Understanding User Needs in Videogame Moment Retrieval

Barrett R. Anderson
UC Santa Cruz
Santa Cruz, California
barander@ucsc.edu

Adam M. Smith
UC Santa Cruz
Santa Cruz, California
amsmith@ucsc.edu

ABSTRACT

Videogames are a rich domain for scholarship, and their dynamic content makes them a new and unique challenge for information retrieval (IR). Recent work has made it possible to cite specific moments in a videogame (like pages in a book), but currently finding those moments to cite is laborious because there are no videogame moment search engines. We conducted an in-depth interview study with ten users across a variety of user profiles: developers, educators, speedrunners, scholars, and streamers. From these interviews, we identify the unique needs of each user profile for retrieving moments from videogames. We outline implications for future research in retrieval and design implications for new tools focused on interactive media.

CCS CONCEPTS

• **Information systems** → **Retrieval tasks and goals**; *Multimedia and multimodal retrieval*; • **Applied computing** → *Computer games*;

KEYWORDS

Information Retrieval, Interviews, Videogame Moments

1 INTRODUCTION

The interactivity of videogames, with their branching paths and dynamic content, presents a new challenge for information retrieval. Conversations about games often need to include references to moments to answer fundamental questions: What can players do in the game, how do they do it, and what effects do those actions have on the game world? The problem of identifying specific moments has been long solved for more traditional media (e.g. a page number in a book, a timestamp for a film), but games present several unique obstacles. For example, because games can be nonlinear, two playthroughs of the same game might see the same moment at wildly different times. Because games are dynamic, and players can

make different choices or have different levels of skill, it is possible for two playthroughs of the same game to contain very different collections of moments with very little overlap. Kaltman et al. [8] recently introduced a system for meeting these challenges, primarily for scholars, which defines a citation format that can identify a unique game moment unambiguously. However, this only helps users communicate about unique moments in a game that they are able to reach by their own play. We believe search and retrieval systems could help them find many other moments they might wish to refer to, replicating the successes of rich-media search in other domains [10]. This study is the first to frame the question of how best to use information retrieval (IR) for videogame moments as a question for human-computer interaction (HCI) research methods. It presents the results of an interview study, partially anchored on a demonstration of current techniques for content-based retrieval of videogame moments, suggesting several new applications both within and beyond game studies.

Enabling rich search capabilities for videogame moments has implications for many different types of users. Speedrunners, streamers, or general consumers might bookmark and share moments in a leisure or hobby context. Individuals in these user profiles are usually interested in well-known commercial games that have been played by many others. Scholars, educators, translators, developers, and critics need to reference specific moments as part of their professional work. This might involve relating moments in less well-known games, or linking other games to the specific game they are developing or critiquing. In an industrial context, the existence and reachability of certain events is important to app store operators and quality assurance testers. Whether certain moments are possible in a soon-to-be-released game may impact the game's release schedule and its later commercial success. This paper aims to explore the needs of individuals in a subset of this broad spectrum of user profiles.

Zhang et al. [17] examined the possible methods and applications of videogame moment search from a technical capability perspective. The present work complements this approach, seeking to identify the stakeholders of videogame moment search and identify their needs with a user-centered design methodology. The way that a scholar, teacher, or hobbyist might want to search for a moment will vary depending on the context and purpose of the search. For example, an educator seeking a moment to best represent a certain game mechanic during a lecture may have different interests than the expert speedrunner using search to review records of their own past play. We interviewed a range of users who could benefit from a videogame moment information retrieval system: speedrunners, streamers, educators, developers, and scholars. For each of these user profiles, we address the following research questions (RQs):

RQ1 What kinds of game moments do various types of users seek?
RQ2 How do they want to identify those moments in a search?

RQ3 What will they do with retrieved moments?

We synthesize answers to these questions into specific recommendations for the next technical steps in interactive media IR research, highlighting a role for systems that work across many game platforms, search with textual queries, ingestion of bulk video data, and applications of retrieval to personal information management.

2 BACKGROUND

This section contextualizes the investigation of human-computer interaction concerns in information retrieval, the newly emerging scholarly study of moments in games, and existing strategies for search used with rich (non-textual) media.

2.1 HCI of IR

There are many possible interfaces to search engines that vary based on their context of use and the type of media they are being used to retrieve. HCI research in IR seeks to identify the distinct patterns of use and the interface designs that best support them. For example Andre et al. [2] analyze differences in use between image and text search engines and recommend new interfaces for searching non-textual media. Work like this is important for setting the direction of new technical research on the basis of user needs discovered through empirical research methods.

Within the human-computer interaction and information retrieval community, recent work by Ince et al. [6] used a semi-structured interview process to understand the search practices of doctoral students in scholarly workflows. Rather than investigating how subjects engaged with a specific retrieval system, they sought to understand patterns of use across different tools. Our present work adopts the same research methodology and similar investigative goals. Where Ince recruits subjects all within a specific early-career academic role, our work recruits subjects spread across several different contexts of use for retrieval technology in an effort to represent the breadth of stakeholders.

In long-form media outside of games, recent work by Bogers et al. [4] considered similarities and differences in how users seek books and movies. The attributes by which these media artifacts are recalled differ from how users seek webpages or research articles. In particular, users often search for the larger book or movie by referring to specific scenes, characters, or even precise lines of dialog within it.

2.2 Indexing Moments in Games

Current games researchers analyze games using a variety of methods, ranging from analysis of pre-recorded video and screenshots to careful investigative play using specially prepared game platform emulators. Even at the level of referring to whole games, the lack of standardization around vocabulary and reference formats has meant scholars work with conventions developed in hobbyist communities, or do the work of developing their own reference rules [1] – it is not enough to reference "*Super Mario World* (1990)" in the face of many versions of the game, some created and modified by fans. Kaltman et al. [8] lead an effort to make it possible to cite specific moments in specific games. In particular, those moments that are captured by the complete state (e.g. of system memory and

processor register contents) of a game platform emulator. This same notion of moments is the focus in our work. In contrast to laboriously described identification of a game moment by text (such as instructions for reproducing it in manual gameplay), these emulator snapshots are compact and precisely replicable.

Zhang et al. [17] introduced the problem of crawling, indexing, and retrieving moments in interactive media as the key processes in a new kind of search engine. Zhan et al. [14] identify several different strategies for extracting a diverse collection of (emulator snapshot) moments from a given videogame on platforms ranging from the 8-bit Atari 2600 to the 32-bit Nintendo 64. Some of these strategies involve replaying expert gameplay traces while others involve running algorithms to automatically explore the space of interaction.

In interviews below, we make contact with some of the individuals who produce these kinds of expert interaction traces (speedrunners). We also interview individuals in a game developer role who might gain insight from moments uncovered via automated exploration that they (or testers on their team) would have been unlikely to find themselves.

2.3 Rich-media Search Engines

A search engine for moments in videogames is situated in the context of previous work on search for rich (i.e. non-textual) media. One component of a game moment is the visual image it presents to the player, and there are several paradigms for visual image search. Users of existing tools like Google Image Search[1] or TinEye[2] search seeking image results. For Google Image Search, this might mean using a descriptive text query, and often involves learning the particular vocabulary necessary to produce the desired results. For TinEye the query is always an image itself, and the sought after result for might be the original web-page source and context of that specific image, variations on the image, or a more canonical or higher resolution example. The Shazam[3] mobile app offers a sound-based search engine where users query the system with a short snippet of recorded audio and expect to see results that list which popular music track they have identified. In part of our interview protocol, we remind subjects of these systems to uncover parallels between audiovisual search and interactive media moment search (which might use rich media artifacts as query or result items).

Another approach to rich-media search considers searching for objects that are not themselves images or audio. In the ProductNet visual search system [3], images of home furnishings are used to retrieve instances of products that are available for purchase. These results are actual products, not merely visually similar images. Many of these rich-media search systems work by training a neural network (or other function representation) that can embed images of similar objects into nearby locations in a latent space while mapping dissimilar objects to distant locations. Zhan et al. [15] introduced a similar technique to represent moments in games by relating screenshot images to savestate memory configurations. Compared to product search, which retrieves possible results from a finite catalog, the space of possible moments that might be returned

---

[1]https://images.google.com/
[2]https://www.tineye.com/
[3]https://www.shazam.com/

by a game moment retrieval system can be astronomically larger. Our interviews suggest it may be useful to retrieve moments that no human player has previously seen or reached via direct (ordinary or tool-assisted) play.

3 METHODOLOGY

To understand the needs of users in videogame moment retrieval, we conducted a series of semi-structured interviews. We contacted individuals who use game moments in a variety of different contexts. Although we do not cover every user profile or context of use sketched in the paper's introduction, we believe the sampling in this initial study on the topic already provides a broad perspective.

3.1 Recruitment

Before recruitment, the authors generated a set of hypothetical user profiles, which was refined and expanded throughout the study. Potential interview subjects needed to have a connection to one or more of these profiles, as judged by the authors. Potential subjects were contacted directly, either via previously established personal connections, through the author's own networks, or through publicly available contact information. Out of twelve individuals contacted, ten agreed to be interviewed. Of those ten, three came from outside of any academic institution. All subjects were recruited for their specific qualifications relevant to each user profile. We were not able to recruit subjects for each profile, and these missing perspectives (e.g., game translator) are considered in Section 6.

Because each subject's individual background is integral to interpreting their needs, their comments are quoted below with attribution. We introduce each subject in the section in which they are first referenced. Of the many contexts of use, our subjects offer coverage of hobby and professional use of videogame moment search, but we were not able to connect with subjects that could speak to industrial uses (e.g. app store operators). All subjects consented to being recorded and to being quoted with identifying attribution. The complete list of subjects and their association with user profiles is shown in Table 1.

3.2 Interview Guide

We followed the semi-structured interview guide outlined in Fig. 1. Interviews (approximately 40 minutes in length) were conducted face-to-face or, at the request of the subject, via video-conferencing software. In one case a subject requested an interview to be conducted via text chat. All other interviews were video or audio recorded. Subjects were encouraged to self-identify among several possible user profiles from a list[5] we provided or to describe their user profile in their own terms. When interview subjects described themselves in ways that related to multiple profiles we prompted them to respond to questions from each relevant perspective. Our questions were designed to elicit answers to our research questions without putting undue emphasis on the specific technology demonstrated by Zhang et al. [17] (detailed below) as an anchor. In its place, we made sure subjects all had been reminded

---

[Q] **Subject identification:** Identify yourself in one or more of the user profiles in the list provided. Describe how your work or other activity relates to videogames.

[Q] **Current search methods:** How do you currently search for and within games now, or instruct others to?

[P] **Current search methods:** Review examples of Google text search, Google Image search, Steam text search and browsing, YouTube text search and browsing, Reverse image search (e.g. TinEye), and Maps search.

[Q] **Speculation:** How do you think a hypothetical videogame moment search engine would work? What are the inputs and outputs? What should it do for you?

[P] **Demonstration:** View a demonstration of Zhang's system and review an explanation of existing technology.

[Q] **Capabilities:** What could / couldn't you do with this system or one like it?

[Q] **Queries and Results:** What game platforms need to be indexed for the system to be useful to you? What needs to be extractable from search results? How would you share results?

[D] **Open-ended discussion**

Figure 1: Outline of our semi-structured interview guide. Q marks questions for subjects, P priming demonstrations intended to guide discussion, and D subject-specific discussion. Items are paraphrased here for brevity.

of a broader array of search technologies already in use for text, image, and video. Later, subjects were asked to specifically react to a demonstration of a prototype moment search engine. Although we used the same template for each interview, the length of topics of discussion varied considerably in response to anecdotes shared by the subjects.

3.3 Anchoring Demonstration

An existing prototype videogame moment search engine was demonstrated about halfway through each interview, followed by an explanation of the underlying technology. This demonstration and explanation served to prompt additional suggestions and ideas, now grounded with a working example. The system is described in detail elsewhere [17], but in brief it allows searching for in-game moments using screenshots as input, currently for a select few games for the Super Nintendo Entertainment System (released in the early 1990s).

In the demonstration video,[6] subjects observed a screenshot from the game *Super Mario World (1990)* being dragged and dropped onto a web-based search engine interface. In response, the system supplies an initial list of potentially relevant search results. Each result item identified what game the moment was from and showed a screenshot thumbnail. Next, the video demonstrates tagging some of the results as examples of relevant and irrelevant results. The system's relevance feedback feature is applied to refine the next

---

[4]Subjects by initials: Nathan Altice (NA), Batu Aytemiz (BA), Bernd Huber (BH), Nick Junius (NJ), Ted Lim (TL), Henry Lowood (HL), Stacey Mason (SCM), Stella Mazeika (SAM), Alexis Ross (AR), and Elizabeth Swensen (ES).
[5]User Profiles: Speedrunner, Streamer, Educator, Scholar, Developer, or Designer.

[6]https://drive.google.com/file/d/1eGkx1mh_Nry1hHP2S4j0p4HVL2pMCzTy

User Profile	Profile Description	Interview Subjects
Speedrunners	Competing to complete games as fast as possible	BH, SAM, AR
Streamers / YouTubers	Broadcasting gameplay video with commentary	NJ, SCM, SAM
Educators	Teaching about the design, history, or interpretation of games	NA, HL, ES
Developers / Designers	Creating new videogames or tools to support videogame creation	BA, TL, ES
Scholars	Doing critical or interpretive work with videogames	NA, HL

Table 1: User profiles and their representative interview subjects.[4]

batch of results under these suggestions. All moments seen in search results are sampled from a recording of non-speedrun gameplay from a student volunteer (rather than from any specific dataset with which the subject might already be familiar).

3.4 Analysis

During the course of the interview, the interviewer made occasional reflective notes. This process of memoing [5, Chap. 4] continued with a subsequent review of the complete recording. In this process the interviewer looked specifically for patterns that could be attached to our broader user profiles as well as for unique insights from individual subjects. These notes and portions of the interview were presented to the research team, who provided additional observations or suggestions for interpretation. The analysis was conducted with an emphasis on identifying specific future directions for technical research (particularly within the artificial intelligence techniques used when constructing new search engines). This marks contrast with analysis purely focused on finding generalizable patterns within a population. Finally, the audio or video recording was reviewed again to find specific quotes supporting the observed patterns or insights.

4 FINDINGS

To report our findings from interviews with subjects in many overlapping roles, we organize this paper by user profiles rather than by subject or research question. The background of specific subjects is included in the first section for which they are introduced. A summary of our findings can be seen in Table 2.

4.1 Speedrunners

This first user profile section examines one category of stakeholder who would use videogame moment search in a hobby context. Speedrunners are an online community of players who compete to complete games as fast as possible. Tracing its roots to the release of id's *Doom* in 1993, this subculture has seen substantial growth, with an increase in dedicated websites (e.g. NESVideos in 2003, which became TASVideos[7] as its scope extended beyond Nintendo games, and the launch of Speedrun.com in 2014). The charity speedrunning convention Games Done Quick (GDQ[8]) has also been growing in attendance, viewership, and fund-raising since its inception in 2010. Speedrunning combines aspects of human performance and technical achievement. Some speedrunners prefer to keep things as close to unaided as possible, while others are interested in using

any mode of play (including the use of automation) to find the fastest times possible. To understand the needs of this community we spoke with individuals who participate in and contribute to different aspects of the hobby.

4.1.1 Speeedrunner Interview Subjects. Alexis "Protomagicalgirl" Ross is a speedrunner, a member of several online speedrunning communities, and one of the organizers of the GDQ convention. Her Speedrun.com profile contains over a hundred runs across a large variety of games,[9] and she has dozens more on a site dedicated specifically to the popular *Mega Man* game series.[10] For GDQ she has been a member of the pre-show and interview teams since 2016, and she's served as staff for a small number of other speedrunning events. She prefers, when possible to do her speedrunning on original hardware.

Bernd "Aran Jaeger / ED" Huber is a speedrunner and a member of the TASVideos community. His contributions there have primarily been related to the game *Super Metroid*, a popular game released for the Super Nintendo platform in 1994. Instead of running the game himself, he looked for exploitable glitches in games that could be used to produce a faster run, which sometimes involves deducing how a game is using the system's memory, and shared his insights on a wiki dedicated to speedrunning the game.[11]

Stella Mazeika is a PhD candidate in computer science at UC Santa Cruz and a hobbyist speedrunner. She is a moderator of a speedrunning forum, and streams her own runs on Twitch.[12]

4.1.2 What kinds of videogame moments do speedrunners seek? Our speedrunners are interested in sharing individual runs, as well as tips, tricks, and useful glitches. Ross and Mazeika explained that a *moment* in games, for a speedrunner, might refer to a specific technique or exploit, and that the speedrunning community has given these tricks names. One example given by Ross was of trying to share a "frame perfect" trick for killing the first boss in *Castlevania*, or finding another specific named trick that would be useful in completing a particular run (e.g. "bat crit" in *Castlevania* or "teleporter room early" in *Paper Mario: Thousand Year Door*). Ross is interested in a system that can handle many types of classic console games (e.g. *Castlevania*), a small number of indie PC games (e.g. *Cave Story*), and arcade games (e.g. *Tetris the Grandmaster*). Huber described a speedrunner sharing on online forums a moment where a runner was stuck, where the game was crashing, or some

---

[7]http://tasvideos.org/
[8]https://gamesdonequick.com/

[9]https://www.speedrun.com/protomagicalgirl
[10]http://megamanleaderboards.net/index.php?page=runner&name=protomagicalgirl)
[11]https://wiki.supermetroid.run/ED%27s_stuff
[12]https://www.twitch.tv/leggystarscream/

User Profile	Moments Sought (RQ1)	Query Formation (RQ2)	Ultimate Goal (RQ3)
Speedrunners	Individual speedruns, useful glitches and tricks	Screenshots, Short Video, Text (with named tricks and descriptions)	Learn from examples and practice for their own runs
Streamers / Youtubers	Moments that support specific arguments or commentary, moments from personal play	Text (with game specific tags and parameters), Screenshots, Video	Resume play from a moment to explore variations or to skip between game sections
Educators	Conceptual and aesthetic examples, Examples of mechanics	Text (with a controlled vocabulary containing examples of game mechanics), Citations, Screenshots	Resume play from a moment to share experiences of game mechanics with students, or produce classroom video demonstrations
Developers / Designers	Moments linked to player reactions, moments where design tools would be useful, bugs and unexpected behavior	Player Reactions (e.g., excitement, boredom, etc.), Text (with an awareness of moment types, e.g., "tutorial," "boss fight," etc.), Screenshots	Finding "fun" or "pain" points for monetization opportunities, communicating game design concepts within development team, finding additional examples of a known bug
Scholars	Rare or unique moments, examples of superior/aesthetic play, historically important moments	Text (with descriptions of known events, or named mechanics), Complex filtering (e.g., "before 1985"), Citations, Code	Validate the possibility of a moment happening in a given game, resume play from a moment for exploration, discovering what game elements specific code enables

Table 2: Overview of findings for our research questions, organized by user profile.

other rare event was occurring, in order to solicit advice. He also suggested that in the course of searching for new paths through the game, a speedrunner might be preferentially interested in moments that are further away from ordinary play.

4.1.3 How do speedrunners want to identify those moments in a search engine? For all of the speedrunners we spoke to, the first idea suggested was using a screenshot to find an example of the trick they were interested in reproducing. This suggestion was offered prior to the anchoring demonstration video. One subject also considered using a short gameplay video as search input. The second common suggestion was using the community's text name for the trick or technique, perhaps with a few additional qualifiers (e.g., with the addition of a location or time in the game) if searching for a particular variation. Speedrunners also expressed interest in making sure they can filter results to those that are relevant to their community specifically. This might mean restricting results to speedrunners, but also the ability to apply additional restrictions (e.g., human-only, tool-assisted, played by a specific person, or played with specific set of additional restrictions or rules). This last set of restrictions is important because even within speedrunning the meaning of completing a game as fast as possible can vary depending on what is considered necessary for a game to be "complete." For example, the popular speedrunning title *Super Mario World* has 11 categories on Speedrun.com that each define different requirements (including the "0 exit" category that requires reaching

the end of the game while passing through none of the typical level exit points).

4.1.4 What will speedrunners do with retrieved moments? A speedrunner working on improving their own runs by mastering a new technique would want to find an example to follow. This example might be a set of sequential images or a short video clip, ideally paired with a walkthrough of the necessary steps. For some tricks, the exact frame-by-frame sequence of button presses is also important. A save state is even more useful, if a player wishes to practice by skipping directly to the point in the game where the technique is relevant. Some games already have systems in place to extract and share specific moves. For example, Ross described the community-created tool Fumen,[13] which can be used to create abstract representations of certain moves, as a useful training and discussion tool for *Tetris* players. For players that are already used to using such tools for *Tetris* or other games, compatibility would be useful. Mazeika also suggested she would be interested in uploading several of her own runs, to allow side-by-side comparison of the moments they contain.

4.2 Streamers and YouTubers

Videogame players who broadcast their play online, often along with their own commentary, are called streamers. These streamers might be trying to showcase a high level of play, or provide entertainment in other ways such as humorous commentary or

---

[13]http://harddrop.com/fumentool

insightful analysis. Streaming is also a hobby that has seen a great deal of growth since the mid-2010s, both on YouTube and the dedicated game streaming website Twitch.com. There are a variety of different kinds of streams, such as esports, "Let's Plays," and academic commentary. Esports are examples of competitive game performance which differ from speedruns because players are competing with each other directly. Let's Plays are generally more casual playthroughs of games which are often unfamiliar to the player. Academic streams are similar, with the primary difference being how they are framed around analysis rather than entertainment.

Separate from but related to this kind of streaming, there are also web-based videos (pre-recorded, rather than broadcast live) which consist of reviews, textual analysis or cultural commentary about games (e.g. the YouTube channels "Extra Credits"[14] or "Tropes vs Women"[15]). These videos often involve demonstrations of a concept that exists across a large variety of games, or less frequently a deeper dive into one specific game. This type of video was popularized on YouTube, which is still its primary venue, so there is not yet an accepted generic term other than YouTuber for these kind of creators.

4.2.1 Streamer and YouTuber Interview Subjects. Stacey Mason is a PhD candidate in computer science at UC Santa Cruz and an academic streamer who provides commentary and analysis on her blog and Twitch channel, both called Cerebral Arcade.[16] She is one of the founders of ScholarsPlay,[17] an academic streaming channel.

Stella Mazeika, mentioned as a hobbyist speedrunner above, is also a contributor to ScholarsPlay.

Nick Junius, another contributor to ScholarsPlay, is the author of various videogame close reading and textual analysis videos available on YouTube.[18]

4.2.2 What kinds of videogame moments do streamers and YouTubers seek? While the primary focus of streamers is usually broadcasting live play, there are circumstances where seeking out particular moments can be beneficial. An esports player might be interested in reviewing their own play. A streamer or YouTuber who is providing an academic analysis or cultural commentary will be interested in searching for moments that serve to support their argument. For example, in preparing his analysis videos Junius looks for moments in games that he is already very familiar with.

Mason seeks moments she has not played through herself: "What I'd most want is a parameterized save state, which is easier to get than by playing myself." She said that for certain kinds of streams she would like to construct demonstrations of play by "Julia Childsing,"—jumping between prepared locations rather than showing the tedious path between them, analogous to how the television chef would skip past time-consuming phases of a recipe demonstration.

4.2.3 How do streamers and YouTubers want to identify those moments in a search engine? Streamers can vary in how familiar they are with the games they are playing, depending on the type of stream they are doing. They could search for play similar to their

own by uploading their own screenshots or video. Searching for evidence to support a particular argument is more likely to be served well by a text search, if the content of the video has been tagged or transcribed. Anything that adds more semantic information about the moment's content can help support performing this kind of analysis.

For Mason's quick navigation, she wants to use game-specific tags and parameters to identify target moments: "my character at a point in the game with particular set of items or with a particular set of bosses killed, etc." For narrative-driven games "like *Detroit[: Become Human]*," she desires a browsable visualization: "a plot map (a tree map of the plot) would be useful." Such a visualization would be useful for both the streamer user of the retrieval system as well as their live audience, to help understand how the space is being navigated.

4.2.4 What will streamers and YouTubers do with retrieved moments? Streamers want to continue playing from the point of a retrieved moment and demonstrate the gameplay possibilities at that point for their audiences. Mason said "it could be useful as a way to play through variations, particularly for story games to look at variations on a particular moment." For those streamers and YouTubers who make arguments, Mason and Junius suggested that finding moments ahead of time can help produce a variety of supporting examples. Junius commented that "When I wanted to talk about an event two-thirds or three-quarters through [*Mass Effect 3*], and I didn't have a save file and couldn't find one that I wanted. It took me about 15-20 hours or play to get 3-4 minutes of footage." It took this amount of time despite Junius using an existing community resource (specific to the game in question) to find an appropriate save file from which to begin again.[19]

4.3 Educators

Educator, here, refers to anyone who is teaching about the design, history, or interpretation of games. To better understand the needs of this community we spoke to several educators who have taught these topics at the university level.

4.3.1 Educator Interview Subjects. Elizabeth Swensen is a game designer and educator. She has taught courses on game design, both digital and physical, and on interactive media at the graduate and undergraduate level at the University of Southern California and at UC Santa Cruz.

Henry Lowood is a digital curator, scholar, and educator. He has taught undergraduate courses on the history of digital games and graduate level courses on the curation of new media at Stanford University, at San José State University, and at UC Santa Cruz.

Nathan Altice is a scholar and educator. He teaches undergraduate courses on game design and games systems as well as graduate courses on media archaeology at UC Santa Cruz.

4.3.2 What kinds of videogame moments do educators seek? Our educators primarily expressed interest in the moments in games that can serve as examples for the concepts that they are teaching. Swensen suggested that for her lessons she might search for a particular kind of jump in a platforming game, systems for tracking

---

[14]https://www.youtube.com/channel/UCCODtTcd5M1JavPCOr_Uydg
[15]https://www.youtube.com/playlist?list=PLn4ob_5_ttEaA_vc8F3fjzE62esf9yP61
[16]http://cerebralarcade.com/ and https://www.twitch.tv/cerebralarcade
[17]https://www.twitch.tv/scholarsplay/
[18]https://www.youtube.com/channel/UCSunwmLFUsUTfovdQ4JyxXQ

[19]http://www.masseffectsaves.com/

the morality of a player's choices in various popular games, or how different games have handled text-input using analog sticks. Altice discussed sharing moments that convey a particular aesthetic experience. When teaching game design, it can be useful to trace an example in several ways. Swensen suggested that an educator might be interested in the most famous or prototypical example of a mechanic as a prototype to present to their class (e.g. jumping and running in *Super Mario Bros. (1985)*), and Altice suggested that they might wish to connect this example to more current incarnations, which might be more familiar to their students (e.g. jumping and running in *Super Meat Boy*). Altice, Swensen, and Lowood all expressed interest in the ability to trace mechanics backwards to find more prior examples (e.g. jumping and running in *Donkey Kong* or *Frogger*).

4.3.3 How do educators want to identify those moments in a search engine? Currently our educators find the moments they are looking for using text-based search engines (e.g. Google or YouTube). Swensen and Altice both said that for more obscure examples this means hoping someone has already done the work of capturing and cataloguing it, or searching for a moment that they know to be close to the one they are looking for and scrubbing through videos to find it. Swensen indicated that she might reach out to her friends or peers for help. Altice said that he has infrequently done the work to capture a moment himself, but often this is impractical because it is a time-consuming process or because it requires rare specialized hardware. Swensen suggested that educators who are interested in searching for mechanics across games might find a controlled vocabulary helpful (e.g. with keywords for jumping, text input, morality systems, action drafting, etc.). Altice and Swensen both suggested (prior to priming in the interview protocol) that a screenshot might be sufficient to connect to a moment, similar to conducting a reverse image search. Although Lowood is one of the contributors to the Game and Interactive Software Scholarship Toolkit (GISST) [8], a component of the GAme MEtadata and CItation Project (GAMECIP) [9], other interview subjects also expressed a desire for the ability to identify moments using citations.

4.3.4 What will educators do with retrieved moments? Swensen and Altice wanted their students to be able to replay the moments they have found, to have the opportunity to experience those moments for themselves. A video alone can be useful, but, according to Swensen video of an interactive moment is more like a description of the experience than it is like the experience itself. She notes: "Unfortunately, what's amazing there is not what you see but what you feel, and how seamless it is." Altice also mentioned referring in a lecture to a sequence in the game *Silent Hill 2 (2001)* where a player rowed a boat to a lighthouse, using analog sticks as a proxy for physical oars, where the difficulty and pace of this sequence contribute to the game's atmosphere. He lamented that classroom facilities typically do not have the capabilities to facilitate this kind of experience, which would require at minimum some kind of emulation system available to all students. He also expressed a concern about finding a clean example of a moment for classroom demonstration, saying: "One of the ironies of teaching game design is that we have to rely on secondary sources, e.g. a Let's Play with some jackass narrating over." Swensen suggested that prototypical or classic "best" examples (from platforms students may not even

have access to) can be useful for conveying the core of a concept to students, while exploring older examples as a way of tracing the origin of an idea or mechanic that students are developing can provide useful context and suggest interesting alternatives to investigate for student projects. Altice offered that modern examples—from platforms that host games actually played by the students—can make the material more immediately relatable. In either case, it is important that students be able to see and experience concrete examples of illustrative moments rather than just read about or hear descriptions of those moments.

4.4 Developers and Designers

Developers and game designers, here, are those whose work includes the creation of videogames and the tools to support videogame creation.

4.4.1 Developer and Designer Interview Subjects. Ted Lim is a game developer for NetEase NA, where his work has included producing feature design documentation, conducting user tests, collecting player feedback, and designing interactions. NetEase NA is focused on developing free to play mobile games, including games in the massively multiplayer online; action role-playing game; card game; endless runner; augmented reality; and puzzle game genres.

Batu Aymetiz a PhD student in computational media at UC Santa Cruz and a game designer. His research involves investigating the way that games are designed, including work on a variety of mixed-initiative tools that help designers do things like create better tutorials, or balance competitive games. He has designed a variety of physical and digital games for entertainment, education, and artistic purposes.

Elizabeth Swensen, first mentioned in the section on educators above, is also a game designer. Her work has included educational strategy games such as *Application Crunch (2011)*, an Indiecade finalist that teaches students about accessing higher education.[20] Her other educational games have focused on metacognitive development, using strategy to teach historical thinking, self-advocacy, and encouraging career aspirations in STEM fields.

4.4.2 What kinds of videogame moments do developers and game designers seek? Lim, as a developer of mobile and free to play games, was interested in understanding specific moments in the player's life-cycle that cause "fun or pain." He also expressed interest in looking for similar moments in competitor's games. Unlike our other examples, these moments are defined not just within the game itself, but by the context of player reactions. Swensen suggested that a developer would be interested in finding moments in their own game in development that were examples of a bug. Aytemiz, whose work includes developing tools to assist in game development, expressed interest in finding moments where those tools could be useful. One example he gave was a tool to help create better tutorials, which would be well served by finding the first moments in a game where the player needs to use a new skill to progress.

4.4.3 How do developers and game designers want to identify those moments in a search engine? Lim offered that developers are often interested in player reactions that are tied to specific moments,

---

[20]http://elizabethmakesgames.com/application-crunch/

and would like to be able to search by these reactions. One way to obtain these reactions would be linking moments to other data recorded during their own playtests. For competitor's games, Lim also suggested collecting reactions from online play (e.g. in Let's Plays on YouTube, or streaming on Twitch), which would then need to be linked to specific moments from the attached gameplay video. These reactions will vary with player experience, so it would also be useful to have information about the player, particularly including their degree of experience with the game (e.g. a new vs. experienced player). Lim also indicated that, for their own games, developers are often well equipped to define moments at a high level of specificity (e.g. at this point in the level progression, with these items in your inventory, etc.). For tracking down a bug, Swensen suggested searching from a screenshot of an example. Swensen and Aytemiz both expressed interest in using a controlled vocabulary to identify specific kinds of moments (e.g. tutorials, or boss fights) in the course of development.

4.4.4   What will developers and game designers do with retrieved moments? Swensen suggested using moments to find more examples of a bug, to understand and eliminate it. Lim cautioned that this approach would need to compete with or complement existing quality assurance practices. Lim also suggested that linking player reactions to specific moments can help a developer decide what needs to be cut, added, or changed. This could mean finding a moment that players enjoy, figuring out why, and changing the game to make similar moments more likely, or finding a moment that players find unnecessarily frustrating, and eliminating it. Alternatively, he offered that for a free-to-play game both kinds of moments can be opportunities to add monetization. On a different note, Lim also suggested that being able to search for specific moments might be helpful for communication within a team during development to make sure everyone has a common understanding of the team's goals, particularly when an older game is being referenced for inspiration. He also offered that easy familiarity with popular moments and reactions from competitor's games might suggest certain directions for development.

4.5   Scholars

For us, scholars are those whose work with games is primarily critical or interpretive. This includes any category of work that seeks to understand the history or future of games, game platforms, or the people who play them.

4.5.1   Scholar Interview Subjects. Nathan Altice, in addition to his role as an educator mentioned above, is a games studies scholar. This work has included writing about the history of computing, platforms, and culture. He is the author of *I AM ERROR* [1], a book about the history of the Nintendo Entertainment System (NES) which emphasized its cultural role and the impact of its physical affordances and limitations.

Henry Lowood, whose work as an educator was mentioned above, is also a digital curator and scholar. He is the curator for History of Science & Technology Collections and Film & Media Collections in the Stanford University Libraries. His scholarship has included writing about the history of games, machinima, and documenting gameplay as a kind of athletic performance.

4.5.2   What kinds of videogame moments do scholars seek? Lowood and Altice indicated that scholars are interested in the moments that can help them develop a deeper critical look at a game. Altice said that this would include moments from game platforms that are not as well studied, and those that do not have good emulators (e.g. the Bally Astrocade). Lowood said that scholars might be interested in moments that are examples of superior or athletic play, or moments that are part of historical events, and those moments that changed how the game was played. One example of such a moment, in a non-videogame context, took place recently in the famous Go match between (human) expert Lee Sedol and (machine) AlphaGo. A certain moment (Move 37) represented a pivotal event, defined by its relation to societal context: the machine chose an action that would have been rated as next to useless by human master players.[21] The challenge for information retrieval is knowing that the relevance of this moment is not captured in the state of the game board but instead almost entirely in the context.

4.5.3   How do scholars want to identify those moments in a search engine? Scholars indicated they would be likely to use (Lowood said) "99% text based search," including text descriptions of a known event (e.g. "the gnome demonstration" for a specific protest that took place inside the game *World of Warcraft*). Text search might also include description of game mechanics, general or specific (e.g "the rocket jump", or "*Super Mario* zero [minus] world glitch"). Altice also provided several examples of ways that scholars might wish to add complex filtering to their searches (e.g. "all examples before 1985," or even "all horizontally scrolling games with a blue background from before 1985, in all consoles and arcades," "Mario on Yoshi in an underwater level with 8 lives remaining," "all games with bosses," or "vertical scrolling in all games before 1985, across consoles and arcade"), and said that he would also be interested in adding platform specific characteristics to his search (e.g. a specific way the audio processor is being manipulated). For some of this work, Altice indicated there are already communities which have existing standards, e.g. junk headers in NES ROMs, while for other less popular platforms creating a search system might mean inventing these standards. Lowood, and Altice expressed an interest in being able to search for a specific moments from screenshots, and Altice also expressed an interest in being able to take video as search input as well. Altice was familiar with the GISST citation system (and Lowood assisted in developing it), where a persistent URL link can connect to a particular moment. Lowood remarked that enabling that kind of search also means having a repository built around the tool. Lowood also suggested that in rare cases a scholar might want to search with code itself, to discover what mechanics or play the code enables.

4.5.4   What will scholars do with retrieved moments? Lowood and Altice expressed interest in being able to validate a moment from a screenshot. This means being able to determine if it represents a possible, achievable state, and if possible linking it not just to the game it came from, but to a particular version (ROM) of that game. Altice gave the example of taking a screenshot from a magazine, and finding out if it came from an early prototype or variant of a game rather than the released version. Altice had also suggested using

---

[21]https://www.wired.com/2016/03/two-moves-alphago-lee-sedol-redefined-future/

specific platform characteristics for search, and one example he gave for how this would be useful was in early Nintendo games, where ways of manipulating the audio processors which were hallmarks of certain (uncredited) composers. Lowood and Altice both expressed interest in being able to continue play directly to a moment in a format as accurate as possible to the original. This could mean emulation in the web browser, or a downloadable save-state which can be loaded into an emulator. For this second possibility Altice noted that there is no canonical or ideal save state, and each save state is specific to an individual emulator. Lowood also noted that for some kinds of games this will not be possible at all, such as a massively multiplayer online games which requires external servers to play. Lowood expressed interest in knowing what specific code enables, and if possible understanding what happens when you change it. Lowood also suggested that a search system that knows how to find a "rocket jump," if it understands it at a semantic level, could also help a researcher find games with mechanical precursors.

Lowood explained that the scholar's need to make new discoveries is in some ways at odds with search. Anything that is made available by a search has, in some sense, already been found. Following the examples given above, if a scholar conducted a text search using the terms "rocket jump" or "gnome protest" to find a particular moment, any moment the search engine returns must have already have been tagged (in inferred to be intended as tagged) as containing that mechanic, or belonging to that event. For this reason, and because they are often looking for unexpected or unknown connections, scholars are interested in having as much information as possible attached to each moment (including information about the social context of the moment for players and for other scholars), and in the ability to conduct the kind of complex compound searches that might never have been anticipated.

5 TECHNICAL RECOMMENDATIONS

Based on our observations above, this section offers suggestions of next steps for technical research in retrieval for interactive media.

5.1 New Platforms

While some speedrunners and scholars are interested in search on classic game consoles, everyone we spoke with was interested in search for other platforms. Our interview subjects were primarily interested in PC games, and in some cases also interested in mobile (e.g. Android and iOS games). For applications to be relevant to new or in-development games, retrieval systems cannot rely on the existence of pre-recorded interaction traces. In web-search parlance, future systems will need to be able to *crawl* the content in full-scale games for modern game platforms automatically.

While there is ample work on automatically playing videogames [13, Chapter 3], the challenge here is to crawl the contents a wide variety of culturally-impactful games (ideally with little to no game-specific engineering) including those for which there is no obvious score to maximize or winning state to achieve (e.g. *Mario Paint(1992)*).

5.2 Textual Queries and Domains-specific Vocabulary

All types of users were interested in text search, although they had different ideas about how they would use it. Speedrunners,

developers, and educators were all interested in search based on vocabulary that was specific to their community, or developed to support this kind of search. Scholars, streamers, and others were also interested in being able to search with free-form descriptions. Altice's suggested query "Mario on Yoshi in an underwater level with 8 lives remaining" presents the specific challenge of building a system that can understand game-specific distinctions that might not have been considered at the time moments were indexed.

Based on the issues raised during our interviews, Xiaoxuan Zhang (whose work introduced the problem of videogame moment search [17]) completed a master's degree with a project specifically focusing on allowing users to search with natural-language queries [16]. This initial work combined ideas from natural language processing and computer vision to extract several different sense of keywords from screenshots taken from gameplay videos (as downloaded from YouTube). While this system could not understand domain-specific vocabulary, it could identify moments referencing everyday concepts (e.g., horses, sunsets, and selfies) for games that did not mention these concepts in their title and description (or other paratext [12]) or in other writings about them (e.g. Wikipedia articles).

5.3 Videos in the Wild

Developers and speedrunners expressed interest in being able to search using video data in the wild. Most speedrunners share their runs via online video. Developers would want to pair this with player engagement and reactions to find out which of their (and their competitor's) ideas are being enjoyed by their audience. Scholars asked for as much information as possible, and this would serve their interest in understanding moments in their cultural context. Transcripts of live game play streams would also be useful for enabling more effective text searches, by making it possible to search for both on-screen text and for player reaction vocalizations. Many videogames do not inherently possess a textual component, so there is the problem of aligning the distribution of moments in a game with the specific vocabulary familiar to specific classes of users. In the case of streaming applications, there is the additional technical challenge of processing ephemeral data as fast as it is produced. Future research should find ways of making use of bulk gameplay video content including on-screen text, player reaction vocalizations or facial expression, and audience comments on/in the video.

5.4 Personal Data Sources

Speedrunners, streamers, and YouTubers were also interested in a form of introspective search, being able to upload data from their own playthroughs. For a speedrunner or streamer this might be a way of evaluating or sharing their own performance, comparing their play to world records or to their own previous attempts. This is similar to how users may want to search within their own photos or other personal information. Future work in this area should on build past studies of Personal Information Management (PIM) [7].

6 BETTER UNDERSTANDING USER NEEDS

This paper describes preliminary work in the interactive media retrieval domain, using an interview-based qualitative research

method. There are several other methods that previous research in the user-needs analysis community suggest would also be fruitful. For example, researchers interested in book and movie recommendations coded the content of relevant forum posts [4]. Researchers examining image search interfaces supplemented their informal interviews with system designers with a large scale analysis of search logs [2]. Additionally, the interdisciplinarity of game studies suggests a range of alternative methods that could potentially be used to supplement these results [11, Chap. 8], which could include those drawn from a range of humanistic, social science, and design research traditions. Even staying within our current methodology, we believe that we have not reached a saturation point and that additional fruitful technical directions might be discovered with additional interviews.

In the course of preparing this work we had to select specific individuals to contact, and were only able to report based on those who were able to respond to our queries. This means we are missing several important stakeholder perspectives. One of these is the game distributors who manage storefronts where potential players can discover and purchase games, along with their consumers. We can speculate that they might be interested in finding the best possible moments or images to sell games, in measures that might make discovery in the context of a massive catalog of games easier, and perhaps in moments that are a better representation of a game than the developer's own carefully curated (but potentially misleading) screenshots. Other stakeholders that should be considered include game translators, quality assurance testers, and even in-game advertisers. Hopefully future work will provide more grounded answers to how these users would use videogame moment retrieval systems.

7 CONCLUSION

We started by asking: What kind of videogame moments do various types of users seek? Our responses were varied. Speedrunners wanted to see examples of tricks, streamers and YouTubers sought out moments they knew well, educators looked for prototypical examples, developers looked for moments with clear positive or negative engagement, and scholars looked for whole categories of moments.

Next, we asked how people wanted to identify those moments to a search engine. Interestingly, there was a good deal of commonality here, with many user categories interested in using screenshots or text search. When the screenshot search suggestion was offered by speedrunners, it was before they had been presented with the anchoring demonstration. Some users also were interested in search via online video, which has some overlap with search by image. There was some variation in how users wanted text search to operate, but generally that was a matter of wanting a controlled vocabulary, and one that understood terms already in use by the community.

Finally, we asked: What will they do with retrieved moments? Here again we saw a great deal of variety. Speedrunners seek moments for self improvement, streamers and YouTubers seek moments to support an argument, educators seek moments to provide context and examples, developers seek moments to understand player reactions, and scholars seek moments to find interesting and unusual connections.

While we saw a great deal of variety in both the kinds of moments that users are interested in, and in what they want to do with those moments once they have found them, we saw more commonality in the ways that users want to search for moments. This is promising for the development of systems for videogame moment retrieval that seek to serve many types of users.

By offering concrete challenges to be addressed in future technical systems, we hope to anchor technical research in this new information retrieval domain in the user-centered design approach.

REFERENCES

[1] Nathan Altice. 2015. *I AM ERROR: The Nintendo Family Computer / Entertainment System Platform* (first edition edition ed.). The MIT Press, Cambridge, Massachusetts.
[2] Paul André, Edward Cutrell, Desney S. Tan, and Greg Smith. 2009. Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage. In *Human-Computer Interaction INTERACT 2009 (Lecture Notes in Computer Science)*, Tom Gross, Jan Gulliksen, Paula Kotzé, Lars Oestreicher, Philippe Palanque, Raquel Oliveira Prates, and Marco Winckler (Eds.). Springer Berlin Heidelberg, 340–353.
[3] Sean Bell and Kavita Bala. 2015. Learning visual similarity for product design with convolutional neural networks. *ACM Transactions on Graphics (TOG)* 34, 4 (2015), 98.
[4] Toine Bogers, Maria Gäde, Marijn Koolen, Vivien Petras, and Mette Skov. 2018. "What was this Movie About this Chick?" A Comparative Study of Relevance Aspects in Book and Movie Discovery. In *International Conference on Information*. Springer, 323–334. https://doi.org/10.1007/978-3-319-78105-1_36
[5] Kathy Charmaz. 2006. *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.
[6] Sharon Favaro Ince, Christopher Hoadley, and Paul A. Kirschner. 2018. A Study of Search Practices in Doctoral Student Scholarly Workflows. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval (CHIIR '18)*. ACM, New York, NY, USA, 245–248. https://doi.org/10.1145/3176349.3176877
[7] Amalie Enshelm Jensen, Caroline Møller Jægerfelt, Sanne Francis, Birger Larsen, and Toine Bogers. 2018. I Just Scroll Through My Stuff Until I Find it or Give Up: A Contextual Inquiry of PIM on Private Handheld Devices. In *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval (CHIIR '18)*. ACM, New York, NY, USA, 140–149. https://doi.org/10.1145/3176349.3176394
[8] Eric Kaltman, Joseph Osborn, Noah Wardrip-Fruin, and Michael Mateas. 2017. Getting the GISST: A Toolkit for the Creation, Analysis and Reference of Game Studies Resources. In *Proceedings of the 12th International Conference on the Foundations of Digital Games (FDG '17)*. ACM, New York, NY, USA, Article 16, 10 pages. https://doi.org/10.1145/3102071.3102092
[9] Eric Kaltman, Noah Wardrip-fruin, Mitch Mastroni, Henry Lowood, Greta De groat, Glynn Edwards, Marcia Barrett, and Christy Caldwell. 2016. Implementing Controlled Vocabularies for Computer Game Platforms and Media Formats in SKOS. *Journal of Library Metadata* 16, 1 (Jan. 2016), 1–22. https://doi.org/10.1080/19386389.2016.1167494
[10] Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. 2006. Content-based Multimedia Information Retrieval: State of the Art and Challenges. *ACM Trans. Multimedia Comput. Commun. Appl.* 2, 1 (Feb. 2006), 1–19. https://doi.org/10.1145/1126004.1126005
[11] Frans Mäyrä. 2008. *An Introduction to Game Studies: Games in Culture*. Sage, London.
[12] Annika Rockenberger. 2014. Video Game Framings. In *Examining Paratextual Theory and its Applications in Digital Culture*, Nadine Desrochers and Daniel Apollon (Eds.). IGI Global, Hershey, PA, 252–286.
[13] Georgios N. Yannakakis and Julian Togelius. 2018. *Artificial Intelligence and Games*. Springer. http://gameaibook.org.
[14] Zeping Zhan, Batu Aytemiz, and Adam M Smith. 2019. Taking the Scenic Route: Automatic Exploration for Videogames. In *Proceedings of the Second AAAI Knowledge Extraction from Games Workshop (KEG-19)*.
[15] Zeping Zhan and Adam M Smith. 2018. Retrieving Game States with Moment Vectors. In *Proceedings of the AAAI Knowledge Extraction from Games Workshop (KEG-18)*.
[16] Xiaoxuan Zhang. 2019. Navigating Moments in Videogames. (March 2019). Project report for MS in Computational Media.
[17] Xiaoxuan Zhang, Zeping Zhan, Misha Holtz, and Adam M. Smith. 2018. Crawling, Indexing, and Retrieving Moments in Videogames. In *Proceedings of the 13th International Conference on the Foundations of Digital Games (FDG '18)*. ACM, New York, NY, USA, 16:1–16:10. https://doi.org/10.1145/3235765.3235786