Interacting with Context Factors in Music Recommendation and Discovery

Article *in* International Journal of Human-Computer Interaction · August 2016

DOI: 10.1080/10447318.2016.1225881

Pirkka Åman
Aalto University

International Journal of Human–Computer Interaction

ISSN: 1044-7318 (Print) 1532-7590 (Online) Journal homepage: http://www.tandfonline.com/loi/hihc20

Interacting with Context Factors in Music Recommendation and Discovery

Pirkka Åman & Lassi A. Liikkanen

To cite this article: Pirkka Åman & Lassi A. Liikkanen (2016): Interacting with Context Factors in Music Recommendation and Discovery, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2016.1225881

To link to this article: http://dx.doi.org/10.1080/10447318.2016.1225881

Accepted author version posted online: 26 Aug 2016.
Published online: 26 Aug 2016.

INTERNATIONAL JOURNAL OF HUMAN–COMPUTER INTERACTION
http://dx.doi.org/10.1080/10447318.2016.1225881

Taylor & Francis
Taylor & Francis Group

Interacting with Context Factors in Music Recommendation and Discovery

Pirkka Åman[a] and Lassi A. Liikkanen[b]

[a]Media Lab Helsinki, Department of Media, School of Art, Design and Architecture, Aalto University, Helsinki, Finland; [b]Helsinki Institute for Information Technology HIIT, Aalto University, Helsinki, Finland

ABSTRACT

The rapid development in mobile computing has brought context sensing and information available for music recommendation as well. We reviewed 19 experimental contextual music systems and found that while the context factors of location, time, activity, and identity are adopted in a wide variety of ways, the systems mainly rely on common UI solutions for interacting with these factors. Specifically, context factors could be employed to offer explanations, transparency, and visualizations of music recommendations in more explorative ways, providing novel user experiences. Based on our review, we provide implications for design and research on future media discovery systems, which we believe can realize the great potential of context-aware content services.

1. Introduction

Access to recorded music is today greater than ever before. However, it is often hard to find new interesting music from the selections of millions of tracks. Numerous recommendation systems for music discovery have been developed to help the selection process. The most common recommendation techniques are collaborative filtering (CF) and content-based analysis.

In CF, user preferences for items are predicted by learning how a user has listened to music earlier. The recommender then identifies other users who have the same kind of preferences with the active user, suggesting items favored by these like-minded users and unseen by the active user (Adomavicius, 2011; Celma, 2010; Park, 2015).

As CF-based systems have several issues, including popularity bias, i.e. the recommendations tend to concentrate on the popular items instead of offering new music, other techniques have been suggested (Fleder & Hosanagar, 2007). Content-based recommendation is based on storing information describing items and offering recommendations that resemble the items the user likes. It is carried out either automatically or manually. In the automatic approach, some of the features (e.g. tempo, harmonics, or signal intensity) of music tracks are analyzed by software (Downie, 2008; Ricci, Rokach, & Shapira, 2011), while in the manual approach, experts or users classify and tag the music tracks according to, for example, genre, mood, or instrumentation (Kaminskas & Ricci, 2012). Many commonly used recommendation systems are hybrids, adopting both content analysis and CF. Popular commercial online services such as iTunes, Spotify, and YouTube typically apply CF, but content analysis and hybrid approaches are also adopted in their music recommendation features (Magno & Sable, 2008). Online personalizable radio Pandora is unique in its content analysis approach of tagging every music track manually by musicians and musicologists (ibid.).

As music choices are heavily affected by the user's mood and other situation-dependent attributes (North, Hargreaves, & Hargreaves, 2004), context factors have recently made their way to music recommenders (Adomavicius & Tuzhilin, 2011; Ricci et al., 2011). While traditional music recommendation approaches suggest users either music (tracks or artists) or other users, contextual recommendation adds situational factors to the equation. For example, adding location as a context factor adds a third dimension, matching music to location.

Several context-aware music recommendation systems have been put forward in research and some introduced into the consumer market, with no commercial breakthroughs yet. Of the popular commercial systems, Spotify has recently started to offer curated, expert-created playlists for various moods and activities (Page & Ning, 2014).

In this review, we present the spectrum of context factors and interaction design features affecting contextual music discovery user experience (UX). Since the promises of contextual music recommenders often center around their potential for serendipitous, i.e. positively surprising, discoveries of new music, we will pay attention to in which ways the systems promote serendipity. We will investigate how the systems under review apply:

(a) context factors
(b) interactive features.

We reviewed 19 research systems that apply context factors for music recommendation, in most of the cases (14) including a user study. In our analysis, we also provide insights related to each context factor from the user evaluation of the systems.

CONTACT Pirkka Åman pirkka.aman@aalto.fi Media Lab Helsinki, Department of Media, School of Art, Design and Architecture, Aalto University, PO Box 31000, FI-00076 AALTO, Helsinki, Finland.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/UTAF.

Our article aims to aid researchers and designers in developing not only better music systems but also other context-aware applications. As there has not yet been a review of contextual music interfaces, the novel contribution of this article is in providing a summary of the ways context factors and the related interaction design features have been adopted for music recommendation so that innovative UX solutions can emerge from what has been suggested earlier.

The article is structured as follows: in Chapter 2, we present the conceptual background. Chapter 3 presents an overview of the groups of the reviewed systems according to their designed purpose. In Chapter 4, we present primary context factors of location, time, activity, and identity, presenting their idiosyncrasies, together with typical and nonstandard examples of the ways context factors and interaction dimensions are adopted in the systems. Observations from the user studies are presented in Chapter 5. In Chapter 6, we discuss the main findings of contextual music interactions thematically and provide concluding remarks.

2. Background: Conceptual Framework

Creating an unambiguous and universal categorization of context types is a research issue in itself (Dourish, 2004; Razzaque, Dobson, & Nixon, 2005). After reviewing several views on context, we opted for Dey & Abowd's (1999) view of seeing *identity, activity, time*, and *location* as the primary context information types, of which the other contextual information types can be derived from. Time, location, and identity answer the fundamental situation-dependent questions of *when, where*, and *who*. Activity answers the question of *what* is occurring in the situation (Dey & Abowd, 1999). We discuss the concept of context in more detail in section 2.1.

Three of the reviewed systems involve *mood* as a context factor, relying on obtaining the mood value explicitly from the user (Mobile Music Genius; Lehtiniemi & Holm, 2012; InCarMusic). While mood could be assigned to any of the four primary categories or some combination of them, we leave it mainly out of our analysis, since only Lehtiniemi and Holm (2012) present a user interface (UI) solution that does not follow common user input values for mood selection (e.g. the binaries of "calm–active" and "happy–sad" in InCarMusic). We will discuss Lehtiniemi & Holm's "mood pictures" UI within the context factor activity.

Throughout the article, we use the terms *music discovery* and *music recommendation*. The former refers to the user's experience, the latter to the techniques adopted for discovery. While contextual approaches have been reviewed earlier, they have focused on recommendation techniques or usability evaluation. Jones and Pu (2007) conducted user evaluation of two popular non-contextual commercial music recommendation systems and their UIs. Kaminskas and Ricci (2012) reviewed recommendation techniques adopted in the fields of music information retrieval (MIR) and recommendation systems research, but did not review how context factors have been made available at the UI level.

Recommendations are often evaluated by their quality (e.g. relevance, novelty, serendipity), which is best evaluated by users (Celma, 2010; Swearingen & Sinha, 2001). Instead of the perceived quality of recommendations, we analyze:

(a) the use of *context factors (discussed in* Section 2.1),
(b) interaction features forming a construct of *interaction adequacy*, adapted from (Jones & Pu, 2007),
(c) *interaction metaphors* and *paradigms*, which can be subjected to expert analysis.

Interaction adequacy covers user effort, presentation, transparency, and explanations as well as scrutability and refinement. Transparency and explanations refer to communicating the user why a recommendation was made and revealing the system logic. Transparency can be seen as a design challenge while providing explanations is a solution to it; however, we use these terms together, following the previous literature on the subject. Refinement features allow users to give feedback to the system. We present both typical and nonstandard examples of how (a), (b), and (c) are applied in the systems. (Åman & Liikkanen, 2010; Jones & Pu, 2007; Konstan & Riedl, 2012; Pu & Chen, 2010; Tintarev & Masthoff, 2007; Tintarev & Masthoff, 2011).

Metaphors and interaction paradigms are manifested on the UI level, specifically as part of the presentation category (Figure 1). Interaction metaphors refer to the ways recommendations are presented to the user. In general, metaphors connect a set of concepts to another set of concepts, transferring meanings from one area of reality to another (Lakoff & Johnson, 1980). In the context of UI design, interaction metaphors are usually visual guides and frameworks familiar from the real world, helping users perform tasks and make sense of the computing environment at the UI level (Carroll, Mack, & Kellogg, 1991). For example, web UIs employ several metaphors from the real-world organization of books and libraries (bookmark, page, tab) for easier interaction with digital objects.

Interaction paradigms describe the ways in which the user actually interacts with music content. Despite the numerous digital music services, interaction with recorded music is still dominated by two interaction paradigms: *curated* and *on-demand* paradigms that hark back to the analog era of physical radio devices and vinyl discs (Liikkanen & Åman, 2015; Liikkanen, Amos, Cunningham, Downie, & McDonald, 2012). The logic of this division is based on dividing interactions between different abstraction levels (Te'eni & Sani-Kuperberg, 2005). On the *operational*, or executive, level, interaction is commonly defined by the

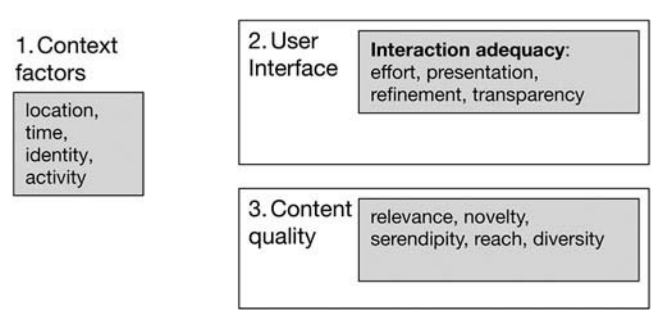

Figure 1. Model of three levels of UX for contextual music recommendation and discovery. In our review, we address context factors and dimensions of interaction adequacy. Legend: white boxes: levels of UX; gray boxes: evaluable properties.

constraints of a technology at hand (Liikkanen & Åman, 2015). For example, modern smartphones are operated by tapping interactive "buttons" on a touch screen, while the phones of early 2000s had still physical buttons for the same functions. On the *strategic* level, interactions function today as they have been functioning from the introduction of the two dominant technologies of mechanical reproduction of recorded music, radio and physical disc media. The currently dominant curated and on-demand strategic interaction paradigms match the operational level content selection modes of *tuning in* and *playlisting* (Liikkanen & Åman, 2015). The former demands only a selection of a station or channel for a continuous stream of music, while the latter is based on the selection of content and the order of play. In our review, we address typical examples and point out uncommon cases of both interaction paradigms and metaphors.

Based on our previous work on the subject (Åman & Liikkanen, 2010, 2013; Åman, Liikkanen, & Hinkka, 2015; Forsblom, Nurmi, Åman, & Liikkanen, 2012), we built an analysis framework, a descriptive model presenting the three levels of UX (context, content, and interactions) that are encountered and interacted upon by users (Figure 1).

The analysis was conducted by examining each system as they were presented in the respective articles by observing the following properties.

Context Factors

If the system applies one or more context factor out of the four primary factors, how is it applied? Drawing from the analyzed systems' features, we present typical and original uses of location, time, activity, and identity. As the context factors fundamentally differentiate from each other, the ways the systems apply them are addressed in the separate sections (Sections 4.1–4.4).

Interaction Adequacy

(1) What kind of *user effort* is required to receive a recommendation?
(2) How is the *presentation of* recommendations implemented? Does it adopt a playlist, stream, or nonstandard presentation? What kinds of *metaphors* and *interaction paradigms* does the system adopt, if any, beyond a playlist or stream?
(3) *Transparency*: How does the system communicate why certain recommendations were offered?
(4) *Refinement*: What kinds of interactions does the system offer when informing the user that the recommendation was bad or providing some other form of feedback?

User Feedback

What did the users think about the system, in terms of the use of context factors and the interactions provided? The results are presented in Chapter 5.

2.1. Context in Music Discovery and Recommendation

Context has been one of the central concepts in HCI since the mid-1990s. In search and discovery tasks, context is "a necessary source of meaning" (Dervin, 2003). Context can be thought of as a snapshot (Beale & Lonsdale, 2004) of all the factors that are relevant to the user and the system in a particular moment. We reviewed five views on context that emphasize and include different aspects in the concept.

For Dourish (2004), context is a relational property between objects and activities, generated through *activity*, and thus something that cannot be categorized outside a certain situation. Dourish does not present a framework for practical design work. In Krippendorff's (2006) product semantics, *meaning* is in the center of the concept of context. For Krippendorff, context is a frame where users create meanings for design artifacts. Like Dourish, he does not offer a categorization of context types.

In a review of techniques employed in context-aware music recommendation, Kaminskas and Ricci (2012) present three context categories: user-related contextual information, environment-related contextual information, and multimedia context. However, their definition is based on MIR research and its approach on context, in which all information surrounding the music track is context, makes the scope of the concept unusefully broad (covering almost all music services) for our analysis. Specifically, we do not consider the category of multimedia context such as metadata (e.g. lyrics and track title) contextual information. For us, these data are part of the content.

Razzaque et al. (2005) reviewed previous definitions of context and provided a comprehensive context categorization. However, their categorization (user context, physical context, network context, activity context, device context, service context) includes many dimensions that are formed from the system point of view (e.g. network connectivity, device's battery lifetime). As the focus of our review is on the context factors that are relevant for recommending music to the users, we tried to find a simpler, but still comprehensive categorization. After reviewing earlier views on context in HCI, Dey and Abowd (1999, 3–4) define context as follows:

> *Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application in a situation, including the user and application themselves.*

This operational definition is directed to designers and developers working in the field of context-aware computing. It helps determine the factors that belong to the application's context of use. If a piece of information is relevant to interaction in that particular use situation, it is part of the context. For example, in a mobile travel app for exploring the sights of a foreign city, location is relevant contextual information while a weather forecast of the user's hometown is not.

We see that Razzaque et al.'s view is useful for technical system design and implementation, while Krippendorff's and Dourish's views are helpful in explaining system use: the activity with the system and the meanings people construct while using it. On the other hand, Dey & Abowd's view is clear, definite, and unequivocal, covering all aspects of context in a compact way. It gives designers and developers a pragmatic tool to analyze existing services and search design alternatives and potential future solutions. Moreover, it is formed by reviewing nine previous definitions of context in the HCI literature and has numerous citations (ACM Digital library citation count, as in May 2016: 259).

There is a hierarchy of importance within context types. According to Dey and Abowd (1999), certain types of context are more important than others. The *primary* context types are location, identity, activity, and time. All the other context types (i.e. secondary) can be derived from these. For example, a user's email address is derived from the primary category of identity. In a similar manner, in a mobile journey planner, distances to various bus stations are derivatives of location. Sometimes, multiple primary categories are needed to map a secondary context factor. For instance, meaningful forecasted weather conditions typically require location and time (Dey & Abowd, 1999). Primary context factors of time, location, and activity are mostly situation-dependent in nature, i.e. transitory and subject to change; however, identity has also more or less stationary dimensions (e.g. birthplace, gender, nationality) that characterize the user regardless of the situation, which could be involved in music recommendation as well.

For the realization of contextual services, information about context is acquired by deducing or retrieving it (1) explicitly, with the user providing the data, (2) implicitly, by means of sensing technologies, or (3) by inference from data (Adomavicius & Tuzhilin, 2011; Schmidt, Beigl, & Gellersen, 1999; Tu et al., 2013). We will apply the concepts "implicit" and "explicit" to distinguish between automated (both implicit and inferred) and user-given (explicit) context data acquisition modes. An example of explicitly entered context data is a user choosing a city (e.g. "London") for personalized weather updates instead of the implicit localization of the user via location services.

Implicit context data can be more reliable and more granular than user-provided explicit data although they lack the meaningfulness of the information given by the user and must be computationally interpreted (Ricci et al., 2012). However, the required user effort can be minimized when automatically acquiring contextual information. In many cases, contextual data acquisition is accomplished both implicitly and explicitly in the same system. While our main focus is on the user interactions, we will also describe the ways in which various context data are acquired for the contextual music recommendations.

3. Overview of the Reviewed Systems

We analyzed 19 music recommendation systems that apply contextual information for personalized music discovery in order to find out the user interactions they provide for contextual music discovery, recommendation, and music listening. The reviewed systems were published between 2004 and 2015 in HCI and related forums (Figure 2). The articles were selected based on the following criteria: they presented a *research prototype* of a *contextual music application* presenting *a UI*. To ensure that all the relevant research was included, we ran multiple searches on ACM Digital library and IEEE Xplore services.

We start with an overview of the systems. Based on their design intent, we grouped the systems into three main clusters: *location-dominant, task-specific*, and *all-inclusive systems*. The clusters inform us about researchers' and industry's understanding of the most lucrative areas for contextual music recommendation from the last decade. Only one system fell out of the clusters. It was a prototype using "mood pictures" to produce playlists matching the depicted mood or activity (Lehtiniemi & Holm, 2012).

3.1. Location-Dominant Systems: Geo-Tagged and Collocated

These systems foremost support encounters of music, places, and people in the city Tables 1 and 2. *Geo-tagged* systems offer music attached to places in urban environments and provide music in conjunction with the exploration of urban spaces. They are targeted for visitors or citizens, providing users with a new information layer of the city, containing music and sometimes other content (e.g. user comments) as well. The systems that use nearby people or devices for music discovery in *real time* are called *collocated* (Jacucci et al., 2009; Lundgren, Fischer, Reeves, & Torgersson, 2015; Seeburger, Foth, & Tjondronegoro, 2012).

3.2. Task-specific systems

These systems use context factors for recommending music for certain activities. In our review, there were two systems targeted to sports and two for driving a car.

3.3. Automated, All-Inclusive Systems

These systems employ multiple context factors, relying heavily on implicit context acquisition to provide an unobtrusive UX in everyday activities. "Automated" refers to implicit context acquisition that reduces the required user effort greatly.

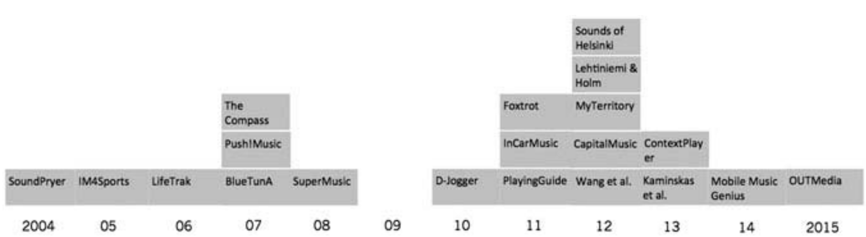

Figure 2. Timeline of the reviewed research prototypes.

Table 1. Collocated systems.

System name	Main features
BluetunA	Pre-smartphone era mobile phone app for music information sharing with nearby people.
Capital Music	Tablet app for showing album art of music currently played by nearby people.
Push!Music	Early tablet app for finding nearby users for sending and receiving music files.
The Compass	Pre-smartphone era mobile phone app showing direction and distance of other users, whose music can be downloaded.

Table 2. Geo-tagged systems.

System name	Main features
Foxtrot (Ankolekar and Sandholm, 2011)	Mobile and desktop app that automatically creates radio-like stream of geo-tagged music, ambient sounds, and comments.
MyTerritory	Mobile phone map-based app. By circumnavigating areas, they can be "conquered" by attaching music to them.
OUTMedia	Mobile phone augmented reality browser app for user-created points-of-interest (POIs) with music.
PlayingGuide	Mobile phone map-based app recommending music for POIs during touristic city itineraries.
Sounds of Helsinki	Mobile phone app with push notification recommendations of music events for urban settings.

Table 3. Task-specific systems.

System name	Main features
D-Jogger	Custom hardware that applies user pace to choose and adapt the music in real time.
IM4Sports	Custom hardware that matches music with the current pace of a walker or jogger.
InCarMusic	Mobile phone app for streaming music channels based on user-input and implicit data for car travel.
Sound Pryer	Custom hardware applying early tablet computer as a shared car stereo for musical traffic encounters with other drivers.

Table 4. Automated systems.

System name	Main features
ContextPlayer	Mobile phone app recommending tracks based on context information.
Lifetrak	Mobile app for early tablet computer for automatic music stream, also user-adjustable through "Context Equalizer".
Mobile Music Genius (Schedl et al., 2014)	Tablet app for automatic real-time contextual playlists.
SuperMusic	Mobile music player with context-aware streaming music feature.
Wang et al., 2012	Mobile app offering automatic playlists for six activities, with a manual mode option.

4. Contextual Music Interactions

In the following sections, we review and analyze the ways in which the primary context factors of location, time, activity, and identity have been adopted in the contextual music UIs.

4.1. Location

Today, millions of people use their mobile devices as the main channel for accessing media content and content recommendation services as well (Ricci et al., 2012; Yang, Lu, Gupta, & Cao, 2012). It has been easy to measure location quite accurately for over a decade. In this light, it is no wonder that most of the systems utilize location (16 of 19).

While location-*aware* is a generic term for systems that utilize location information, location-*sensitivity* is a subcategory, meaning that content can be discovered or accessed *only* when close or on location (Åman et al., 2015). Examples of location-sensitive systems include OUTMedia (Åman et al., 2015) and PlayingGuide (Braunhofer, Kaminskas, & Ricci, 2011), which allow users to listen to music only when nearby or on the location.

Another distinction used in association with location is the conceptual pair of *collocated* and *co-present* users. Collocated refers to technical connectedness, e.g. by means of Wi-Fi or Bluetooth (Jacucci et al., 2009; Lundgren et al., 2015; Seeburger et al., 2012). Co-presence refers to a shared space and relationships with a social position (e.g. English pub institution), relating closely to the concept of neighborhood (de Certeau, 1984; de Certeau, Giard, & Mayol, 1998; Silverstone, 2003). A public space can be conquered at least temporarily for private use. This dimension is utilized for example in MyTerritory app, letting users "conquer" areas with music (Lehtiniemi & Ojala, 2012). By allowing users to share song selection information between collocated strangers in public urban places, Capital Music aims to result in experiences that are mixed in terms of private and shared reality. It is deliberately designed to employ the nuances along the axis of public and private space. Through music discovery, it aims to facilitate communication between strangers (Seeburger et al., 2012).

At the UI level, maps are a common way to satisfy an aspect of interaction adequacy, *transparency*, i.e. telling users how the system works or why a certain recommendation was made. While most recommenders are black boxes, not revealing their system logic, transparency and explanations have proved to be crucial trust-evoking features improving recommender UX (Åman & Liikkanen, 2010; Sinha & Swearingen, 2002; Tintarev & Masthoff, 2007; Tintarev & Masthoff, 2011).

Some collocated systems (Capital Music, BluetunA) allow users to discover nearby strangers' music, providing a potential source for serendipitous discoveries. The Compass, one of the collocated systems, provides a rare example of nonstandard UI (Figure 3). It employs a simple navigation metaphor for discovering music from nearby users' devices (Tanaka, Valadon, & Berger, 2007). The UI has a compass needle that displays the direction and distance to

Figure 3. The Compass: Rare example of an unconventional UI metaphor.

Figure 4. OUTMedia mobile augmented reality music browser.

walk to reach the range of a Wi-Fi or Bluetooth network to exchange music files with another user. Contrary to geographic map interfaces, The Compass shows only the direction and allows user to stray, in the spirit of Situationist *dérive* or Baudelaire's *flaneur* (Baudelaire, 1863/1964; Debord, 1955/2006; Tanaka et al., 2007), potentially leading to serendipitous discoveries of not only music but places as well.

A nonstandard modification of the playlist paradigm for discovering music is exemplified by OUTMedia, a geo-tagged system that offers music discovery through fixed music–place combinations and mobile augmented reality UI. Music discovery happens by scanning the user's surroundings through the mobile phone camera with a layer of augmented reality objects hovering upon the camera view (Figure 4). It is one of the seven systems that do not rely on the basic playlist visualization of recommended music offered for discovery. The other six systems (Capital Music, ContextPlayer, D-Jogger, Lifetrak, SoundPryer, The Compass) adopt either playlist that is not visualized (following the *tuning in* paradigm) or present the playlist as an album art mosaic. Besides being an easy way to *refine* recommendations just by looking in another direction, OUTMedia UI enables the user to look toward the world, and not the screen, offering an alternative to the common mobile UX. This also provides transparency for interaction.

4.2. Time

Time affects music choices directly as music preferences commonly vary according to the time of day, weekday, and between work and leisure (Bull, 2008; North et al., 2004). For example, some people want to listen to energetic music right after waking up, and many prefer slow tempo music in the late evening (Levitin, 2007). As certain activities have a close relationship to time, time is closely linked to the context factor of activity, and many activities can be described in terms of time (e.g. evening run, Sunday dinner; Reddy & Mascia, 2006). The most typical granularities of time adopted in the systems are time of day (e.g. morning, evening) and weekday (e.g. in Lifetrak, Mobile Music Genius, ContextPlayer, InCarMusic).

In the reviewed sports systems (D-Jogger, IM4Sports) time is adopted together with other context factors such as identity, heartbeat, or pace to provide recommendations to match the desired intensity of the exercise, a common sports-related function of music (Karageorghis et al., 2009). D-Jogger employs the user's pace to dynamically choose and adapt the music in real time. In IM4Sports, user's speed and music have a two-way relationship: the desired intensity of exercise is achieved by music tempo and, in another mode, time stretching of music is adopted to match the user pace. *Spotify running* (www.spotify.com/us/running) is a commercial example of an app that employs phone accelerometer data for suggesting music that matches the user pace.

The most important dimensions of time for music recommendation include the categories of "most recently", "right now", and "near future". These functions resemble the uses of media studied in media and cultural studies (Fiske, 1990). Most recently happened events are interesting because they have news-like appeal. Events happening right now resemble live events, and part of their appeal is due to the ephemeral nature of the "live": once it has happened, it may never happen again exactly the same way. Events or media happening in the near future provide a potential agenda of what to do (events) or follow (media) next (for an analysis of time as a part of UX, see Liikkanen & Gómez, 2013). Out of these, most of the reviewed systems (e.g. Capital Music, Lifetrak, InCarMusic, OUTMedia, SoundPryer) adopt the approach of recommending music for the actual, real-time moment and situation the user is in. In several reviewed systems, music is recommended between collocated users, where not only location, but also temporal proximity is considered. Collocated systems such as Push!Music, BluetunA, and Capital Music all make use of the real-time aspect, thus providing potential for serendipitous encounters with music and other users. Sounds of Helsinki recommends festival events happening in the coming hours of the same day.

Although filtering features at the UI level are rare in the reviewed systems (for an exception, see Figure 8), more control over content may be achieved by introducing time-based filtering of recommendation histories through chronological layers of users, locations, and content genres. None of the reviewed systems offer this kind of filtering. OUTMedia (Åman et al., 2015) offers users added value by including the "most recently" UI feature for showing a playlist of the four latest tracks played near the user's location.

4.3. Identity: Social Features

Identity can be looked through two lenses: (a) social media and related features that typically connect friends or friends of friends; and (b) systems that employ collocated people connect strangers. The potential for serendipity varies along the continuum of the familiarity of users: stranger–familiar stranger–friend of a friend–friend. In the other end, there are purely random encounters with strangers' music, and in the other end, friends' music. All positions along the continuum may provide serendipitous discoveries, or relevant but more familiar music. Pure randomness often decreases relevance and increases chance for novelty. In music recommender system research, this is called the trade-off between novelty and relevance (Celma, 2010). While friend's music often matches the user's taste (relevant), stranger's music has greater potential for novelty and serendipity.

Phone number, email address, birthday, social media connections, and the user's relationships to nearby people are all part of identity (Dey & Abowd, 1999; Theimer & Schilit, 1994). At first, it may appear that all subcategories of identity (e.g. gender, email, age) are not "real" contextual factors since they are not transitory. However, identity includes dimensions that are at least subject to change, such as social media connections, not to mention a user's relationships to the surrounding people in a certain situation. In our analysis, we focus on social features since the majority of identity-related interactions in the systems center around them.

Previous research on music recommender systems shows that social features increase users' loyalty to a system and increase social cohesion (Jones & Pu, 2007; Zhou, Xu, Li, Josang, & Cox, 2012). For example, personalized Internet radio Last.fm received positive perception in a study by Jones and Pu (2007). By letting 64 participants use two systems (Last.fm and Pandora) for an hour a piece, followed by an online questionnaire about item quality, interaction adequacy, and other UX dimensions, the study showed that the wide variety of social features (blog entries, tagging, sharing, etc.) of Last.fm resulted in more satisfying UXs.

In the context of social media features, identity can be seen as a construct of status updates, likings, and the shared content over time (Silfverberg et al., 2010). The most important dimensions here are the *credibility* of a person (e.g. friend, friend of a friend, or a popular DJ) or, on the other hand, the *familiarity* of a person (stranger, familiar stranger) and his/her music preferences.

Many of the reviewed systems include social media functionalities. While the most recent systems have sharing features similar to Facebook or Twitter (see ContextPlayer carousel, Figure 5), the older ones such as BluetunA employ text messaging (Baumann et al., 2007). By using different icons, BluetunA shows the user if the other users detected nearby are strangers, friends, or familiar strangers; people whom the user does not know but who have already crossed her path. BluetunA works as a vehicle for communicating one's musical identity for collocated people within proximity of some meters (Bluetooth radius), finding people with similar music taste, and if desired, to start communication by text messaging.

InCarMusic (Baltrunas et al., 2011) implements the social dimension by requiring users to create profiles. For a specific car trip, accompanying passengers who have profiles are selected from the application preferences. The system then offers music tracks based on the personal preferences of all users combined with contextual conditions such as traffic, weather, and mood by user input. All the other systems featuring social dimensions (see Figure 9) that employ identity are based on either collocated random encounters or common social media features such as sharing and liking.

Capital Music designers believe that sharing song choices may result in further social interactions between strangers through a messaging feature (Seeburger et al., 2012). It is a collocated system, supporting anonymous sharing of music choices between collocated people and visualizes songs that are currently played in the vicinity. It aims to expand the common practice of "cocooning" with one's phone in the public space toward a location-sensitive socializing.

Taking a look at the ways social information is implemented at the UI level, ContextPlayer's "recommendation carousel" shows related information and media in diverse ways using social media services for nearby events, songs other users listen to in similar situations, videos, and the playback history (Figure 5). It works on a touch display and sideways swiping gestures (horizontal carousel) for a multifaceted UI for supporting music discovery. It is an example of *presenting details* (beyond title, artist, and album) about the recommended items as another form of transparency and explanations (Jones & Pu, 2007).

Anonymous and socially agnostic (not differentiating between friends and strangers) social interaction was also present. This feature was received positively for its potential for serendipity particularly by the users of a collocated system of Push!Music and Capital Music, although privacy issues were noted as well. For example, Capital Music users commented that music might be used as an icebreaker in communications between strangers in public spaces. However, some users were concerned about giving their location information to strangers, or receiving offending messages (Seeburger et al., 2012).

4.4. Activity

People commonly choose different music to accompany different activities (Bull, 2008; Levitin, 2007; North et al., 2004). Several systems that detect user activities and automatically choose music have been suggested to overcome the often laborious task of manually creating and switching between playlists (e.g. Cunningham, Caulder, & Grout, 2008; Han, Rho, Jun, & Hwang,

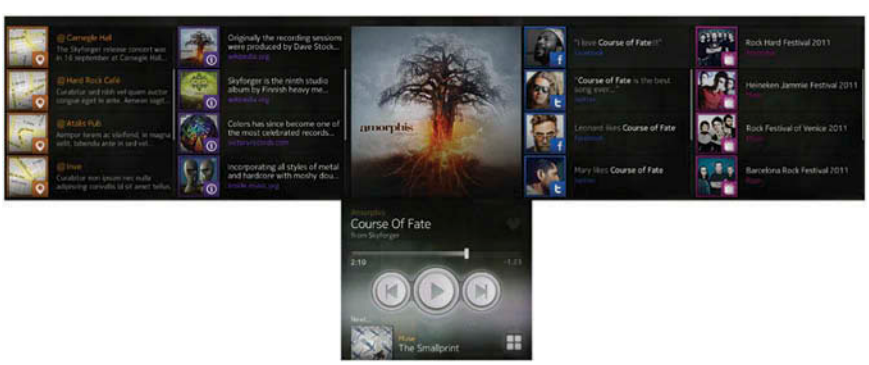

Figure 5. ContextPlayer's current playing song screen and "recommendation carousel" providing details along several dimensions, including various social media.

2010; Park, Yoo, & Cho, 2006; Su, Yeh, Yu, & Tseng, 2010; Wang, Rosenblum, & Wang, 2012; Wijnalda, Pauws, Vignoli, & Stuckenschmidt, 2005). The standard way for reasoning user activity in these systems is via mobile device sensor measurements. The older reviewed systems that employ accelerometers for activity measurement often rely on using external devices (e.g. DJogger by Biehl et al., 2006, IM4Sports), while the recently designed systems apply the smartphone or tablet built-in sensors (e.g. Wang et al., 2012). Sensor data typically includes orientation and location measured by GPS, accelerometer, and gyroscope data. Beyond movement- and location-based sensors, Wang et al. (2012) apply ambient audio and lighting conditions measured by a microphone and a light sensor for producing activity-based recommendations targeted specifically at everyday activities such as working, studying, and shopping (Figure 6).

Activity is the primary context factor for sports. D-Jogger adopts the phenomenon of entrainment, i.e. the synchronization of two rhythmical processes, in this case music and user's pace, for making exercise more motivational and fun. Data are gathered from the phone accelerometer. Music changes its tempo according to the user pace, and in another mode, recommendations are made to match the user's pace (Moens, Van Noorden, & Leman, 2010). Another sports system, IM4Sports, employs heart rate and pedometer signals for synchronizing music tempo for motivation and perseverance during exercising (Wijnalda et al., 2005).

Our main observation is that many systems provide users with no interactions for controlling activity. Beyond a binary selector for choosing between auto and manual modes, two examples from the reviewed systems show how to offer user control for activity: choosing from a list of activities (Figure 6) and adjusting the amount of effect activity has on recommendations (Figure 8).

Figure 6. UI with buttons for choosing from six activities (manual mode) or showing the measured activity (auto mode) for a playlist generation (Wang et al., 2012).

The system by Lehtiniemi and Holm (2012) does not rely on sensor data at all. Instead, it lets users choose a photographic image of seven different activities as playlist seed. The activities include situations such as a sports event, car driving, and evening dinner. As we can see from the provided activities and the related pictures (Figure 7), activity is closely related to mood.

Often the purpose of including contextual factors in the system design is to minimize user effort. For example, Lifetrak employs automated, context-driven play to minimize interaction needs. The authors talk of "a cleaner, simpler interface that still offers powerful control" (Reddy & Mascia, 2006). In automated systems (e.g. ContextPlayer, Lifetrak, Wang et al., 2012), the application gathers context data and gives recommendations typically without user effort, updating the playlist of music stream automatically as the context changes.

Lifetrak's Context Equalizer provides a rare example of a UI that allows for adjusting the effect that context factors have on music recommendations (Figure 8). The user-adjustable mix of context factors influence the selection of tracks to be played next. Besides being an unconventional UI widget, Context Equalizer is a good example of including both automated and user-requested recommendations in a balanced way. To override the implicitly detected context factors, the user can choose to adjust the values to suit his/her situation.

While recommenders often require frequent input from the user, previous research has shown that users prefer interfaces that require little initial effort for requesting recommendations (Jones & Pu, 2007; Sinha & Swearingen, 2002). Eyes- or hands-busy activities, such as car driving or sports, naturally call for UIs that require minimal visual or haptic user effort. For instance, in Sound Pryer, minimal UI is employed for the busy activity of car driving to offer music through traffic encounters. The guiding aspects of the design process included a device mounted on the dashboard, a modest visual interface, and mostly audio-based operation (Östergren, 2004).

Figure 7. List view of pictures used for user-chosen playlist seeds (Lehtiniemi & Holm, 2012).

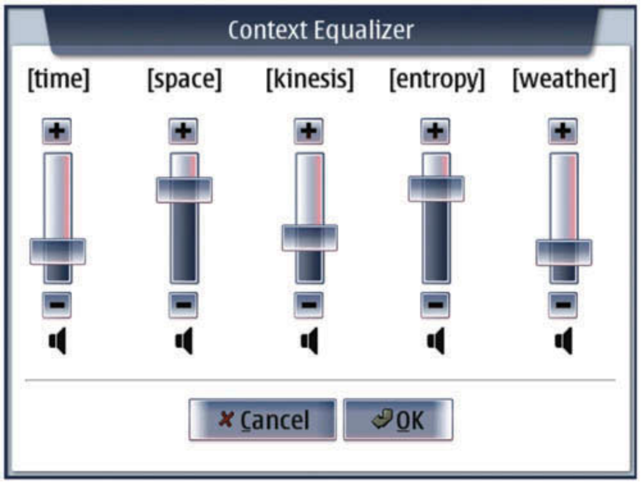

Figure 8. Context Equalizer allows overriding implicitly acquired context values.

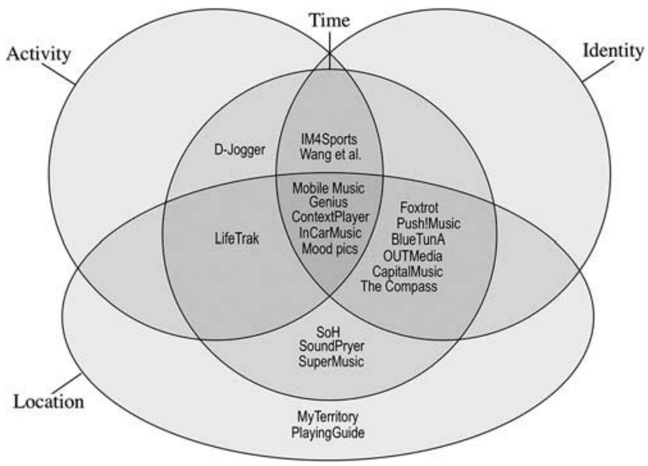

Figure 9. The primary context types as employed in the reviewed systems, showing that time and location are the most common factors, while location-time-identity is the most frequently applied combination.

4.5. Summary of the Adoption of Context Factors in the Systems

Location is not only the dominant context factor, but it is the only factor that appears by itself in the reviewed systems. Typically, systems use location together with at least one other factor (Figure 9). Systems not applying location include two sports systems (IM4Sports, D-Jogger) that adopt user pace or heartbeat (activity) as their central contextual factor; and a system relying on playlist creation through photos of various situations (Lehtiniemi & Holm, 2012). *Time* does not function as the sole contextual factor in any of the reviewed systems. Compared with location, which was the only context factor employed alone, time typically plays a supporting role. While time is often not presented at the UI level, in many systems it is paramount to adding precision and personal feel to recommendations. Almost all of the reviewed systems (17 of 19) involve the time factor in some way (Figure 9).

Most of the systems apply context factor *identity* (13/19), and location, time, and identity is the most common combination (15/19). Like time, identity is employed not alone, but with other primary context types for added value for music discovery. *Activity* is employed in nearly half of the systems (8 out of 19). Task-specific systems (sports, driving a car) obviously focus on activity as the main context factor for producing recommendations (Table 3). The rest of the systems involving activity apply it as one factor in a combination of context factors for automated contextual music experience (Table 4).

4.6. Comparison of the Systems' Presentation of Recommendations; Visualization and User Control of Context Factors

Table 5 shows a comparison of the reviewed systems, showing (in bold) that (1) only three of the 19 systems present music tracks in a nonstandard way; (2) nine let users control context factors on the UI level (typically with common social media features, listed here as "identity"); and (3) nine offer some kind of visualization of context factors (typically location on a geographical map). The three nonstandard presentation of music tracks included CapitalMusic and ContextPlayer, presenting tracks as a variation of a list, an album art mosaic. One system (InCarMusic) let the users control all four context factors. Other than geographical map, visualizations included mobile AR view (OUTMedia, Figure 4), a compass metaphor for location (The Compass, Figure 3), and an equalizer for controlling various activity, location, and time-related dimensions (Lifetrak, Figure 8).

5. User Reactions to Contextual Recommendations

User studies were conducted on 14 of 19 systems. In general, the systems received positive feedback in usability and UX questionnaires. However, we focus on the results of the open-ended user data such as interviews and group discussions that are richer in conveying the user opinions. Another reason for leaving out quantitative questionnaires is that as we did not have the access to the set of questions in most of the cases, it was hard to know if the questions targeted the dimensions we were interested in, i.e. the use of context factors and the user interactions with the factors. Usually, the authors mentioned that a questionnaire was administered and then provided the results, or only a selected part of them. In the case of qualitative results, it was easy to see when the user feedback concerned the reviewed dimensions.

Table 6 gives an overview of the user studies. It highlights the number of participants, evaluation techniques, and the evaluation site. We were delighted to see that the majority of systems (10 of 14) were field studied, as studies in the wild are often necessary for mobile, contextual applications (Tamminen, Oulasvirta, Toiskallio, & Kankainen, 2004). In some studies, the concept was evaluated in a laboratory setting before the field study.

5.1. Social Interaction through Music and Context Factors

Social interaction enabled by music and context factors was clearly the most commented aspect in the open-ended user data.

Table 5. Comparison of the reviewed systems.

System	System type	1. Playlist, stream, or nonstandard UI for presenting music tracks	2. User control of context factors, which?	3. Visualization of context factors, which?
BlueTuna	Collocated	Playlist	None	None
Capital Music	Collocated	**Nonstandard list: Album art mosaic**	**Identity**	None
ContextPlayer	Automated	**Nonstandard list: Album art mosaic**	None	**Location, Identity**
D-Jogger	Task-specific	Stream	None	None
Foxtrot (Ankolekar and Sandholm, 2011)	Geo-tagged	Stream	**Location, identity**	**Location, identity**
IM4Sports	Task-specific	Stream	None	None
InCarMusic	Task-specific	Playlist	**Location, time, activity, identity**	None
Lifetrak	Automated	Stream	**Location, time, activity on an EQ UI**	**EQ: context factor levels**
Mobile Music Genius (Schedl et al., 2014)	Automated	Playlist	None	None
Lehtiniemi & Holm, 2012	Mood based	Playlist	**Activity, choice of 7 photos**	**Photos of various situations**
MyTerritory	Geo-tagged	Map & playlist	**Location, identity**	**Location**
OUTMedia	Geo-tagged	**Nonstandard list: mobile augmented reality**	**Location, identity**	**Location, identity on augmented reality UI**
PlayingGuide	Geo-tagged	Map & playlist	None	**Location**
Push!Music	Collocated	Playlist	None	None
SoundPryer	Task-specific	Stream	None	None
Sounds of Helsinki	Geo-tagged	Map & playlist	**Identity**	**Location**
SuperMusic	Automated	Playlist	None	None
The Compass	Collocated	Playlist	None	**Location through a compass UI**
Wang et al., 2012	Automated	Playlist	**Activity**	None

Social Media Features: Creating, Sharing, and Commenting Content with Context Factors

Word-of-mouth, together with radio, are still the main music discovery channels (Nielsen, 2012). The importance of the context factor identity is emphasized by the social media practices of sharing, liking, and commenting on content. Commenting functions as a popular electronic word-of-mouth mechanism. It is not surprising that users appreciate the social media features, even more so with an additional twist brought by other context factors such as location. Users of several studies stated that sharing one's music and music taste is important for satisfying UX, emphasizing the need for self-expression by making their musical identities available for others (see OUTMedia, Capital Music, MyTerritory, Push! Music; Silfverberg et al., 2011).

For example, MyTerritory users appreciated the low threshold for contextual content creation and sharing. It seems that the key element in the systems relying on user-created content

Table 6. Comparison of user evaluation approaches.

System	Users	User evaluation techniques	Study site
Capital Music	5 + 13	2-stage: focus group, questionnaire	lab & field
ContextPlayer	10 + 60	2-stage: interview, questionnaire	lab & field
D-Jogger	33	questionnaire	lab
Foxtrot (Ankolekar and Sandholm, 2011)	100	questionnaire	lab
IM4Sports	6	questionnaire	lab
Lehtiniemi & Holm, 2012	40	observation, interview, questionnaire	lab & field
MyTerritory	15	interview, questionnaire	field
OUTMedia	18	questionnaire, interview	field
PlayingGuide	26	questionnaire	field
Push!Music	5	questionnaire, focus group	field
SoundPryer	13	video recording, interview	field
Sounds of Helsinki	8 + 15	2-stage: User workshop, questionnaire	lab & field
SuperMusic	42	interview, questionnaire	field
Wang et al., 2012	10	questionnaire	lab

(OUTMedia, MyTerritory, Capital Music) was motivating users to contribute by making content creation and sharing easy and user identity visible.

Users voiced the importance of appropriating systems for their own purposes. For example, Capital Music users repurposed the nickname functionality by emoticons or short messages, explaining that they wanted to describe their motivation, mood, or activity behind their music choices.

Real-Time Social Interaction through Music and Context Factors

Music discovery linked with real-time social interaction was particularly satisfying. For example, Push!Music allows users to send music files to other users. This function was highly appreciated, as users described receiving previously unknown music as serendipitous "treats." Suddenly receiving new music also triggered spontaneous social interaction. The general opinion was that Push!Music was the most engaging when used in a social way, when "*something was happening*", i.e. when the playlist was updating in real time with music tracks moving between participants (Hakansson, Rost, Jacobsson, & Holmquist, 2007).

Capital Music user study found that music was an ideal icebreaker for mediated social interactions (Seeburger et al., 2012). While music is very personal on the experiential level, music-related information such as album art and metadata are not personal. Sharing music-related information anonymously let the users remain in their private bubble and at the same time connect with collocated people in public spaces.

When asked how Capital Music would change their commuting experience, several users commented that they liked the app since it adds an extra level of community to the music listened through interacting people who are physically close but still anonymous. The authors emphasize that digital augmentations and mobile-mediated interactions between collocated strangers must be designed carefully in order to balance between anonymity and communal aspects (Seeburger et al., 2012).

Another real-time-related observation from the user studies concerns the common "status checking" behavior of smartphones (Oulasvirta, Rattenbury, Ma, & Raita, 2012). During a 2-week user study, participants monitored Push! Music frequently and switched views between the playlist, the transfer list, and the list of online nearby users. User data shows that the participants enjoyed keeping track of what was happening in the application. Real-time activities such as transfers of songs triggered discussions in the group.

5.2. Communicating the Current Context Is Important

User data shows that unless system logic is communicated, users often feel confused about the recommendations. While the logic of sports systems is easily understood by users since the systems rely on user's activity levels measured from physical activity, the cluster of automated systems have several context factors and typically do not tell the reason for recommendation. A music stream without explanations may result in an enjoyable and effortless UX, but providing users with some transparency could result even in better UX (Lehtiniemi, 2008; Okada, Karlsson, Sardinha, & Noleto, 2013; Wang et al., 2012).

The need for transparency is illustrated by the SuperMusic (Lehtiniemi, 2008) users, stating that they had difficulties in understanding how the concept of "situations" affected the suggested recommendations. This is because the system did not reveal how the system had produced the recommendations. Around 63% of the users felt that it was important to see the recommendation criteria and that the explanations should be more descriptive and detailed than those provided by the system. Almost half of the ContextPlayer users suggested that notifications of context changes or the reason for receiving specific recommendations should be reflected in the UI (Okada et al., 2013). It seems that it is vital to show or tell why recommendations were made, even more so in the automated contextual systems playing music without request.

SoundPryer user study found that visualizing the source of music recommendation was rewarding. Users appreciated being able to identify the source of recommendation by seeing which nearby car was the source of music. In Wang et al.'s (2012) system, there were interactive buttons for six different activities. Some users wanted more categories for showing the current context more accurately and for interacting with them as seeds for playlists.

Summing up, observations from the user data show that users specifically appreciated features that utilize context factors for enabling social interaction and explaining the system logic on the interface level.

6. Discussion

In this article, we reviewed 19 contextual music recommendation systems and analyzed how primary context factors (time, location, identity, and activity) were adopted for interactive music recommendation and discovery. Location was the dominant context factor and location-time-identity the most common combination. While finding out that some systems help users by communicating their system logic, most of the systems could have been improved by developing the interfaces in terms of transparency, explanations, visualizations, and refinement.

We also reviewed user research results related to the 19 systems. Based on the review findings, we present design implications in Section 6.3. The main findings can be summarized as follows:

- Maps are the typical way of presenting location-based recommendations, be it discovery through collocated users or static geo-located music.
- Social media features, together with context factors, help in supporting serendipitous discoveries.
- Automated, unobtrusive contextual listening experience must be maintained by recommendation techniques that provide contextually relevant results and explain the system logic before the use.
- Filtering features help in browsing content accumulated over time.
- The systems could be improved with novel, interactive ways for transparency and explanations, presentation, and refinement of recommendations.
- Although list view presentation of recommendation results prevails, more information could be shown by using innovative visualizations.
- Multifunctional, interactive UI features are an effective way of interacting with context factors underlying the recommendations.

6.1. System Types and Their Characteristics

Four clusters were formed from the reviewed systems: automated, collocated, geo-tagged, and task-specific. Each group has its idiosyncrasies in terms of the adoption of context features and interaction dimensions that support their designed purpose.

Automated systems typically aim for unobtrusive UX by acquiring the context data implicitly. They provide the user with an automatic stream of music tailored to the current situation in a proactive way. These systems usually adopt multiple context factors in an all-inclusive approach, aiming to cover the user's situation completely. Interactions are typically minimal, so that the user can go on with his everyday activities uninterrupted. However, users often find it hard to understand why a certain recommendation was made. Therefore, transparency and explanations should be provided at the beginning of the system use, and, in order to maintain the "automated" unobtrusive listening UX, the recommender should be made technically so good that users are able to easily understand the recommendation results without needing to make new requests regularly.

Collocated systems exploit nearby people and their music to enable context-sensitive experiences that can lead to further social interaction with other users. Music-related encounters with other people function as an alternative to the common practice of cocooning with one's phone in public places. In these systems, location-time-identity is the dominant context factor combination.

Systems offering geo-tagged music differ from collocated systems in that they allow finding places with music tagged on them. They are location-dominant and often offer map visualization for recommendation transparency. A challenge for these systems, if made commercially available, is finding interesting content from the potentially great volume of content accumulated over time. We suggest including filtering features based on, for instance, time or identity, which allows means for browsing the accumulated content.

In our review, there were two task-specific systems for both car driving and sports. They aimed to offer added value for the task at hand by increasing motivation by activity inference (sport systems), music that matches the overall driving situation by combining various implicitly and explicitly acquired context data (InCarMusic), or sharing music stream during temporary traffic encounters (Sound Pryer).

6.2. Observations about Interaction Adequacy

As with traditional recommendation systems, transparency and explanations of recommendations have a key role in contextual systems. Since some of the systems reveal their system logic immediately and intuitively (e.g. MyTerritory, The Compass), thus explicating how the context factors are employed for offering recommendations, some may argue that contextual systems do not need explanations to the same extent as non-contextual systems. This is true, for example, when the logic is communicated through an obvious feature, such as a map telling that an item was recommended because of the user's close proximity. However, when transparency is not intuitively revealed, it needs to be communicated to the users.

User data provides more evidence on the importance of transparency, suggesting that unless system logic is communicated, users often feel confused about the recommendations. For instance, the users of SuperMusic (Lehtiniemi, 2008) and ContextPlayer (Okada et al., 2013) stated they had difficulty understanding how the concept of "situations" affected the suggested recommendations, since the system did not show the current context factors explicitly. In many cases, the effect of context factors could be communicated to the users by applying interaction and visual features such as timelines, sliders, dials, or non-geographic maps. However, it is surprising that most of the systems do not use the opportunity of utilizing context factors for advanced and nonstandard visualizations that go beyond the typical list view (Table 5).

The contextual factors themselves provide a wide array of information and great potential for visual explanations that allow for much interesting UX. Most systems employ several context factors simultaneously, so visualizations could show the ratio of context factors currently affecting the recommendations as well (see Context EQ Figure 8). Giving the user explanation about why a piece of music was recommended in that particular situation and to that particular person, the perceived quality of the recommendation could be increased. This is because often the why is as important as what is recommended (Celma, 2010).

Presentation of recommendations is crucial for the acceptance of the system (Konstan & Riedl, 2012; Ricci et al., 2012). Explanations must be presented and well-presented recommendations also give the user clues why the recommendation was made. A list view of query results dominates the design of recommenders in general (Ricci et al., 2012) and is a de facto standard in digital music UIs as well (cf. the playlist paradigm). The reviewed contextual systems mostly follow the dominant interaction paradigms and content selection modes of tuning in to either a curated stream or a playlist (Liikkanen & Åman, 2015). However, in the common list view presentation, a lot of information is lost. More information can be shown, for instance, by using (non-geographic) two- or three-dimensional maps that visualize the relative similarity of items, grouping similar items close to each other (Gretarsson et al., 2010). None of the reviewed systems tried to embed additional explanations within the list format.

Refinement and feedback features inspire trust in a system (Jawaheer, Weller, & Kostkova, 2014; Kay, 2006; Tintarev & Masthoff, 2007; Tintarev & Masthoff, 2011). As in music recommender UIs in general (Åman & Liikkanen, 2010), only a few of the reviewed contextual systems have interactive features for scrutability or refinement (e.g. Lifetrak).

Well-designed UI features can successfully serve several purposes. For example, Lifetrak's Context Equalizer (see Figure 8) functions as a combined refinement, scrutability, user feedback and transparency aid. Furthermore, it gives user-controlled means to acquiring potentially serendipitous music recommendations through more informative visualization than a mere list.

Although modern devices and software techniques allow for experimenting with a wide variety of nonstandard interactions and visualizations, we found out that most of the systems rely on standard UI features for interacting with contextual factors (Table 5). Nonstandard and experimental UIs for music and context factor interaction may draw inspiration for example from game interfaces that commonly deploy real-time data. Another area of potential future research is in lifestyle and health applications. For example, visualizations of biophysical data could be used in media discovery as well. Refinement features could involve 2D maps wherein context factors could be placed in weighted order as search keywords, as Andolina et al. (2015) suggest. Another refinement strategy would be to use a query result as seed for a new query, as in some studies (Baur, Boring, & Butz, 2010; Klouche et al., 2015). By communicating in which ways the contextual features work and actually letting users play with them, better UX could be achieved.

Our observations of music interaction paradigms replicate the findings of Liikkanen and Åman's (2015) study on the case of online music services. Contextual music recommendations systems rely on the dominant interaction paradigms of playlist and radio metaphors. The rare exceptions include compass and equalizer metaphors in two reviewed systems. Although dominant paradigms are well-established, there are hundreds of millions of users with unique use situations, music tastes, cultures, needs, and desires; it is thus surprising and inspiring that the potential for innovative visualizations, and interactions with music through context factors, is yet to be realized.

6.3. Implications for Design

We encourage designers to experiment with the following design implications.

Visual and Interactive Explanations

Explore alternatives to textual explanations by offering intuitive visual and interactive UI elements that communicate the system logic or explain why a recommendation was made. Offer users interactions for adjusting the effect of context factors driving the recommendations (for an example, see Figure 8).

Automated Contextual Music Stream Needs Explanations Too

Automated, unobtrusive contextual listening experience must be maintained by recommendation techniques that offer precise context-aware recommendation results and by explaining the system logic before the use.

Filtering Accumulated Content

Include filtering features, such as timelines, sliders, or dials, based on for example time or identity for the faceted browsing of the content accumulated over time.

Multifunctional UI Features

Explore with UI elements and metaphors that can take care of several interactional functions. By combining several functions in one UI feature, simpler, more rewarding UX may be provided. Multifunctional UI features are also an economic option for interacting with context factors.

7. Conclusion

Our main contribution in this study has been in pointing out that while context factors have been included in music recommendation in a number of ways that follow standard playlist and radio-like stream paradigms, innovative and nonstandard interactions and visualizations for context factors are mostly nonexistent. Reviewing user feedback showed that by applying context factors for enabling social interaction and explaining the system logic were especially appreciated by the users.

Beyond contextual music services, our findings can be applied in designing and studying other contextual media services, including various media content such as video, text, audio notes, spoken word, or photos, either curated or user-created. There have been relatively few studies about how to adopt context factors and the related interactions for other media content discovery. For example, Bentley & Basapur (2012), Vihavainen, Mate, Liikkanen, & Curcio (2012), Vihavainen, Mate, Seppälä, Cricri, & Curcio (2011), and Zhang et al. (2012) have studied media discovery (video, photos) and the related social dimensions, but with little focus on interactions or the visualizations of context factors.

Our contribution complements the research on context-aware media applications and services by pointing out the need for exploring novel ways of interaction and visualization. Our findings are applicable to the design of other areas of contextual services, aiming for the recommendation and discovery of video, photos, and other, either user-created or commercially provided content. In conclusion, we hope that our review can help in realizing the great potential in context-aware, personalizable media services.

One limitation concerning the study is that it was conducted as a review. This is because we did not have access to the working prototypes, only articles on them. If the systems were available for user evaluation, that would potentially yield interesting results. Another limitation concerns the systems being not fully comparable with each other, e.g. there were systems designed to accompany a certain task (car driving, sports) and others targeted for non-specified activities (moving about in the city). Therefore, it may be argued that our framework of analysis did not give a balanced treatment to the systems. Still, within our selection criteria (context-aware music recommender featuring a UI and user-studied), we believe that we have produced a fruitful review, mapping out the ways in which context factors were applied for music recommendation and discovery.

7.1. Future Work

Interaction features that did not occur in the reviewed systems include browsing content by histories of use, places, users, and activities, but we encourage designers to experiment with them, e.g. by adopting interactive timelines. An example of a potential time-based concept for music discovery would be filtering content by the release date of a track combined with location to promote local artists. Timelines or other visualizations can help in filtering content and fulfill the functions of each of the dimensions of interaction adequacy (user effort, transparency, presentation, refinement). Real-time recommendations and interactions have great potential for rewarding UXs, as they make the users feel that something is happening right now and they are part of it. This resembles the feel of following a live event, and it is one of the dimensions we encourage to experiment with and support in contextual media systems.

In the modern ubiquitous media environment, it is gradually becoming common to anticipate that recommenders detect the context and offer recommendations when needed and without request. Therefore, proactive recommendations (Ricci 2012) that utilize context factors is a lucrative direction for future work.

While stream and playlist paradigms will prevail also in the future, by exploring modifications such as mobile-augmented reality, compass, or equalizer metaphors reviewed in this article, more interesting UX with music recommendations and music interactions could be provided.

References

Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In F. Ricci, F. L. Rokach, B. Shapira, & P. Kantor (Eds.), *Recommendation systems handbook* (pp. 217–256). New York, NY: Springer.

Åman, P., & Liikkanen, L. A. (2013). Painting the city with music: Context-aware mobile services for urban environment. *Continuum*, *27*(4), 542–557.

Åman, P., & Liikkanen, L. A. (2010). A survey of music recommendation aids. *CEUR Workshop Proceedings*, *633*, 25–28.

Åman, P., Liikkanen, L. A., & Hinkka, A. (2015). Enriching user experiences with location-sensitive music services. *Journal of Location Based Services*, *9*(3), 167–186.

Andolina, S., Klouche, K., Peltonen, J., Hoque, M., Ruotsalo, T., Cabral, D., … Jacucci, G. (2015, March). *IntentStreams: Smart parallel search streams for branching exploratory search*. Proceedings of the 20th International Conference on Intelligent User Interfaces - IUI ‘15. Andolina, GA.

Ankolekar, A., & Sandholm, T. (2011, August). *Foxtrot: A soundtrack for where you are*. Proceedings of Interacting with Sound Workshop:

Exploring Context-Aware, Local and Social Audio Applications - IwS '11. Stockholm, Sweden.

Baltrunas, L., Kaminskas, M., Ludwig, B., Moling, O., Ricci, F., Aydin, A., Lüke, K. H., et al. (2011). InCarMusic: Context-aware music recommendations in a car. In C. Huemer and T. Setzer (Eds.). *Lecture notes in business information processing* (Vol. 85, pp. 89–100). Berlin, Germany: Springer Verlag.

Baumann, S., Jung, B., Bassoli, A, & Wisniowski, M. (2007, April-May). *BluetunA: Let your neighbour know what music you like*. Proceedings of ACM CHI 2007 Conference on Human Factors in Computing Systems (1941–1946). San Jose, CA.

Baudelaire, C. (1863/1964). *The painter of modern life*. Originally published in Le Figaro 1863. New York, NY: Da Capo Press.

Baur, D., Boring, S., & Butz, A. (2010, February). *Rush: Repeated recommendations on mobile devices*. Proceedings of Intelligent User Interfaces - IUI '10, ACM (pp. 91–100). Hong Kong, China.

Beale, R., & Lonsdale, P. (2004, September). *Mobile context aware systems: The intelligence to support tasks and effectively utilise resources*. Mobile Human–Computer Interaction - MobileHCI 2004 (pp. 240–251). Glasgow, UK.

Bentley, F., & Basapur, S. (2012, May). *StoryPlace.Me: The path from studying elder communication to a public location-based video service*. CHI'12 Extended Abstracts on Human Factors in Computing Systems (pp. 777–792). Austin, TX.

Biehl, J. T., Adamczyk, P. D., & Bailey, B. P. (2006). *DJogger: A mobile dynamic music device*. CHI '06 Extended Abstracts on Human Factors in Computing Systems - CHI EA '06 (pp. 556–561). ACM Press.

Braunhofer, M., Kaminskas, M., & Ricci, F. (2011, October). *Recommending music for places of interest in a mobile travel guide*. Proceedings of the fifth ACM Conference on Recommender Systems (pp. 253–256). Chicago, IL.

Bull, M. (2008). *Sound moves: IPod culture and urban experience*. New York, NY: Routledge.

Carroll, J., Mack, R., & Kellogg, W. (1991). Interface metaphors and user interface design. In M. Helander (Ed.), *Handbook of human–computer interaction* (2nd ed., pp. 67–86). Amsterdam, Netherlands: North-Holland.

Celma, Ò. (2010). *Music recommendation and discovery*. Heidelberg, Germany: Springer.

Cunningham, S., Caulder, S., & Grout, V. (2008, October). *Saturday night or fever? context aware music playlists*. Proceedings of the Audio Mostly Conference - A Conference on Interaction with Sound (pp. 64–71). Piteå, Sweden.

de Certeau, M. (1984). *The practice of everyday life*. Berkeley, CA: University of California.

de Certeau, M., Giard, L., & Mayol, P. (1998). *The practice of everyday life* (Vol. 2). Minneapolis, MN: University of Minnesota Press.

Debord, G. (1955/2006). Situationist international anthology-revised and expanded edition. In K. Knabb (Ed.), *Introduction to a critique of urban geography*. Berkeley, CA: Bureau of Public Secrets.

Dervin, B. (2003). *Sense-making methodology reader: Selected writings of Brenda Dervin*. Hampton Press.

Dey, A. K., & Abowd, G. D. (1999). Towards a better understanding of context and context-awareness. *Computing Systems*, *40*(3), 304–307.

Dourish, P. (2004). What we talk about when we talk about context. *Personal and Ubiquitous Computing*, *8*(1), 19–30.

Downie, J. S. (2008). The music information retrieval evaluation exchange (2005–2007): A window into music information retrieval research. *Acoustical Science and Technology*, *29*(4), 247–255.

Fiske, J. (1990). *Introduction to communication studies*. London, UK: Routledge.

Fleder, D. M., & Hosanagar, K. (2007, June). *Recommender systems and their impact on sales diversity*. Proceedings of the 8th ACM Conference on Electronic Commerce, ACM (pp. 192–199). San Diego, CA.

Forsblom, A., Nurmi, P., Åman, P., & Liikkanen, L. (2012). Out of the bubble - Serendipitous event recommendations at an urban culture festival. *International Conference on Intelligent User Interfaces, Proceedings IUI*, 253–256.

Gretarsson, B., O'Donovan, J., Bostandjiev, S., Hall, C., Höllerer, T. (2010). Smallworlds: Visualizing social recommendations. *Computer Graphics Forum* 29(3), 833–842.

Hakansson, M., Rost, M., Jacobsson, M., & Holmquist, L.-E. (2007, January). *Facilitating mobile music sharing and social interaction with Push!Music*. Proceedings of 40th Annual Hawaii International Conference on System Sciences (pp. 87). Waikoloa, HI.

Han, B. J., Rho, S., Jun, S., & Hwang, E. (2010). Music emotion classification and context-based music recommendation. *Multimedia Tools and Applications*, *47*(3), 433–460.

Jacucci, G., Peltonen, P., Morrison, A., Salovaara, A., Kurvinen, E., & Oulasvirta, A. (2009). Ubiquitous media for collocated interaction. In K. S. Willis, G. Roussos, K. Chorianopoulos, & M. Struppek (Eds.), *Shared encounters* (pp. 23–45). Heidelberg, Germany: Springer.

Jawaheer, G., Weller, P., & Kostkova, P. (2014). Modeling user preferences in recommender systems: A classification framework for explicit and implicit user feedback. *ACM Transactions on Interactive Intelligent Systems*, *4*(2), 1–26.

Jones, N., & Pu, P. (2007, September). *User technology adoption issues in recommender systems*. Proceedings of the 2007 Networking and Electronic Commerce Research Conference (pp. 379–394). Lake Garda, Italy.

Kaminskas, M., & Ricci, F. (2012). Contextual music information retrieval and recommendation: State of the art and challenges. *Computer Science Review*, *6*(2–3), 89–119.

Karageorghis, C. I., Mouzourides, D., Priest, D.-L., Sasso, T., Morrish, D. J., & Walley, C. J. (2009). Psychophysical and ergogenic effects of synchronous music during treadmill walking. *Journal of Sport & Exercise Psychology*, *31*(1), 18–36.

Kay, J. (2006, June). *Scrutable adaptation: Because we can and must*. Proceedings of 4th International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems - AH 2006 (pp. 11–19). Dublin, Ireland.

Klouche, K., Ruotsalo, T., Cabral, D., Andolina, S., Bellucci, A., & Jacucci, G. (2015, April). *Designing for exploratory search on touch devices*. Proceedings of the ACM CHI'15 Conference on Human Factors in Computing Systems (Vol. 1, pp. 4189–4198). Seoul, Korea.

Konstan, J. A., & Riedl, J. (2012). Recommender systems: From algorithms to user experience. *User Modeling and User-Adapted Interaction*, *22*(1–2), 101–123.

Krippendorff, K. (2006). *The semantic turn*. Boca Raton, FL: Taylor & Francis.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. London, UK: University of Chicago Press.

Lehtiniemi, A. (2008, December). *Evaluating SuperMusic: Streaming context-aware mobile music service*. Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology (pp. 314–321). Yokohama, Japan.

Lehtiniemi, A., & Holm, J. (2012, July). *Using animated mood pictures in music recommendation*. 16th International Conference on Information Visualisation, IEEE (pp. 143–150). Montpellier, France.

Lehtiniemi, A., & Ojala, J. (2012, December). *MyTerritory - Evaluation of outdoor gaming prototype for music discovery*. Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia - MUM '12. Ulm, Germany.

Levitin, D. (2007). *Life soundtracks: The uses of music in everyday life*. Montreal, Canada: McGill University.

Liikkanen, L., Amos, C., Cunningham, S. J., Downie, J. S., & McDonald, D. (2012, May). *Music interaction research in HCI*. Proceedings of the 2012 ACM annual conference extended abstracts on Human Factors in Computing Systems Extended Abstracts - CHI EA '12 (pp. 1119). Austin, TX.

Liikkanen, L. A., & Åman, P. (2015). Shuffling services: Current trends in interacting with digital music. *Interacting with Computers*, *28*, 352–371.

Liikkanen, L. A., & Gómez, P. G. (2013, September). *Designing interactive systems for the experience of time*. Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces - DPPI '13 (pp. 146). Newcastle upon Tyne, UK.

Lundgren, S., Fischer, J. E., Reeves, S., & Torgersson, O. (2015, March). *Designing mobile experiences for collocated interaction*. Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing - CSCW '15 (pp. 496–507). Vancouver, Canada.

Magno, T., & Sable, C. (2008, September). *A comparison of signal-based music recommendation to genre labels, collaborative filtering,* *musicological analysis, human recommendation, and random baseline*. Proceedings of the 2008 International Conference on Music Information Retrieval - ISMIR '08 (pp. 161–166). Philadelphia, PA.

Moens, B., Van Noorden, L., & Leman, M. (2010, July). *D-Jogger: Syncing music with walking*. Proceedings of Sound and Music Computing Conference. Barcelona, Spain.

Nielsen SoundScan. (2012). *Music discovery still dominated by radio, says Nielsen Music 360 report*. Retrieved May 4, 2016, from http://www.nielsen.com/us/en/press-room/2012/music-discovery-still-dominated-by-radio–says-nielsen-music-360.html

North, A. C., Hargreaves, D. J., & Hargreaves, J. J. (2004). Uses of music in everyday life. *Music Perception: An Interdisciplinary Journal*, *22*(1), 41–77.

Okada, K., Karlsson, B. F., Sardinha, L., & Noleto, T. (2013, November). *ContextPlayer*. SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications on - SA '13 (pp. 1–7). Hong Kong, China.

Östergren, M. (2004, September). *Sound Pryer: Adding value to traffic encounters with streaming audio*. Proceedings of the 3rd International Conference on Entertainment Computing - ICEC '04 (pp. 541–552). Eindhoven, The Netherlands.

Oulasvirta, A., Rattenbury, T., Ma, L., & Raita, E. (2012). Habits make smartphone use more pervasive. *Personal and Ubiquitous Computing*, *16*(1), 105–114.

Page, W., & Ning, K. (2014). *Anatomy of a hit: How Meghan Trainor made the UK chart without selling downloads*. In Spotify Insights. Retrieved May 4, 2016, from https://insights.spotify.com/us/2014/12/16/anatomy-of-a-hit-how-meghan-trainor-made-the-uk-chart-without-selling-downloads/

Park, H.-S., Yoo, J.-O., & Cho, S.-B. (2006, September). *A context-aware music recommendation system using fuzzy Bayesian networks with utility theory*. Proceedings of Fuzzy systems and knowledge discovery (pp. 970–979). Xi'an, China.

Park, J. (2015). Improving user interaction in interactive TV based on ethnographic insights from real-life problems. *International Journal of Human–Computer Interaction*, *31*(1), 17–32.

Pu, P., & Chen, L. (2010). A user-centric evaluation framework of recommender systems. *CEUR Workshop Proceedings*, *612*, 14–21.

Razzaque, M., Dobson, S., & Nixon, P. (2005, August). *Categorisation and modelling of quality in context information*. Proceedings of IJCAI 2005 Workshop on AI and Autonomic Communications. Edinburgh, Scotland.

Reddy, S., & Mascia, J. (2006, October). *Lifetrak: Music in tune with your life*. International Multimedia Conference (pp. 25–34). Santa Barbara, CA.

Ricci, F. (2012, April). *Context-aware music recommender systems: Workshop keynote abstract*. Proceedings of the 21st international conference companion on World Wide Web - WWW '12 (pp. 865–866). Lyon, France.

Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor (Eds.), *Recommendation systems handbook*. New York, NY: Springer.

Schedl, M., Breitschopf, G., & Ionescu, B. (2014). *Mobile music genius: Reggae at the beach, metal on a Friday night?* Proceedings of International Conference on Multimedia Retrieval - ICMR '14 (pp. 507–510).

Schmidt, A., Beigl, M., & Gellersen, H.-W. (1999). There is more to context than location. *Computers and Graphics*, *23*(6), 893–901.

Seeburger, J., Foth, M., & Tjondronegoro, D. (2012, December). *The sound of music: Sharing song selections between collocated strangers in public urban places*. Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia - MUM '12 (pp. 1). Luleå, Sweden.

Silfverberg, S., Liikkanen, L. A., & Lampinen, A. (2010, March). *I'll press Play, but I won't listen": Profile work in a music-focused social network service*. Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (pp. 1–10). Hangzhou, China.

Silverstone, R. (2003). Private reveries and public spaces. *Proboscis Cultural Snapshots*, *4*. Retrieved May 4, 2016, from http://proboscis.org.uk/projects/2000-2005/private-reveries-public-spaces/

Sinha, R., & Swearingen, K. (2002, April). *The role of transparency in recommender systems*. CHI 02 Extended Abstracts on Human Factors in Computing Systems - CHI 02 (pp. 830). New Orleans, LA.

Su, J.-H., Yeh, -H.-H., Yu, P. S., & Tseng, V. S. (2010). Music recommendation using content and context information mining. *IEEE Intelligent Systems*, *25*, 16–26.

Swearingen, K., & Sinha, R. (2001). *Beyond algorithms: An HCI perspective on recommender systems*. ACM SIGIR 2001 Workshop on Recommender Systems (pp. 1–11).

Tamminen, S., Oulasvirta, A., Toiskallio, K., & Kankainen, A. (2004). Understanding mobile contexts. *Personal Ubiquitous Computing*, *8*(2), 135–143.

Tanaka, A., Valadon, G., & Berger, C. (2007, May). *Social mobile music navigation using the compass*. Proceedings of 4th International Mobile Music Workshop. Amsterdam, Netherlands.

Te'eni, D., & Sani-Kuperberg, Z. (2005). Levels of abstraction in designs of human–computer interaction: The case of e-mail. *Computers in Human Behavior*, *21*(5), 817–830.

Theimer, M. M., & Schilit, B. N. (1994). Disseminating active map information to mobile hosts. *IEEE Network*, *8*(5), 22–32.

Tintarev, N., & Masthoff, J. (2011). *Designing and Evaluating Explanations for Recommender Systems*, Pages 479-510. In F. Ricci, L. Rokach, B. Shapira, & P. Kantor (Eds.), *Recommendation systems handbook*. New York, NY: Springer.

Tintarev, N., & Masthoff, J. (2007, April). *A survey of explanations in recommender systems*. Proceedings of International Conference on Data Engineering (pp. 801–810). Istanbul, Turkey.

Tu, Y., Chen, L., Lv, M., Ye, Y., Huang, W., & Chen, G. (2013). iReminder: An intuitive location-based reminder that knows where you are going. *International Journal of Human–Computer Interaction*, *29*(12), 838–850.

Vihavainen, S., Mate, S., Liikkanen, L., & Curcio, I. (2012, May). *Video as memorabilia: User needs for collaborative automatic mobile video production*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '12 (pp. 651–654). Austin, TX.

Vihavainen, S., Mate, S., Seppälä, L., Cricri, F., & Curcio, I. (2011, May). *We want more: Human–computer collaboration in mobile social video remixing of music concerts*. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '11 (pp. 287–296). Vancouver, Canada.

Wang, X., Rosenblum, D., & Wang, Y. (2012, October-November). *Context-aware mobile music recommendation for daily activities*. Proceedings of the 20th ACM international conference on Multimedia - MM '12 (pp. 99). New York, NY.

Wijnalda, G., Pauws, S., Vignoli, F., & Stuckenschmidt, H. (2005). A personalized music system for motivation in sport performance. *IEEE Pervasive Computing*, *4*(3), 26–32.

Yang, S., Lu, Y., Gupta, S., & Cao, Y. (2012). Does context matter? The impact of use context on mobile internet adoption. *International Journal of Human–Computer Interaction*, *28*(8), 530–541.

Zhang, L., Liu, Y., Zhan, X., Yang, X., Chi, X., & Zhao, S. (2012). Campus view: An innovative location and context-aware video sharing application on smart phone. *Wireless Personal Communications*, *66*(3), 493–509.

Zhou, X., Xu, Y., Li, Y., Josang, A., & Cox, C. (2012). The state-of-the-art in personalized recommender systems for social networking. *Artificial Intelligence Review*, *37*(2), 119–132.

About the Authors

Pirkka Åman (@pirkkaaman) is a doctoral candidate in Media Lab Helsinki, School of Arts, Design and Architecture, Aalto University, Finland. His current research includes music discovery and recommendation through contextual and experimental interfaces. He also has research interests in augmented urban environments and energy and health applications.

Lassi A. Liikkanen (@lassial) is an adjunct professor at Aalto University and design analyst at Futurice, Finland. His PhD in Product Development was awarded by the Aalto University, Finland, with a postdoctorate period partially spent at Stanford University. His interest lies in design practices and psychology. He has authored 80+ peer-reviewed publications.