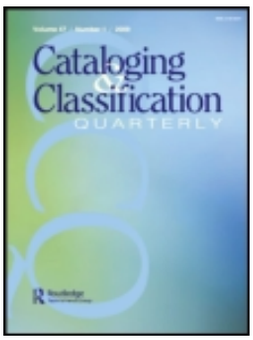

Cataloging & Classification Quarterly

ISSN: 0163-9374 (Print) 1544-4554 (Online) Journal homepage: http://www.tandfonline.com/loi/wccq20

# Full Steam Ahead: A Conceptual Analysis of User-Supplied Tags on Steam

Travis W. Windleharth, Jacob Jett, Marc Schmalz & Jin Ha Lee

**To cite this article:** Travis W. Windleharth, Jacob Jett, Marc Schmalz & Jin Ha Lee (2016) Full Steam Ahead: A Conceptual Analysis of User-Supplied Tags on Steam, Cataloging & Classification Quarterly, 54:7, 418-441, DOI: 10.1080/01639374.2016.1190951

**To link to this article:** http://dx.doi.org/10.1080/01639374.2016.1190951

CATALOGING & CLASSIFICATION QUARTERLY
2016, VOL. 54, NO. 7, 418–441
http://dx.doi.org/10.1080/01639374.2016.1190951

Routledge
Taylor & Francis Group

# Full Steam Ahead: A Conceptual Analysis of User-Supplied Tags on Steam

Travis W. Windleharth[a], Jacob Jett[b], Marc Schmalz[a], and Jin Ha Lee[a]

[a]Information School, University of Washington, Seattle, Washington, USA; [b]School of Information Sciences, University of Illinois at Urbana-Champaign, Champaign, Illinois, USA

**ABSTRACT**
This article describes a conceptual analysis of user-generated tags applied to video games in the Steam video game distribution system. The research team scraped all user-generated tags available on Steam and then conducted a conceptual analysis of the tags, sorting them into categories and comparing them to the current version of the Video Game Metadata Schema. This analysis allowed the team to identify new metadata elements and terms useful to game players. We present a discussion covering the major issues in organizing the terms, as well as the implications for the future work in the area of video game metadata.

**ARTICLE HISTORY**
Received February 2016
Revised May 2016
Accepted May 2016

**KEYWORDS**
Video games; user tags; Steam; metadata; video game metadata schema

## Introduction

The consumption of video games is growing rapidly among the public and has become a major source of media consumption, with 42% of Americans playing video games at least three hours per week.[1] Video games have grown into a large industry, achieving $22.41 billion in sales in 2014.[2] Despite the popularity of games and their global scale of production, systems that enable game users to search and browse for titles remain fragmented and incomplete. One reason for this is a gap between the information users desire when searching or browsing for games and the common terms employed to describe that information. This study is one effort to reduce that gap—to help researchers understand the tags and terms game users find useful—in an effort to inform game metadata systems. The objective of our study is to compare user-contributed Steam tags with metadata elements in the Video Game Metadata Schema (VGMS), developed by the Game Metadata Research (GAMER) Group at the University of Washington and Seattle Interactive Media Museum (SIMM), to inform revisions to it.

Steam is one of the most popular online digital game distribution platforms. It offers over 4,500 computer games, and serves over 100 million active users.[3] Steam utilizes a tagging system, whereby users can generate descriptive "tags" and apply

**CONTACT** Travis W. Windleharth travisw@uw.edu Information School, University of Washington, 4200 Mary Gates Memorial Drive, Seattle, WA 98105, USA.

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

them to game titles, in order to organize and describe them. This article reports the results of a conceptual analysis of Steam's user-generated tags.

The research team collected all of the tags applied to each publicly accessible game on Steam—which resulted in 294 unique tags—and organized them into categories as part of the conceptual analysis. This process allowed the team to increase understanding of what types of game metadata are important to users, and to gain insight into how these tags and terms are being used. The results of this conceptual analysis are also compared and contrasted with existing work in user-centered video game metadata, represented by the VGMS.

To this end, we sought to answer the following research questions:

1. Which dimensions of information are represented in Steam user tags?
2. What kinds of challenges exist in organizing Steam user tags?
3. How well do the Steam user tags align with the GAMER Group's VGMS? Based on the analysis of tags, what information can be added to the VGMS?

## Background and related work

Starting in the early 2000s, a number of websites in a variety of domains started to offer social tagging features similar to those used by Steam today. These features enable users to describe and annotate different information objects or resources using free-form tags.[4] Some examples include del.ici.ous (web bookmarks), Flickr (images), Amazon (commercial products), Last.fm (music), LibraryThing (books), and CiteULike (scholarly publications). While some of these websites have abandoned or limited the use of tags (e.g., del.ici.ous and Amazon), the practice of user tagging is still thriving and being adopted in new domains. User tags represent both opportunities and challenges for catalogers. Several prior studies on user tags compared user generated information to officially curated metadata and found a number of differences regarding the specificity, recency, and uniqueness of the data.[5,6,7]

Despite the common quality concerns with user-generated tags due to lack of oversight and control over tagging behavior, many still perceive tags as a potentially useful source for creating metadata in a cost-effective and efficient manner.[8] When we consider the ever-growing volume of digital resources on the Web,[9] it appears that a number of advantages can be achieved for metadata and vocabulary development and enrichment by utilizing user-generated metadata such as these tags.

### *Video game metadata schema*

Since 2011, the GAMER Group at the University of Washington Information School, in partnership with the SIMM, has been developing a formal representation and description framework for video games and interactive media. The latest version of the VGMS contains 56 elements, many of which are accompanied by controlled vocabularies developed through user studies and evaluation.[10] The

GAMER Group and SIMM continue to actively develop the schema to improve its accuracy, consistency, and applicability. Prior studies employed various methods—including semi-structured interviews and online surveys—to discover user needs and behaviors pertaining to video games.[11] Their direct feedback on the schema, cataloging exercises using the VGMS and its controlled vocabularies, and card-sorting activities have helped the GAMER Group build vocabularies that closely reflect user understanding of the domain. As an expansion of such efforts, we sought to collect and analyze user-generated tags that describe video games in order to discover additional information about what users find important to describe about video games.

Steam's user tags are an ideal resource for this analysis for a number of reasons. First, Steam tags are contributed to by a substantially larger population than those involved in previous VGMS-related user studies. Second, Steam games are digitally distributed and, while the VGMS has been revised multiple times with digitally distributed games in mind, it was initially built with an emphasis on physically distributed games.[12] The most recent sales data from the Entertainment Software Association show that in 2014 digital distribution of games surpassed physical distribution for the first time, with 52% of games digitally distributed.[13] Identifying the missing dimensions of information from analyzing Steam tags, and incorporating those aspects into the next version of the metadata schema will help improve its applicability to digitally distributed games.

### *User-generated tags*

Unsworth describes the set of activities that all scholars engage in as they conduct their research as *primitives* and counts the activity of *annotating* as one of them.[14] Scholars are not the only people who employ these primitives as they engage with text and other media. Ordinary people frequently affix labels, especially genre labels and adjectives, to media as an aid when describing them to others or to help them locate an object of interest again. Tagging, a labeling practice that has emerged as the definitive method of collective classification on the web, is a vitally important annotation task.

Golder and Huberman describe tagging as "the practice of allowing anyone—especially consumers—to freely attach keywords […] to content."[15] The practice of tagging has become so ubiquitous as an annotation practice on the web today that the W3C's Web Annotation Working Group (WAWG) has developed a specialized vocabulary identifying tags serving as the content-bearing portion of web annotations.[16]

From this web-centric perspective, tagging can be seen as a fundamental way in which users interact with information objects. Tagging bypasses traditional notions of library practice, which mediates between users and information objects through the formalized indirections of descriptive metadata records and their attendant controlled vocabularies. Steam's service offers a

hybrid of library-centric and web-centric practices: an official, extensive catalog data supplemented with user tags.

As Veres points out, these kinds of crowdsourced metadata "can challenge the role of established taxonomies."[17] For this reason, many studies have examined the level of agreement between official curated metadata and crowdsourced user tags with a goal of evaluating the quality of these tags. For instance, Carman conducted a quantitative review of user tags from LibraryThing—an online personal book cataloging service—and compared them to Library of Congress subject headings in terms of agreement and divergence.[18] Carman found that while "cataloguers are good at identifying the basic categories to which most books belong," they are less effective at identifying more specific details and, as a result, some access points end up being neglected. In a parallel to this idea, Spiteri concludes that using broader facets (or categories) to help clarify the context of tags can "create more efficient and structured browsing mechanisms for tags. If these benefits are to be realized, however, it is important that methodological frameworks be designed by which to create, evaluate, and monitor facets."[19]

In a study of terms assigned to academic articles, Kipp also found pronounced differences between keyword sets generated by users, authors, and official indexers ("intermediaries") on CiteULike.org. These variations demonstrated differences in "views of the concept space of the articles analysed."[20] User-generated tags were often related to author and intermediary tags, but employed synonymous terms or newer, emergent terminology not present when the items were added to the catalog. Kipp concluded that user tagging could provide additional access points, supplementing those provided by traditional controlled vocabularies at a lower cost of production as well as presenting novel associations for users. Spiteri concludes that "facets may help clarify the meaning or context of tags, and create more efficient and structured browsing mechanisms for tags. If these benefits are to be realized, however, it is important that methodological frameworks be designed by which to create, evaluate, and monitor facets."[21]

Williams also analyzed the user-contributed tags for a set of identical catalog items in two different library systems.[22] She found large differences between the two sets of tags, which she attributed to the unique traits of each library environment. These traits included: the specific affordances of the catalog software, the involvement of the user base, and the extent to which official catalogers performed curation of their respective user-generated folksonomies. Williams' study highlights the usefulness of tags in complementing existing descriptive frameworks.

Merholz compares social tagging to "the foot-worn paths that appear in landscape over time."[23] He goes on to say,

> Called "desire lines," these trails demonstrate how a landscape's users choose to move, which is often not on the paved paths. A smart landscape designer will let wanderers create paths through use, and then pave the emerging walkways, ensuring optimal utility.

His point is that there are essentially three ways to deal with such paths: ignore them, rope them off and enforce the plan, or pave the worn paths and add them to the recognized system of walkways. In our work, we aim for the third approach, seeking to identify the desire lines through analysis of user-generated tags to help us make decisions regarding changes to the VGMS.

## Study design and methods

While Valve Corporation, which owns and maintains Steam, does make a list of "popular tags" available on the Steam website, the team desired a comprehensive list of all tags in use, as well as information on which tags were applied to which games. The Steam database provides a view of one particular sample of games, all of which are games only playable on personal computers. Therefore the key limitation of this data source is that any tags that might be uniquely relevant to mobile, handheld, or console games might not be represented in this study. The full desired data set was collected on March 11, 2015 through a series of custom Ruby scripts, which used XPath queries to analyze the HTML returned by Hypertext Transfer Protocol (HTTP) calls to the site. These scripts used two different methods—scanning the site's available lists of game titles as well as a brute force test of all IDs in the range that appeared in the former scan—to find all games available through the website and record their user-generated tags. From this data, a total of 294 unique tags were identified.

The tags were organized through a card-sorting activity using Trello, a free web-based project management application. We employed conceptual analysis as a guiding method for the card-sorting activity. Best known as a method employed by analytical philosophers, conceptual analysis is the practice of breaking complex concepts down into their constituent concepts. This kind of analysis facilitates the discovery and description of ambiguous, context-sensitive, and emergent terms. It has been used in past studies to examine the concept of *information* and the development of ontologies.[24,25] We used it to group and categorize the tags harvested from Steam.

This study is not without its limitations. Because users only tagged games that were available on Steam, not all types of video games are fully represented in the dataset. Although Steam has an impressive game library, a significant number of console-based video games are missing. Thus, the dataset lacks any descriptors for video games that are unique to consoles. Nevertheless, the findings provide useful insights and serve as a benchmark for evaluating our metadata schema and associated controlled vocabularies, making them more comprehensive by expanding the scope to cover digitally distributed games. Another potential limitation is that user tags go through an initial filtering process by Steam, which removes inappropriate and offensive tags. Therefore, some users' perspectives may not be represented in the dataset.

## Data and discussion

The final data set obtained from the Steam website was acquired on March 11, 2015, and contained a total of 294 tags, applied to the 4,495 games. With the dataset in hand, we went through the list of tags one at a time, sorting them into categories. In an effort to understand the context and use of specific tags, the team frequently searched for all games associated with a specific tag (an example of such a search appears in Figure 1) during the sorting process. These searches allowed the team to develop a more informed interpretation of tag context and use. At the end of the process, the result was a set of 29 categories into which all the tags were sorted. The number of tags per category ranged from one (e.g., platform) to 95 (e.g., gameplay genre). Most of the categories mapped closely to the VGMS's existing categories with some exceptions (further discussed in "Recommendation for video game metadata schema").

During the discussion of tags in this article, Steam user tags will be presented as they are on the site, despite inconsistencies and spelling and grammatical errors. For example, "e-sports" is the only un-capitalized tag on the site, and title case is inconsistently enforced (e.g., "Real-Time with Pause," "Based On A Novel," etc.). Punctuation such as hyphenation ("Third Person" and "First-Person") and apostrophes ("1980s" and "1990's"), and conjunctions (e.g., "Hack and Slash" and "Point & Click") were also inconsistently formatted.

### *Types of information represented in Steam tags*

A total of 29 categories emerged from the conceptual analysis. Rather than provide details on each and every one of the categories, we have selected a few of the most interesting ones for our discussion. These were all categories that possessed a broad selection of tag terms or that triggered a great deal of internal debate among the

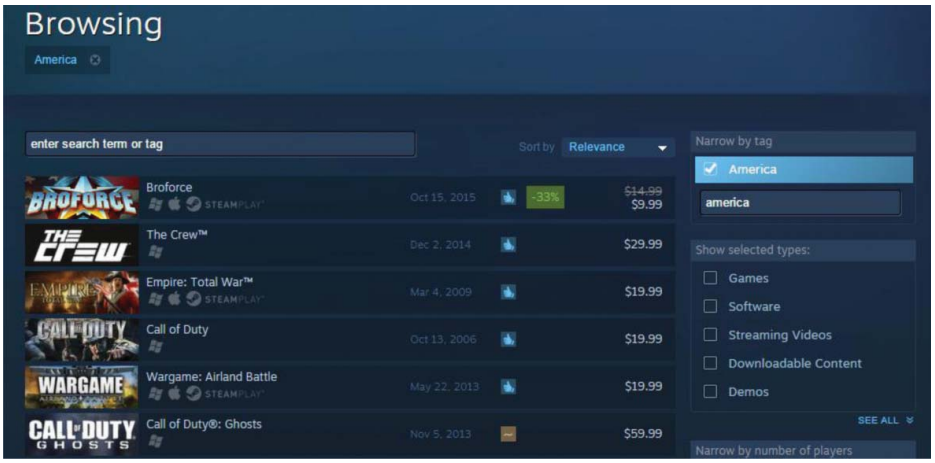


**Figure 1.** Search results showing games users tagged with the term "America."

authors as we sought to find the appropriate conceptual container for each user-provided term. The categories are organized into three sections focusing on tags describing different aspects of video game information: features of games, narrative/presentation of games, media types, and other categories of information.

Two lists are provided in the Appendix of this article to show all of the Steam tags, and the categories into which they were organized. Table A1 shows Steam tags sorted by categories that are currently represented in the VGMS. Table A2 shows the list of Steam tags organized into new categories.

### *Features of games*

***Gameplay genre.*** The VGMS defines the gameplay genre element as "the overall nature of a video game's interactivity based on its objectives, types of rules, distinctive characteristics, modes of action, and manners of gameplay."[26] Of all the categories of Steam tags that we identified, tags that aligned with gameplay genre were most commonly represented. This may be because Steam implements a fairly simple list of 13 broad gameplay genre categories to catalog each game. Steam users expanded this categorical information by adding tags that refined a category's granularity or that highlighted other gameplay genres present in the game. Ten of the 96 terms categorized under gameplay genre overlapped with Steam's internal gameplay genre categories. Granularity within a particular gameplay genre seemed to depend on its popularity as well as the availability of games that fell within its scope.

Steam tags proved to be a good source for obtaining potential lead-in terms as we were able to find several terms that were nearly synonymous to each other (e.g., "Visual Novel," "Interactive Fiction," "Choose Your Own Adventure"). We were also able to find genre terms that are relatively new and thus are often not included in commercial game-related websites and online game stores, such as "MOBA," "Parkour," and "Physics." While most of the terms could be clearly categorized as gameplay genre terms, some were more contentious; examples include "Lemmings" and "Walking Simulator."

In case of "Lemmings," the term seemed to have been applied to games that feature similar kind of gameplay mechanics as the original *Lemmings*, published in 1991 (e.g., *Monomino, Flockers, So Many Me*, and *Clones*). In *Lemmings*, the objective is to guide a group of little creatures called lemmings safely to the exit while overcoming a variety of obstacles. We verified that all of the games tagged with the "Lemmings" descriptor shared this common objective and distinctive style of gameplay. However, the authors had a hard time agreeing that "Lemmings" actually belonged in the gameplay genre category, due to the fact that the term was relatively unfamiliar to some of them and was associated with a specific video game rather than broad characterization of a style of gameplay. After much discussion, we came to the conclusion that certain games have special significance in the history of gaming due to their novelty and popularity, and thus become the

prototypical examples for explaining a group of games with similar characteristics. A similar example of this phenomena can be seen in the tag "Metroidvania," a portmanteau coined to refer to side-scrolling platformers with the free-roaming mechanic, inspired by the games *Castlevania* and *Metroid*.[27]

***Mechanics and input.*** Mechanics are the specific rules and behaviors of the components of a game that defines how the game is played. In chess, for example, the movement of pieces on an 8 × 8 grid, the manner in which the different pieces move, and rules for capturing the opponent's' pieces are all mechanics of the game. The concept of mechanics is central to all games, including video games, and is one of the three components of the Mechanics, Dynamics, and Aesthetics (MDA) framework for analyzing video games.[28] Due to the interactive nature of video games and the use of multimedia, a variety of different types of mechanics are possible. Currently, there is no element for describing mechanics in the VGMS, although a considerable amount of discussion has taken place about adding one.

There are a number of Steam tags that represent mechanics, which were sorted into this category with other elements that did not immediately appear to fit the VGMS. Examples of Steam tags representing mechanics include "6DOF" (six degrees of freedom), "Match 3," and "Economy." These three examples represent different types of mechanics.

The "6DOF" tag names a mechanic that describes how movement occurs in the game world, and several other tags seem to be related, such as "Grid-Based Movement," "Hex Based," and "Parkour." The number of these tags seems to indicate that movement mechanics are important to users. Several tags appear to represent puzzle mechanics, or interactions with elements intended to complete puzzle-like activities.

The "Match 3" tag describes a common puzzle mechanic where players try to match three similar items in a row, and is used in in games like *Bejeweled* and *Puzzle Quest*. Another example of a puzzle mechanic tag is "Inventory Tetris," which is a mechanic where users optimize object placement in some sort of storage or grid space.

The "Economy" tag is related to the key game concept "Resource Management" and the idea of in-game inventory and items. The tag appears to represent the presence of a working in economy in the game to facilitate the transfer currency and items within the game world. Many video games have in-game economies, and many have equipment, items, and objects that can be found, purchased, or taken. Steam users are tagging games based on these inventory and resource mechanics.

In addition to tags that represent mechanics within the game, we identified a cluster of closely related terms describing how the game is played relating to its inputs. Examples of machine "input" items among the Steam tags include "Controller" (using a handheld controller to interact with the game), "Point & Click" (using a mouse to click on items on the screen to interact with the game), and

"Mouse only" (using a mouse is the exclusive method to interact with the game). These input terms appear to be closely related to game mechanics.

***Progression.*** This was a newly added element in version 3.0 of the VGMS. While the previous versions of the schema had the pacing element, a new element was necessary to express what is represented through user tags such as "Open World" and "Linear." These tags refer to branching options or user choices afforded in the game rather than other temporal information. This element is defined as "a description of how the player progresses or advances through the video game."[29]

A key point of discussion involved distinguishing between types of progression in the game. The VGMS "linear" progression denotes that there is one path through the game, and all players progress through each element of that path in turn until completing the sequence. As a Steam tag, it appears that "Linear" is used in a similar manner, in that the game unfolds along a single set path. For instance, the game *Final Fantasy XIII* is the only item from the franchise tagged with "Linear" by users on the Steam website, as the game is notorious for differing from other Final Fantasy games in that progression is completely linear.

The terms "Open World" and "Sandbox" are also employed as tags on Steam, and also seem to bear some relation to how a game progresses or unfolds. Based on our examination of the games marked with these tags, there did not seem to be clear criteria for applying one or both of these tags; we observed no coherent pattern other than the games noted have some degree of freedom afforded to the player in exploring in the world.

The VGMS includes both terms as a part of its controlled vocabulary (CV) for gameplay genres. Within the Gameplay Genre CV, "Open World" is used as a lead-in term for the preferred term "Sandbox." After much discussion, the authors agreed that the two terms in fact convey different meanings. "Open World" is a world in which players are free to explore with few physical or programming limitations as to what areas and spaces in the game the player may move. A "Sandbox" game is similar to Open World in that the player has many affordances, but is given freedom to use many tools, techniques, and mechanics to complete more actions in the game space. As such, "Open World" refers to freedom and affordances in space, and "Sandbox" refers to freedom and affordances in game mechanics and creativity.

### Narrative and presentation of games

***Narrative genre.*** Tags that described the narrative genre of the game were the second most common kind represented. The narrative genre "describes the type of game world and plot, similar to literary genres such as science fiction and fantasy."[30] The strong narrative component of the vast majority of video games makes this category an important factor for their representation through descriptive metadata. Modern literary genres such as "Science Fiction," "Horror," and

"Action-Adventure" are well represented in the video game medium by such games as *Kerbal Space Program, Left 4 Dead 2*, and *Batman: Arkham Asylum*, respectively.

Like gameplay genre and several of the other categories, the tags categorized as expressing narrative genre information possessed a broad range of granularity. Super-genres as broad as "Comedy" were noted for such games as *Deponia, The Sims 3*, and *Goat Simulator*. Conversely, some tags evoked extremely specific contemporary sub-genres, such as "Lovecraftian" and "Steampunk." Examples of these last two tags included *Sunless Sea* and *Bioshock*, respectively.

The notion of genre itself—especially with regard to notions of film genre, literary genre, and music genre—remains a rather poorly understood concept. The category of "narrative genre" is intended to separate the notion of story from the specifics of the media used to express it, as all too often medium-specific factors are conflated into contemporary genre definitions. Thematic, mood-specific, and setting-specific factors are also frequently conflated or deeply integrated into traditional library-centric notions of genre, as can easily be seen in the various genre definitions put forth in Saricks' work on readers' advisory.[31,32]

Video games are a complex medium that employ images, sound, and text to interactively tell stories, which has created a growing need to tease apart information specific to all of the factors that are so frequently conflated into genre. Our attempt to follow this principle has raised some interesting questions regarding genres which, while we do not have the space to explore here, we plan to engage more thoroughly as we develop a controlled vocabulary for narrative genre.

Among the questions posed are:

- Does the tag "Detective" represent a sub-genre of the mystery genre or is this a theme used in the game?
- Does the tag "Dark Humor" represent a sub-genre of the comedy genre or is it evidence of the game's mood?
- Does the tag "Post-apocalyptic" describe a sub-genre of science fiction or is it communicating something with regards to the game's setting?

***Theme, setting, and mood.*** A video game theme is a descriptor denoting a set of objects, concepts, and events that represent a main subject in the game narrative. In the VGMS, the definition of theme is adapted from the Getty Art & Architecture Thesaurus (GAAT), and is defined as "a common thread, motif, subject, or idea that recurs in the game."[33] Some examples include "zombies," "demons," and "conspiracy theories."

During the sorting and analysis of the theme tags, we discovered that many of the terms used in the tags were frequently ambiguous and clearly possessed multiple meanings. We also noted that many of the tags described different dimensions of theme, leading us to hypothesize that there may be multiple facets to theme.

A group of six candidate dimensions of theme emerged from our analysis. One dimension captured themes that best described recurring visual motifs, such as

"Blood," "Gore," and "Nudity." A second facet focused on activities, including behaviors and occupations, such as "Trading," "Mining," and "Hunting." A third cluster of terms centered on kinds of entities, such as "Robots," "Zombies," and "Aliens." The fourth facet addressed central concepts or ideas that drove the game's narrative forward. These included "War," "Dystopian Futures," and "Conspiracy." The fifth facet related to some kind of domain or subject matter, such as "Science," "Education," or "Agriculture." The last facet described a temporal coverage in the video game content, or gameplay mechanics that were representative of those decades, including terms like "1980s." The unifying factor behind all of these themes is that they are entities, motifs, objects, ideas, and subjects that are central to the narrative or the visual style of the game, that help set the tone of the game, and/or that drive the game's story forward.

We experienced a great deal of disagreement when trying to determine if specific tags were representative of themes or settings. This difficulty arose because theme and setting are frequently intertwined, further complicating an already complex analysis. For instance, our first response to the tag "America" was to consider it as a setting. However, further examination of the games tagged with "America" revealed that it was used in a wide variety of ways, implying a great deal of breadth and versatility in how the term is understood and used by game players. For instance, the theme of America engaged in warfare or American patriotism is prominently represented among games such as *Call of Duty: Modern Warfare* and *America's Army*. The "America" tag is also used for games that users appear to believe represent themes that are closely aligned with ideas about America, such as guns (*World Guns: Gun Disassembly*), hunting (*Cabela's African Adventures*, which is actually set in Africa), and the mafia (Mafia II). It is interesting to note that other countries and geographical regions do not have tags representing them—*Cabela's African Adventures* is only tagged "Action," "Hunting," and "America"—but we must leave a deeper exploration of the semantics of thematic tags to a future publication. As illustrated by cases like this, Setting is defined in VGMS as "the type of world, location and time period in which the video game takes place."[34] Steam tags categorized into "setting" represent the primary periods and places where the narrative of the game appears to occur, such as "Rome," "Medieval," or "World War I." In some cases, they refer to a state of affairs, such as "Post-apocalyptic." In the VGMS, three aspects of setting are separately described: world, place, and time period. In Steam tags, one type of setting tag names historical time periods in which games take place, such as "Medieval." This tag carries with it implied themes such as knights, castles, and European visual motifs. The tag "Rome" seems to name a specific geo-political entity, but all uses of the tag are for games set during the Roman Empire, another historical time period, and some have additional locations. *Rome: Total War* is a historical war simulation that takes place "from Macedonia to Greece and into the heart of the ancient Persian Empire"[35] and has "Rome" as a setting tag. None of the Steam tags, however, represent the third type of setting in VGMS: geographical locations linked to specific kinds of environments, such as "Swamp" or "Ocean." This

conspicuous absence may be a clue to reevaluate the value and use cases for this type of setting information.

Like theme and setting, mood is also a vital aspect of narratives. The VGMS defines mood as "the pervading atmosphere or tone of the video game which evokes or recalls a certain emotion or state of mind."[36] Again, the mood of the game can stem from multiple aspects, including the visual style or presentation (e.g., "Cinematic," "Atmospheric," "Realistic"), setting (e.g., "Futuristic," "Gothic"), or game action (e.g., "Competitive"). As a result, several of the mood terms were also categorized as visual style, setting, or theme; for instance, the tag "Political" was categorized under both mood and theme.

***Tropes.*** In video games as in literature and film, tropes are "generally recognizable narrative devices or conventions, which rely on culturally mediated expectations in order to expedite the development of events, characterization, or narrative."[37] Tropes are utilized to quickly develop characters and establish plot lines in the game narrative. An example of a trope in the VGMS is "the chosen one," a character "that has been chosen by divine forces or their people to make a difference in the world." Along these lines, Steam users employed three tags that represent common character tropes: "Silent Protagonist," "Villain Protagonist," and "Female Protagonist."

The tags indicating these tropes on the Steam website map closely to similar terms in the VGMS controlled vocabulary for tropes. There are two categories of tropes in the controlled vocabulary: character tropes and narrative tropes.[38] We were able to map all three of the tags indicating tropes on Steam to character tropes. The characters in the games tagged with the term "Silent Protagonist" (e.g., the protagonist of *Half-Life*) are similar to those with the definition "silent hero" in the controlled vocabulary. The main characters in games tagged with the term "Female Protagonist" (e.g., the protagonist of *Portal 2*) closely match the definition of the "strong female lead" trope from the controlled vocabulary, defined as "a strong, independent female character that is central to the story." The Steam term "Villain Protagonist" (e.g., the protagonist of *Overlord II*) maps closely to the VGMS term "anti-hero."

It is notable that there are only three Steam tags that are classified as tropes in the context of the VGMS, whereas the VGMS's Trope CV defines 90 character tropes and 26 narrative tropes, for a total of 116 defined tropes. One potential reason for this is that many of the tropes in the schema intentionally describe racial and gender stereotypes, and other potentially offensive plot devices whereas these tags would most likely be removed through Steam's filtering process. Some examples of potentially offensive tropes from the VGMS's Trope CV include "the token black person," "the fat bastard," and "the dumb blonde."

***Visual style.*** The visual style element in VGMS is defined as "the predominant and recognizable visual appearance of a video game as originally intended by its

creator, and/or determined in the context of creation."[39] Many of the tags categorized as visual style were also sorted into other categories, such as mood (e.g., "Surreal," "Cinematic," "Cute," "Dark") or media (e.g., "Anime," "Cartoon"). The separation of visual style and mood, in particular, seems challenging. Considering that the use of particular visual styles is one of the most effective ways of creating a certain mood or tone in a game, separation of the two categories is extremely difficult. Additionally, we determined that two Steam tags were also overlapped with regards to visual style and a game's mechanics at the same time: "Hex Grid" and "Grid-Based Movement."

Our examination of the other visual style tags indicated that users were trying to capture multiple dimensions of visual style. These dimensions broke down into five broad categories:

1. Visual looks or appearances, describing a variety of visual styles created by the varying use of colors and shapes with a different degree of abstraction (e.g., "Stylized," "Minimalist," "Colorful").
2. Visual mood, describing the mood or tone that can be affected by the way particular visual style is used (e.g., "Cute," "Dark," "Surreal").
3. Visual representation of gameplay mechanics, including tags that are describing some kind of game action or movement within certain visual frame (e.g., "Hex Grid," "Grid-Based Movement").
4. Visual techniques, explaining different kinds of technical methods used to create a particular look (e.g., "Hand-drawn").
5. Visual motifs, such as "Blood" or "Gore," as explained in the previous "Theme" section.

Tags in sub-categories representing visual looks or mood were more common than the others, such as techniques sub-category that contained only one tag. This is consistent with our finding from previous user testing of the VGMS's Visual Style CV, which showed that users did not fully understand and/or were unable to describe different kinds of visual techniques (e.g., low poly, ray traced, rotoscoped).[40]

### *Media types*

Since its purpose is to describe video game media, the media category itself is not a formal category or element of the VGMS. However, a number of Steam tags appeared to invoke particular kinds of media (e.g., "Anime," "Music," "Comic Book," and "Board Game"). The presence of these tags seemed to highlight relationships between the tagged video game and the particular media type indicated by the tag's text.

A variety of such relationships can exist between a video game and these different types of media. For instance, a video game may be accompanied by "Music" or a "Great Soundtrack," and may also be based on a "Comic Book" or made into an "Anime" after its commercial success, and so on. As previously mentioned, the only relationship that was specifically described as "based on" was the tag "Based

On A Novel." In other cases, the tags were simply referring to different types of media objects without clear specification on the relationships between video games and those objects. This led us to create a separate category for media rather than speculating about the different types of relationships.

The tag "Anime" was a particular challenge for us. The complicating factor was determining whether or not the tag was intended to remark on the video game's media type or visual style. Since the context is frequently ambiguous in cases where the video game shares a name with existing anime (film or serial) or manga, it is difficult to definitively determine which category the tag falls into. It may even be the case that the end user's intention is to indicate both media type and visual style at the same time. This last notion is explored more fully in the section "Challenges and issues."

### *Other categories*

The other categories that we did not explore in this article form a long tail of concepts that either simply had less mass (i.e., contained fewer tags) or which we found less controversial. Examples of categories with fewer tags include tools (e.g., "Video Production"), production (e.g., "Indie"), type of ending (e.g., "Multiple Endings"), estimated time for completion (e.g., "Short"), price (e.g., "Free to Play"), corporate body (e.g., "Games Workshop"), additional content (e.g., "Mod"), and platform (e.g., "Arcade"). Examples of categories with numerous tags but that caused less internal debate include point of view (e.g., "Third Person"), dimension (e.g., "3D"), evaluative terms (e.g., "Moddable"), relationships indicators (e.g., "Remake"), and pacing (e.g., "Turn-based").

## *Challenges and issues*

One of the most challenging factors in the tag sorting task, or in any natural language classification task, is interpreting the meaning of a particular word's usage. In philosophy of language problems, such as substitutivity, these issues are frequently examined in terms of whole propositions.[41,42,43] Unfortunately, tags typically take the form of unigrams or bigrams that lack the supporting context that sentences typically provide.

### *Ambiguous and nebulous meanings*

It is the nature of unigrams (i.e., single word strings of text) that they are frequently used to communicate different meanings in different contexts. This is natural; it makes language much more efficient as the overall vocabulary requirements for participants are reduced. This ambiguity is usually alleviated by the context of the sentences and circumstances that particular unigrams are part of. This is not the case with tags (user-generated or otherwise).

Some help can be gained from the essential nature of the objects being tagged, but ambiguities persist. Among the most ambiguous terms that we encountered

during our card sorting exercise were the tags "Abstract," "Adventure," "Blood," "Crafting," "Detective," "Gore," "Historical," "Lego," "Linear," "Open World," and "Political." In each of these cases there was clear evidence—in the form of the games bearing the tags—that the users employing the tags did not agree to any particular authoritative usage. A particular term's meaning was frequently ambiguous, as in the case of "Abstract."

Taking the tag "Abstract" as an example, it was evident that in some cases it was being employed in order to refer to the game's gameplay genre (e.g., as in abstract games such as chess which can be played in the mind alone, with a board and pieces, or on a computer). In other cases, it seemed to refer to the overall visual style of the video game's imagery (e.g., *Mirror Moon EP*). In a third case, the video game *Trauma*, the tag "Abstract" seemed to be employed to describe a thematic element possessed by the game.

During the course of our analysis, we focused on providing a category for every possible meaning for a tag, but the VGMS's controlled vocabularies do not have the flexibility to reuse terms in the same manner that our analysis did. As development of the vocabularies continues, some method of managing term ambiguity and vagueness needs to be explored by the GAMER Group. Linked data practices suggest one manner in which this problem could be resolved, but we leave such a discussion to a future paper.[44,45]

### *Emergent terms*

A handful of tags were new to our vernacular and we found ourselves frequently disagreeing over how these new terms were being employed by the Steam users. The most challenging of these terms was "e-sports." The crux of our debate revolved around whether or not we should consider "e-sports" to be a kind of media in the same manner we were considering music and video to be distinct media types, or if it was an exemplar of a yet-undefined competition related category. The argument for it being a kind of media was extended on the basis that:

1. Games are a kind of media.
2. Sports are a kind of game.
3. E-sports are a kind of sport.
4. Therefore, e-sports are a kind of media.[46,47,48]

An alternate argument that focuses on the use of the term "e-sports" as an indicator of activities or events that involve particular games during particular periods of time can be made. This interpretation would seem to indicate that a new category should be created to represent game-related events and which may include terms such as "e-sports" or "game conventions." Like so many of the other terms we examined, it may be the case that terms like "e-sports" are also ambiguous and have multiple meanings.

Ambiguous word meaning and emergent terms provide a number of challenges for controlled vocabulary development. Ordinarily, controlled vocabulary development calls for the selection of terms that uniquely describe concepts that will help

users select desired materials. If the goal of the vocabulary is to accommodate terms suggested by the users themselves, then a certain amount of ambiguity—and the costs associated with constant cycles of revision—must be accepted as part of its development.

The ambiguity problem is particularly vexing for the precision of information retrieval systems that rely on indexing the controlled vocabularies for optimized performance. While we have remarked upon the difficulties we faced during our card-sorting exercises above, many of the ambiguities that we were dealing with were very subtle, similar to cases of homonyms (e.g., the string "wind", as in, *the movement of the air* or *a twisting motion*) or ambiguous terms (e.g., the Chinese pictogram "羊" that could refer to *goats* or *sheep*).

End users are uniquely suited to determine word meanings based on the context in which they actively employ them. However, we were frequently at a loss for how best to determine what sense of a term was meant by the Steam users, which led us to reuse many of the terms in multiple, disparate categories. This semantic indifference to the sense of a word can lead to underperformance by information retrieval systems that depend on controlled vocabularies that employ ambiguous terms. For instance, if the user searches for "lemmings" will they be satisfied with results that conflate games with *Lemmings*-like gameplay with those featuring small rodents?

## Recommendations, conclusion, and future work

### *Recommendation for video game metadata schema*

Based on our analysis of Steam tags, we recommend adding three elements to the VGMS: Mechanics, User Interactions, and Evaluation.

#### *Mechanics*

Of the suggested additions for the VGMS, the Mechanics category is the one that was most commonly expressed through user tags on Steam. As we noted above, the types of mechanics represented among the tags include both specific game rules and operations, and broader gameplay concepts. Examples include the different methods for interacting with the game state (e.g., "Match 3," "Quick-Time Events"), the design of the game flow or narrative (e.g., "Choices Matter," "Procedural Generation"), and victory and loss conditions (e.g., "Perma Death") were all relevant to Steam users.

#### *User interactions*

Two of the categories that emerged in the conceptual analysis, "user inputs" and "number of players," each describe the interactions of users with the game itself, as well as how they interact with other players that may be engaged with the game. Collectively, these categories can be unified under the broader concept of User Interactions. User inputs such as "Mouse Only" and "Controller" were discussed in an earlier section, and describe how players interact with the game state, and are

related to game mechanics. Tags in the category "number of players," along with the related set of tags indicating types of player interaction, were well represented in the data set, and include such tags as "PvP" (player vs. player), "PvE" (player vs. environment), "Co-op" (cooperative), and "Local Multiplayer." While Steam has official "Multi-player" and "Single-player" descriptors for games, user tags were used to provide an additional level of detail, such as "Team-Based," "Asynchronous Multiplayer," and "Local Multiplayer." Tags such as "Team-Based" represent different ways users can interact with each other within the gaming environment, and impact the structure of those interactions. The tag "Local Multiplayer" also indicates a specific mode of user connection, which is a feature shared by several similar tags such as "Local Co-op" and "Online Co-op." The importance of multiplayer interactions to players suggested by the Steam data is corroborated by recent Entertainment Software Association data, showing the average gamer spends 6.5 hours per week playing with others online, and 5 hours per week playing with others in person.[49] The conceptual analysis revealed that these concepts, important to users, are currently not well represented in the VGMS, and the team recommends adding User Interaction as an element.

### *Evaluation*

Last, while we did not include it in the discussions above due to limited space, there was ample evidence that the VGMS could benefit from the inclusion of a category for evaluative terms. The Evaluation category would afford the VGMS with the flexibility to capture personal judgments that users share with one another regarding replay value, novelty (e.g., "Experimental"), importance of narrative (e.g., "Story Rich," "Lore-Rich"), and similar video game characteristics.

Many of these evaluative tags map well to the concept of core aesthetics (sometimes referred to as appeals, motivations, or gratifications). Hunicke, LeBlanc, and Zubek define eight core aesthetics of a game: (1) sensation, (2) fantasy, (3) narrative, (4) challenge, (5) fellowship, (6) discovery, (7) expression, and (8) submission, representing core reasons why certain people are attracted to playing certain games.[50] For instance, tags like "Story Rich" or "Lore-Rich" could be mapped to narrative, "Great Soundtrack" to sensation, and "Experimental" to expression, and so on. While this information can be more challenging to describe, especially for catalogers who have not played the games themselves, we believe it can be highly useful for selecting or recommending video games, and can be relatively easily sourced from game players and enthusiasts.

Considering our discussion media types, it would seem that such a category would be helpful to capture information indicating relationships between video games and other media. We are not making such a recommendation at this time because such a category would require the VGMS to express ontological relationships to other media, which seems out of scope for the VGMS per se.

In addition to the new elements proposed here, some deep thought regarding the development of CVs is indicated. The employment of ambiguous and emerging

terms is likely to lead to a cascading series of additional complications for the planning, design, and implementation of information retrieval systems designed solely on CVs. Such systems are not equipped to navigate ambiguities by relying on context, leading to additional implementation and maintenance costs for the systems in question as their indexed CVs inexorably drift out of alignment with the vernacular employed by their users.

Contextual problems can be mitigated by indexing and building thesauri, but the internal indexing and thesauri must grow with the vocabulary. This practice places serious burdens on technical infrastructure. Fortunately, such burdens can be alleviated if controlled vocabularies are built according to linked data principles.[51,52,53,54] Using unique Uniform Resource Identifiers (URIs) for each sense of a term is not a new concept in libraries. Sutton and Tennis were among the earliest proponents of library linked data, since it allows computers to treat indexed terms unambiguously.[55] When engineered correctly, vocabularies optimized for machines are just as usable by humans, and may come with system performance benefits.

There are a number of additional benefits that can be realized by using linked data principles. One such benefit is the ability to share the costs of developing and maintaining the controlled vocabulary with a much larger community of stakeholders. Another benefit is the ability to seamlessly change from language to language—a major boon that has been realized in multinational digital libraries like Europeana.[56] Some portions of the array of controlled vocabularies being developed in support of the VGMS have already been realized in the form of linked data vocabularies.[57]

## Conclusion and future work

While Steam appears to be an ideal first step in using folksonomies to expand and refine the GAMER VGMS, it should not be the last. Despite representing one specific subset of games—digitally distributed games—the tags used on Steam cover a broad range of information that apply to all types of games. The study provides significant insight into Steam users' tagging behavior, including which concepts and game descriptors are important to users, how users employ these terms, and how prevalent each tag is among games in the Steam library.

Quantitatively, we found that tags are most frequently used to denote some aspect of the "aboutness" of a game, such as using terms to describe gameplay genre and theme. Tags are also very frequently used to describe aspects about the experience of gameplay, such as "mood" and "visual style." Additionally, the analysis shows that while Steam users are generating tags to denote the commonly described and sought after characteristics of games, they are also using them to highlight something unique or unexpected about a given game, which would not otherwise be explained with the genre label alone.

Observations of Steam's tags reveal some challenges of describing games employing this tagging method. Many tags are duplicated, used inconsistently, or

unclear in their meaning. In addition, users do not always use tags to explain all of the major characteristics of a game, and it is unclear if or how various tags are curated by users or managers. In order to gain a better sense of how these challenges can be overcome we plan to continue exploring additional sources. We also hope to conduct a deeper investigation on Steam tags regarding how Steam users make use of these tags, as well as a comparative analysis of Steam tags in multiple languages to understand cross-cultural similarities and differences in describing games.

## Notes

1. Entertainment Software Association, "Essential Facts About the Computer and Video Game Industry" (Washington, DC: Entertainment Software Association, 2015), http://www.theesa.com/wp-content/uploads/2015/04/ESA-Essential-Facts-2015.pdf
2. Ibid.
3. Steam Store, "Welcome to Steam," http://store.steampowered.com/about/ (accessed February 10, 2016).
4. Caimei Lu, Jung-ran Park, and Xiaohua Hu, "User Tags Versus Expert-Assigned Subject Terms: A Comparison of LibraryThing Tags and Library of Congress Subject Headings," *Journal of Information Science* 36, no. 6 (2010): 763–779. doi:10.1177/0165551510386173
5. Margaret E. I. Kipp, "Complementary or Discrete Contexts in Online Indexing: A Comparison of User, Creator and Intermediary Keywords," *Canadian Journal of Information and Library Science* 29, no. 4 (2005): 419.
6. Nicholas Carman, "LibraryThing Tags and Library of Congress Subject Headings: A Comparison of Science Fiction and Fantasy Works" (MLIS thesis, Victoria University of Wellington, 2009), http://hdl.handle.net/10063/1272
7. Rachel D. Williams, "A Tangle of Tags: The Impact of User-generated Tagging in Public Library Catalogs," *Proceedings of the American Society for Information Science and Technology* 50, no. 1 (2013): 1, doi:10.1002/meet.14505001079
8. Lu, Park, and Hu, "User Tags Versus Expert-Assigned Subject Terms," 763.
9. Jung-ran Park and Caimei Lu, "Application of Semi-Automatic Metadata Generation in Libraries: Types, Tools, and Techniques," *Library & Information Science Research* 31, no. 4 (2009): 225. doi:10.1016/j.lisr.2009.05.002
10. Jin Ha Lee, Andrew Perti, Rachel Ivy Clarke, and Travis W. Windleharth, *Video Game Metadata Schema. Version 3.1*, June 22, 2015, http://gamer.ischool.uw.edu/wp-content/uploads/2015/07/UWSIMMschemaVersion3.1_20150710.pdf
11. Jin Ha Lee, Rachel Ivy Clarke, and Andrew Perti, "Metadata for Digitally Distributed Video Games at the Seattle Interactive Media Museum" (paper presented at the annual conference of Museums and the Web, Baltimore, Maryland, April 2–5, 2014), http://mw2014.museumsandtheweb.com/paper/metadata-for-digital-video-games-at-the-seattle-interactive-media-museum
12. Lee et al., *Video Game Metadata Schema*.
13. Entertainment Software Association, "Essential Facts."
14. John Unsworth, "Scholarly Primitives: What Methods do Humanities Researchers Have in Common, and How Might Our Tools Reflect This?" (presentation at the Symposium on Humanities Computing: Formal Methods, Experimental Practice, London, United Kingdom, May 13, 2000), http://people.brandeis.edu/~unsworth/Kings.5-00/primitives.html.
15. Scott A. Golder and Bernardo A. Huberman, "User Patterns of Collaborative Tagging Systems," *Journal of Information Science* 32, no. 2 (2006): 198, doi:10.1177/0165551506062337

16. World Wide Web Consortium, "Web Annotation Working Group," https://www.w3.org/annotation/ (accessed Janiuary 19, 2016).
17. Csaba Veres, "The Language of Folksonomies: What Tags Reveal About User Classification," in *Natural Language Processing and Information Systems*, eds. Christian Kop, Günther Fliedl, Heinrich C. Mayr, and Elisabeth Métais (New York: Springer, 2006), 58.
18. Carman, "LibraryThing."
19. Louise F. Spiteri, "Incorporating Facets into Social Tagging Applications: An Analysis of Current Trends," *Cataloging & Classification Quarterly* 48, no. 1 (2010): 105, doi:10.1080/01639370903338345
20. Kipp, "Complementary or Discrete Contexts in Online Indexing," 419.
21. Spiteri, "Incorporating Facets into Social Tagging Applications," 105.
22. Williams, "A Tangle of Tags," 1.
23. Peter Merholz, "Metadata for the Masses," *Adaptive Path*, October 19, 2004, http://adaptivepath.org/ideas/e000361 (accessed January 19, 2016).
24. Jonathan Furner, "Conceptual Analysis: A Method for Understanding Information as Evidence, and Evidence as Information," *Archival Science* 4, no. 3/4 (2004): 233–265. doi:10.1007/s10502-005-2594-8
25. Karen M. Wickett, Simone Sacchi, David Dubin, and Allen H. Renear, "Identifying Content and Levels of Representation in Scientific Data," *American Society for Information Science and Technology* 49, no. 1 (2012): 1, doi:10.1002/meet.14504901199
26. Jin Ha Lee, Andrew Perti, Andrew Weaver, and Travis Windleharth, *GAMEPLAY GENRE*.
27. Lee, Clarke, and Perti, "Metadata for Digitally Distributed Video Games at the Seattle Interactive Media Museum."
28. Robin Hunicke, Marc LeBlanc, and Robert Zubek, "MDA: A Formal Approach to Game Design and Game Research," in *Challenges in Game Artificial Intelligence: Papers from the 2004 AAAI Workshop* (Technical Report WS-04-04) (Menlo Park, CA: The AAAI Press, 2004): 1.
29. Lee, Perti, Clarke, and Windleharth, *Video Game Metadata Schema*.
30. Ibid.
31. Joyce G. Saricks, *The Readers' Advisory Guide to Genre Fiction* (Chicago, IL: American Library Association, 2001).
32. Joyce G. Saricks, *The Readers' Advisory Guide to Genre Fiction*. 2nd edition (Chicago, IL: American Library Association, 2009).
33. Lee, Perti, Clarke, and Windleharth, *Video Game Metadata Schema. Version 3.1*.
34. Ibid.
35. Steam Store, "Rome: Total War™—Alexander," http://store.steampowered.com/app/4770 (accessed February 10, 2016).
36. Lee, Perti, Clarke, and Windleharth, *Video Game Metadata Schema*.
37. Steam Store, "Rome: Total War—Alexander."
38. Ibid.
39. Lee, Perti, Clarke, and Windleharth, *Video Game Metadata Schema*.
40. Andy Donovan, Hyerim Cho, Chris Magnifico, and Jin Ha Lee, "Pretty as a Pixel: Issues and Challenges in Developing a Controlled Vocabulary for Video Game Visual Styles," in *JCDL '13: Proceedings of the 13th ACM/IEEE-CS Joint Conference on Digital Libraries* (New York: ACM, 2013). doi:10.1145/2467696.2467747
41. Peter Ludlow, "Descriptions," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Fall 2013 Edition, http://plato.stanford.edu/archives/fall2013/entries/descriptions. (accessed February 10, 2016).

42. Harold Noonan and Ben Curtis, "Identity," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Summer 2014 edition, http://plato.stanford.edu/archives/sum2014/entries/identity (accessed February 10, 2016).
43. Marga Reimer and Eliot Michaelson, "Reference," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Winter 2014 edition, http://plato.stanford.edu/archives/win2014/entries/reference (accessed February 10, 2016).
44. Tim Berners-Lee, "Linked Data," *World Wide Web Consortium*. (2006) Last modified June 18, 2009, http://www.w3.org/DesignIssues/LinkedData.html.
45. Christian Bizer, Tom Heath, and Tim Berners-Lee, "Linked Data—The Story So Far," *International Journal on Semantic Web and Information Systems* 5, no. 3 (2009): 1, doi:10.4018/jswis.2009081901
46. Klaus V. Meier, "On the Inadequacies of Sociological Definitions of Sport," *International Review for the Sociology of Sport* 16, no. 2 (1981): 79–102. doi:10.1177/101269028101600206
47. Klaus V. Meier, "Triad Trickery: Playing with Sport and Games," *Journal of the Philosophy of Sport* 15, no. 1 (1988): 11–30. doi:10.1080/00948705.1988.9714458
48. Deborah P. Vossen, "The Nature and Classification of Games," *Avante* 10, no. 1 (2004): 53.
49. Entertainment Software Association, "Essential Facts."
50. Robin Hunicke, Marc LeBlanc, and Robert Zubek, "MDA."
51. Tim Berners-Lee, "Linked Data."
52. Bizer, Heath, and Berners-Lee, "Linked Data—The Story So Far."
53. Jennifer B. Bowen, "Moving Library Metadata Toward Linked Data: Opportunities Provided by the eXtensible Catalog," in *DC-2010: International Conference on Dublin Core and Metadata Applications*, ed. Diane I. Hillmann and Michael Lauruhn (Pittsburgh, PA, 2010): 44–59. http://dcpapers.dublincore.org/pubs/article/view/1010
54. Gillian Byrne and Lisa Goddard, "The Strongest Link: Libraries and Linked Data," *D-Lib Magazine* 16, no. 11/12 (2010). doi:10.1045/november2010-byrne
55. Stuart A. Sutton and Joseph T. Tennis, "Creating and Managing Controlled Vocabularies for Use in Metadata," tutorial presented at the International Conference on Dublin Core and Metadata Applications, Shanghai, China, October 11–14, 2004, http://dublincore.org/resources/training/dc-2004/english/DC-2004_Tutorial_4_en.pdf
56. "Europeana Collections," Europeana Foundation, http://www.europeana.eu/portal/ (accessed June 3, 2016).
57. "GAMECIP—Computer Game Platforms," Open Metadata Registry, http://metadataregistry.org/vocabulary/show/id/354.html (accessed June 3, 2016).

# Appendix A

**Table A1.** List of Steam tags by existing VGMS category.

| | | | | |
|---|---|---|---|---|
| **Gameplay genre** | | | | |
| 2d fighter | 3d platformer | 4x | action | action rpg |
| action-adventure | adventure | Arcade | arena shooter | base-building |
| beat 'em up | board game | Bowling | building | bullet hell |
| bullet hell | card game | Casual | character action game | chess |
| choose your own adventure | city builder | Classic | crafting | crpg |
| cult classic | dating sim | Driving | dungeon crawler | education |
| exploration | fight | Fighting | football | fps |
| god game | grand strategy | hack and slash | hidden object | interactive fiction |
| jrpg | lemmings | management | massively multiplayer | metroidvania |
| mmorpg | moba | Music | offroad | on-rails shooter |
| otome | parkour | party-based rpg | physics | pinball |
| platformer | programming | Puzzle | puzzle-platformer | racing |
| rail shooter | real time tactics | resource management | rhythm | rogue-like |
| rogue-lite | rpg | Rts | sandbox | shoot 'em up |
| shooter | side scroller | Simulation | sniper | soccer |
| space sim | spectacle fighter | Sports | stealth | strategy |
| strategy rpg | survival | survival horror | tactical | tactical rpg |
| third-person shooter | top-down shooter | tower defense | trading card game | turn-based combat |
| turn-based strategy | turn-based tactics | twin stick shooter | visual novel | walking simulator |
| wargame | fight | parkour | | |
| **Theme** | | | | |
| illuminati | horses | historical | heist | hacking |
| robots | science | space sim | star wars | swordplay |
| tanks | time travel | Trading | trains | vampire |
| war | zombies | Magic | memes | mechs |
| military | naval | Mining | ninja | pirates |
| politics | demons | dinosaurs | hunting | physics |
| martial arts | agriculture | Aliens | assassin | blood |
| gore | capitalism | conspiracy | cold war | crafting |
| dragons | crime | destruction | detective | diplomacy |
| dystopian | economy | education | exploration | nudity |
| flight | fighting | Loot | | |
| **Narrative genre** | | | | |
| horror | psychological horror | romance | satire | sci-fi |
| sci-fi | steampunk | superhero | supernatural | western |
| war | heist | lovecraftian | martial arts | mystery |
| crime | noir | Parody | post-apocalyptic | alternate history |
| comedy | dark fantasy | dark humor | detective | documentary |
| **Visual style** | | | | |
| hex grid | hand-drawn | Retro | stylized | surreal |
| voxel | gothic | minimalist | grid-based movement | pixel graphics |
| cinematic | colorful | Cute | dark | |
| **Mood** | | | | |
| gothic | funny | relaxing | realistic | surreal |
| violent | psychological | futuristic | atmospheric | political |
| cinematic | competetive | Cute | dark | dark humor |
| **Setting** | | | | |
| rome | warhammer 40k | world war i | world war ii | futuristic |
| modern | post-apocalyptic | 1980s | 1990s | medieval |
| mars | america | | | |

(*Continued on next page*)

| | | | | |
|---|---|---|---|---|
| **Number of players** | | | | |
| singleplayer | team-based | local co-op | local multiplayer | massively multiplayer |
| multiplayer | online co-op | 4 player local | asynchronous multiplayer | co-op |
| co-op campaign | | | | |
| **Pacing** | | | | |
| real-time | real-time with pause | time manipulation | turn-based | bullet time |
| fast-paced | episodic | | | |
| **Point of view** | | | | |
| split screen | third person | top-down | isometric | first-person |
| **Dimension** | | | | |
| 2d | 2.5d | 2d fighter | 3d platformer | 3d vision |
| **Customization options** | | | | |
| gun customization | character customization | class-based | difficult | |
| **Rating** | | | | |
| mature | nudity | family friendly | violent | |
| **Franchise** | | | | |
| lego | lara croft | batman | star wars | |
| **Tropes** | | | | |
| silent protagonist | female protagonist | villain protagonist | | |
| **Progression** | | | | |
| open world | linear | | | |
| **Release date** | | | | |
| 1980s | 1990s | | | |
| **Type of ending** | | | | |
| multiple endings | | | | |
| **Relationships** | | | | |
| remake | | | | |
| **Estimated time for completion** | | | | |
| short | | | | |
| **Price/MSRP** | | | | |
| free to play | | | | |
| **Corporate body** | | | | |
| games workshop | | | | |
| **Additional content** | | | | |
| mod | | | | |
| **Platform** | | | | |
| arcade | | | | |

# Appendix B

**Table B1.** List of Steam tags in new categories.

| | | | | |
|---|---|---|---|---|
| **Mechanics** | | | | |
| 6dof | choices matter | dynamic narration | match 3 | inventory tetris |
| procedural generation | music-based procedural generation | quick-time events | score attack | time attack |
| grid-based movement | perma death | narration | hex grid | pvp |
| pve | | | | |
| **Media** | | | | |
| anime | music | card game | cartoon | |
| comic book | movie | board game | great soundtrack | soundtrack |
| based on a novel | software | e-sports | | |
| **Tools** | | | | |
| video production | photo editing | web publishing | utilities | rpgmaker |
| animation & modeling | audio production | benchmark | game development | design & illustration |
| level editor | gamemaker | software training | | |
| **Evaluation** | | | | |
| abstract | story rich | lore-rich | replay value | experimental |
| moddable | great soundtrack | | | |
| **Inputs** | | | | |
| touch-friendly | mouse only | controller | point & click | trackit |
| **Production** | | | | |
| crowdfunded | Indie | kickstarter | | |