the presence of recommender systems with communication interfaces as conversational agents in voice assistants.

To the best of our knowledge there is no previous SLR in this specific domain and we consider that this work can support developers and researchers in music recommendation and context-aware recommender systems by helping them in their initial documentation, in the identification of the most used and referenced datasets in the literature, and in the exploration of new research approaches in the music domain.

Future lines of work following this systematic review of the literature include the construction of a dataset that combines information from the previously reviewed datasets with social and content information of the songs in order to perform automatic contextual labelling of new songs without previous contextual information, the use of new interaction methods to extract contextual information such as in the case of recommender systems in voice assistants and conversational systems, or to explore the context information in recommender systems for groups in the music domain, a kind of recommender system not addressed in this review.

Author Contributions: Conceptualization and methodology Á.L.M. and M.N.M.-G.; software, A.V.R. and D.M.J.-B.; validation, Á.L.M. and M.N.M.-G.; data curation, Á.L.M. and J.F.D.P.S.; writing—original draft preparation, A.V.R., Á.L.M., M.N.M.-G. and D.M.J.-B.; writing—review and editing, A.V.R., J.F.D.P.S., Á.L.M. and D.M.J.-B.; visualization, A.V.R., J.F.D.P.S., Á.L.M. and M.N.M.-G.; supervision, Á.L.M. and M.N.M.-G. All authors have read and agreed to the published version of the manuscript.

Funding: Álvaro Lozano research work is supported by a postdoctoral fellowship “Programme II: Postdoctoral Contracts” from the University of Salamanca. The research of Diego M. Jiménez-Bravo has been co-financed by the European Social Fund and Junta de Castilla y León (Operational Programme 2014–2020 for Castilla y León, EDU/574/2018 BOCYL). This research has also been funded by the Department of Education of the Junta de Castilla y León, Spain, (ORDEN EDU/667/2019). Project code: 2019/00266/001.

Acknowledgments: We are grateful to the University of Salamanca for supporting this work through the use of its facilities and providing access to bibliographic sources.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ricci, F.; Shapira, B.; Rokach, L. *Recommender Systems Handbook*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2015; ISBN 9781489976376.
2. Machado, G.M.; Maran, V.; Dornelles, L.P.; Gasparini, I.; Thom, L.H.; de Oliveira, J.P.M. A systematic mapping on adaptive recommender approaches for ubiquitous environments. *Computing* **2018**, *100*, 183–209.
3. Context Aware Recommender Systems—CARS Workshop. Available online: <https://cars-workshop.com/> (accessed on 19 May 2021).
4. Jassy, A. 2012 re:Invent Day 1 Keynote: Andy Jassy. Available online: <https://www.youtube.com/watch?v=8FJ5DBLSFe4> (accessed on 19 May 2021).
5. Colombo-Mendoza, L.O.; Valencia-García, R.; Rodríguez-González, A.; Alor-Hernández, G.; Samper-Zapater, J.J. RecomMetz: A context-aware knowledge-based mobile recommender system for movie showtimes. *Expert Syst. Appl.* **2015**, *42*, 1202–1222.
6. Chen, C.; Meng, X.; Xu, Z.; Lukasiewicz, T. Location-aware personalized news recommendation with deep semantic analysis. *IEEE Access* **2017**, *5*, 1624–1638.
7. Korakakis, M.; Mylonas, P.; Spyrou, E. Xenia: A context aware tour recommendation system based on social network metadata information. In *Proceedings of the 11th International Workshop on Semantic and Social Media Adaptation and Personalization, SMAP 2016*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2016; pp. 59–64.
8. Jalan, K.; Gawande, K. Context-aware hotel recommendation system based on hybrid approach to mitigate cold-start-problem. In *Proceedings of the 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing, ICECDS 2017*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2018; pp. 2364–2370.
9. Bentaleb, A.; El Bouzekri, Y.; Lahcen, A.A.; Boulmalf, M. Context Aware Recommender Systems for Tourism: A Concise Review. In *Proceedings of the Colloquium in Information Science and Technology, CIST*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2018; pp. 149–154.
10. Gallego, D.; Barra, E.; Rodriguez, P.; Huecas, G. Incorporating proactivity to context-aware recommender systems for e-learning. In *Proceedings of the 2013 World Congress on Computer and Information Technology, WCCIT, Sousse, Tunisia, 22–24 June 2013*.
11. Hong, W.; Li, L.; Li, T. Product recommendation with temporal dynamics. *Expert Syst. Appl.* **2012**, *39*, 12398–12406.
12. Wu, K.K.; Liu, P.; Helen, M.; Yam, Y. An embedding approach for context-aware collaborative recommendation and visualization. In *Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2016*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2017; pp. 3457–3462.
13. Maia, R.; Ferreira, J.C. Context-aware food recommendation system. *Lect. Notes Eng. Comput. Sci.* **2018**, *2237*, 349–356.
14. Zhu, Q.; Wang, L. Context-Aware Restaurant Recommendation for Group of People. In *Proceedings of the 2020 IEEE World Congress on Services, SERVICES 2020*; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2020; pp. 51–54.
15. Ji, K.; Sun, R.; Li, X.; Shu, W. Improving matrix approximation for recommendation via a clustering-based reconstructive method. *Neurocomputing* **2016**, *173*, 912–920.
16. Schedl, M.; Knees, P.; McFee, B.; Bogdanov, D.; Kaminskas, M. Music Recommender Systems. In *Recommender Systems Handbook*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 453–492.
17. Burland, K. The Social and Applied Psychology of Music by Adrian, C. North & David, J. Hargreaves. *Br. J. Music Educ.* **2010**, *27*, 101–103.
18. Kitchenham, B.; Charters, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Joint Report; Durham University: Durham, UK, 2007.
19. Petticrew, M.; Roberts, H. *Systematic Reviews in the Social Sciences: A Practical Guide*; Blackwell Publishing Ltd: Worcestershire, UK, 2008; ISBN 1405121106.
20. Keele, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*; Version 2.3; EBSE Technical Report; University of Durham: Durham, UK, 2007.
21. Wohlin, C.; Runeson, P.; Höst, M.; Ohlsson, M.C.; Regnell, B.; Wesslén, A. Systematic Literature Reviews. In *Experimentation in Software Engineering*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 45–54.
22. Parsifal. Available online: <https://parsif.al/> (accessed on 19 May 2021).
23. Hidasi, B.; Tikk, D. Fast ALS-based tensor factorization for context-aware recommendation from implicit feedback. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2012; Volume 7524, pp. 67–82. ISBN 9783642334856.

24. Li, M.; Jiang, W.; Li, K. When and what music will you listen to? Fine-grained time-aware music recommendation. In Proceedings of the 15th IEEE International Symposium on Parallel and Distributed Processing with Applications and 16th IEEE International Conference on Ubiquitous Computing and Communications, ISPA/IUCC 2017, Guangzhou, China, 12–15 December 2017; pp. 1091–1098.
25. Wang, D.; Xu, G.; Deng, S. Music recommendation via heterogeneous information graph embedding. In Proceedings of the International Joint Conference on Neural Networks; IEEE: Piscataway, NJ, USA, 2017; Volume 2017, pp. 596–603.
26. Pichl, M.; Zangerle, E.; Specht, G. Towards a Context-Aware Music Recommendation Approach: What is Hidden in the Playlist Name? In ICDMW '15: Proceedings of the 2015 IEEE International Conference on Data Mining Workshop (ICDMW); IEEE: Piscataway, NJ, USA, 2016; pp. 1360–1365.
27. Sánchez-Moreno, D.; Zheng, Y.; Moreno-García, M.N. Time-aware music recommender systems: Modeling the evolution of implicit user preferences and user listening habits in a collaborative filtering approach. *Appl. Sci.* **2020**, *10*, 1–32.
28. Katarya, R.; Verma, O.P. Efficient music recommender system using context graph and particle swarm. *Multimed. Tools Appl.* **2018**, *77*, 2673–2687.
29. Cai, Z.Q.; Hu, H. Session-aware music recommendation via a generative model approach. *Soft Comput.* **2018**, *22*, 1023–1031.
30. Darapisut, S.; Suksawatchon, J. Comparison of the constant prediction time of collaborative filtering algorithms by using time contexts. In Proceedings of the 2014 11th International Joint Conference on Computer Science and Software Engineering: “Human Factors in Computer Science and Software Engineering”—e-Science and High Performance Computing: eHPC, JCSSE 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 302–306.
31. Darapisut, S.; Suksawatchon, U.; Suksawatchon, J. The constant time of predictive algorithm for music recommendation with time context. In Proceedings of the 2015 12th International Joint Conference on Computer Science and Software Engineering, JCSSE 2015; IEEE: Piscataway, NJ, USA, 2015; pp. 63–68.
32. Rumiantcev, M. Emotions and activity recognition system using wearable device sensors. In Proceedings of the Conference of Open Innovation Association, FRUCT; IEEE: Piscataway, NJ, USA, 2021.
33. Tao, Y.; Zhang, Y.; Bian, K. Attentive context-aware music recommendation. In Proceedings of the 2019 IEEE 4th International Conference on Data Science in Cyberspace, DSC 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 54–61.
34. Sanchez-Moreno, D.; Zheng, Y.; Moreno-Garcia, M.N. Incorporating Time Dynamics and Implicit Feedback into Music Recommender Systems. In Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence (WI 2018), Santiago, Chile, 3–6 December 2018; pp. 580–585.
35. Dhahri, C.; Matsumoto, K.; Hoashi, K. Mood-aware music recommendation via adaptive song embedding. In Proceedings of the 2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 135–138.
36. Lee, D.; Park, S.E.; Kahng, M.; Lee, S.; Lee, S.G. Exploiting contextual information from event logs for personalized recommendation. In *Studies in Computational Intelligence*; Springer: Berlin/Heidelberg, Germany, 2010; Volume 317, pp. 121–139. ISBN 9783642154041.
37. Chen, C.M.; Tsai, M.F.; Liu, J.Y.; Yang, Y.H. Using emotional context from article for contextual music recommendation. In Proceedings of the 21st ACM International Multimedia Conference, (MM'13), Barcelona, Spain, 21–25 October 2013; pp. 649–652.
38. Dias, R.; Fonseca, M.J. Improving music recommendation in session-based collaborative filtering by using temporal context. In Proceedings of the International Conference on Tools with Artificial Intelligence, ICTAI; IEEE: Piscataway, NJ, USA, 2013; pp. 783–788.
39. Chen, C.M.; Tsai, M.F.; Liu, J.Y.; Yang, Y.H. Music recommendation based on multiple contextual similarity information. In Proceedings of the 2013 IEEE/WIC/ACM International Conference on Web Intelligence, WI 2013; IEEE: Piscataway, NJ, USA, 2013; Volume 1, pp. 65–72.
40. Naser, I.; Pagare, R.; Wathap, N.; Pingale, V. Hybrid music recommendation system: Enhanced collaborative filtering using context and interest based approach. In Proceedings of the 11th IEEE India Conference: Emerging Trends and Innovation in Technology, INDICON 2014; IEEE: Piscataway, NJ, USA, 2015.
41. Hosseinzadeh Aghdam, M. Context-aware recommender systems using hierarchical hidden Markov model. *Phys. A Stat. Mech. Appl.* **2019**, *518*, 89–98.
42. Shen, T.; Jia, J.; Li, Y.; Ma, Y.; Bu, Y.; Wang, H.; Chen, B.; Chua, T.-S.; Hall, W. PEIA: Personality and Emotion Integrated Attentive Model for Music Recommendation on Social Media Platforms. *Proc. AAAI Conf. Artif. Intell.* **2020**, *34*, 206–213.
43. Ospina-Bohórquez, A.; Gil-González, A.B.; Moreno-García, M.N.; de Luis-Reboredo, A. Context-aware music recommender system based on automatic detection of the user's physical activity. In Proceedings of the Advances in Intelligent Systems and Computing; Springer: Berlin, Germany, 2021; Volume 1237, pp. 142–151.
44. Yousefian Jazi, S.; Kaedi, M.; Fatemi, A. An emotion-aware music recommender system: Bridging the user's interaction and music recommendation. *Multimed. Tools Appl.* **2021**, *80*, 13559–13574.
45. Zangerle, E.; Chen, C.M.; Tsai, M.F.; Yang, Y.H. Leveraging Affective Hashtags for Ranking Music Recommendations. *IEEE Trans. Affect. Comput.* **2021**, *12*, 78–91.
46. Shen, T.; Jia, J.; Li, Y.; Wang, H.; Chen, B. Enhancing Music Recommendation with Social Media Content: An Attentive Multimodal Autoencoder Approach. In Proceedings of the International Joint Conference on Neural Networks; IEEE: Piscataway, NJ, USA, 2020.
47. Hansen, C.; Hansen, C.; Maystre, L.; Mehrotra, R.; Brost, B.; Tomasi, F.; Lalmas, M. Contextual and Sequential User Embeddings for Large-Scale Music Recommendation. In Proceedings of the 14th ACM Conference on Recommender Systems, Online, 22–26 September 2020; pp. 53–62.

48. Moscato, V.; Picariello, A.; Sperli, G. An emotional recommender system for music. *IEEE Intell. Syst.* **2020**, doi:10.1109/MIS.2020.3026000.
49. Zhou, C.; Jin, Y.; Wang, X.; Zhang, Y. Conversational music recommendation based on bandits. In *Proceedings of the 11th IEEE International Conference on Knowledge Graph, ICKG 2020*; IEEE: Piscataway, NJ, USA, 2020; pp. 41–48.
50. Lozano Murcigo, Á. Context-Aware Recommender Systems in the Music Domain: A Systematic Literature Review. Available online: <https://bit.ly/3oXyHHT> (accessed on 26 January 2021).
51. elloza/SLR-Context-Aware-Recommenders-Music-Domain: This repository contains all the information about the SLR Conducted about Context Aware Recommender Systems in the Music Domain. Available online: <https://github.com/elloza/SLR-Context-Aware-Recommenders-Music-Domain> (accessed on 28 May 2021).
52. Villegas, N.M.; Sánchez, C.; Díaz-Cely, J.; Tamura, G. Characterizing context-aware recommender systems: A systematic literature review. *Knowl. Based Syst.* **2018**, *140*, 173–200.
53. Villegas, N.M.; Müller, H.A. Managing dynamic context to optimize smart interactions and services. *Lect. Notes Comput. Sci.* **2010**, *6400*, 289–318.
54. Adomavicius, G.; Mobasher, B.; Ricci, F.; Tuzhilin, A. Context-aware recommender systems. *AI Mag.* **2011**, *32*, 67–80.
55. Delic, A.; Masthoff, J. Group recommender systems; In *SpringerBriefs in Electrical and Computer Engineering*; Springer International Publishing: Cham, Switzerland, 2018; ISBN 9781450355896.
56. Codina, V.; Ricci, F.; Ceccaroni, L. Exploiting the semantic similarity of contextual situations for pre-filtering recommendation. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 7899, pp. 165–177. ISBN 9783642388439.
57. Kaminskas, M.; Ricci, F. Emotion-Based Matching of Music to Places. In *Human-Computer Interaction Series*; Springer International Publishing: Cham, Switzerland, 2016; pp. 287–310.
58. Dridi, R.; Zammali, S.; Arour, K. Fuzzy rule-based situational music retrieval and recommendation. In *Proceedings of the of IEEE/ACS International Conference on Computer Systems and Applications, AICCSA*; IEEE: Piscataway, NJ, USA, 2018; Volume 2017, pp. 549–556.
59. Dridi, R.; Zammali, S.; Arour, K. Situation-aware rating prediction using fuzzy rules. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2016; Volume 9983, pp. 209–221, ISBN 9783319476490.
60. Cheng, Z.; Shen, J. Just-for-Me: An adaptive personalization system for location-aware social music recommendation. In *Proceedings of the ACM International Conference on Multimedia Retrieval*, Glasgow, UK, 1–4 April 2014; pp. 185–192.
61. Cheng, Z.; Shen, J. On effective location-aware music recommendation. *ACM Trans. Inf. Syst.* **2016**, *34*, 1–32.
62. Magara, M.B.; Ojo, S.; Ngwira, S.; Zuva, T. MPList: Context aware music playlist. In *Proceedings of the 2016 IEEE International Conference on Emerging Technologies and Innovative Business Practices for the Transformation of Societies, EmergiTech 2016*; IEEE: Piscataway, NJ, USA, 2016; pp. 309–316.
63. Mastrika Giri, G.A.V.; Harjoko, A. Music recommendation system based on context using case-based reasoning and self organizing map. *Indones. J. Electr. Eng. Comput. Sci.* **2016**, *4*, 459–464.
64. Su, J.H.; Yeh, H.H.; Yu, P.S.; Tseng, V.S. Music recommendation using content and context information mining. *IEEE Intell. Syst.* **2010**, *25*, 16–26.
65. Baltrunas, L.; Kaminskas, M.; Ludwig, B.; Moling, O.; Ricci, F.; Aydin, A.; Lüke, K.H.; Schwaiger, R. InCarMusic: Context-aware music recommendations in a car. In *Lecture Notes in Business Information Processing*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 85, pp. 89–100, ISBN 9783642230134.
66. Braunhofer, M.; Kaminskas, M.; Ricci, F. Recommending music for places of interest in a mobile travel guide. In *Proceedings of the 5th ACM Conference on Recommender Systems, RecSys'11*, Chicago, IL, USA, 23–27 October 2011; pp. 253–256.
67. Hariri, N.; Mobasher, B.; Burke, R. Using social tags to infer context in hybrid music recommendation. *Int. Conf. Inf. Knowl. Manag. Proc.* **2012**, doi:10.1145/2389936.2389946.
68. Schedl, M. Ameliorating music recommendation: Integrating music content, music context, and user context for improved music retrieval and recommendation. In *Proceedings of the ACM International Conference Proceeding Series*, St. Andrews, UK, 6–9 October 2013; pp. 3–9.
69. Domingues, M.A.; Rezende, S.O. The impact of context-aware recommender systems on music in the Long Tail. In *Proceedings of the Brazilian Conference on Intelligent Systems (BRACIS)*, Fortaleza, Brazil, 19–24 October 2013; pp. 119–124.
70. Yang, Y.H.; Teng, Y.C. Quantitative study of music listening behavior in a smartphone context. *ACM Trans. Interact. Intell. Syst.* **2015**, *5*, 1–30.
71. Ojagh, S.; Malek, M.R.; Saeedi, S.; Liang, S. An Internet of Things (IoT) Approach for Automatic Context Detection. In *Proceedings of the 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2018*; IEEE: Piscataway, NJ, USA, 2019; pp. 223–226.
72. Wang, M.; Kawamura, T.; Sei, Y.; Nakagawa, H.; Tahara, Y.; Ohsuga, A. Music recommender adapting implicit context using “renso” relation among Linked Data. *J. Inf. Process.* **2014**, *22*, 279–288.
73. Schedl, M.; Breitschopf, G.; Ionescu, B. Mobile music genius: Reggae at the beach, metal on a friday night? In *Proceedings of the ACM International Conference on Multimedia Retrieval*, Glasgow, UK, 1–4 April 2014; pp. 507–510.
74. Vahidi Ferdousi, Z.; Colazzo, D.; Negre, E. CBPF: Leveraging context and content information for better recommendations. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2018; Volume 11323, pp. 381–391. ISBN 9783030050894.

75. Zhao, V.N.; Moh, M.; Moh, T.S. Contextual-Aware Hybrid Recommender System for Mixed Cold-Start Problems in Privacy Protection. In Proceedings of the IEEE 2nd International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), New York, NY, USA, 9–10 April 2016; pp. 400–405.
76. Dixit, V.S.; Jain, P. Weighted percentile-based context-aware recommender system. In *Advances in Intelligent Systems and Computing*; Springer: Singapore, 2019; Volume 697, pp. 377–388.
77. Jain, P.; Dixit, V.S. Recommendations with context aware framework using particle swarm optimization and unsupervised learning. *J. Intell. Fuzzy Syst.* **2019**, *36*, 4479–4490.
78. Kaminskas, M.; Fernández-Tobías, I.; Cantador, I.; Ricci, F. Ontology-Based Identification of Music for Places. In *Information and Communication Technologies in Tourism 2013*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 436–447.
79. Zheng, Y.; Mobasher, B.; Burke, R. CSLIM: Contextual SLIM recommendation algorithms. In Proceedings of the 8th ACM Conference on Recommender Systems, RecSys’14, Foster City, CA, USA, 6–10 October 2014; pp. 301–304.
80. Abdul, A.; Chen, J.; Liao, H.Y.; Chang, S.H. An emotion-aware personalized music recommendation system using a convolutional neural networks approach. *Appl. Sci.* **2018**, *8*, 1103.
81. Wang, M.; Kawamura, T.; Sei, Y.; Nakagawa, H.; Tahara, Y.; Ohsuga, A. Context-aware music recommendation with serendipity using semantic relations. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2014; Volume 8388, pp. 17–32. ISBN 9783319068251.
82. Kaminskas, M.; Fernández-Tobías, I.; Ricci, F.; Cantador, I. Knowledge-based identification of music suited for places of interest. *Inf. Technol. Tour.* **2014**, *14*, 73–95.
83. Gillhofer, M.; Schedl, M. Iron maiden while jogging, debussy for dinner?: An analysis of music listening behavior in context. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2015; Volume 8936, pp. 380–391. ISBN 9783319144412.
84. Sain Dixit, V.; Jain, P. A Proposed Framework for Recommendations Aggregation in Context Aware Recommender Systems. In Proceedings of the 8th International Conference on Cloud Computing, Data Science and Engineering (Confluence), Noida, India, 11–12 January 2018; pp. 231–236.
85. Ben Sassi, I.; Ben Yahia, S. How does context influence music preferences: A user-based study of the effects of contextual information on users’ preferred music. *Multimed. Syst.* **2021**, *27*, 143–160.
86. Rho, S.; Song, S.; Nam, Y.; Hwang, E.; Kim, M. Implementing situation-aware and user-adaptive music recommendation service in semantic web and real-time multimedia computing environment. *Multimed. Tools Appl.* **2013**, *65*, 259–282.
87. Wang, D.; Deng, S.; Xu, G. GEMRec: A graph-based emotion-aware music recommendation approach. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2016; Volume 10041, pp. 92–106. ISBN 9783319487397.
88. King, J.; Imbrasaite, V. Generating music playlists with hierarchical clustering and Q-learning. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2015; Volume 9022, pp. 315–326. ISBN 9783319163536.
89. Yang, Q.; Zhan, L.; Han, L.; Zhang, J.; Bi, Z. Recommending more suitable music based on users’ real context. In *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*; Springer International Publishing: Cham, Switzerland, 2019; Volume 268, pp. 124–137. ISBN 9783030129804.
90. Roy, S.; Biswas, M.; De, D. iMusic: A session-sensitive clustered classical music recommender system using contextual representation learning. *Multimed. Tools Appl.* **2020**, *79*, 24119–24155.
91. Takama, Y.; Zhang, J.C.; Shibata, H. Context-aware music recommender system based on implicit feedback. *Trans. Jpn. Soc. Artif. Intell.* **2021**, *36*, 1–10.
92. Wang, D.; Zhang, X.; Yu, D.; Xu, G.; Deng, S. CAME: Content- And Context-Aware Music Embedding for Recommendation. *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *32*, 1375–1388.
93. Polignano, M.; Narducci, F.; de Gemmis, M.; Semeraro, G. Towards Emotion-aware Recommender Systems: An Affective Coherence Model based on Emotion-driven Behaviors. *Expert Syst. Appl.* **2021**, *170*, 114382.
94. Ibrahim, K.M.; Royo-Letelier, J.; Epure, E.V.; Peeters, G.; Richard, G. Audio-Based Auto-Tagging with Contextual Tags for Music. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, Virtual Conference, 4–9 May 2020; pp. 16–20.
95. Cheng, Z.; Shen, J. VenueMusic: A venue-aware music recommender system. In Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR’15, Santiago, Chile, 9–13 August 2015; pp. 1029–1030.
96. Wang, D.; Deng, S.; Zhang, X.; Xu, G. Learning to embed music and metadata for context-aware music recommendation. *World Wide Web* **2018**, *21*, 1399–1423.
97. Jin, L.; Yuan, D.; Zhang, H. Music recommendation based on embedding model with user preference and context. In *Proceedings of the 2017 IEEE 2nd International Conference on Big Data Analysis, ICBDA 2017*; IEEE: Piscataway, NJ, USA, 2017; pp. 688–692.
98. Shah, F.; Desai, M.; Pati, S.; Mistry, V. Hybrid Music Recommendation System Based on Temporal Effects. In *Advances in Intelligent Systems and Computing*; Springer: Singapore, 2020; Volume 1034, pp. 569–577. ISBN 9789811510830.
99. Selvi, C.; Sivasankar, E. An Efficient Context-Aware Music Recommendation Based on Emotion and Time Context. In *Lecture Notes on Data Engineering and Communications Technologies*; Springer: Singapore, 2019; Volume 16, pp. 215–228.

100. Yang, H.; Zhao, Y.; Xia, J.; Yao, B.; Zhang, M.; Zheng, K. Music playlist recommendation with long short-term memory. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2019; Volume 11447, pp. 416–432. ISBN 9783030185787.
101. Zheng, L.; Zhu, F.; Huang, S.; Xie, J. Context Neighbor Recommender: Integrating contexts via neighbors for recommendations. *Inf. Sci.* **2017**, *414*, 1–18.
102. Alhamid, M.F.; Rawashdeh, M.; Dong, H.; Hossain, M.A.; Alelaiwi, A.; El Saddik, A. RecAm: A collaborative context-aware framework for multimedia recommendations in an ambient intelligence environment. *Multimed. Syst.* **2016**, *22*, 587–601.
103. Pichl, M.; Zangerle, E. Latent feature combination for multi-context music recommendation. In *Proceedings of the International Workshop on Content-Based Multimedia Indexing*; IEEE: Piscataway, NJ, USA, 2018; Volume 2018.
104. Alhamid, M.F.; Rawashdeh, M.; Al Osman, H.; Hossain, M.S.; El Saddik, A. Towards context-sensitive collaborative media recommender system. *Multimed. Tools Appl.* **2015**, *74*, 11399–11428.
105. Zhou, C.; Jin, Y.; Zhang, K.; Yuan, J.; Li, S.; Wang, X. MusicRoBot: Towards conversational context-aware music recommender system. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2018; Volume 10828, pp. 817–820. ISBN 9783319914572.
106. Lee, W.P.; Chen, C.T.; Huang, J.Y.; Liang, J.Y. A smartphone-based activity-aware system for music streaming recommendation. *Knowl. Based Syst.* **2017**, *131*, 70–82.
107. Vall, A.; Quadrana, M.; Schedl, M.; Widmer, G. Order, context and popularity bias in next-song recommendations. *Int. J. Multimed. Inf. Retr.* **2019**, *8*, 101–113.
108. Pichl, M.; Zangelere, E.; Specht, G. Improving context-aware music recommender systems: Beyond the pre-filtering approach. In *Proceedings of the International Conference on Multimedia Retrieval, ICMR'17*, Bucharest, Romania, 6–9 June 2017; pp. 201–208.
109. Kim, H.H. A semantically enhanced tag-based music recommendation using emotion ontology. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 7803, pp. 119–128. ISBN 9783642365423.
110. Sánchez-Moreno, D.; Pérez-Marcos, J.; Gil González, A.B.; Batista, V.L.; Moreno-García, M.N. Social influence-based similarity measures for user-user collaborative filtering applied to music recommendation. *Adv. Intell. Syst. Comput.* **2019**, *801*, 267–274.
111. Sánchez-Moreno, D.; Batista, V.L.; Vicente, M.D.M.; Lázaro, Á.L.S.; Moreno-García, M.N. Exploiting the user social context to address neighborhood bias in collaborative filtering music recommender systems. *Information* **2020**, *11*, 439.
112. Wijonarko, A.; Nurjanah, D.; Kusumo, D.S. Hybrid recommender system using random walk with restart for social tagging system. In *Proceedings of the 2017 International Conference on Data and Software Engineering, ICoDSE 2017*; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 1–6.
113. Nanopoulos, A.; Rafailidis, D.; Symeonidis, P.; Manolopoulos, Y. MusicBox: Personalized Music Recommendation Based on Cubic Analysis of Social Tags. *IEEE Trans. Audio Speech Lang. Process.* **2010**, *18*, 407–412.
114. Yang, T.; Cui, Y.; Jin, Y.; Song, M. A new algorithm for multi-mode recommendations in social tagging systems. IEEE 2nd International Conference on Cloud Computing and Intelligence Systems (CCIS 2012), Hangzhou, China, 30 October–1 November 2012; Volume 2, pp. 696–700.
115. Hariri, N.; Mobasher, B.; Burke, R. Context-aware music recommendation based on latent topic sequential patterns. In *Proceedings of the 6th ACM Conference on Recommender Systems*, Dublin, Ireland, 9–13 September 2012; pp. 131–138.
116. Mo, J.; Feng, Y.; Jia, A.; Huang, S.; Qin, Y.; Zhao, D. Community-based matrix factorization for scalable music recommendation on smartphones. In *Proceedings of the IEEE International Conference on Multimedia and Expo*; IEEE: Piscataway, NJ, USA, 2014.
117. Rosa, R.L.; Rodríguez, D.Z.; Bressan, G. Music recommendation system based on user's sentiments extracted from social networks. In *Proceedings of the 2015 IEEE International Conference on Consumer Electronics, ICCE 2015*; IEEE: Piscataway, NJ, USA, 2015; pp. 383–384.
118. Mao, M.; Lu, J.; Zhang, G.; Zhang, J. Multirelational Social Recommendations via Multigraph Ranking. *IEEE Trans. Cybern.* **2017**, *47*, 4049–4061.
119. Fathan, G.; Adji, T.B.; Ferdiana, R. Impact of matrix factorization and regularization hyperparameter on a recommender system for movies. In *Proceedings of the International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*; IEEE: Piscataway, NJ, USA, 2018; Volume 2018, pp. 113–116.
120. Chen, J.; Ying, P.; Zou, M. Improving music recommendation by incorporating social influence. *Multimed. Tools Appl.* **2019**, *78*, 2667–2687.
121. Mao, K.; Shou, L.; Fan, J.; Chen, G.; Kankanhalli, M.S. Competence-based song recommendation: Matching songs to one's singing skill. *IEEE Trans. Multimed.* **2015**, *17*, 396–408.
122. Ilarri, S.; Hermoso, R.; Trillo-Lado, R.; Del Carmen Rodríguez-Hernández, M. A Review of the Role of Sensors in Mobile Context-Aware Recommendation Systems. *Int. J. Distrib. Sens. Networks* **2015**, doi:10.1155/2015/489264.
123. Del Carmen Rodríguez-Hernández, M.; Ilarri, S. AI-based mobile context-aware recommender systems from an information management perspective: Progress and directions. *Knowl. Based Syst.* **2021**, *215*, 106740.
124. Braunhofer, M.; Kaminskis, M.; Ricci, F. Location-aware music recommendation. *Int. J. Multimed. Inf. Retr.* **2013**, *2*, 31–44.
125. Hong, J.; Hwang, W.S.; Kim, J.H.; Kim, S.W. Context-aware music recommendation in mobile smart devices. In *Proceedings of the Symposium on Applied Computing*, Gyeongju, Korea, 24–28 March 2014; pp. 1463–1468.
126. Chang, H.Y.; Huang, S.C.; Wu, J.H. A personalized music recommendation system based on electroencephalography feedback. *Multimed. Tools Appl.* **2017**, *76*, 19523–19542.
127. Verbert, K.; Duval, E.; Lindstaedt, S.N.; Gillet, D. Context-aware recommender systems. *J. Univ. Comput. Sci.* **2010**, *16*, 2175–2178.

128. Schedl, M.; Schnitzer, D. Hybrid retrieval approaches to geospatial music recommendation. In Proceedings of the 36th International ACM SIGIR Conference on Research and Development in Information Retrieval, Dublin, Ireland, 28 July–1 August 2013; pp. 793–796.
129. Fernández-Tobías, I.; Campos, P.G.; Cantador, I.; Díez, F. A contextual modeling approach for model-based recommender systems. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; Volume 8109, pp. 42–51. ISBN 9783642406423.
130. Teng, Y.C.; Kuo, Y.S.; Yang, Y.H. A large in-situ dataset for context-aware music recommendation on smartphones. In Proceedings of the IEEE International Conference on Multimedia and Expo Workshops (ICMEW), San Jose, CA, USA, 15–19 July 2013.
131. Schedl, M.; Vall, A.; Farrahi, K. User geospatial context for music recommendation in microblogs. In Proceedings of the 37th International ACM SIGIR Conference on Research and Development in Information Retrieval, Gold Coast, Queensland, Australia, 6–11 July 2014; pp. 987–990.
132. Schedl, M.; Schnitzer, D. Location-aware music artist recommendation. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2014; Volume 8326, pp. 205–213. ISBN 9783319041162.
133. Deng, S.; Wang, D.; Li, X.; Xu, G. Exploring user emotion in microblogs for music recommendation. *Expert Syst. Appl.* **2015**, *42*, 9284–9293.
134. Yang, J.; Chae, W.J.; Kim, S.Y.; Choi, H. Emotion-aware music recommendation. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2016; Volume 9747, pp. 110–121. ISBN 9783319403540.
135. Wang, X.; Rosenblum, D.; Wang, Y. A daily, activity-aware, mobile music recommender system. In Proceedings of the ACM Multimedia Conference, Nara, Japan, 29 October–2 November 2012; pp. 1313–1314.
136. Hao, M.; Zhou, D.; Liu, C.; Lyu, M.R.; King, I. Recommender systems with social regularization. In Proceedings of the 4th ACM International Conference on Web Search and Data Mining, Hong Kong, China, 9–12 February 2011; pp. 287–296.
137. Pichl, M.; Zangerle, E. User models for multi-context-aware music recommendation. *Multimed. Tools Appl.* **2020**, doi:10.1007/s11042-020-09890-7.
138. Wang, D.; Deng, S.; Xu, G. Sequence-based context-aware music recommendation. *Inf. Retr. J.* **2018**, *21*, 230–252.
139. Kulkarni, S.; Rodd, S.F. Context Aware Recommendation Systems: A review of the state of the art techniques. *Comput. Sci. Rev.* **2020**, *37*, 100255.
140. Chen, M.; Liu, Pan. Performance Evaluation of Recommender Systems. *IJPE* **2017**, *13*, 1246–1256.
141. Ge, M.; Delgado-Battenfeld, C.; Jannach, D. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. In Proceedings of the 4th ACM Conference on Recommender Systems, Barcelona, Spain, 26–30 September 2010; pp. 257–260.
142. Quadrana, M.; Cremonesi, P.; Jannach, D. Sequence-aware recommender systems. *ACM Comput. Surv.* **2018**, *51*, 66.
143. Monti, D.; Palumbo, E.; Rizzo, G.; Morisio, M. Sequeval: An Offline Evaluation Framework for Sequence-Based Recommender Systems. *Inf.* **2019**, *10*, 174, doi:10.3390/info10050174.
144. 2nd International Workshop on Information Heterogeneity and Fusion in Recommender Systems (HetRec 2011). Available online: <http://ir.ii.uam.es/hetrec2011/> (accessed on 31 May 2021).
145. HetRec 2011 | GroupLens. Available online: <https://grouplens.org/datasets/hetrec-2011/> (accessed on 31 May 2021).
146. LastFM1K. Available online: <http://ocelma.net/MusicRecommendationDataset/lastfm-1K.html> (accessed on 30 May 2021).
147. LastFM360K. Available online: <http://ocelma.net/MusicRecommendationDataset/lastfm-360K.html> (accessed on 30 May 2021).
148. LFM-1b Dataset. Available online: <http://www.cp.jku.at/datasets/LFM-1b/> (accessed on 31 May 2021).
149. Schedl, M.; Ferwerda, B. Large-Scale Analysis of Group-Specific Music Genre Taste from Collaborative Tags. In *Proceedings of the 2017 IEEE International Symposium on Multimedia, ISM 2017*; IEEE: Piscataway, NJ, USA, 2017; Volume 2017; pp. 479–482.
150. Bertin-Mahieux, T.; Ellis, D.P.W.; Whitman, B.; Lamere, P. The million song dataset. In *ISMIR 2011 Proceedings of the 12th International Society for Music Information Retrieval Conference*; University of Miami, Miami, FL, USA, 2011; pp. 591–596.
151. Brian McFee: Art of the Mix. 2011. Available online: <https://brianmcfree.net/data/aotm2011.html> (accessed on 31 May 2021).
152. Poddar, A.; Zangerle, E.; Yang, Y.-H. #Nowplaying-RS: A New Benchmark Dataset for Building Context-Aware Music Recommender Systems. In Proceedings of the 15th Sound and Music Computing Conference (SMC2018), Limassol, Cyprus, 4–7 July 2018.
153. #Nowplaying | Zenodo. Available online: <https://zenodo.org/record/2594483#.YLQTZqgzaUk> (accessed on 31 May 2021).
154. Pichl, M.; Zangerle, E. Spotify Playlists Dataset Available online: <https://zenodo.org/record/2594557> (accessed on 31 May 2021).
155. Hauger, D.; Kepler, J.; Schedl, M.; Košir, A.; Tkaličič, M. The Million Musical Tweets Dataset: What Can We Learn from Microblogs, In Proceedings of the annual Conference of the International Society for Music Information Retrieval (ISMIR), Curitiba, PR, Brazil, 4–8 November 2013.
156. MMTD—Million Musical Tweets Dataset. Available online: <http://www.cp.jku.at/datasets/MMTD/> (accessed on 31 May 2021).
157. Dror, G.; Koenigstein, N.; Koren, Y.; Weimer, M.; Dror, G.; Koren, Y.; Weimer, M. The Yahoo! Music Dataset and KDD-Cup'11. *Work. Conf. Proc.* **2012**, *18*, 3–18.
158. Webscope | Yahoo Labs. Available online: <https://webscope.sandbox.yahoo.com/catalog.php?datatype=c> (accessed on 31 May 2021).
159. Liu, J.Y.; Liu, S.Y.; Yang, Y.H. LJ2M dataset: Toward better understanding of music listening behavior and user mood. In *Proceedings of the IEEE International Conference on Multimedia and Expo*; IEEE Computer Society: Washington, DC, USA, 2014.
160. Datasets—Erheng Zhong. Available online: <https://sites.google.com/site/erhengzhong/datasets> (accessed on 31 May 2021).
161. Last.fm | Escucha Música, Encuentra Temás y Descubre Artistas. Available online: <https://www.last.fm/es/> (accessed on 31 May 2021).
162. Schedl, M.; Zamani, H.; Chen, C.W.; Deldjoo, Y.; Elahi, M. Current challenges and visions in music recommender systems research. *Int. J. Multimed. Inf. Retr.* **2018**, *7*, 95–116.

-
163. Ferraro, A. Music cold-start and long-tail recommendation: Bias in deep representations. In *Proceedings of the RecSys 2019 13th ACM Conference on Recommender Systems*; Association for Computing Machinery, Inc: New York, NY, USA, 2019; pp. 586–590.
 164. Idrissi, N.; Zellou, A. A systematic literature review of sparsity issues in recommender systems. *Soc. Netw. Anal. Min.* **2020**, *10*, 15.
 165. Celma, Ö. *Music Recommendation and Discovery: The Long Tail, Long Fail, and Long Play in the Digital Music Space*; Springer: Berlin, Germany, 2010; ISBN 9783642132865.
 166. Kowald, D.; Schedl, M.; Lex, E. The unfairness of popularity bias in music recommendation: A reproducibility study. In *Lecture Notes in Computer Science*; Springer International Publishing: Cham, Switzerland, 2020; Volume 12036, pp. 35–42. ISBN 9783030454418.
 167. Mansoury, M.; Burke, R.; Mobasher, B. Flatter Is Better: Percentile Transformations for Recommender Systems. *ACM Trans. Intell. Syst. Technol.* **2021**, *12*, doi:10.1145/3437910.
 168. Sánchez-Moreno, D.; López Batista, V.F.; Muñoz Vicente, M.D.; Gil González, A.B.; Moreno-García, M.N. A Session-Based Song Recommendation Approach Involving User Characterization along the Play Power-Law Distribution. *Complexity* **2020**, *2020*, doi:10.1155/2020/7309453.
 169. MovieLens|GroupLens. Available online: <https://grouplens.org/datasets/movielens/> (accessed on 31 May 2021).
 170. Zhao, G.; Fu, H.; Song, R.; Sakai, T.; Chen, Z.; Xie, X.; Qian, X. Personalized Reason Generation for Explainable Song Recommendation. *ACM Trans. Intell. Syst. Technol.* **2019**, *10*, doi:10.1145/3337967.
 171. Zhao, G.; Fu, H.; Song, R.; Sakai, T.; Xie, X.; Qian, X. Why you should listen to this song: Reason generation for explainable recommendation. In *Proceedings of the IEEE International Conference on Data Mining Workshops, ICDMW*; IEEE: Piscataway, NJ, USA, 2019; Volume 2018; pp. 1316–1322.
 172. Google Activity Recognition API|Google Developers. Available online: <https://developers.google.com/location-context/activity-recognition/> (accessed on 21 June 2021).
 173. Apple CMMotionActivity|Apple Developer Documentation. Available online: <https://developer.apple.com/documentation/coremotion/cmmotionactivity> (accessed on 21 June 2021).