GameNet and GameSage: Videogame Discovery as Design Insight

[ANONYMIZED]

.

.

.

ABSTRACT

The immense proliferation of videogames over the course of recent decades has yielded a discoverability problem that has largely been unaddressed. Though this problem affects all videogame stakeholders, we limit our concerns herein to the particular context of game designers seeking prior work that could inform their own ideas or works in progress. Specifically, we present a tool suite that solicits text about a user's idea for a game to generate an explorable listing of the existing games most related to that abstract idea. From a study in which 182 novice game designers used these tools to find games related to their own, we observe a demonstrated utility exceeding that of the current state of the art, which is the coordinated usage of assorted web resources. More broadly, this paper provides the first articulation of videogame discovery as an emerging application area.

Author Keywords

Games; Videogame Discovery; Visualization; Exploratory Search; Design Support.

ACM Classification Keywords

H.5.0. Information Interfaces and Presentation: General

INTRODUCTION

Attendant to the immense proliferation of videogames over the course of recent decades is a *discoverability problem* that has gone largely unaddressed. Though it launched only seven years ago, there are over 400,000 games available on the iTunes App Store and 376 new titles are being added each day [68]. Google Play features a comparable number of games [93, 88], and there remains a considerable array of titles published for consoles and personal computers. Regarding the latter, more games were released on the digital-distribution platform Steam [61] in the first twenty weeks of 2014 than in the entirety of 2013 [75]. These staggering numbers point to a simple fact: there are so many games in existence that players, scholars, critics, and designers need dedicated tools if they are to successfully encounter titles related to their particular discovery needs. There are only a handful of such tools, however, and so a discoverability problem persists.

A major symptom of this problem is a strong *bandwagon effect* [3] that has yielded a blockbuster-driven games industry. Because a lack of resources makes games largely undiscoverable, the best-marketed titles tend to reach the largest audiences, which in turn persuades industry prime movers that only games like those *can* reach wide audiences, which causes more such games to be well-marketed, and so forth. At its most extreme, this feedback loop produces endless instantiations of the same formulaic series, and so we find that the *Call of Duty* franchise, for instance, has generated 26 titles in twelve years [31]. This notion is well-understood in the games community and has been decried by luminaries such as Warren Spector [60] and Simon Carless [6].

Of course, other media also have discoverability problems (and show bandwagon effects [18, 26, 49]), but in games the problem is compounded by the high dimensionality of their ontologies as media artifacts. Games incorporate text, sound, moving images, and nontrivial interaction, each axis of which may subsume several areas of concern. The only way to alleviate this discoverability problem is to develop dedicated tools for game discovery, but such tools must index games according to their composition along *all* these dimensions, so that a user can operate over these concerns while exploring the system. We contend that this represents a harder task than building discovery systems for other media that are lower-dimensional in their ontological makeup, which means that we cannot simply adapt to this domain successful discovery tools that were built for other domains.

While all videogame stakeholders are susceptible to the discoverability problem we have so far outlined, let us consider the particular case of how it affects game designers. Like any designer, a person who makes games may seek out prior work that could inform his or her own process, or more specifically the development of a new design. But lacking tools that meet this need, especially ones that may facilitate the discovery of novel or historically notable games, we find game designers whose practices are clearly unaware of related games, and often of the rich history of the medium. This is certainly one respect in which videogames show their immaturity relative to more established media. In this paper, we will limit our focus to this design-centric use case for a game-discovery system as a *creativity support tool* [85]. Before doing so, however, we would like to further articulate what exactly we mean by 'game discovery'.

First, game discovery is not game recommendation. Recommender systems (sometimes called *recommenders*) are often part of larger commercial applications such as online retailers [41, 72], and a prototypical task of these systems is product recommendation [67]. In contrast, *discovery tools* (also called *exploratory search systems* [56, 36]) promote user learning above user purchasing. While in the recommendation task there is often an explicit notion of the correctness or accuracy of a recommendation [74, 84], the parallel concern in discovery is the *usefulness* of an item. (Of course, this distinction has made evaluating discovery tools a much trickier issue [95].) Finally, while recommenders are susceptible to a *popularity bias* by which a small proportion of the item space is recommended exponentially more often [35, 50, 103, 8]—a phenomenon that may actually aggravate the bandwagon effect we have just outlined—discovery tools aim to provide diverse, serendipitous offerings [96, 70].[1]

Neither then should game discovery be considered as merely game information retrieval, primarily for the reason that a game-discovery tool will index *games*, not information about them. Moreover, in game discovery, getting back results is not a resolution—the offerings provided by a discovery tool are meant to be analyzed and explored [96]. Finally, game discovery should always be user-centric, while in information retrieval ancillary concerns (namely algorithmic nuances) tend to take center stage, a truth that bears out in the offline, batch-style evaluation methods that have prevailed in the field for decades [95].

So, let us distill what distinguishes discovery tools from related applications into a succinct specification for what a good game-discovery tool should do: it should index games along all dimensions of their high-dimensional ontologies, afford queries that may operate over the same dimensions, and present diverse, serendipitous, explorable offerings.

In this paper, we present a system for videogame discovery that we hope meets these criteria; it is specifically one that may allow game designers to find related prior work that could inform the development of an idea for a game, or a game that is already in development. This system comprises two interconnected components: GameNet is a network of nearly 12,000 games that are linked to one another according to how related they are, and GameSage is an interface to GameNet that solicits a description of an idea for a game (or a game that already exists) and uses natural language processing (NLP) to generate an explorable listing in GameNet of the existing games most related to the user's idea. Influenced by the notion of *task-based evaluation* of exploratory search systems [40], in which a system is evaluated for its adequacy in the natural context of its user task, we conducted a user experiment in which novice game designers sought out games related to their own using GameNet and GameSage and also a baseline method in which they were permitted to use any resources available online. We chose this baseline method because we believe it represents the (lack of a) state of the art in game discovery today. The primary variables of interest to our analysis are the number of games discovered using both methods, the diversity of games discovered (*i.e.*, percentage that were unique), and the proportion of discovered games that were unfamiliar to users prior to the experiment. From these, we constructed three major hypotheses, which we return to below in the discussion of our results:

H1 Participants will discover more related games using GameNet and GameSage.

H2 Participants will discover a greater diversity of games using GameNet and GameSage.

H3 Participants will discover a greater proportion of unfamiliar games using GameNet and GameSage.

Our main contribution to HCI is threefold: we present the first articulation of videogame discovery as an emerging application area; we describe what to our knowledge is only the second game-discovery system to have appeared in the literature; we carry out the first controlled user study of such a system. Further, we believe that the interaction method of the system we present—in which a fully formed description of an idea for an artifact is matched against a database of real artifacts—could be adapted to different media (*e.g.*, film, music, literature) or to altogether different domains (*e.g.*, a system that takes an abstract for a prospective paper and matches it against the existing literature). It is our hope that this work will encourage the development of subsequent tools that may help to alleviate the videogame discoverability problem that we have outlined above.

A TERMINOLOGICAL NOTE

We acknowledge that the term 'discovery tool' has been used with other connotations in fields such as library science (an electronic resource for exploring library holdings [94, 24]), bioinformatics (a tool for automatic drug discovery [11] or domain-specific knowledge discovery [15]), and data mining (automated knowledge discovery [110]). While 'videogame exploratory search system' would perhaps better follow existing academic terminology, we prefer 'videogame discovery tool' for its relative brevity and especially for the reason that it is already a term of use among games researchers [47, 38] and industry practitioners [6, 23, 69].

RELATED WORK

Videogame Discovery

Videogame discovery is an emerging application area for which very little work has yet been done. In what would appear to be the first published effort in this domain, Lee et al. present Vizmo, a discovery tool that indexes games by visual style and mood [47]. Influenced by earlier work in constructing browsers for other media [97, 37, 14, 58], the tool was built using a *faceted-metadata* approach [22]. Specifically, Vizmo is underpinned by a database of games that have been manually annotated for their visual style and mood using a subset of a larger videogame metadata schema that was created by the same group (see [48, 46]). Users can browse the games in this database by setting filters for different combinations of visual style and mood, and the results are displayed

[1] We do note that work on recommendation systems has begun to also value such properties [29, 84, 1].

in a chronologically oriented stacked area chart. From an expert evaluation conducted with nineteen game professionals, Lee et al. found Vizmo to be an aesthetically pleasing tool with potential use for game discovery along aesthetic or historical concerns [47]. Though promising, Vizmo is an early prototype that currently houses only 604 titles, many of which are platform variants of the same game (*e.g.*, *Rampart* [82] as released on the Sega Genesis and Super Nintendo Entertainment System platforms respectively). That so few games are yet included is not surprising given that each must be manually annotated prior to being indexable by the tool. Until more games are added, it appears that Vizmo will not offer extensive practical use for game discovery (a point we return to later in discussing the results of our experiment).

If Vizmo may be thought of as taking a top-down approach to game discovery, one in which humans handcraft indexable representations of games (which a discovery tool may then operate over), recent work in game discovery by Ryan et al. takes a decidedly bottom-up approach. Specifically, they submit large quantities of text describing games to a bottom-up technique from statistical NLP called *latent semantic analysis* [21]. By this method, they semiautomatically derived indexable representations for nearly 12,000 games, which they used to build a suite of visualizations and game-discovery tools, namely GameNet and GameSage [77, 78, 79, 80]. Because we describe their model and these tools in detail below, here we will only discuss the apparent trade-offs between these two approaches. Employing a semiautomated method, Ryan et al. could quickly build full-fledged discovery tools comprising several thousand games. But relying on a statistical procedure to derive indexable game representations, they were left with tools that operate over rather opaque notions of what games are made of—as they discuss in [80], their tools reason about games in terms of arcane statistical features of their textual descriptions. Vizmo, on the other hand, reasons over games purely in terms of human-crafted specifications of them. As such, it will always be clear how Vizmo indexes games, because its indexable representations are simply human annotations.[2] So while Vizmo's game indexing is perhaps more reliable and certainly more transparent, GameNet appears to offer reliable-enough indexing for twenty times as many games [80].

Videogame Recommender Systems

As we have argued above, game discovery is different than game recommendation. However, the tasks are similar enough that we might still look to game recommenders for promising high-level approaches that could inform new work in game discovery. Here, we again find a dearth of systems: to our knowledge, there are only four reported instances of videogame recommender systems. In [86], player data from the game-distribution platform Steam [61] is submitted to multiple machine-learning techniques, namely *archetypal analysis* [13], to build a series of recommenders. The authors of [7] present a domain-specific system that uses case-based reasoning to recommend rehabilitation games to patients. The recommender system of [59] is fueled by a model that was trained by submitting a corpus of user-submitted game reviews to *information-theoretic co-clustering* [19]. Finally, in [78], a variant of the latent semantic analysis model underpinning GameNet fuels a recommender that is used to test the intuitive notion that people tend to like related games.

[2]One perhaps unexpected advantage of *unsupervised* reasoning, however, is that it may reveal unexpected associations among games that counter problematic human presuppositions [80].

HCI and Crowdsourcing

The discovery tools central to our current project, GameNet and GameSage, rely on indexable game representations that were derived from encyclopedic game descriptions found on Wikipedia. In the sense that this collection of source text was authored by thousands of individuals, this paper may be thought of as extending the recent surge of work in the HCI community on crowdsourcing, which has leveraged the crowd for data analysis [39, 44], system prototyping [45], media production [83, 25], and many other applications (*e.g.*, [99, 98, 65]).

NLP on Videogame Text Corpora

An emerging bottom-up approach to *game studies*, the field in which videogames are critically studied, submits large text corpora about games to techniques from machine learning and NLP, including *cluster analysis* [100, 102], *sentiment analysis* [71, 9], and other approaches [101, 33]. Other related work, concerning the harvesting of playability heuristics from large-scale text corpora, has appeared in HCI venues [109, 104, 106, 107, 105, 108]. In [80], Ryan et al. provide a detailed overview of these projects that connects them to the specific NLP approach underpinning GameNet and GameSage, which we describe in the next section.

GAMENET AND GAMESAGE

GameNet and GameSage are components of a videogame-discovery system that solicits text concerning a user's idea for a game and, using natural language processing, generates an interactive listing of the existing games most related to that abstract idea. Both tools are underpinned by a *latent semantic analysis* model trained on Wikipedia articles describing games; the tools are hosted online as freely available web apps (see link below). In this section, we explain latent semantic analysis and the construction of the model underpinning these discovery tools before proceeding to describe the tools themselves.

Latent Semantic Analysis

Latent semantic analysis (LSA) is an NLP technique by which words are attributed vectorial semantic representations according to their contextual distributions across a large collection of text [17, 21, 42, 43].[3] From such a corpus, a *co-occurrence matrix* of its words and documents is built; this matrix specifies which words occurred in which documents (and thereby which documents words occurred in). Crucially, the columns and rows in this matrix can be thought of as vectors that represent the meanings, in an approximate sense, of the words and documents that they correspond to—this is called a *vector space model* of semantics [81, 92]. LSA

[3]LSA is sometimes called *latent semantic indexing*.

is an example of such a model, but its hallmark is that it reduces the dimensionality of these vectors by a matrix factorization algorithm called *singular value decomposition* [32]. Remarkably, doing this allows the model to infer semantic associations that are not encoded in the full co-occurrence matrix (*i.e.*, *latent* semantic associations). For instance, an LSA model may capture that two terms that do not appear together in any document—perhaps dialectal variants that denote the same thing, like 'gas' and 'petrol'—are in fact highly semantically related [21]. By the same token, LSA may infer the semantic relatedness of two documents that have no terms in common. This ability to learn global associations from local co-occurrences is the achievement of LSA and what led to it becoming one of the major NLP techniques of the last twenty years. Having an LSA model, one can easily calculate how semantically related any of its documents are by taking the cosine between their LSA vectors—this measure is called *cosine similarity* [92]. In corpora in which each document pertains to a specific individual concept, such as a corpus of encyclopedia articles, these relatedness scores can reasonably be utilized as a measure of the relatedness of the *concepts themselves*. As we illustrate below, GameNet and GameSage crucially rely on this notion.

Videogame LSA Model

The LSA model underpinning GameNet and GameSage was trained by submitting 11,829 Wikipedia articles about videogames to latent semantic analysis [80]. Each of these articles described an individual game and, at the time of extraction (May 2014), was at least 250 words in length and not marked by Wikipedia as being a stub. Ryan et al. preprocessed the corpus (which totals some 14.5M words) using a standard pipeline and, by the term-comparison technique of [5], settled on a reduced model dimensionality of 207; these procedures are described in depth in [80]. Using cosine similarity, described above, this model affords the direct calculation of how related any two of its 11,829 games are—this is what fuels GameNet and GameSage. In [78], Ryan et al. argue that this computation operationalizes an intuitive notion of how games may be alike, which they call *game relatedness*. Beyond the tools under consideration here, the group has built a suite of interactive visualizations of this model, which they briefly introduce in [77].

GameNet

GameNet is a tool for game discovery in the form of a network in which related games are linked; it houses entries for the 11,829 games that are included in the LSA model we have just described. Each game's entry includes links to GameNet entries for other games that are related to that game, as well as to gameplay videos and other informative sources found elsewhere on the web. Figure 1 shows excerpts from the GameNet homepage and its entry for the Nintendo Entertainment System game *Wall Street Kid* [87].

Usage

At the GameNet home page, the user indicates which game he or she wishes to start at and is brought to that game's entry. Here, in a header, the game's title and year of release are prominent, as well as links to the game's Wikipedia article

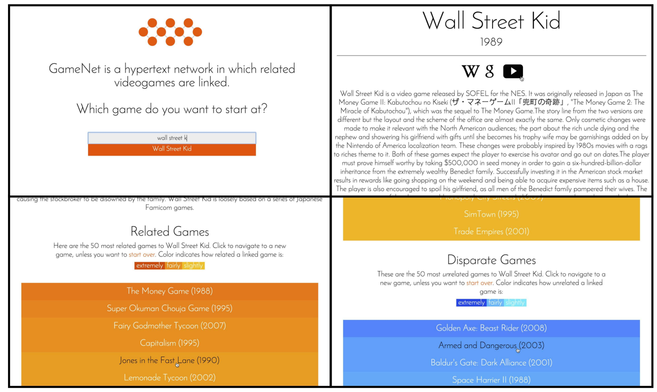

Figure 1. GameNet entry for *Wall Street Kid*.

and Google Images and YouTube search results for the game found using autogenerated queries; included as well is a summary of the game that was extracted from Wikipedia. Below these elements is the core of the entry, which is a color-coded listing of the fifty most related games to the game at hand (warmer colors indicate greater relatedness). As alluded to above, GameNet judges how related any two games are by taking the cosine between their Wikipedia articles' LSA vectors. To promote exploration, the related games are stylized as hyperlinks to their own GameNet entries. In this way, the user can navigate the medium of digital games (inasmuch as it is represented by the several thousand included in the system) by following links between associated games at each jump. Finally, below the listing of related games is a listing of the most *unrelated* games to the game at hand. These links can potentially serve as portals to corners of the medium that were previously unknown to the user.

Evaluation

In [80], Ryan et al. reported on an expert evaluation of GameNet as a tool for game scholars to discover games related to their research topics. Specifically, six published game scholars used the tool for fifteen minutes (starting from a game central to a recent research project) and answered a series of questions about the experience. The respondents indicated that GameNet is especially useful for the scholar who wishes to explore a relatively unfamiliar area of games ("[it was] useful to get broad connections in a space I wasn't as familiar in"), but that it may also be used to discover unforeseen cases related to topics that have already been thoroughly researched ("[I came upon] one game I had not thought about much since childhood and seeing it described now made me realize that it had some interesting features relevant to my research") [80].

GameSage

GameSage is an interface to GameNet that solicits free-text input describing an *idea* for a game (or alternatively a game that already exists) and lists the existing games that are most related to that idea. This is done by utilizing the notion in LSA of *folding in*, whereby a new document that was not used during model training is fitted with a representation in the semantic space derived by the model. By treating the user's input text (which specifies his or her game idea) as a

Figure 2. Excerpts from a GameSage session; the user describes the indie-game *Ice-Bound* and obtains a list of related games.

corpus document (on par with the videogame Wikipedia articles used to train the videogame LSA model) and folding it into the model, the tool is able to derive an LSA vector for the idea. From here, GameSage determines which existing games (from among GameNet's 11,829 games) are most related to the game idea by using cosine similarity, just as GameNet already does in determining which existing games are most related.

Usage

At the GameSage home page, the system is personified as a sage character who proceeds to ask the user eight questions about his or her game idea, each pertaining to a particular aspect of the (prospective) game. The eight aspects that the sage asks about are *genre*, *setting*, *plot*, *gameplay*, *mechanics*, *aesthetics*, *unique aspects*, and *miscellaneous*. After the final question is answered, the system concatenates the responses and preprocesses this text using the same procedure that was enacted during corpus preprocessing (see [80]).[4] From here, the preprocessed text is attributed an LSA vector by folding it into the model underpinning GameNet, whose 11,829 games are then ranked (using cosine similarity) according to how related they are to the user's idea. Finally, GameSage makes a request to GameNet to generate an entry for the game idea, to which the user is then taken. Here, the idea's most related and unrelated games are listed just as they are for all other GameNet entries. This entire process of folding in, from idea submission through generation of a GameNet entry, takes only a few seconds.

Evaluation

GameSage did not yet exist at the time of GameNet's expert evaluation, reported in [80], though it was briefly introduced in a conference demonstration [79]. As such, the current paper represents its formal presentation and evaluation, the latter of which we report below.

Example Session

In this section, we will work through an example session with the tool in which a user describes the in-development indie game *Ice-Bound* [73] using text from the game's official website; Figure 2 shows excerpts from this session. At the GameSage home page (shown in the first frame of Figure 2), the user clicks the sage character and proceeds to answer his eight questions (responses to the questions on *setting* and *gameplay* are shown in Figure 2), clicking the character to submit each:

Genre First, tell me a bit about the genre of this game.

Ice-Bound was born from our love of interactive narratives and our frustration that hand-authored branching paths are still state of the art.

Setting Now tell me about the setting that this game takes place in.

the legendary Carina Station, an abandoned polar base sinking into ice, layers of frozen history stretching down to unmapped depths.

Plot Next, tell me about the narrative or plot of this game.

A nested-doll narrative that reveals more depth the more it's explored. A cunning publisher commissions an AI simulacrum of a long-dead author, to finish his famously incomplete masterpiece. The AI is neurologically identical to its human predecessor, but as a constructed sentience, has no human rights. And a curious book begins appearing, stacked on street corners and left on busses: a real paper book, filled with glitched transmissions, contradictory drafts, distorted photos and vicious secrets. It's a book only one person was ever meant to see.

Gameplay Okay, now tell me about its gameplay.

Gameplay driven by cutting-edge interactive story research. By scanning pages from a companion printed book, the Ice-Bound Compendium, readers reveal information to a digital collaborator, who develops a sense of what themes the reader thinks are important to the overall story. The story will begin to emphasize the same themes the reader consistently focuses on.

Mechanics Next, tell me some of the specific actions the player may take in this game.

Dynamic conversation with an intriguing, reactive character.

Aesthetics Now, tell me about the visual and aesthetic style of this game.

Text that shimmers and changes as you read it. A gorgeous full-color print book. A nested, recursive story inspired by writers like Borges and Nabokov and books like House of Leaves.

[4] As such, multiple questions are asked of the user only so that he or she will adequately describe the game; *i.e.*, the tool does not reason differently about the individual aspects the user describes.

Unique aspects Now, what makes this game unique?

Unique merger of physical and digital storytelling. Thousands of sculptable permutations of each story.

Miscellaneous Lastly, tell me anything else about this game that you'd like to add.

The game's constructive narrative system is part of our ongoing PhD research. An academic paper describing our combinatorial narrative system is available to those interested in more technical details.

After the final question is answered, the sage squints in thought for a moment before the user is delivered to a generated GameNet entry for the game (shown in the last frame of Figure 2).[5] Here, the user finds an explorable list of fifty related games that includes other unique indie games (*Dinner Date* [89], *Jazzpunk* [62], etc.), story-driven games (*I Have No Mouth, and I Must Scream* [20], *Anchorhead* [30], *Gone Home* [27], etc.), and games that were developed, or are discussed in, narrative-technology research (*Façade* [57], *Dwarf Fortress* [2]). Among the list of *unrelated* games is Japanese cult classic *I'm Sorry* [10]; while probably not insightful with regard to the design of *Ice-Bound*, it nonetheless may be an interesting title that the user had not known prior. Having this generated GameNet entry, he or she is then free to explore that tool in the way we have described above.

EXPERIMENT

We conducted a user study to evaluate the utility of GameNet and GameSage as a tool suite with which game designers may discover existing games related to their own works in progress (as a way of gathering insight during the early stages of game design). Given the lack of dedicated game-discovery tools in existence, a problem we have outlined above, the state of the art in game discovery is best represented by the coordinated usage of assorted web resources. As such, in this experiment we assessed the utility of these tools, with regard to this specific task context, relative to that of this current (lack of a) state of the art.

Participants

182 participants (20% female) took part in the experiment, with ages ranging between 18 and 27 (M = 19.45). The participants were all undergraduate students enrolled in an introductory game-design course. This course fulfills a general-education requirement and has no prerequisites; as such, the students hailed from diverse academic backgrounds encompassing 42 different degree programs. In a preliminary questionnaire, we asked the participants about their level of game-development expertise: 34% reported no game-development experience prior to enrolling in the class, 57% claimed novice-level expertise, and the remaining 9% called themselves experienced.

Experimental Design

We employed a within-subjects design in which all participants searched for existing games related to a game they had made using both of two experimental conditions: a *baseline condition*, in which they could freely use a web browser to utilize any resources available online, and a *tools condition*, in which they described their game to GameSage and then searched GameNet starting from the entry generated for the game. All participants carried out the baseline phase first. Though we realize this configuration could have led to order effects, we clarify later in the paper that these would only have made the test of the system's efficacy a more rigorous one.

[5]We invite curious readers to submit the example text to GameSage to see the full listing of related and unrelated games that it provides.

Figure 3. Screenshot from one student's GameMaker game, *BreadQuest*.

Experimental Task

Prior to us conducting the experiment, each of the students had completed an assignment in which he or she used GameMaker, a game-development platform for novice programmers [66], to create a game that would emphasize exploration in an unusual or metaphorical space; Figure 3 shows a screenshot from one of these student games. For both experimental conditions, the task was the same: participants were provided with an online form and asked to spend fifteen minutes finding games related to their respective GameMaker games, entering their titles into the online form as they did. On this form, the participant was asked to place games, as they discovered them, into one of two categories—*familiar games*, games they were already familiar with prior to the experiment, and *unfamiliar games*, games they were not familiar with prior. We let the participants operationalize their own criteria for both familiarity and relatedness. During the second phase, in which the participants used GameNet and GameSage, we allowed them to repeat games from the first phase, but these repetitions were not counted during data analysis.

Materials

As we explain next, experimental sessions were held in a university computer lab reserved for a class discussion section. Participants were provided with their own desktop computers that were each equipped with Mozilla Firefox web browsers.

Procedure

As noted above, all participants were enrolled in an introductory game-design course. Experimental sessions were held during the course's six discussion sections (one session for each section) across a two-day period that was bookended by the course's first two lectures that week. Sections were led by three teaching assistants in total (each led two sections), but all experimental sessions were led by the same experimenter (who was not affiliated with the class) in the same university computer lab, which was reserved for class sections. Session start times ranged between 9:00 and 16:00, and participant totals were between 25 and 35.

Pre-Experiment

Each experimental session was preceded by a short preliminary phase in which the section leader discussed class logistics, introduced the experimenter, and explained the basic purpose of the study. Participants were told that the anonymous procedure would last roughly 45 minutes (about half of the section's allotted time), was not mandatory, and would not be graded, but that it could assist in the completion of an upcoming extra-credit assignment to compare one's GameMaker game to related games that had not been discussed during lecture. The experimenter then detailed the basic protocol of the study and asked participants to log into their respective university desktop computers and start up their preinstalled Mozilla Firefox browsers. After this prelude, the participants were asked to follow a link to a survey hosted on Google Forms and to fill out a preliminary page there with demographic questions. Once everyone had filled out this preliminary page, which also asked the participant to provide a username (whose purpose we explain momentarily), the participants were asked to proceed to the second page of the survey.

Baseline Phase

Here, participants were presented two boxes for free-text entry, one labeled 'Familiar games' and the other 'Unfamiliar games'. At this point, they were asked to spend the next fifteen minutes using any web resources available to find games related to their respective GameMaker games, entering the titles of any discovered games into the appropriate box (depending on whether the participant was familiar with the game at hand prior to beginning the experiment). After the fifteen minutes had elapsed, the experimenter asked the participants to stop searching and to proceed to the next page of the survey, which was identical to the previous one.

Tools Phase

At this stage, the participants were given a link to a special experimental GameSage instance and were asked to follow the link, spend no more than ten minutes there describing their GameMaker game, and spend the remainder of the fifteen-minute phase exploring GameNet to discover related games (entering their titles into the survey form just as they had done in the previous step). Participants were permitted to repeat titles found during the previous phase, though these were not counted during data analysis, as we discuss in the next section. Our experimental GameSage instance required login using a username; the experimenter instructed participants to use the same username they had entered on the online form (to allow for cross-referenced analysis between usage data and survey responses).[6]

[6] We were forced to exclude data from 28 participants who were not consistent in the usernames they provided across the survey and our experimental GameSage instance.

Post-Experiment

Finally, after this second phase, the participants were asked to proceed to and complete the final page of the survey, which was a post-experiment questionnaire in which we asked participants what resources they used in the baseline phase and how likely they would be to reuse GameNet and GameSage in the future. Additionally, we invited freeform feedback about the tools.

Measures and Instruments

This experimental procedure yields three primary variables of interest that correspond to the hypotheses we postulated above:

- **Mean total games discovered.** Mean number of games discovered across all participants in a given condition.

- **Percentage of games discovered that were unique.** Percentage of games discovered across all participants in a given condition that were *unique*. That is, with *total count* being the total number of game-discovery instances in a condition and *unique count* being the number of unique games discovered by all participants in that condition (*i.e.*, we do not count repeated games already tallied due to discovery by another participant), the percentage of unique games for the condition is simply $(\textit{unique count}/\textit{total count}) * 100$. We use this a measure of the *diversity* of games discovered in each condition.

- **Percentage of games discovered that were unfamiliar.** Percentage of games discovered across all participants in a condition that were unfamiliar (prior to beginning the experiment) to the participants who discovered them. As mentioned above, we let participants develop and operate over their own criteria for game familiarity.

Additionally, we have another variable of interest for which data was collected in a post-experiment questionnaire (discussed below):

- **Likelihood to reuse tools.** After being informed that GameSage is freely available online, we asked participants to indicate, using a four-point Likert scale, how likely it is that they would revisit the tools to use them again. Available responses were 'very unlikely', 'somewhat unlikely', 'somewhat likely', 'very likely'.

Potential Order Effects

At this point, the reader may be concerned with potential order effects that this study design could yield, given that we had all participants first undergo the baseline condition before the tools condition. We acknowledge that this probably generated a small practice effect, in that participants would have been more primed to the game-discovery task by the time they underwent the latter condition. However, we believe that, overall, our ordering privileges the baseline condition for two reasons. First, it is likely that a *fatigue effect* is also at play here, *i.e.*, that participants who had already spent the first experimental phase searching for games may have been less diligent in their search during the tools phase. Second, because we did not allow repetition in the tools phase

	Total*	% Unique	% Unfamiliar*
Baseline	6.49 (6.02)	58.2	47.0
Tools	14.16 (20.11)	60.1	80.3

Table 1. Number of games discovered by participants (n = 182) in both conditions (including repetitions across multiple participants; means with standard deviations in parentheses), and percentages of those that were unique and familiar respectively. *Asterisks indicate statistically significant differences between conditions at $p < 0.0001$.

of games discovered during the baseline phase, the space of potential related games for each participant was reduced in the tools phase. Moreover, this reduction was not constituted by the removal of arbitrary games from the search space, but games that may have immediately sprung to a participant's mind upon beginning the task. For these reasons, we contend that, all other things being equal, the tools condition would be expected to produce fewer discovered games.

Data Canonicalization

Because we allowed participants to freely type in titles of the games they discovered, we had to first canonicalize the responses. To do this, we went through each of the responses by hand and replaced each title instance with a persistent identifier; if no identifier yet existed for the game title, we assigned one to it. For example, title variants for *The Legend of Zelda* [64] such as 'The Legend of Zelda', 'Legend of Zelda', 'Zelda 1', and 'the first zelda' would be canonicalized to the unique identifier `865`. If we were unsure of whether a given title was a variant of some title that had already been attributed a code, we assigned it a new code. Variants of the same game on different platforms—*e.g.*, *The Legend of Zelda* as released for the Nintendo Entertainment System and for Game Boy Advance—were merged by the assignment of a single identifier. This preprocessing step was necessary for two reasons. First, though we allowed repetition in the second phase of games discovered during the first phase of the experiment, we elected not to include these in our data analysis. In order to weed out these repetitions, we first had to be able to detect them, which necessitated the canonicalizing of title variants. Second, this preprocessing step was further necessitated by our desire to investigate the respective *distributions* of games discovered during each phase of the experiment. Specifically, we were interested in the diversity of titles found during each phase, as we have outlined above.

RESULTS

Our major results are shown in Table 1. Because the two conditions exhibited unequal variances on mean total games discovered, and because we do not have huge sample sizes (n = 182 for both samples), we tested for statistical significance between these condition means using Welch's t-test, which is an adaptation of the standard t-test that better suits data with these characteristics [76]. Differences between condition proportions were tested using a standard two-proportion z-test. Our results relative to our variables of interest (outlined above) are as follows:

- **Mean total games discovered.** Many more games were found in the tools condition (14.16 vs. 6.49 per participant); this difference was statistically significant ($t(213)$ = 4.92; $p < 0.0001$).
- **Percentage of games discovered that were unique.** A greater percentage of games discovered in the tools condition were unique (60.1 vs. 58.2), but this difference was not statistically significant ($z(362)$ = -1.336; p = 0.09).
- **Percentage of games discovered that were unfamiliar.** A far greater percentage of games discovered in the tools condition were unfamiliar to participants (80.3 vs. 47.0); this difference was statistically significant ($z(362)$ = 6.6; $p < 0.0001$).
- **Likelihood to reuse tools.** 7% of participants reported they were "very likely" to reuse GameNet and GameSage, 42% indicated they were "somewhat likely" to do so, 31% responded "somewhat unlikely", and the remaining 20% chose "very unlikely".

DISCUSSION

In this section, we discuss these results with regard to our initial hypotheses before providing some additional analysis.

Revisiting Our Hypotheses

Here, we will revisit each of our initial hypotheses, given above, in light of the results we have just presented.

H1: Participants will discover more related games using GameNet and GameSage

Our results strongly support this hypothesis. Indeed, participants discovered more than twice as many games using the tools (nearly one per minute) than they did using assorted web resources. This difference is especially remarkable when considering that none of the games participants discovered in the baseline condition could be counted toward their discoveries in the tools condition. For our baseline condition, we chose to allow participants to use any available web resources because we believe this best represents the (lack of a) state of the art in game discovery. This may seem curious in that we have also outlined the existence of another game-discovery system, Vizmo, which perhaps could have served as a better method to compare to. The reason we did not do this is that Vizmo currently only contains 604 games, many of which are platform variants (which we merged in this experiment during data canonicalization). In both experimental conditions, participants cumulatively discovered more titles than exist in the entire database that undergirds Vizmo, and so we believe this decision is vindicated.

H2: Participants will discover a greater diversity of games using GameNet and GameSage

The results did not support this hypothesis, as the difference in condition proportions was not statistically significant. While we had anticipated that the proportion of unique titles

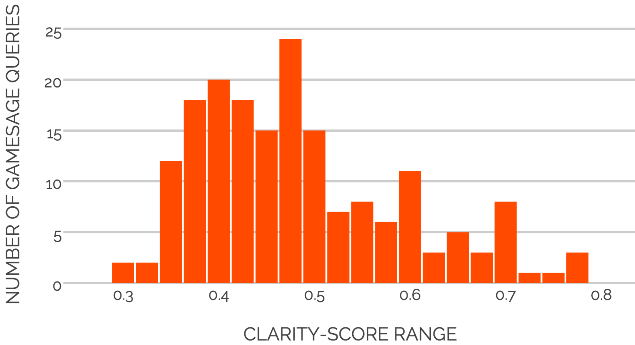

Figure 4. Distribution of *clarity scores* for the GameSage queries submitted during our experiment (M = 0.5, σ = 0.11). A clarity score for a query is its average cosine similarity relative to the fifty most related games provided for it by GameNet.

would be significantly higher in the tools condition, the fact that it even approximates that of the baseline condition is still remarkable. Because the effective search space yielded by all resources on the web is orders of magnitude greater than that represented by the 12,000 games in GameNet, one might expect that, all other things being equal, a much larger proportion of games discovered in the baseline condition would be unique. This was indeed not the case, however, and in fact 47% of the unique titles discovered during the baseline condition were games that are included in GameNet. This suggests that, had the tools condition been ordered first in our study design, participants might have discovered more games with the tools, as well as a greater diversity of games (since games included in GameNet that were already discovered by a participant in the baseline condition could not be discovered by that participant in the tools condition).

H3: Participants will discover a greater proportion of unfamiliar games using GameNet and GameSage

Our results strongly support this hypothesis. We made this prediction from the intuition that, in lieu of a dedicated game-discovery tool, participants would tend to seek out (as a sort of scaffolding) games they could already name as related to their own. Indeed, roughly half of the games discovered by participants in the baseline condition were already known to them. We believe that the demonstrated facility of GameNet and GameSage to provide users with extensive listings of games that were previously unfamiliar to them is perhaps the tool suite's greatest strength.

Tool Appraisal and Query Performance

Given that participants discovered significantly more games using the dedicated discovery tools, it was surprising to find that only 49% indicated that they would be likely to reuse the tools in the future. As such, we decided to further investigate this aspect of our data, particularly with regard to whether the participants' GameSage queries themselves were correlated with those participants' reported likelihood to reuse the tools. This notion relates to the common observation in information retrieval that, in any search task, some queries will be more likely than others to yield good results [12]. This variance can be accounted for by three properties of a query: its well-formedness, the degree to which it accommodates the matching operations enacted by the search engine to retrieve indexed items in the search space,[7] and the degree to which there are actually valid matches among those indexed items. In information retrieval, a *clarity score* is often used as a metric for predicting query performance that captures (by an entropy calculation) the cumulative effect of these sources of variance in query performance [12, 34, 90].

For our purposes here, we propose a variant of the clarity score calculated as the average cosine similarity between a GameSage submission and its fifty most related games as provided in its generated GameNet entry. Figure 4 plots the distribution of clarity scores for the 182 GameSage queries submitted in our experiment, which ranged between 0.32 and 0.79, with a mean and standard deviation of 0.5 and 0.11 respectively. To test whether participants' clarity scores correlated with desire to reuse GameNet and GameSage, we placed the participants into two groups—those who indicated they were "somewhat likely" or "very likely" to reuse the tools, and those who reported otherwise—and calculated mean clarity scores for each, which were 0.51 and 0.49 respectively. We tested for whether the former group's mean was significantly higher using a standard two-sample t-test (because the groups had roughly equal sample sizes and variances), but the test did not support this ($t(362) = 1.44$, $p < 0.08$).

Web Resources Used

During the baseline condition, participants reported using a total of 21 web resources, which speaks to the lack of dedicated game-discovery systems that we have outlined herein. The most frequently cited resources were Google search (used by 96% of participants), Google Images (27%), Wikipedia (23%), YouTube (18%), GameFAQs (10%), Steam (9%), and Giant Bomb (9%). A similar array of web resources was reported in [80]. We note that materials from Google Images, Wikipedia, and YouTube are already integrated into GameNet.

Limitations of the Tools

In our estimation, the two major limitations of these tools are both rooted in the LSA model underpinning them, particularly that model's reliance on Wikipedia text. First, Wikipedia authoring practices, and in fact the very nature of Wikipedia's notability standards [91], disadvantage a class of novel games that may live at the fringes of the medium. As we have outlined above, the bandwagon effect at play in videogames today works to privilege tried-and-true designs above more adventuresome ones, which limits the visibility of games showcasing the latter. As such, novel games tend to be known to less people, which makes it less likely that such games will earn Wikipedia authors or that they would even meet the website's notability standards if they did. This is especially troublesome for the reason that obscure, novel games are exactly the type that could lend the greatest insight to designers exploring similar concepts. Indeed, in research on recommender systems (and in work on search more generally) there is a notion that obscure recommendations from the

[7] This is the central concern in *search engine optimization* [16].

long tail are often the most valuable to a user [29], but in the case of GameNet and GameSage, the long tail is unfortunately truncated. Second, while videogame concerns surrounding, *e.g.*, the platform, development, or critical reception of a game may be useful to the GameSage user describing a game whose title he or she has forgotten (which is one potential usage of the tool), only concerns pertaining to elements of game design (*e.g.*, gameplay, aesthetics, narrative) are likely to interest most users. But the encyclopedic nature of Wikipedia leads its articles about games to often include description pertaining to any and all notable aspects of their ontologies. As such, these extraneous concerns creep their way into the model, and while they are useful in many of GameNet's specific use cases, users of GameSage would likely prefer the model to incorporate only concerns that are central to the game-design process.

Post-Experiment Usage

GameNet and GameSage were made publicly available in June 2015. At the time of writing, the system has logged 4,397 sessions by 3,811 unique users (*i.e.*, 15.2% revisited), with 3.84 GameNet entries viewed per session. This indicates that approximately 45 new users try the tools each day.

IMMEDIATE REVISIONS

In their freeform feedback, participants broadly indicated that GameSage asks too many questions and that submitting a game idea takes too long. As an immediate revision to the tool, its default mode of interaction has already been changed into one in which the user submits a full description of his or her game idea by filling out a single text-entry field. For users that seek a more directed experience, a button labeled 'Guide Me' may be clicked to engage a mode that asks most (but not all) of the same questions that were asked in the version described in this paper. We invite the reader to try out GameSage as it is currently structured by following the link given below.

CONCLUSION AND FUTURE WORK

The immense proliferation of videogames over the course of recent decades has yielded a discoverability problem that has gone largely unaddressed. Though this problem affects all videogame stakeholders, in this paper we have limited our concerns to the specific context of game designers seeking prior work that could inform their own ideas or works in progress. Specifically, we have presented GameNet and GameSage, components of a game-discovery system that solicits text about a user's idea for a game to generate (using an NLP technique called *latent semantic analysis*) an explorable listing of the existing games most related to that abstract idea. In a user study in which 182 novice game designers used both this system and unrestricted web resources (the apparent state of the start in game discovery today) to find games related to their own, participants discovered more than twice as many titles using GameNet and GameSage, and a far greater proportion of these games were previously unfamiliar to the participants who discovered them. This suggests that this system currently represents the state of the art in videogame discovery, though we acknowledge that sadly there is currently little in the way of competition. More broadly, we have provided in this paper the first articulation of videogame discovery as an emerging application area.

Looking ahead, we envision several avenues of future work, some of which we are actively planning or already carrying out. Earlier, we noted that one of the few reported videogame recommender systems employs *archetypal analysis*, a machine learning technique whose application to games is also being explored by Chong-U Lim and D. Fox Harrell [51, 52, 53, 54, 55]. By this method, instances in a data set are represented as mixtures of *pure types*, or *archetypes*, which themselves are represented as mixtures of the instances [13]. This approach is alluring because the representations that it yields are *human-interpretable* (unlike purely vectorial representations), as they simply specify mixtures of other instances (in game discovery, this would be mixtures of games). We are enticed by the prospect of an *archetypal game-discovery tool* that could afford both the semiautomated game indexing of GameNet and the interpretability of Vizmo.

While Ryan et al. evaluated GameNet in the context of game scholars seeking games related to their research topics [80], that study preceded the creation of GameSage. We plan to evaluate the full tool suite in the context of this use case by having scholars explore GameNet entries generated by submitting abstracts of their research topics to GameSage. Further, as we have alluded to above, we believe this system could show utility as a tool with which game historians, scholars, and enthusiasts may discover games that are related to topics of concern in videogame history. Currently, we are making preparations for a user study that will evaluate GameNet and GameSage in this context. As an additional evaluation, we might also directly compare this game-discovery system to the other such existing system, Vizmo, though this presents a challenge in that the latter has far fewer games; perhaps we could limit GameNet's library of games to a subset that is roughly coextensive with Vizmo's.

Lastly, we have nearly completed development of a variant of GameNet and GameSage whose underlying LSA model is trained using text from videogame walkthroughs extracted from the website GameFAQs [28]. Because walkthroughs [63, 4] describe gameplay exclusively and systematically, often in exhaustive detail, we believe that the second major system limitation that we have outlined above—that game designers will prefer to operate over a subset of the concerns represented in Wikipedia encyclopedic description, specifically *gameplay* concerns—could conceivably be defeated in this revision. In a future study, we will directly compare this revised system to the version trained on Wikipedia text.

We hope that this project will encourage the development of new tools that may work to better alleviate the videogame discoverability problem that we have outlined herein.

LINKS

GameNet and GameSage are freely available web apps hosted online at `https://gamecip.soe.ucsc.edu/projects`.

REFERENCES

1. Panagiotis Adamopoulos and Alexander Tuzhilin. 2011. On unexpectedness in recommender systems: Or how to expect the unexpected. In *Proc. Novelty and Diversity in Recommender Systems*. ACM, 11–18.
2. Tarn Adams and Zach Adams. 2006. *Slaves to Armok: God of Blood Chapter II: Dwarf Fortress*. Bay 12 Games.
3. Moshe Adler. 1985. Stardom and talent. *The American Economic Review* (1985), 208–212.
4. Daniel Ashton and James Newman. 2011. Slow Play Strategies: Digital Games Walkthroughs and the Perpetual Upgrade Economy. *Transformations* 20 (2011).
5. Roger B Bradford. 2008. An empirical study of required dimensionality for large-scale latent semantic indexing applications. In *Proc. Information and Knowledge Management*. ACM, 153–162.
6. Simon Carless. 2014. Why game discovery is vital. *Gamasutra* (2014).
7. Laura Catalá, Vicente Julián, and José-Antonio Gil-Gómez. 2014. A CBR-based game recommender for rehabilitation videogames in social networks. In *Proc. Intelligent Data Engineering and Automated Learning*. Springer, 370–377.
8. Òscar Celma and Pedro Cano. 2008. From hits to niches?: or how popular artists can bias music recommendation and discovery. In *Proc. Large-Scale Recommender Systems*. ACM.
9. Chaochang Chiu, Re-Jiau Sung, Yu-Ren Chen, and Chih-Hao Hsiao. 2013. App Review Analytics of Free Games Listed on Google Play. (2013).
10. Coreland. 1985. *I'm Sorry*. Sega.
11. Sophie Courtois, Carmela M Cappellano, Maria Ball, Francois-Xavier Francou, Philippe Normand, Gérard Helynck, Asuncion Martinez, Steven J Kolvek, Joern Hopke, Marcia S Osburne, and others. 2003. Recombinant environmental libraries provide access to microbial diversity for drug discovery from natural products. *Applied and Environmental Microbiology* 69, 1 (2003), 49–55.
12. Steve Cronen-Townsend, Yun Zhou, and W Bruce Croft. 2002. Predicting query performance. In *Proc. Information Retrieval*. ACM, 299–306.
13. Adele Cutler and Leo Breiman. 1994. Archetypal analysis. *Technometrics* 36, 4 (1994), 338–347.
14. Raimund Dachselt and Mathias Frisch. 2007. Mambo: a facet-based zoomable music browser. In *Proc. Mobile and Ubiquitous Multimedia*. ACM, 110–117.
15. Allan Peter Davis, Cynthia G Murphy, Cynthia A Saraceni-Richards, Michael C Rosenstein, Thomas C Wiegers, and Carolyn J Mattingly. 2009. Comparative Toxicogenomics Database: a knowledgebase and discovery tool for chemical–gene–disease networks. *Nucleic Acids Research* 37 (2009), D786–D792.
16. Harold Davis. 2006. *Search engine optimization*. O'Reilly.
17. Scott C. Deerwester, Susan T Dumais, Thomas K. Landauer, George W. Furnas, and Richard A. Harshman. 1990. Indexing by latent semantic analysis. *JASIS* 41, 6 (1990).
18. Chrysanthos Dellarocas and Ritu Narayan. 2007. Tall heads vs. long tails: Do consumer reviews increase the informational inequality between hit and niche products? *Robert H. Smith School of Business Research Paper* 06-056 (2007).
19. Inderjit S Dhillon, Subramanyam Mallela, and Dharmendra S Modha. 2003. Information-theoretic co-clustering. In *Proc. Knowledge Discovery and Data Mining*. ACM, 89–98.
20. Dreamers Guild, The. 1995. *I Have No Mouth, and I Must Scream*. Cyberdreams.
21. Susan T Dumais. 2004. Latent semantic analysis. *Annual Review of Information Science and Technology* (2004).
22. Jennifer English, Marti Hearst, Rashmi Sinha, Kirsten Swearingen, and Ka-Ping Yee. 2002. Hierarchical faceted metadata in site search interfaces. In *Proc. CHI*. ACM, 628–639.
23. Everyplay & Unity Technologies. 2014. *Mobile gaming: Social motivations*. Technical Report.
24. Jody Condit Fagan, Meris A Mandernach, Carl S Nelson, Jonathan R Paulo, and Grover Saunders. 2012. Usability Test Results for a Discovery Tool in an Academic Library. *Information Technology and Libraries* 31, 1 (2012), 83–112.
25. Martin D Flintham, Raphael Velt, Max L Wilson, Edward J Anstead, Steve Benford, Anthony Brown, Timothy Pearce, Dominic Price, and James Sprinks. 2015. Run Spot Run: Capturing and Tagging Footage of a Race by Crowds of Spectators. In *Proc. CHI*. ACM, 747–756.
26. W Wayne Fu and Clarice C Sim. 2011. Aggregate bandwagon effect on online videos' viewership: Value uncertainty, popularity cues, and heuristics. *American Society for Information Science and Technology* 62, 12 (2011), 2382–2395.
27. Fullbright. 2013. *Gone Home*.
28. GameFAQs. GameFAQs.com. http://www.gamefaqs.com/. (????).
29. Mouzhi Ge, Carla Delgado-Battenfeld, and Dietmar Jannach. 2010. Beyond accuracy: evaluating recommender systems by coverage and serendipity. In *Proc. Recommender Systems*. ACM, 257–260.
30. Michael S Gentry. 1998. *Anchorhead*.

31. Giant Bomb. 2015. Call of Duty franchise. http://www.giantbomb.com/call-of-duty/3025-82/games/. (19 Sept. 2015).

32. Gene H Golub and Christian Reinsch. 1970. Singular value decomposition and least squares solutions. *Numer. Math.* 14, 5 (1970), 403–420.

33. Lindsay D Grace. 2014. A Linguistic Analysis of Mobile Games: Verbs and Nouns for Content Estimation. In *Proc. Foundations of Digital Games*.

34. Ben He and Iadh Ounis. 2004. Inferring query performance using pre-retrieval predictors. In *String Processing and Information Retrieval*. Springer, 43–54.

35. Dietmar Jannach, Lukas Lerche, Fatih Gedikli, and Geoffray Bonnin. 2013. What recommenders recommend–an analysis of accuracy, popularity, and sales diversity effects. In *User Modeling, Adaptation, and Personalization*. Springer, 25–37.

36. Tingting Jiang. 2014. Exploratory Search: A Critical Analysis of the Theoretical Foundations, System Features, and Research Trends. In *Library and Information Sciences*. Springer, 79–103.

37. Amy K Karlson, George G Robertson, Daniel C Robbins, Mary P Czerwinski, and Greg R Smith. 2006. FaThumb: a facet-based interface for mobile search. In *Proc. CHI*. ACM, 711–720.

38. Maija Koivisto. 2015. Discoverability Problem of Free-to-Play Mobile Games. (2015).

39. Nicholas Kong, Marti A Hearst, and Maneesh Agrawala. 2014. Extracting references between text and charts via crowdsourcing. In *Proc. CHI*. ACM, 31–40.

40. Wessel Kraaij and Wilfried Post. 2006. Task based evaluation of exploratory search systems. In *Proc. Evaluating Exploratory Search Systems*. 24–27.

41. Shyong K Lam and John Riedl. 2004. Shilling recommender systems for fun and profit. In *Proc. WWW*. ACM, 393–402.

42. Thomas K Landauer and Susan T Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104, 2 (1997).

43. Thomas K Landauer, Peter W Foltz, and Darrell Laham. 1998. An introduction to latent semantic analysis. *Discourse Processes* 25, 2-3 (1998).

44. Walter S Lasecki, Mitchell Gordon, Winnie Leung, Ellen Lim, Jeffrey P Bigham, and Steven P Dow. 2015a. Exploring Privacy and Accuracy Trade-Offs in Crowdsourced Behavioral Video Coding. In *Proc. CHI*. ACM, 1945–1954.

45. Walter S Lasecki, Juho Kim, Nicholas Rafter, Onkur Sen, Jeffrey P Bigham, and Michael S Bernstein. 2015b. Apparition: Crowdsourced User Interfaces That Come To Life As You Sketch Them. In *Proc. CHI*. ACM, 1925–1934.

46. Jin Ha Lee, Rachel Ivy Clarke, and Andrew Perti. 2015a. Empirical evaluation of metadata for video games and interactive media. *Association for Information Science and Technology* (2015).

47. Jin Ha Lee, Sungsoo Ray Hong, Hyerim Cho, and Yea-Seul Kim. 2015b. VIZMO Game Browser: Accessing Video Games by Visual Style and Mood. In *Proc. CHI*. ACM, 149–152.

48. Jin Ha Lee, Natascha Karlova, Rachel Ivy Clarke, Katherine Thornton, and Andrew Perti. 2014. Facet analysis of video game genres. *Proc. iConference* (2014).

49. Young-Jin Lee, Kartik Hosanagar, and Yong Tan. 2015c. Do I Follow My Friends or the Crowd? Information Cascades in Online Movie Ratings. *Management Science* (2015).

50. Mark Levy and Klaas Bosteels. 2010. Music recommendation and the long tail. In *Proc. Music Recommendation And Discovery*.

51. Chong-U Lim and D Fox Harrell. 2015a. Comparing Clustering Approaches for Modeling Players' Values through Avatar Construction. In *Proc. Player Modeling*. AAAI.

52. Chong-U Lim and D Fox Harrell. 2015b. A Data-Driven Approach for Computationally Modeling Players' Avatar Customization Behaviors. In *Proc. Artificial Intelligence and Interactive Digital Entertainment*. AAAI.

53. Chong-U Lim and D Fox Harrell. 2015c. The Marginal: A Game for Modeling Players' Perceptions of Gradient Membership in Avatar Categories. In *Proc. Experimental AI in Games*. AAAI.

54. Chong-U Lim and D Fox Harrell. 2015d. Revealing Social Identity Phenomena in Videogames with Archetypal Analysis. In *Proc. AISB Symposium on AI and Games*. AISB.

55. Chong-U Lim and D Fox Harrell. 2015e. Understanding Players Identities and Behavioral Archetypes from Avatar Customization Data. In *Proc. Computational Intelligence and Games*.

56. Gary Marchionini. 2006. Exploratory search: from finding to understanding. *Commun. ACM* 49, 4 (2006), 41–46.

57. Michael Mateas and Andrew Stern. 2003. Façade: An experiment in building a fully-realized interactive drama. In *Proc. Game Developers Conference*.

58. Yevgeniy Medynskiy, Mira Dontcheva, and Steven M Drucker. 2009. Exploring websites through contextual facets. In *Proc. CHI*. ACM, 2013–2022.

59. Michael Meidl, Steven Lytinen, and Kevin Raison. 2014. Using game reviews to recommend games. *Proc. Games and NLP* (2014).

60. Patrick Miller. 2013. What's Next? Spector: 'Design innovation is where the future lies'. *Gamasutra* (2013).

61. Christopher Luke Moore. 2009. Digital games distribution: the presence of the past and the future of obsolescence. *M/C* 12, 3 (2009).

62. Necrophone Games. 2014. *Jazzpunk*. Adult Swim Games.

63. James Newman. 2011. (Not) Playing Games: Player-Produced Walkthroughs as Archival Documents of Digital Gameplay. *Digital Curation* 6, 2 (2011), 109–127.

64. Nintendo R&D4. 1986. *The Legend of Zelda*. Nintendo.

65. Jahna Otterbacher. 2015. Crowdsourcing Stereotypes: Linguistic Bias in Metadata Generated via GWAP. In *Proc. CHI*. ACM, 1955–1964.

66. Mark Overmars. 2004. Teaching computer science through game design. *Computer* 37, 4 (2004), 81–83.

67. Deuk Hee Park, Hyea Kyeong Kim, Il Young Choi, and Jae Kyeong Kim. 2012. A literature review and classification of recommender systems research. *Expert Systems with Applications* 39, 11 (2012), 10059–10072.

68. Pocket Gamer. 2015. App Store Metrics. http://www.pocketgamer.biz/metrics/app-store/. (19 Sept. 2015).

69. Catherine Quinton. 2014. Oscar Clark: The social interaction frontier. *Gamesauce* (2014).

70. Tammera M Race, MP Popp, and D Dallis. 2012. Resource discovery tools: Supporting serendipity. *Planning and Implementing Resource Discovery Tools in Academic Libraries* (2012), 139–152.

71. Kevin Raison, Noriko Tomuro, Steve Lytinen, and José P Zagal. 2012. Extraction of User Opinions by Adjective-Context Co-clustering for Game Review Texts. In *Proc. Advances in NLP*.

72. K Nageswara Rao. 2010. Application domain and functional classification of recommender systemsa survey. *Library & Information Technology* 28, 3 (2010), 17–35.

73. Aaron A Reed, Jacob Garbe, Noah Wardrip-Fruin, and Michael Mateas. 2014. Ice-bound: Combining richly-realized story with expressive gameplay. *Proc. Foundations of Digital Games* (2014).

74. Francesco Ricci, Lior Rokach, and Bracha Shapira. 2011. *Introduction to recommender systems handbook*. Springer.

75. Mike Rose. 2014. More games have released on Steam so far in 2014 than all of last year. *Gamasutra* (2014).

76. Graeme D Ruxton. 2006. The unequal variance t-test is an underused alternative to Student's t-test and the Mann–Whitney U test. *Behavioral Ecology* 17, 4 (2006), 688–690.

77. James Owen Ryan, Eric Kaltman, Andrew Max Fisher, Timothy Hong, Taylor Owen-Milner, Michael Mateas, and Noah Wardrip-Fruin. 2015a. Large-Scale Interactive Visualizations of Nearly 12,000 Digital Games. (2015).

78. James Owen Ryan, Eric Kaltman, Timothy Hong, Michael Mateas, and Noah Wardrip-Fruin. 2015b. People Tend to Like Related Games. In *Proc. Foundations of Digital Games*.

79. James Owen Ryan, Eric Kaltman, Michael Mateas, and Noah Wardrip-Fruin. 2015c. Tools for Videogame Discovery Built Using Latent Semantic Analysis. (2015).

80. James Owen Ryan, Eric Kaltman, Michael Mateas, and Noah Wardrip-Fruin. 2015d. What We Talk About When We Talk About Games: Bottom-up Game Studies Using Natural Language Processing. In *Proc. Foundations of Digital Games*.

81. Gerard Salton, Anita Wong, and Chung-Shu Yang. 1975. A vector space model for automatic indexing. *Commun. ACM* 18, 11 (1975), 613–620.

82. John Salwitz. 1986. *Rampart*. Atari Games.

83. Guy Schofield, Tom Bartindale, and Peter Wright. 2015. Bootlegger: Turning Fans into Film Crew. In *Proc. CHI*. ACM, 767–776.

84. Guy Shani and Asela Gunawardana. 2011. Evaluating recommendation systems. In *Recommender systems handbook*. Springer, 257–297.

85. Ben Shneiderman. 2007. Creativity support tools: Accelerating discovery and innovation. *Commun. ACM* 50, 12 (2007), 20–32.

86. Rafet Sifa, Christian Bauckhage, and Anders Drachen. 2014. Archetypal game recommender systems. *Proc. KDML-LWA* (2014).

87. SOFEL. 1990. *Wall Street Kid*. SOFEL.

88. Statista. 2015. Google Play: number of available apps 2009-2015. http://www.statista.com/statistics/266210/number-of-available-applications-in-the-google-play-store/. (19 Sept. 2015).

89. Stout Games. 2010. *Dinner Date*. Stout Games.

90. Aixin Sun and Sourav S Bhowmick. 2009. Image tag clarity: in search of visual-representative tags for social images. In *Proc. SIGMM Workshop on Social Media*. ACM, 19–26.

91. Dario Taraborelli and Giovanni Luca Ciampaglia. 2010. Beyond notability: Collective deliberation on content inclusion in Wikipedia. In *Proc. Self-Adaptive and Self-Organizing Systems*. IEEE, 122–125.

92. Peter D Turney, Patrick Pantel, and others. 2010. From frequency to meaning: Vector space models of semantics. *Artificial Intelligence Research* 37, 1 (2010), 141–188.

93. VentureBeat. 2015. Comparing Apples and Googles: The App Store vs. Google Play (infographic). http://venturebeat.com/2013/07/17/comparing-apples-and-googles-the-app-store-vs-google-play-infographic/. (19 Sept. 2015).

94. Doug Way. 2010. The impact of web-scale discovery on the use of a library collection. *Serials Review* 36, 4 (2010), 214–220.

95. Ryen W White, Gary Marchionini, and Gheorghe Muresan. 2008. Evaluating exploratory search systems: Introduction to special topic issue of information processing and management. *Information Processing & Management* 44, 2 (2008), 433–436.

96. Ryen W White and Resa A Roth. 2009. Exploratory search: Beyond the query-response paradigm. *Synthesis Lectures on Information Concepts, Retrieval, and Services* 1, 1 (2009), 1–98.

97. Ka-Ping Yee, Kirsten Swearingen, Kevin Li, and Marti Hearst. 2003. Faceted metadata for image search and browsing. In *Proc. CHI*. ACM, 401–408.

98. Lixiu Yu, Aniket Kittur, and Robert E Kraut. 2014a. Distributed analogical idea generation: inventing with crowds. In *Proc. CHI*. ACM, 1245–1254.

99. Lixiu Yu, Aniket Kittur, and Robert E Kraut. 2014b. Searching for analogical ideas with crowds. In *Proc. CHI*. ACM, 1225–1234.

100. José P Zagal and Noriko Tomuro. 2010. The aesthetics of gameplay: A lexical approach. In *Proc. Academic MindTrek Conference*.

101. José Pablo Zagal and Noriko Tomuro. 2013. Cultural differences in game appreciation: A study of player game reviews.. In *Proc. Foundations of Digital Games*.

102. José P Zagal, Noriko Tomuro, and Andriy Shepitsen. 2011. Natural language processing in game studies research: An overview. *Simulation & Gaming* (2011).

103. Xiangyu Zhao, Zhendong Niu, and Wei Chen. 2013. Opinion-based collaborative filtering to solve popularity bias in recommender systems. In *Proc. Database and Expert Systems Applications*. Springer, 426–433.

104. Miaoqi Zhu and Xiaowen Fang. 2013. Using Lexicons Obtained from Online Reviews to Classify Computer Games. In *Proc. AIS-SIGHCI*.

105. Miaoqi Zhu and Xiaowen Fang. 2014a. Developing Playability Heuristics for Computer Games from Online Reviews. In *Design, User Experience, and Usability: Theories, Methods, and Tools for Designing the User Experience*.

106. Miaoqi Zhu and Xiaowen Fang. 2014b. Introducing a Revised Lexical Approach to Study User Experience in Game Play by Analyzing Online Reviews. In *Proc. Interactive Entertainment*.

107. Miaoqi Zhu and Xiaowen Fang. 2014c. What nouns and adjectives in online game reviews can tell us about player experience?. In *Proc. CHI*.

108. Miaoqi Zhu and Xiaowen Fang. 2015. A Lexical Approach to Study Computer Games and Game Play Experience via Online Reviews. *Human-Computer Interaction* (2015).

109. Miaoqi Zhu, Xiaowen Fang, Susy S Chan, and Jacek Brzezinski. 2013. Building a dictionary of game-descriptive words to study playability. In *Proc. CHI*.

110. Wojciech Ziarko, Robert Golan, and Donald Edwards. 1993. An application of datalogic/R knowledge discovery tool to identify strong predictive rules in stock market data. In *Proc. Knowledge Discovery in Databases*. 89–101.