TUM

TECHNISCHE UNIVERSITÄT MÜNCHEN
INSTITUT FÜR INFORMATIK

SerenCast: Towards Understanding the Occurrence of Serendipity During the Consumption of Digital Content

Wessam Abdrabo, Yonata Andrelo Asikin, Wolfgang Wörndl

TUM-I1413

Technischer Bericht

Technische Universität München

Institut für Informatik

SerenCast: Towards Understanding the Occurrence of Serendipity During the Consumption of Digital Content

Wessam Abdrabo
Department of Informatics
TU München
wessam.abdrabo@tum.de

Yonata Andrelo Asikin
BMW Group
yonata-andrelo.asikin@bmw.de

Wolfgang Wörndl
Department of Informatics
TU München
woerndl@in.tum.de

ABSTRACT
The goal of this research is to investigate the factors and processes that lead to the occurrence of serendipity during the consumption of digital content. In the context of recommender systems, serendipity is defined as the act of finding valuable content that is unexpected and unsought for. Hence, the inclusion of serendipity is claimed to challenge the user's curiosity to try the new and unfamiliar which could yield to more enriching and novel experiences. Given the complex and challenging aspects of serendipity (randomness, chance, openness to experience, creativity), various studies are aiming at capturing its nature and in devising methods that can model and induce it in recommender systems. This study focuses on investigating the influence of context information such as location, time, and users' state-of-mind on the occurrence of serendipity. For this purpose, we developed SerenCast; a podcast evaluator system installed on users' mobile devices to elicit a number of factors that capture serendipitous experiences. Later, these factors are analyzed in order to explain which factors lead to the occurrence of serendipity, whether the experience was valuable to the user, and whether it could backfire at times.

Keywords
Recommender systems, serendipity, novelty, metrics, digital content, discovery

1. INTRODUCTION
Recommender systems (RS) are the software tools that generate personalized recommendations which aid users in making decisions during the consumption of digital content such as music, podcasts, news, or movies. There is a gap between how recommender systems are expected to work and how they actually work. The core reason for this gap originates from one of RS's main features: its high predictability. Consider the following phrase: *"If you liked Silent Sea by KT Tunstall, you might also like..."*. If ten people were asked to finish this sentence, there will possibly be ten different answers. However, if ten recommender systems were consulted to finish the same sentence, it is likely that there will not be much variation. The recommendations would probably be one of the following: another song by a british singer, another popular or recent song by the same artist, or another song/artist in the same genre. While this type of recommendation can be considered accurate and useful, it tends to limit the user's experience. The main reason for this behavior is that most of the techniques used in RSs follow the same patterns for recommendations generation. Even an RS that is capable of inducing novel and diverse recommendations is still providing them in the area of what the user expects. What RSs have failed to consider for a long time is the value of unexpected yet useful discoveries, in other words, serendipitous discoveries, which not only makes recommendations less predictable, but is also believed to expand and enrich the user's experiences.

This study focuses on identifying the factors that affect the occurrence of serendipity during the consumption of digital content. For this purpose, we developed SerenCast; a podcast evaluator system installed on users' mobile devices to elicit a number of factors that capture serendipitous experiences. Later, these factors are analyzed in order to explain which factors lead to the occurrence of serendipity, whether the experience was valuable to the user, and whether it could backfire at times.

The remainder of this paper is structured as follows: In section 2, we start by giving background information about recommender systems including their purpose, techniques and evaluation methods. Then we introduce the concept of serendipity, its definitions and implications, as well as the role it plays in the area of recommender systems. Section 3 discusses related studies. In section 4 we explain our application design and the user study setup. Later in section 5, we discuss how we conducted and evaluated SerenCast, and then state the experiment's findings. Finally, we present a discussion of our user study's findings and end with a conclusion and future work.

2. BACKGROUND
2.1 Basics and Motivation
The Long Tail
The consumption of digital content entails activities such as receiving movie suggestions from online stores (e.g. NetFlix), reading online news articles on news portals (e.g. Huffington Post), and listening to music using digital music libraries (e.g. Spotify). A global paradigm shift occurred as the consumption of digital media replaced in all practical sense the consumption of items in the physical world. Today, world-wide consumers resign to the use of digital content which includes a vast array of options, rather than, for example, visiting the local bookstore which may or may not have their desired items, but would definitely not have as much variety as the thousands of choices that can be suggested by an online counterpart. However, this new pattern of con-

sumption developed throughout the years to reach a predicament. As the amount of available digital content started to reach an explosive dimension, it has become impossible to overwhelm the user with all possible choices. While the power of choice is understood to be desirable, overwhelming choices hinder users' ability to make a decision. That pressed the need for tools that can filter available choices and aid users with their decisions. The Long Tail phenomena [2] explains that while the physical world could only make suggestions based on popularity of items, digital media should present consumer-specific recommendations that filters the numerous choices consumers face.

Recommender Systems: Purpose and Techniques

Recommender Systems (RS) is the term used to define the set of software tools that aid users in making informed decisions during the consumption of digital content [7]. Recommender systems are for instance used by Amazon to recommend "what others also bought" on purchasing a similar item. Recommender systems are also used by Spotify to suggest a new song in the user's geographical area, by Google search engine to sort retrieved search results based on the user's history, and by YouTube to suggest which videos to watch among millions of available videos.

One of the early activities in the process of creating recommendations is the elicitation of the user's preferences that help understand what the user needs and expects. This step can be done explicitly (e.g. asking the user to rate items), or implicitly by learning the user's habits (e.g. tracking which items were recently purchased by the user, or which links were recently clicked) [7].

Recommender systems use a number of techniques the most well-known of which are:

- **Content-based**: this technique focuses on the similarity between newly recommended items and the set of items that were previously seen or rated by the user. For example, a content-based approach would record that a given user usually selects movies from the action genre, therefore would keep recommending items from the same genre. Content-based approaches, despite being accurate, suffer from over-specialization of recommendations [7].
- **Collaborative filtering**: the technique filters recommendations based on similar users ' search history and profiles. The most famous example is Amazons' "Customers who bought items in your recent category also bought" section. The technique is considered one of the most widely used in recommender systems [7].
- **Context-aware**: the two previous techniques define a recommendation by means of item-item or user-user similarity. Contextual information such as time of day, weather, location, the user's company, the user's mood or state of mind [1], is however the base for context-aware recommender systems. An example recommendation based on context is recommending a news article to a user in a weekday's morning, stock market on a weekday's afternoon, while recommending movies reviews on a weekend [1]. Recommending items nearby a specific location is also another example. Context can be elicited explicitly (e.g. asking the user specific context related questions), implicitly (e.g. tracking the user's location), or by inferring using data mining methods [1]. More intricate context-aware systems can depict the user's mood and make recommendations base on it. Context-aware recommendations are therefore considered to be more interesting than recommendations solely based on item or user similarity. [1]

What makes for "Good" Recommendations?

The question remains: how can the value of recommendations be evaluated? There are various metrics defined for the evaluation of recommender systems the most important of which is user satisfaction. However, there is no consensus on what exactly constitutes a satisfactory user experience. A user experience is subjective and therefore what makes for a good experience for one user might not necessarily be valid for another. Some recommender systems evaluation metrics are defined below.

- **Accuracy** is one of the traditional measures for evaluating recommendations. An accurate recommendation is one that can predict user's ratings and preferences accurately. In content-based approach, an item that scores low similarity to previously recommended items is considered to be an inaccurate recommendation. For instance, if a user is known to like action movies, a recommendation to watch a comedy movie may be considered inaccurate by a standard content-based RS.
- **Coverage** is defined in [8] as the percentage of items for which effective recommendations can be given. In the content-based analogy, if the system is expected to retrieve all available similar action movies, but only succeeds to retrieve half of this set, the system can not be considered to have full coverage of all possible outcomes.
- **Popularity** is one of the more straightforward and easy to evaluate aspects of recommendations. Popular recommendations are usually computed from lists of the most familiar or trending items, for example, in the area of music retrieval systems, it is given by a list of the most trending songs in a given country, or by a given artist. Popularity provides a valid explanation for recommendations that tend to appeal to wide range of users.
- **Recentness** is used to push more up-to-date items rather than older items on the top of recommendation lists. A recent recommendation does not have to be a popular one. For example, a recent recommendation of a book by the user's favorite author might be more appealing than a recommendation of the most popular book of the year.

The reasons why more metrics for the evaluation of recommender systems are needed is described next.

Escaping the Filter Bubble
While similar, accurate, popular, and recent recommendations could be of great relevance to what the user is seeking, they might also be too relevant that they create what has come to be known as the user's "Filter Bubble" [13]. What a filter bubble creates for users is a space of highly homogenous information which matches their believes and ideologies, and consecutively, separates them from contents that are different or opposing to their believes. It can be argued that this is in itself the sole purpose of recommender systems: the creation of personalized recommendations based on a user's profile. However, there are three aspects that are challenged with this argument. The effect of filter bubbles tend to:

- **Hinder the user's curiosity**: if the user is always presented with items that appeal to their taste, their curiosity to learn about different topics which could be more novel and enriching, will be hindered.
- **Bore the user**: many studies have revealed that users tend to be bored with highly similar recommendations [11, 14, 17]. Users are not just seeking relevant information. Interesting recommendations that tend to expand the user's knowledge are found to be more valuable by a wide variety of users [16].
- **Assume that useful information is only that which is recognized as useful**: according to [16], there are two types of useful information; information that is recognized as useful, and information that is actually useful but is not recognized as useful. To elaborate on this, imagine a user who is a fan of pop music and has never listened to jazz music. A new, popular pop song is going to be recognized as a useful recommendation, and hence will be presented to the user at the top of the list. While a popular or highly rated jazz track is less likely to be on the list of recommendations. However, this user might come accidentally across a jazz track and happen to find it interesting. This later type of discovery is what will never show if recommender systems follow approaches that result in filter bubbles, as accidental discoveries are disregarded. More about accidental/serendipitous discoveries is given in section 2.2.

To escape the filter bubble, aspects such as diversity and novelty of recommendations are integrated in recommendations. These aspects can be defined as:

- **Diversity** reflects the variety included in a recommendation [17]. It is an important factor that help alleviate the homogeneity of recommendation lists. A diverse list of recommendations would include items from dissimilar sets. For example, a fan of Science Fiction would value recommendations of books by Stephen King, however, if the recommendations only include items related to Stephen King, that might not be considered valuable. Other variety of authors might be more valuable. Ziegler et al. [4] define the metric Intra List Similarity on which diversity of a recommendation can be computed. Studies have shown that users would occasionally value variety than similar recommendations [9].
- **Novelty** can be reached by recommending items that users would expect and like, but that they have never seen before. Drawing from the previous example, a novel recommendation would be a Science Fiction movie or book that can not be found in the user's search history. A system that recommends the most accurate recommendations but repeat them over time, such that the user have the same experiences every time she uses the system, could lead to no added value, and hence would negatively affect the user's experience.

While diverse and novel recommendations challenge the limitations of users filter bubbles, there is a more key aspect that radically changes how users discover digital content and whose effect has not yet been thoroughly examined: serendipity. Diversity and novelty are related to serendipity but do not induce the same effect. The focus of this study is the examination of which factors makes for serendipitous experiences.

2.2 Introducing Serendipity

A rather philosophical definition of serendipity is given by Queau in 1986 as "the art of finding what we are not looking for by looking for what we are not finding" [6]. Many discoveries throughout history have been the fruit of "fortunate mistakes" in which a solution to an original problem diverged into a solution to a different problem that was not intentionally sought for. Take for example Columbus's fortunate miscalculations that lead him to the unintentional discovery of America and not to the original intention of his quest to reach Asia. Many other examples from the history of scientific discoveries show that valuable encounters could happen unintentionally during the course of seeking for something else.

Thus, serendipitous encounters emerge from a place where there is no a priori intention, or where the original goal is different from the attained goal. Hence, serendipity can be simply defined as the act of "finding without seeking" [16].

This implies that the nature of serendipitous discoveries is twofold. For serendipity to happen, the encounter needs to be unexpected yet useful. The element of surprise, if not coupled with an added value, would then not lead to serendipity. Methods that has been so far used to trigger serendipitous discoveries include [16]:

- **Role of chance** using random node generation.
- **Pasteur's principle** "Chance favors the prepared mind". In this context, this refers to a n elicited user's profile.
- **Anomalies and exceptions** using poor similarity measures.

This very idea of serendipity, if effectively applied to recommender systems' techniques, could alter how discovering digital content occur. An unexpected and useful recommendation would not just offer a diverse and novel experience to users, but would effectively engage the user's curiosity to explore more enriching content.

To elaborate on how serendipity can happen during the use of recommender systems, consider user X who is a fan of indie-rock music and uses a system that suggests to play a radio based on her preferred genre. Despite the fact that most recommendations accurately matches her profile, user X is bored with the recommendations as most of them have already been suggested before and as they are all basically similar. A recommendation to listen to a new, unexplored genre, such as jazz, might be surprising, but could also end up being more satisfactory to the user than more indie-rock recommendations. User X may end up liking the jazz song and exploring more about the genre. However, this might not necessarily be the case at all times, in all circumstance, for all users.

To device a system that can induce such an effect and result in a positive user experience is a non-trivial task. There are many questions to answer about the implication of serendipity in recommender systems among which are: Is serendipity simply blind chance? Is serendipity random? What value does serendipity add to recommender systems? Can serendipity be effectively modeled and induced? Which factors trigger the occurrence of serendipity? Is serendipitous discovery valued by a recommender system's user? Are the preconditions for the occurrence of serendipity, as well as its value, subjective and relative from one user to another? Can serendipity back fire? Our study focuses on studying the perceived value of serendipity as well as the factors which influence its occurrence.

2.3 SerenCast

While a number of studies have been conducted to define models and methods to induce serendipity in recommender systems, less studies aim at providing a thorough understanding of the factors which lead to the occurrence of serendipity during the consumption of digital content. SerenCast is an iOS application for playing and evaluating podcasts that was developed to explore the factors which could lead to the occurrence of serendipity, as well as to identify the added value of serendipity. SerenCast is a mobile application installed on the devices of a number of participants who were asked to use the application to evaluate a limited number of prepared podcasts. Each evaluation consists of a number of rating criteria. Additionally, contextual information (time and location) were fetched implicitly during the experiment. By the end of the experiment, participants were asked their opinions through a questionnaire to confirm the occurrence of serendipity and to validate the collected data. Data from both the experiment and questionnaire was analyzed to answer the research questions.

3. RELATED WORK

Studies of the aspects of serendipitous discoveries were not always closely related to information and digital content retrieval. Shortcomings of techniques based entirely on accuracy of recommendations lead to the need for methods that create more interesting recommendations. Such realization was the key to studying the effects of factors such as serendipity, novelty and diversity.

The reason why accuracy is not enough is the main concern addressed in [17]. Again, the value of diversity, novelty and serendipity are assumed to be more valuable to a user's experience than highly accurate recommendations. In their attempt to validate this assumption, [17] developed four different variations of their Auralist Framework: Basic Auralist, Community-Aware Auralist, Bubble Auralist, and Full Auralist. Basic Auralist used a content-based approach which is a variation of the topic-modeling technique, Latent Dirichlet Allocation (LDA), that [17] calls the Artist-Based LDA. While the accuracy of Basic Auralist recommendations was high, it failed at addressing the issues of diversity and novelty. As an improvement to the Basic Auralist, Artist-based LDA was combined with Listener Diversity (which aims at the diversification of the suggested artist recommendations) to produce Community-Aware Auralist, and Bubble Auralist was developed to recommend items outside of the user's "musical-bubble". Both Community-Aware and Bubble Auralists recorded a tradeoff between accuracy levels and diversity, novelty, and serendipity levels of recommendations. Full-Auralist (combining approaches of the 3 versions), was developed for a user study to test the effectiveness of the approach in a real-world experiment. The experiment showed lower levels of accuracy than Basic Auralist. However, users' comments showed higher levels of overall satisfaction and the reasons given were all related to a higher level of diversification and serendipitous discoveries that users experienced using Full Auralist than the other versions of the framework.

Several related approaches have been developed to study the effect of serendipity during the consumption of different media [5,10–12,15,17]. There is a general consensus among the reviewed studies that useful yet unexpectedness/serendipitous discoveries introduce a level of diversification and enrichment to the recommendations, that despite the trade-off with accuracy, improves the overall user experience.

In [11], a content-based approach is taken for the development of a model for serendipitous music retrieval. According to [11], the regular approach of basing music recommendations solely on the similarity of newly recommended items to a seed recommendation lead to the creation of highly homogenous lists, and consequently results in a boring experience for the user. Thus, including factors such as diversity, popularity, hotness, recentness and novelty is assumed to improve the user's experience. To introduce a level of diversity, aspects of a recommendation (e.g. rhythm, artist, lyrics, tags) are divided into 2 sets such that similar items would be recommended according to one set, while dissimilar items would be recommended according to the other. Popularity and hotness are measured according to how well-known a song or an artist is in a specific geographical region, which is identified using music charts such as "Billboard", or by inspecting social media feeds such as Twitter. The novelty measure is approximated to be a binary value of wether or not a recommendation has been previously suggested to a user, while recentness is computed as a function of the release date and year of a given track.

Two metrics are given in [12] for the evaluation of recommender systems: coverage (the ability of the system to generate recommendations for as many items as possible), and serendipity. The study suggests that coverage alone is not enough to evaluate a recommender system, and that serendipity is key to measuring how novel and valuable a recommendation is to the user. As with SerenCast evaluation, in [12] the key metrics for evaluating serendipitous experiences are unexpectedness and usefulness. A set of unexpected recommendations is given by UNEXP = RS/PM, which is the set of items belonging to RS (the set of items suggested by a recommender system), but not to PM (the set given by a primitive prediction model). Usefulness of an unexpected recommendation is given by u(RSi), which is either 0 indicating an unexpected yet useless recommendation, or u(RSi) = 1 indicating a useful unexpected item. Overall, a serendipity model is defined as the sum of all unexpected yet useful recommendations, divided by the total number of recommendations. To validate the defined metric, an experiment was conducted where a group of users were asked to listen to a list of song recommendations, then they were asked whether they knew the song before (expectedness measure), and whether they liked it or not (usefulness measure).

The importance of unexpectedness in the evaluation of recommender systems is also emphasized by [15]. Unexpectedness is measured as the deviation from the set of items predicted by a primitive prediction model (PPM), taking in consideration that an item is not considered unexpected if it is unrelated to the user's profile. This model of unexpectedness is refined by defining unexpectedness_r, which additionally takes in consideration the ranking of items in the recommendation list. To validate the suggested models, an empirical study has been conducted on television program recommendation, in which 2 surveys were conducted. The first survey aimed at collecting user's preferences, while the 2nd is used to know which items were considered serendipitous to the user. Three different bayesian networks techniques (based on user favorites, based on favorites and viewing habits, and based on keywords weighted by Graham's method) were used to analyze the data collected from the surveys. The best method that resulted in high unexpectedness and unexpectedness_r values was the last method, which scored the highest when it comes to both accuracy (precision and recall) and serendipity.

Another interesting practical user-study was conducted by [10] to identify strategies that could be used to induce serendipity in a content-based recommender system. Serendipity heuristics were integrated with the content-based approach to avoid over-specialized recommendations. Participants were asked to rate a set of paintings, which were arranged using the developed hybrid approach, in a museum hall. Three serendipity functions were used with different thresholds for randomness (5%, 10%, 15%). The results showed that better ratings were given when the threshold of randomness was higher.

Taking a rather different approach, [5] followed the rules of lateral thinking defined by De Bono [3] in developing an information retrieval system (Max) that aims at acquiring information in a serendipitous manner. Rather than programming serendipity, their aim was to "program for serendipity". Max was developed taking in consideration the lateral thinking concepts of not excluding on basis of immediate relevance, delaying judgements, and offering new entry points.

It is noteworthy that none of the previously mentioned studies started with actual user data regarding the user's actual experiences, but were rather based on assumptions. In SerenCast, a qualitative approach is given, which is believed to be more suitable for studying a rather illusive concept such as serendipity. SerenCast base the analysis entirely on actual user input from which the most important factors of serendipity are deduced. Moreover, context-aware aspects affecting the occurrence of serendipitous discoveries were addressed by none of the studies.

4. APPLICATION DESIGN AND STUDY SETUP

The identification of the context in which serendipity occurs during the consumption of digital content is twofold. One part entails the identification of factors which influence the occurrence of serendipity. This set of factors collectively define the "context" of a serendipitous encounter. A hypothetical context might be "a serendipitous encounter was in place when the user exhibited an open state-of-mind to accept new and unexpected recommendations". Hence, the user's openminded-ness could be identified as a factor, or as part of the context in which serendipity occurs. The other part of this process is the identification of measures that can be used to validate the occurrence of serendipity. Our experiment conducted using SerenCast aims at gathering a set of factors which are believed to contribute to the occurrence of serendipity. On the other hand, using the defined measures, the occurrence of serendipity would be identified. As the measures confirm the occurrence of serendipity, factors gathered at such point would be analyzed to show which of them could have lead to the occurrence of the serendipitous encounter.

This section defines how we came about identifying which factors would the analysis be based on, the identification of the measures for which serendipity is validated, defining working definitions for serendipity, and last how SerenCast was designed and used to conduct the user study.

4.1 Identification of Factors Leading to Serendipity

The factors focused on in our study of serendipity are contextual factors including the user's environment (e.g. location and time) and the user's mood which is studied partially through a post-experiment questionnaire. The basic assumption that leads us to focus our analysis on contextual factors rather than other factors is that the users' context affects greatly how they perceive the world.

The specific contextual information that are gathered using the experiment are:

- **Location** (longitude, latitude, city, country, district)
- **Time** (exact time and time of day)
- **Place identity** (e.g. university lab, workspace, house, etc...)
- **Identity of people nearby** (e.g. friends, family, partner, etc...)

- **The user's mood** (evaluated in post-experiment questionnaire)

Contextual data are gathered during the user's consumption of SerenCast's podcasts. Some of the factors are gathered during the experiment, and some separately after. Two methods are used for gathering contextual data:

- **Implicit elicitation**: during the consumption of the podcasts, the user's exact location, and time are automatically gathered and stored.
- **Explicit elicitation**: after the experiment, users are asked through a survey to validate the identities of the different locations they have been to during the experiment. They are also asked questions related to the identity of the people whom they have been with, and questions related to their general mood during the experiment.

Using the collected contextual information, the aim is to show if during the consumption of digital content, the user's location, this location's identity, the time of day, the company of a specific person, or the user's mood and state-of-mind could influence the occurrence of serendipity.

4.2 Identification of Serendipity Measures

Prior to the identification of the factors affecting the occurrence of serendipity, serendipity measures need to be defined. Such measures would be used to answer the question: did serendipity in fact occur in a given situation? For this purpose, four criteria are considered:

- **User's Satisfaction**: measures the degree to which the user enjoys (likes/dislikes) the presented content.
- **Unexpectedness**: measures the degree of the user's familiarity with the presented content, or in other words, how surprised the user is with the content presented.
- **Novelty**: a measure of how new the presented content is to the user. This aspect can be used to identify wether the content is new and enriching or boring.
- **Usefulness**: a measure of whether the user perceive the presented content as useful.

To identify the weights assigned to each measure, we consider the main definition of serendipity which indicates that a serendipitous encounter is a surprising one which should leave the user in a positive state. From this definition the following can be deduced:

- For an encounter to be serendipitous, it needs to be surprising. Thus the occurrence of serendipity is strictly conditional on unexpectedness.
- For an encounter to be serendipitous, it needs to be a positive experience. Thus user's satisfaction is another conditional aspect for serendipity to happen.

Measure	Criteria	Conditions
rUSat	User's Satisfaction	Cond_S1:0.66<=rUSat<= 1
rUnexp	Unexpectedness	Cond_S2:0<=rUnexp<=0.5
rNov	Novelty	Cond_S3:0.41<=rNov<=1
rUse	Usefulness	Cond_S4:0.33<= rUse<=1

Table 1: Defining Serendipity Conditions Based on Serendipity Measures

- Positivity can also be drawn from how novel and useful the content is. However, the occurrence of serendipity is not strictly conditional on the these two factors. A user can find surprising content satisfactory, which needs not be entirely new or useful. Thus, the weights given to novelty and usefulness are less than the weights for unexpectedness and user's satisfaction.

A rating system is designed to elicit the value the user assigns to each measure. During the experiment, the user is asked to rate each podcast on a scale from 1 to 7 with respect to: user satisfaction, unexpectedness, novelty and usefulness. The user's rating criteria with respect to the given measures are presented in Table 1.

4.3 Working Definitions of Serendipity

Based on the defined measures and conditions presented in Table 1, a number of working definitions for serendipity are given on which our analysis is based:

Definition 1: Strict Serendipity
"A serendipitous encounter is in place iff conditions Cond_S1, Cond_S2, Cond_S3, and Cond_S4 are true."

That is to say that a serendipitous encounter occurs if the user likes the content, finds the content unexpected, novel, and useful.

Definition 2: Mild Serendipity
"A serendipitous encounter is in place iff conditions Cond_S1, Cond_S2 and Cond_S3 are true."

This definition is less restrictive than definition 1. According to definition 2, serendipitous encounters occur if the user likes the content, finds the content unexpected and novel, but not necessarily useful.

Definition 3: Weak Serendipity
"A serendipitous encounter is in place iff conditions Cond_S1 and Cond_S2 are true."

Definition 3 is the least restrictive. The only criteria for the evaluation of serendipity is given by the user's satisfaction and the unexpectedness of the content.

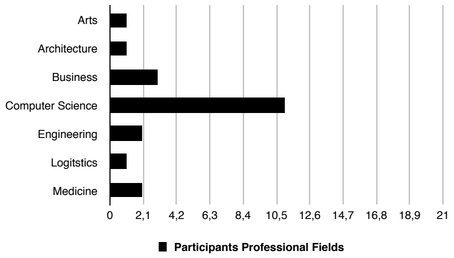
The definitions will be used in section 5 to evaluate the findings of the experiment

4.4 Content

50 podcast episodes were selected at random from online catalogues of different podcasts providers including NPR,

Scientific American, TED, and The Economist. Data about each episode is gathered for the analysis. A sample of the parameters gathered for each episode is shown in Appendix B. Durations of the episodes range from 1 to 10 minutes. A distribution of the podcasts' topics is presented in Figure 1.

Figure 1: Distribution of Podcasts Topics.



4.5 Experiment

The experiment starts with a profile creation step. Participants are asked to enter basic information such as email, age, gender, and occupation (Figure 2, 3).

Figure 2: Initial Screens from SerenCast Showing Profile Creation Step 1 .

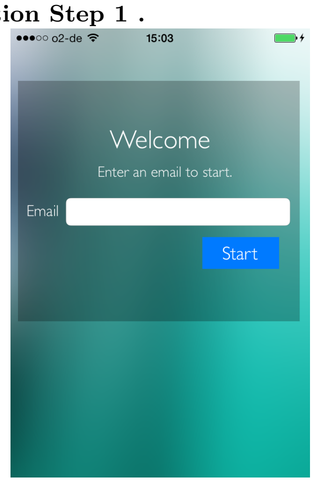
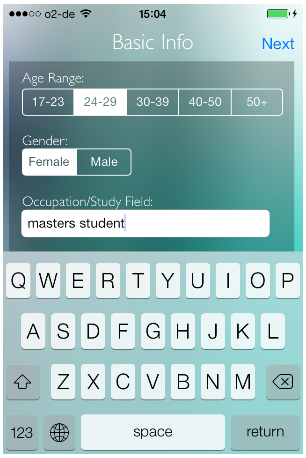
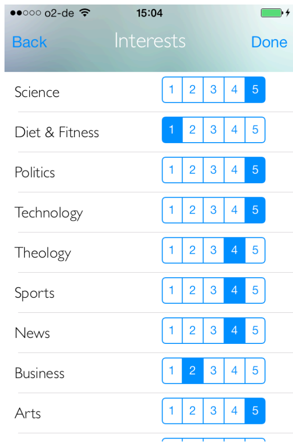


Figure 3: Profile Creation Step 2.



Later, participants are asked to rate a list of interests (Figure 4). This profile information does not influence the selection of podcasts presented to the user, however, they give an insight about the user's preferences which is a valuable input to the analysis process.

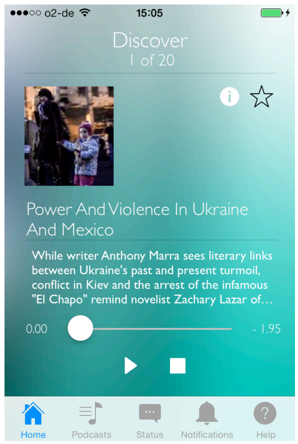
Figure 4: Interests Rating



The main activity consists of the podcasts listening and rating. All 50 podcasts episodes are fixed for all the participants. There are two different modes for listening to the episodes:

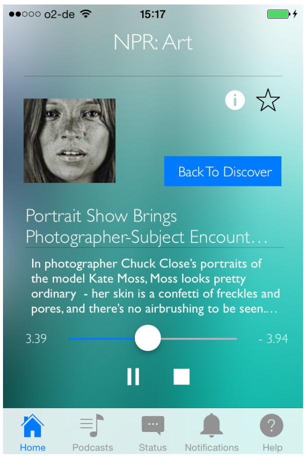
- **Discover mode** (Figure 5) in which 20 episodes are preselected for the participant and presented in a radio fashion. The 20 episodes are fixed for all participants. The selection is not based on participants interests, but is rather random. The aim of this mode is present diversified content.

Figure 5: SerenCast Discover Mode



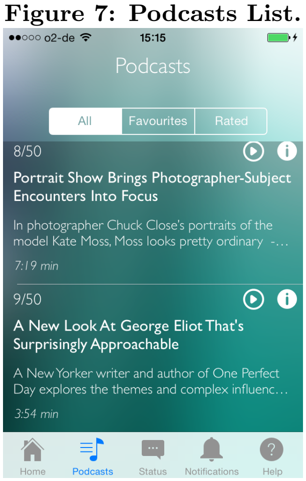
- **Podcasts List mode** (Figure 6) in which participants can select their desired episodes from a list containing all 50 episodes including the 20 episodes from Discover mode.

Figure 6: SerenCast Podcasts List Mode



The podcasts list (Figure 7) is added to track the interaction of participants with the application.

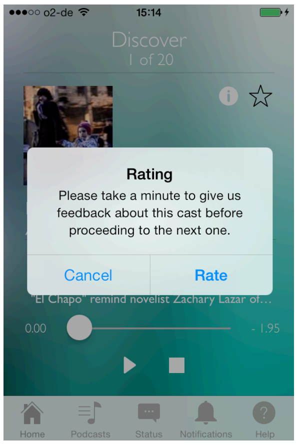
Figure 7: Podcasts List.



There is a constraint on the number of episodes a participant can listen to per day. A participant is only allowed to complete (listen to and rate) 2 different episodes from either modes. The reason for this restriction is that data needed to be gathered at different locations, different times of day, and perhaps different participant's moods. Thus, to achieve a variety of experiences per participant, this constraint is added.

After listening to each episode, participants are prompted to rate it (Figure 8). If the episode is not rated, it is not counted out of the 20 episodes needed to finish the experiment. Rating questions are formulated to reflect the four criteria described in section 4.2 which are user satisfaction as well as unexpectedness, novelty and usefulness of the podcasts (Figure 9).

Figure 8: Rating Prompt Message

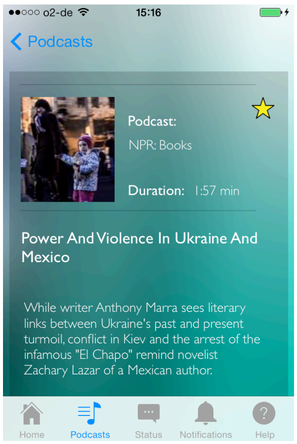


The experiment ends when 20 different podcasts are listened to and rated by each participant. Thus, an initial estimate of a minimum of 10 days is given to finish the experiment. However, the actual duration depends on the number of episodes participants rate per day, and the frequency of their participation.

Figure 9: Rating Questions

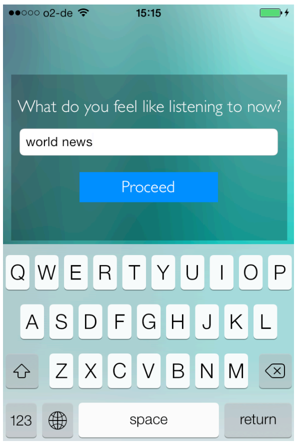
Participants interaction with the application, including which modes they choose to play the podcasts, is also considered as valuable input to the experiment. Therefore, they are given the option to add podcasts to favorites list. Participants' favorites also give an insight about their preferences. Detailed information (Figure 10) of each of the podcasts episodes is also presented to participants including detailed description, ratings, podcast source, duration and main actors.

Figure 10: Podcasts Detailed Information Screen.



To depict participants' moods and preferences during the experiment, they are presented with a status dialogue (Figure 11) inside the application that they are asked to use as frequently as possible to input what they would like to listen to at the moment.

Figure 11: Status Input Screen.



To encourage participation, notifications are presented to participants on daily basis to remind them to listen to the podcasts. Also, if they are not active for more than 3 days, another type of notification is presented to remind them to use the application.

Since usability factors such as learnability, memorability, and efficiency influence how users experience applications, participants were presented with tutorial screens on performing different task for the first time such as profile creation, doing the first rating, switching between modes, and status input. Moreover, to guide participants, a help section is provided in the application.

5. EVALUATION RESULTS

SerenCast is developed as a data gathering tool for collecting the defined serendipity factors. From the different types of digital content available for user consumption, we chose podcasts as a subject for being an area that is less frequently tackled by studies. Podcasts are easy to consume, and can be provided in a fashion similar to radio listening. Since the factors we are interested in are based on contextual information such as the user's location, we decided to develop SerenCast as a mobile application to be able to track the user in different context. The following subsection describes how SerenCast was used in a user study that aimed at eliciting users actual input and impressions during the consumption of the podcasts, which is the base of the later analysis.

5.1 Study Participants

20 participants voluntarily took part of the user study. The majority's age ranged between 22 to 29, 1 senior citizen, and 1 teenager, of which 9 are females and 11 are males (Figure 12). The geographic distribution of the participants

Figure 12: Participants Age Distribution.

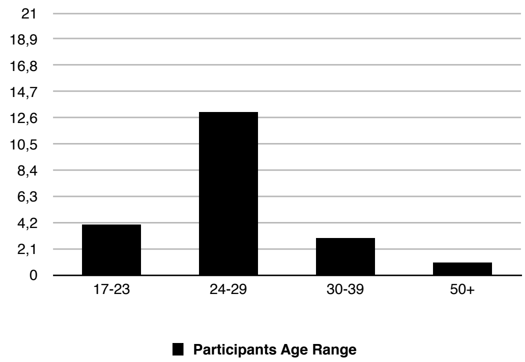
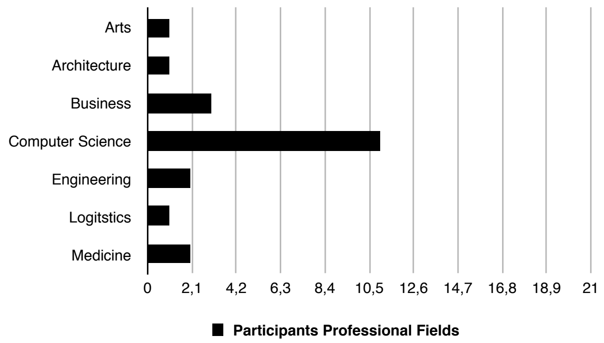


Figure 13: Participants Professional Fields Distribution.

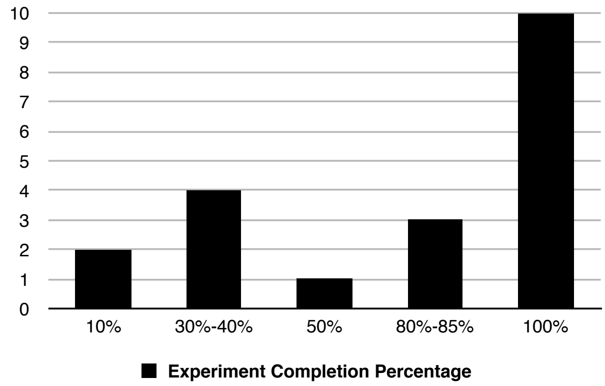


varied between Germany, Egypt, Canada, and Singapore. The majority of users are computer science and engineering professionals and graduate students (Figure 13). Other participants are professionals in the areas of arts, business, marketing, architecture, and medicine.

5.2 Data

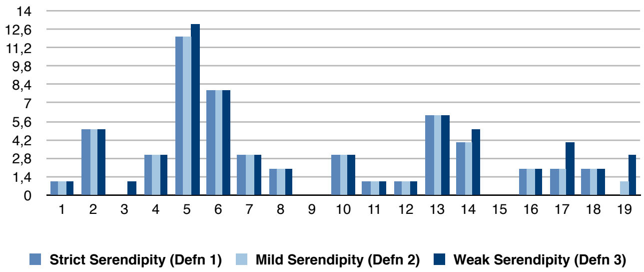
For a period of 4 and half weeks, the experiment's data has been gathered. 10 participants among the 19 total participants have completed the experiment (rated 20 or more podcasts). Figure 14 shows the experiment's completion percentage with respect to the number of participants.

Figure 14: Experiment completion percentage with respect to the number of participants.



A total of 278 ratings have been submitted during the experiment. Table 2 shows the number of serendipitous encounters which have occurred per participant with respect to the 3 definitions of serendipity defined in the previous section.

Figure 15: Serendipity encounters per participant with respect to serendipity definitions 1, 2 and 3.



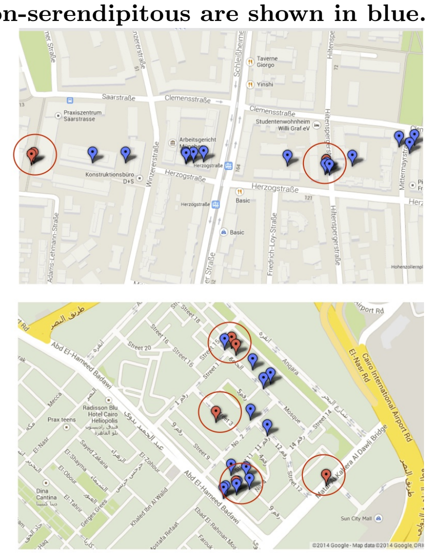
The data shows that 17 participants made 15 serendipitous encounters according to the Strict Serendipity definition or definition 1, 16 participants made 56 serendipitous encounters according to Mild Serendipity definition (definition 2), and 17 participants have made 63 serendipitous encounters according to Weak Serendipity definition (definition 3). Only 2 out 19 active participants have not made any serendipitous encounters, one of which have only rated2 episodes.
A Total Serendipity Score is computed from the 3 serendipity scores presented in table 2. The Total Serendipity Score is given by the following formula:
((3 Strict Serendipity Score) + (2 * Mild Serendipity Score) + (Weak Serendipity Score)) / (6 * Total Number of Ratings Per User)*

	Total Ratings	Strict Serendipity	Mild Serendipity	Weak Serendipity	Total Serendipity Score
1	10	0.1	0.1	0.1	0.1
2	20	0.25	0.25	0.25	0.25
3	2	0.5	0	0	0.08
4	20	0.15	0.15	0.15	0.15
5	21	0.6	0.6	0.6	0.6
6	20	0.4	0.4	0.4	0.4
7	20	0.15	0.15	0.15	0.15
8	7	0.3	0.3	0.3	0.3
9	2	0	0	0	0
10	22	0.1	0.1	0.1	0.1
11	6	0.2	0.2	0.2	0.2
12	17	0.06	0.06	0.06	0.06
13	20	0.3	0.3	0.3	0.3
14	7	0.7	0.6	0.6	0.6
15	21	0	0	0	0
16	8	0.25	0.25	0.25	0.25
17	16	0.25	0.125	0.125	0.1
18	19	0.1	0.1	0.1	0.1
19	20	0.15	0.05	0	0.04

Table 2: Serendipity Scores Per Participant

Interactions with the application such as the number of switches between Discover and Podcasts Lists modes, and the number of times each mode was used was gathered during the experiment. Locations of participants were tracked during the experiment. Figures 16 show different locations for a number of participants where they listened to and rated different episodes. Non serendipitous encounters are shown using blue map marker, while locations where serendipitous encounters occurred (based on definition 3) are shown using red map-markers. It's noteworthy that gathered location's data were insufficient to be used in our analysis.

Figure 16: Locations of serendipitous encounters mapped for 2 different participants shown in red, while non-serendipitous are shown in blue.



The following section shows how the gathered data was analyzed to deduce correlations between serendipity and the different examined factors.

5.3 Analysis

Our analysis focuses on showing dependencies between participants' total serendipity scores (Table 2, Section 5.2) and their ratings, preferences and interactions with the application. Correlations are based on the computation of the following 3 measures.

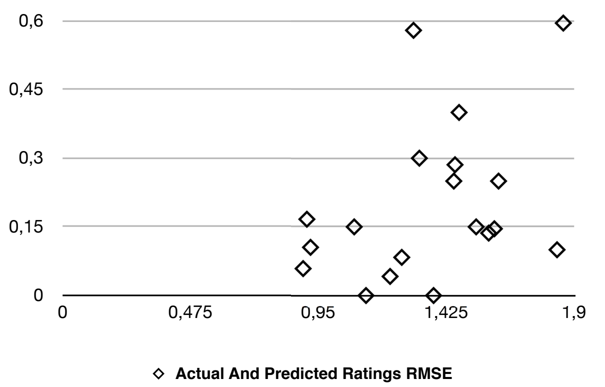
RMSE for Predicted and Actual Ratings

Root Mean Square Error (RMSE) for participants' actual podcasts ratings and their predicted ratings is computed as a measure for the occurrence of serendipity. A high RMSE indicates a high deviation between the predicted and actual ratings, meaning that the actual ratings are significantly higher than the predicted ratings. This can be interpreted as an indication of participants' satisfaction with content that was originally predicted to be unrelated to their preferences and therefore predicted to receive low ratings. Hence, RMSE can be considered as a measure for the occurrence of serendipity.

Predicted ratings are calculated per participant for each of the 50 available podcasts based on participants' rated profile interests (Table 2, Section 5.2). Profile interests contains 15 different categories which are used to categorize the podcasts. Each podcast is categorized with a vector of 15 values by respectively assigning 0.5, 0.3, and 0.2 to the top 3 categories related to the podcast's topic. Predicted rating is given by the sum product of participants' interests' ratings and the 50 podcasts' category vectors. RMSE of the computed predicted ratings and actual rating is computed are then computed.

Results presented in Figure 17 show a positive correlation of 0.393 between the computed RMSE and the total serendipity scores (section 5.2), which shows that our computed total serendipity scores is a good indicator to the occurrence of serendipity.

Figure 17: Correlation between RMSE and Total Serendipity Score Per User.



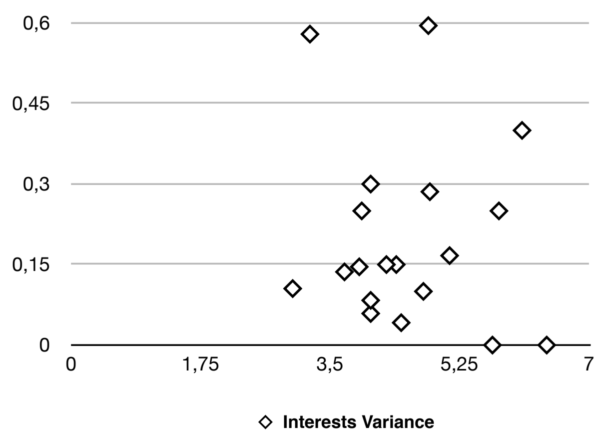
Variance of Participants' Interests

The variance of participants' interests is an indicator to the variety of topics that participants would appreciate or find interesting. A low variance means that participants rated their preferences' topics similarly, which is an indicator to the openness of participants to a wide range of interests. A participant with such profile is more likely to be satisfied with new and surprising content rather than a participant with a high variance of interests, who would be more inclined towards a specific number of topics. Hence, low variance would indicate the participant's appreciation of serendipitous encounters.

The interests' variance is calculated per participant as the sum of the variance and the average of the interests' ratings given per participant. The average rating is considered in the calculations since low variance with low ratings can not alone be considered a measure for openness to a variety of topics, since participants in such case would not be interested in most topics.

Figure 18 shows the results of a negative correlation of -0.1095 between the computed interests variance and the total serendipity scores per participant. The negative correlation shows that as the variance decreases (indicating openness to new topics), serendipity is more likely to happen.

Figure 18: Correlation between Interests Variance and Total Serendipity Scores.

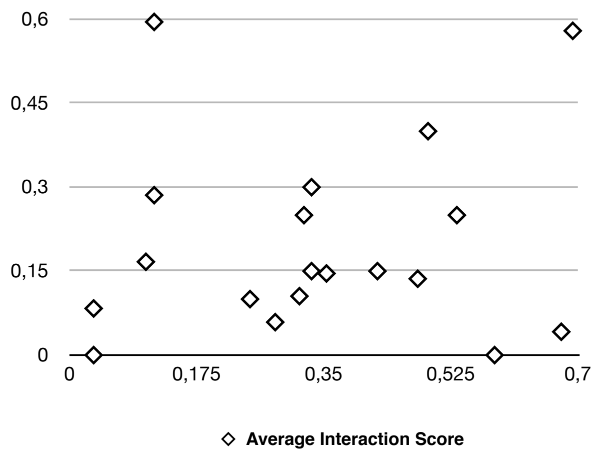


Interactions with SerenCast

As described in section 4.5, participants could interact with SerenCast using Discover mode or Podcasts List mode. The interactions with the application is used as a measure for the user's curiosity to listen to new topics. Curiosity metrics is then compared with total serendipity scores in order to show if there is any dependency between the two measure.

Interactions are computed as the number of switches between the 2 modes, as well as the number of interactions with each mode (counts of the number of times each mode was used per participant). The curiosity score per participant is computed as the average of the 3 counts. Correlations between the average curiosity scores and total serendipity scores per participant is computed. Figure 19 shows the resulting positive correlation of +0.117 which is an initial indication that appreciation of serendipitous encounters could be dependent on increased participant's curiosity levels. Our

Figure 19: Correlation between curiosity scores and total serendipity score.



analysis is partially based on a qualitative evaluation which is used to validate the conclusions we reached from the computed correlations, and which provides more insight about other factors that the experiment's data was insufficient to deduce. The following section describes the post-experiment questionnaire that we used as a basis for our qualitative evaluation.

5.4 Qualitative Evaluation

A post-experiment online questionnaire consisting of 30 questions was conducted to elicit participants' impressions and experiences with SerenCast. 15 of SerenCast's participants took part in the questionnaire, out of which 10 participants completed all sections.

In the beginning of the questionnaire, participants were debriefed about the content and purpose of each section. However the word "serendipity" as well as the purpose of the study were not mentioned in the questions' descriptions to avoid influencing participants' answers. Instead, we used terms such as "new and enriching", "surprising", "unfamiliar" to describe the experiences they were asked to evaluate.

The questionnaire consisted of 5 sections which were described to participants as follows:

1. **Basic Information**
2. **Impressions**: 6 questions were presented in this section to get participants' feedback about the different factors that could have influenced their experience with SerenCast.
3. **User Experience**: 8 questions were presented in this section to understand in details the user's experience with serendipitous encounters.
4. **Location and Time**: in this section, participants are asked to mention at least one location and time-of-day where they frequently used SerenCast.
5. **Feedback**: participants are asked to give their general feedback about their experience with SerenCast.

Factor	Average Rating (/5)
Location	2.7
Time-of-day	3.6
Mood	3.5
State-of-mind	3.6
Company of people	2.7
Surrounding Objects	1.8
Weather	1.9

Table 3: Average ratings for question "How do you think the following factors affected how you liked/disliked new and surprising podcasts?"

The questions cover different measurements that we were aiming to elicit. A list of all the questions with respect to which measurements they covered is presented in Appendix A.

Serendipity Factors

Location, time-of-day, mood, state-of-mind, company of people, surrounding objects, and weather are the factors whose influence on the occurrence of serendipity was investigated by the questions.

Table 3 summarizes the average ratings of participants when asked to rate on a scale from 1 to 5 (1 being the worst, 5 being the best) the factors which affected their enjoyment of podcasts which were new and surprising. Most of the participants believe that the time-of-day, their state-of-mind, and mood highly affected how they reacted to new and surprising content. Location and company of people were rated with an average of 2.7 out of 5 which indicates that both factors are not believed to be highly influential by most of the participants. While weather and surrounding objects gained the lowest ratings. The following were participants' comments when asked to elaborate on the ratings they gave for each factor:

> *"Places or surroundings in general didn't have any effect on the listening process, unless there was some distractions which would then hinder the process. However, my mood basically defined my eagerness to listen to any podcast not only surprising ones; i.e if I'm in a good mood I'd most probably listen to anything surprising or not, but If I'm not in a good mood I would certainly react differently to surprising pod-casts or refrain from listening."*

> *"I was [satisfied] when I listened while I'm all alone in my cosy living room."*

> *"Three conditions affected my experience differently: - Beginning of work: at office, between*

un_openness	Down to Earth, conventional, practical, traditional, plan, straightforward, prefer familiarity resistant to change.*	30%
openness	Imaginative, creative, intellectually curious, appreciative of art and sensitive to beauty, aware of my feelings, unconventional, individualistic. At an extreme may be prone to following fads.*	70%

Table 4: Openness to Experience Scores. Source for personality traits:Cambridge Personality Traits Test - Openness.

> *07:30 - 09:00. Normally i had a lot of plans to do on the day, and wanted to listen to something short and important. Appreciate something new in my expertise, but not something completely new. - End of the work: at the office, between 17:00 and 19:00. I was normally more relaxed in this situation, and ready for any kind of stories (short preferable). - Weekend or at home (everywhere, anytime): ready for everything."*

More questions were specifically asked to investigate factors affecting the occurrence of serendipity more individually. The influence of the sate-of-mind was investigated by directly asking participants if they believe that their state-of-mind affected how why reacted to surprising content by rating on a scale from 1 to 5. An average rating of 3.7 was given. 85% of participants said that their state-of-mind would make them more willing to accept surprising recommendations. One of the participants made the following comments:

> *"As I'm open to new topics or experiences, my state of mind certainly played a part in keeping me eager to listen to even subjects that are of no interest to me."*

To investigate openness as a personality trait and its effect on perceiving serendipity, participants were asked to select between 2 groups of personality traits (Table 4) that characterize openness and un-openness to new experiences. 70% of participants identified themselves as open to new experiences, while 30% identified themselves as un-open to new experiences. However, this did not coincide with the ratings given for valuing and enjoying serendipitous content, as some of the participants who chose the personality traits for un-openness did show high interest in new and surprising content.

Furthermore, the influence of usability factors was investigated by asking users to rate how much they think such aspects influenced how they perceived surprising content on scale of 1 to 5 (5 being the best). Table 5 shows the average ratings given to each usability factor. In general, all factors were given an average rating above 4 which indicates that the application's usability did not negatively influence

Learnability	4.3
Ease Of Use	4.7
Memorability	4.3
Efficiency	4.2
User Interface Design	4.2

Table 5: Usability Factors Scores.

the user's experience. Participants were asked to identify identities of places and times of day where and when they frequently listened to SerenCast's podcasts. This part of the questionnaire was added since location data gathered by the experiment was not sufficient to conclude the influence of location the occurrence of serendipity. Moreover, place identity (home, work, university...etc) as well as the type of activity carried out in a different time frame were added to the questionnaire. How participants identified their preferred places and times of day is shown receptively in figures 20 and 21.

Figure 20: Preferred places identities.

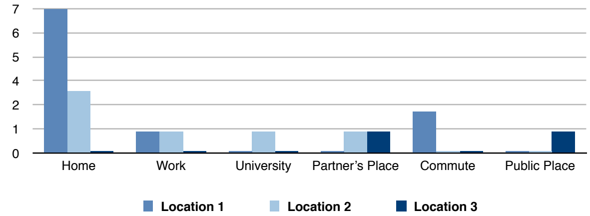
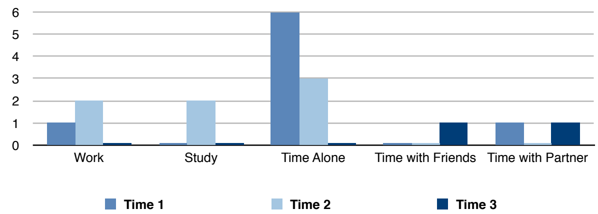


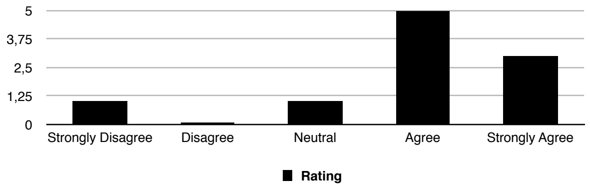
Figure 21: Preferred times of day.



Perceived Value of Serendipitous Encounters

The evaluation of the perceived value of serendipitous encounters and how it varied among different participants was elicited by asking participants whether or not they would rate new and surprising content as more enriching than content based on their personal profiles (Table 6). An average rating of 4.2 out of 5 was given by participants which shows a high inclination towards valuing more new and surprising content (Figure 22). To understand more about how participants value serendipitous encounters and which factors affect how they perceive them, participants were asked to rate on scale from 1 to 5 (5 being the best) if most of the podcasts they enjoyed and found new and enriching were also unfamiliar and surprising to them which resulted in an average rating of 3.6. This confirms that participants tend to value surprising content. Among the reasons that were given high ratings were openness to listen to new topics which do not necessarily match interests (4.2), and valuing surprising content rather than familiar content (3.6). Other reasons

Figure 22: Ratings for question "Would you generally rate new and surprising recommended content as more enriching than content based on your personal interests?".



High: Usefulness. Low: Unexpectedness, Novelty	Useful and familiar to my taste	5%
High: Novelty, Unexpectedness. Low: Usefulness	New, Enriching, and unexpected even if they were not very useful	30%
High: Usefulness, Unexpectedness. Low: Novelty	Both useful and unexpected	60%
High: Usefulness, Novelty. Low: Unexpectedness	Both useful and new but still relevant to my interests	5%

Table 6: Importance of Serendipity Metrics

included seeking to try new content out of boredom (2.8) or when having free time (3.4).

On the other hand, disliking unfamiliar and surprising content was given low average rating (1.6). Only 1 participant, who was also identified earlier from studying the experiment's data as having no serendipitous encounters, strongly agreed to disliking new and surprising content. The most important reasons were expecting to receive recommendations based on user profile and preferring to listening to topics of interest which were not presented, which could indicate that accuracy-based recommendations are still important.

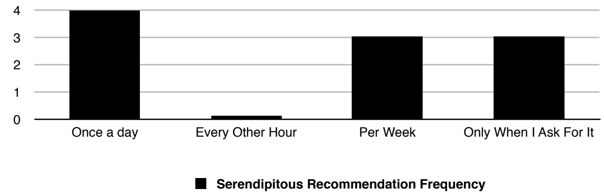
Importance of Serendipity Metrics

Participants were asked the reasons they liked the podcasts with respect to usefulness, novelty and unexpectedness. Unexpectedness and usefulness were the highest rated metrics by 60% of participants. 30% still thought unexpectedness is important if coupled with new and enriching (novelty) with low priority to usefulness. Only 10% preferred content that are useful (or useful and new) but also related to their interests and taste.

When Should Serendipitous Recommendations Be Offered?

To know how frequent participants think serendipitous content should be recommended, we asked them to select among: Once a day, every other hour, occasionally per week, only when I ask for it. Results are shown in figure 23. It's noteworthy that while feedback varied between once a day, or a week, or conditionally per request, none of the participants chose to have serendipitous recommendations made to them more than once per day or "every other hour" which can be an indicator that despite of valuing serendipitous encounters, too much serendipity could back fire.

Figure 23: Number of ratings given to the frequency of serendipitous recommendations.



Overall Feedback

This sections shows a number of general feedback comments about participants' experiences with SerenCast. The comments show general openness to this type of discovery and willingness to try the unfamiliar and surprising.

> *"Overall the experience was an enriching and useful one. I found it fun to return home and listen to something new that would benefit me. It made me think about subjects that I wouldn't normally come up with in my everyday life, and I started learning more about some of the subjects too."*

> *"My two favorite podcasts related to economics and technology, although they're not my favorite subjects. I liked the economics one because it helped me in my economics class and presentation as it talked generally about GDP and how Italy manipulated its calculation. The other one talked about Apple inc. and their next step in revolutionizing computers and tablets. One important factor was the way the content was presented in both. It was easy to follow and direct to the subject."*

> *"I found the whole experience entertaining and enriching. I enjoyed listening to the different podcasts of various subjects especially those far from my interest; that I've lost count of the number of podcasts that I've listened to, so maybe a counter as a reminder for the number of remaining podcasts. The app itself was comprehensible and very easy to use."*

> *"I really enjoyed the experiment as a whole. Some of the [podcasts] were new to me and so interesting....The one specifically I was not comfortable listening to was a political [podcast] showing only one point of view that [which showed a certain bias]. Still I listened carefully and I gained important information which made me accept listening to any ideas from anybody to reach my own."*

"I, generally, liked the idea of having a suggestion engine for podcasts, with both discovery and normal modes. And asking me what I feel like listening today is cool! However, I didn't like the content. Most of them were really boring. I usually tried to pick the shortest ones for that reason. I think, in general, that the shorter the podcasts are and the more they are to the point is the better."

6. DISCUSSION

The most significant conclusion from the results of our conducted user study was that participants who exhibited an increased level of curiosity and openness to new experiences are those who experienced serendipity the most. Both results from the post-experiment questionnaire and the analyzed experiment data are consistent with this conclusion.

However, the same conclusion could not be easily drawn for the influence of other factors. The importance of location as a factor both received low ratings in the questionnaire, and could not be computed from the experiment due to the sparseness of collected location data. However, analyzing the location and time section of the questionnaire did show that there were patterns to participants' preferred places and times-of-day during their consumptions of the podcasts, which indicates that the influence of such factors can not be excluded. It is possible that with enough data, such patterns could be analyzed and related to serendipity.

Among the factors that our experiment lacked and was later considered in the post-experiment questionnaire is the mood. The questionnaire's results show a positive indication to the importance of mood as elicited from participants' high ratings to mood related questions.

Perhaps the most important contribution of our study is that it provides the ground truth of serendipity. Our results show a high perceived value for serendipitous encounters and that serendipity can be associated with high levels of user's satisfaction and more enriched experiences. Consequently, our study takes the same stand as other related studies [11, 15, 17] that call for the inclusion of serendipity in recommender systems' techniques for the sake of escaping the limitation of filter bubbles and for the creation of more engaging and enriching user experiences.

It's noteworthy that our computed serendipity correlations are not highly significant. We root that down to the use of a limited dataset and simplistic computational models, and not to an implication that such correlations are weak or insignificant. Hence, our study lays a foundation for future work that can further enhance our method to use more comprehensive models and larger datasets to lead to better results.

7. CONCLUSION AND FUTURE WORK

Our study provide an initial indicator to the factors which trigger the occurrence of serendipity during the consumption of digital content. We developed SerenCast as a tool for the purpose of eliciting and interpreting such factors from participant's preferences, interactions and ratings.

To lay a foundation for a qualitative evaluation, we gave three serendipity definitions; Strict, Mild and Weak Serendipity, which were based on a combination of four measures: User satisfaction, unexpectedness, novelty, and usefulness. Based on the definitions, we developed a formula for a Total Serendipity score.

Through analyzing the gathered experiment's data, and despite of the limitation of the dataset, we were able to validate the occurrence of serendipitous encounters with most of the study's participants. We were also able to show sufficient correlations between exhibited user's curiosity, openness to new topics and total serendipity scores.

To validate the occurrence of serendipity and to further investigate factors that could not be elicited during the experiment, we conducted a post-experiment questionnaire. The questionnaire's results were consistent with the results of the experiment's data and gave further insights about how users value serendipitous experiences.

We believe our method can be further developed in a number of ways. The data gathered during our study could be used in further studies for quantitative evaluation. Furthermore, SerenCast could possibly be turned into a real serendipitous recommendation system.

Finally we believe that with the use of more comprehensive computational models and a larger dataset, SerenCast can be used in future experiments related to serendipity to achieve more significant results.

8. REFERENCES

[1] Gediminas Adomavicius and Alexander Tuzhilin. Context-aware recommender systems. *Proceedings of the 2008 ACM Conference on Recommender Systems*, 2008.

[2] C. Anderson. The long tail: Why the future of business is selling less of more. *Hyperion Books*, 2006.

[3] E. De Bon. Lateral thinking. *Penguin Books.*, 1990.

[4] G. Lausen C.-N. Ziegler and S.-T. Lars. Taxonomy-driven computation of product recommendations. *ACM CIKM.*, 2004.

[5] J. Campos and A. Dias de Figueiredo. Searching the unsearchable: Inducing serendipitous insights. *Weber, R. and Gresse, C. (2001), Proceedings of the Workshop Program at the Fourth International Conference on Case-Based Reasoning*, 2001.

[6] A. Dias de Figueiredo and Jose Campos. The serendipity equations. *Weber, R. and Gresse, C. (2001), Proceedings of the Workshop Program at the Fourth International Conference on Case-Based Reasoning*, 2001.

[7] Lior Rokach Francesco Ricci and Bracha Shapira. Introduction to recommender systems handbook. *Springer-Verlag New York, Inc.*, 2010.

[8] Shani G. and Gunawardana A. Evaluating recommendation systems. *Technical report, No. MSR-TR- 2009-159.*, 2009.

[9] D. Quercia J. An and J. Crowcroft. Why individuals seek diverse opinions. *Technical Report of the University of Cambridge*, 2012.

[10] Giovanni Semeraro Leo Iaquinta Marco de Gemmis, Pasquale Lops and Piero Molino. Can a recommender system induce serendipitous encounters? *In Tech(check)*, 2010.

[11] David Hauger Markus Schedl and Dominik Schnitzer. A model for serendipitous music retrieval. *Proceedings of the 2Nd Workshop on Context-awareness in Retrieval and Recommendation*, 2012.

[12] Carla Delgado-Battenfeld Mouzhi Ge and Dietmar Jannach. Beyond accuracy: Evaluating recommender systems by coverage and serendipity. *Proceedings of the fourth ACM conference on Recommender systems(RecSys 10)*, 2010.

[13] E. Pariser. The filter bubble: What the internet is hiding from you. *Penguin Group USA*, 2011.

[14] Kiyohiro Obara Takayuki Akiyama and Masaaki Tanizaki. Proposal and evaluation of serendipitous recommendation method using general unexpectedness. *Proc. Workshop on the Practical Use of Recommender Systems Algorithms and Technologies*, 2010.

[15] Koichiro Mori Tomoko Murakami and Ryohei Orihara. Metrics for evaluating the serendipity of recommendation lists. *Proceedings of the 2007 conference on New frontiers in artificial intelligence(JSAI 07)*, 2007.

[16] E. G. Toms. Serendipitous information retrieval. *ACM*, 2000.

[17] D. Quercia Yuan Cao Zhang, D. Seaghdha and T. Jambor. Auralist: Introducing serendipity into music recommendation. *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, 2011.

APPENDIX
A. POST-EXPERIMENT QUESTIONNAIRE

Measurement	Question
Serendipity Factors (Location, time-of-day, state-of-mind, mood, company of people, weather)	- How did the following factors influence how you liked/disliked new and surprising podcasts? - Please add any details you have about your answers in the previous question. For example, you can mention specific places, time of day, or mood that have influenced the experience described in the question.
Perceived value of serendipitous encounters	- Would you generally rate new and surprising recommended content as more enriching than content based on your personal interests? - Most of the podcasts I liked and found new and enriching were unfamiliar and surprising to me./In your opinion, how did the following factors influence the experience described in the previous question?... - I did not like most of the podcasts that were unfamiliar and surprising/ How did the following factors in your opinion influence the experience described in the previous question?...
Influence of State-of-mind on occurrence of serendipity	- Do you believe your state of mind had an effect on how you reacted to surprising podcasts? - Did your state of mind made you more or less willing to explore new content despite of its relevance? - Which of the following best describes your character?
Influence of Usability on occurrence of Serendipity	- How did the following usability factors influence your experience with SerenCast?
Frequency of serendipitous recommendations	- If you liked having surprising recommendations, how often do you think such recommendations should be suggested to you?
Importance of Serendipity Metrics (Unexpectedness, novelty, usefulness)	I mostly liked the podcasts which were: Useful and familiar to my taste/ New, enriching, and unexpected even if they were not very useful/ Both useful and unexpected/ Both useful and new but still relevant to my interests
Influence of Location and Place Identity	- Please locate where you listened to most of the podcasts - This location is: home, work, university, partner's place, commute, other.
Influence of Time-of-day	- Please mention a specific time when you usually listened to the podcasts everyday - What do you usually do at that time of the day? Work, Study, Leisure time alone, Spend time with friends, Spend time with my partner, Other.

B. PODCASTS SAMPLE CONTENT

Title	'Portrait Show Brings Photographer-Subject Encounters Into Focus''
Description	'In photographer Chuck Close's portraits of the model Kate Moss, Moss looks pretty ordinary - her skin is a confetti of freckles and pores, and there's no airbrushing to be seen. Moss Trusted Close' are, but as an exhibit at the Washington's Phillips Collection demonstrates that isn't always the case.''
Podcast Provider	NPR
Podcast Show	NPR: Art and Life
Release Date	December 26, 2013 3:16 am ET
URL	http://www.npr.org/2013/12/26/255447261/portrait-show-brings-photographer-subject-encounters-into-focus
Image URL	http://media.npr.org/assets/img/2013/12/19/chuck-close_custom-4df9b80610bc947031d1a5b770fbd38241f5c142-s2-c85.jpg
Category	Art
Sub-Category 1	Photography
Sub-Category 2	Culture
Author	Susan Stamberg
Number of comments	21
Sample Comment	'Stripped of protective make up, Kate Moss bore, at the time, the face of a younger woman who never saw a need to protect it and had lived a very fast life that caught up to her. I wish NPR post the Freda Kahlo picture and included the portrait of the homeless man separately, as well. I have seen plenty of Kate Moss.''