

Summer 2016

Illusions of a 'Bond': Tagging Cultural Products across Online Platforms

Nadine Desrochers

University of Montreal, nadine.desrochers@umontreal.ca

Audrey Laplante

University of Montreal, audrey.laplante@umontreal.ca

Anabel Quan-Haase

The University of Western Ontario, aquan@uwo.ca

Kim Martin

Western University, kmart5@uwo.ca

Louise Spiteri

Dalhousie University, louise.spiteri@dal.ca

Follow this and additional works at: <https://ir.lib.uwo.ca/fimspub>



Part of the [Communication Technology and New Media Commons](#), [Information Literacy Commons](#), [Mass Communication Commons](#), and the [Scholarly Communication Commons](#)

Citation of this paper:

Desrochers, N., LaPlante, A., Quan-Haase, A., Spiteri, L., & Martin, K. (2016). Illusions of a "Bond": Tagging cultural products across online platforms. *Journal of Documentation*.

Illusions of a ‘Bond’: Tagging Cultural Products across Online Platforms¹

Nadine Desrochers (corresponding author)

École de bibliothéconomie et des sciences de l’information, Université de Montréal

nadine.desrochers@umontreal.ca

Audrey Laplante

École de bibliothéconomie et des sciences de l’information, Université de Montréal

audrey.laplante@umontreal.ca

Kim Martin

Faculty of Information and Media Studies University of Western Ontario

kmart5@uwo.ca

Anabel Quan-Haase

Faculty of Information and Media Studies University of Western Ontario

aquan@uwo.ca

Louise Spiteri

School of Information Management, Dalhousie University

Louise.spiteri@dal.ca

Desrochers, N., LaPlante, A., Quan-Haase, A., Spiteri, L., & Martin, K. (2016). Illusions of a “Bond”: Tagging cultural products across online platforms. *Journal of Documentation*. doi :

Structured Abstract

Purpose

Most studies pertaining to social tagging focus on one platform or platform type, thus limiting the scope of their findings. This study explores social tagging practices across four platforms in relation to cultural products associated with the book *Casino Royale*, by Ian Fleming.

Design/methodology/approach

A layered and nested case study approach was used to analyze data from four online platforms: Goodreads, Last.fm, WordPress, and public library social discovery platforms. The top-level case study focuses on the book *Casino Royale*, by Ian Fleming, and its derivative products. The analysis of tagging practices in each of the four online platforms is nested within the top-level case study. ‘Casino Royale’ was conceptualized as a cultural product (the book), its derived

¹ An earlier version of this work was presented at the 76th Annual Meeting of the Association for Information Science and Technology (ASIS&T 2013), in Montreal and published in the proceedings: Desrochers, N., Laplante, A., Martin, K., Quan-Haase, A., Rasmussen Pennington, D. and Spiteri, L. (2013). “Beyond the playlist: Looking at user-generated collocation of cultural products through social tagging”, in *Proceedings of the American Society for Information Science and Technology*, Vol. 50, No. 1, pp. 1-4.

products (e.g., movies, theme songs), as well as a keyword in blogs. A qualitative, inductive, and context-specific approach was chosen to identify commonalities in tagging practices across platforms whilst taking into account the uniqueness of each platform.

Findings

The four platforms comprise different communities of users, each platform with its own cultural norms and tagging practices. Traditional access points in the library catalogues focused on the subject, location, and fictitious characters of the book. User-generated content across the four platforms emphasized historical events and periods related to the book, and highlighted more subjective access points, such as recommendations, tone, mood, reaction, and reading experience. Revealing shifts occur in the tags between the original book and its cultural derivatives: Goodreads and library catalogues focus almost exclusively on the book, while Last.fm and WordPress make additional cross-references to a wider range of different cultural products, including books, movies, and music. The analyses also yield apparent similarities in certain platforms, such as recurring terms, phrasing and composite or multifaceted tags, as well as a strong presence of genre-related terms for the book and music.

Originality/value

The layered and nested case study approach presents a more comprehensive theoretical viewpoint and methodological framework by which to explore the study of user-generated metadata pertaining to a range of related cultural products across a variety of online platforms.

Introduction

Use of social media continues to grow across all platforms in the U.S. and about half of Americans have adopted more than one social media platform (Pew Research Center, 2014). This heavy reliance on social media and the volume of user-generated content (henceforth referred to as UGC) create a real need to better understand the role and meaning of social tagging in these platforms. Social tagging is usually defined in opposition to standardized ontologies and taxonomies, as it is “the process by which many users add metadata in the form of keywords to shared content” (Golder and Huberman, 2006, p. 198). Problems with social tagging abound, with users tagging content in often idiosyncratic ways without giving much consideration to existing best practices, obtained from years of research in library and information science (e.g., Bates and Rowley, 2011; Gerolimos, 2013; Moulaison, 2008; Thomas *et al.*, 2010). Social tagging has also been criticized for several additional reasons: synonymy (multiple terms for the same concept) (Merholz, 2004), being susceptible to ‘gaming’ (Kroski, 2005), lacking hierarchy (Smith, 2004; Kroski, 2005), and for being ‘narrow’ (when most of the tagging is done by a limited number of users, usually through permissions settings, with owner-only tagging being the narrowest option) (Vander Wal, 2005; Peters, 2009).

Despite all its problems, social tagging is an important and popular means of organizing the vast amounts of UGC available on the Web. For example, hashtags on Twitter have been identified as critical for highlighting and organizing content linked to specific topics, revealing trends, and establishing relevance (Huang *et al.*, 2010; Yang, Quan-Haase, & Rannenberg, 2016). Social tagging can also contribute to browsing, allowing users who are not looking for a pre-determined information resource to explore what information is available by clicking on tags. Social tagging may therefore be an alternative means of accessing information in an environment where users do

not only want to search for content via keywords, but are also interested in exploring new and unexpected content (Quan-Haase, Martin, & McCay-Peet, 2015).

Tagging may be thought of as a homogenous process, an action likely to be performed similarly by various communities on the Web. While such an assertion is tempting, this has not been studied extensively, although some comprehensive reviews of folksonomic practices across various platforms, such as the one done by Peters (2009), have laid foundational groundwork for such comparative studies. Studies pertaining to tagging are often limited to investigating one platform or platform type, with the obvious and usually stated limitations of restricting the generalizability of findings. This has two noteworthy effects on the perceptions of tagging practices in social media. First, it creates a sense that user communities are separate and distinct; in other words, it creates a perhaps illusory boundary between types of platforms. Second, and conversely, the body of research shaped by single-platform studies creates a sense that tagging on various platforms presents commonalities, thereby shaping similar relationships to, for example, cultural products such as books or music. However, since few studies have actually done this kind of comparative work (that is, contrasting tagging practices across various sites), this remains more impressionistic than evidence-driven, creating a parallel, and perhaps illusory, bond between platforms.

This study explores the tension between commonalities and differences in the tagging of cultural products across platforms by presenting the results of a layered and nested case study approach, performed in parallel but independently. Its shared starting point is a franchised theme, namely the book *Casino Royale*, by Ian Fleming. This book, along with some of its derivative cultural products, is popular in English-speaking regions, thereby providing the case studies with good coverage across platforms. Indeed, *Casino Royale* allowed the researchers to draw upon many types of content; for if the Bond franchise starts with books, which are bought, catalogued, borrowed, and read by a wide community, it also includes films and accompanying soundtracks that provide listeners with instant classics. As a result, *Casino Royale*, outreaching the book itself, is a phenomenon and reference that shows up in a wide range of online content, including blog posts. Added to this was the timely fact that 2013 marked the 60th anniversary of the book (the first instalment in the James Bond series), sparking renewed interest.

Four platforms were chosen to reflect the range of the franchise and the researchers' expertise: Goodreads and public library catalogues (books), Last.fm (music) and WordPress blog entries. This paper is structured to reflect the layered and nested case study approach utilized by first presenting each case study individually. The researchers adopted methodologies best suited to their object of study for each case study, which will be presented along with case-specific results before those results are compared and combined in the discussion. As will be shown, the analyses of tags for *Casino Royale* and related cultural products studied here yield apparent similarities in certain platforms, such as recurring terms, phrasing and composite or multifaceted tags, as well as a strong presence of genre-related terms for the book and music. The context of use and the relationship with the root product *Casino Royale*, however, indicate that identical terms create potentially very different meanings. The closer the relationship (tagging the book), the less likely the tags point outwards to another product, while tags pertaining to the franchise, but used in broader contexts (blog posts), are more likely to refer back to it. This shows that if the bond—pun fully intended—amongst tagging practices on the various platforms is evident on the formal

front, it creates an illusion of equivalency where the relationship to the object tagged is concerned.

Background

Bischoff and colleagues (2008) compared the tag distribution by category on Last.fm (music), Delicious (websites), and Flickr (images). They used a sample of 300 tags stratified by popularity. Their analysis showed that the tag distribution by category varied considerably across platforms. The most popular tag category in Delicious and Flickr was “Topic,” while in Last.fm it was “Type,” which includes musical genres. Last.fm was also the platform with the highest proportion of subjective/opinion tags.

Abel, Araújo, Gao, and Houben (2011) found that individual users tended to use a greater variety of tags in Delicious than in Flickr. Iofciu, Fankhauser, Abel, and Bischoff (2011) found that tagging practices varied considerably, even between platforms dealing with the same type of resources. Indeed, for a majority of users, the overlap between the tags they used in Delicious and in StumbleUpon, two social bookmarking platforms, was less than 20%. Both studies concluded that there was little overlap between the tags used in each platform.

Strohmaier, Körner, and Kern (2012) examined the social tags applied by individual users (‘personomies’) in ESP Game, Flickr, Bibsonomy, and Delicious, and found that user motivation for tagging varied substantially across platforms, but also within each platform. Heckner, Heilemann, and Wolff (2009) surveyed 142 users of Flickr, YouTube, Delicious, and Connotea and found that while Flickr and YouTube users were more likely to tag to facilitate the sharing of digital resources, Delicious and Connotea users tagged mainly for the personal management of digital content.

Social cataloguing platforms and library catalogues

Practice-based studies of social cataloguing platforms such as LibraryThing, Goodreads, or Shelfari, which allow readers to document, discuss, and share their reading interests, often focus on comparisons between sites (Jeffries, 2008), potential pedagogical use (Marcotte, 2011), or use for cataloguing (Hvass, 2008). The use of social cataloguing sites has generally been perceived favourably by practitioners (Braun, 2013; Jonker, 2013; O’Leary, 2012; Starr, 2007; Starr, 2008; Wyatt, 2007). Stover (2009) described the rise of social cataloguing platforms as an opportunity for librarians to engage with the reading public; she noted that some of the language on these sites echoed the “lingo” (p. 246) of readers’ advisors and that, reciprocally, professionals were integrating more folksonomic terms to their practice. Naik (2012) chose five books from five Goodreads lists and analyzed thirty reviews per book, searching for terms associated with the Readers’ Advisory (RA) concept of appeal, which, in its most basic form, is the combination of characteristics which a reader may find interesting in a book, such as pacing, characterization, storyline, setting, mood, or language and style (Saricks, 2009). Naik reported positive feelings towards the titles, as well as what she coined “repel terms” (p. 321) or expressions of negative feelings.

In comparison, academic research often focuses on user indexing practices as they compare to library catalogues and controlled vocabularies (Bates and Rowley, 2011; Lu *et al.*, 2010), interface design (Chang, 2009), or problems inherent to the lack of standards social indexing practices might exhibit (Thomas *et al.*, 2010). Kathuria (2011) looked at the social tagging of

books in LibraryThing pertaining to Asian women. Ratings on Goodreads have recently been posited as a “unique altmetric data source” for the study of the impact of scholarly books, in this case from the field of History (Zuccala *et al.*, 2015).

Šauperl’s 2012 study of three sites (LibraryThing, Amazon.com, and Primorci beremo) revealed that users’ descriptions of novels extend beyond standard bibliographic description: Characteristics frequently mentioned include the literary genre, the time frame of the story, and character names. Users also often mention awards received, movie adaptations, and related novels. Šauperl urged library professionals to seek and encourage the participation of patrons in describing novels in the library catalogue.

One possible solution for including UGC in library catalogues is to use LibraryThing for Libraries (LTFL), which allows libraries to import LibraryThing tags and user reviews in their catalogue for a fee. Mendes, Quinonez-Skinner, and Skaggs (2009) examined the use of LTFL in the Oviatt Library at California State University over 170 days. The authors found that for every new book a user discovers using LCSH, they would discover four books using LTFL tags, and suggested that the addition of user-generated metadata to catalogue records enhances resource discovery, for example, for those titles lacking subject headings, which is sometimes the case for works of fiction.

Music recommender services

Music indexing in library catalogues is also guided by the desire to provide neutral descriptions of works, mostly by assigning subject headings representing broad musical genres (e.g., popular music, jazz) or forms (e.g., lullabies, sonatas). Laplante (2008) found that library patrons considered this type of indexing of little help for discovering music: The description is too thin and the genres, particularly for popular music, are too broad and therefore not discriminative enough to allow for browsing. In contrast, the descriptions of music works that emerge from tags assigned by users in music recommendation platforms are rich, multifaceted, and subjective. A few studies have focused on the distribution of tags by category in the music recommendation service Last.fm, and genre was consistently found to be the category with the higher proportion of tags, especially amongst the most popular tags, where it accounts for more than 60% of the tags (Lamere, 2008; Laplante, 2015; Thompson, 2008). Other popular tag categories are the place of origin of the artist, the instrumentation, the mood of the song, and the opinion of the user (Lamere, 2008; Laplante, 2015). Geleijnse, Schedl, and Knees (2007) analyzed the tags applied to 1,995 artists in Last.fm and found that 56% of tags had been applied to only one artist in their sample and that fewer than 2% had been applied to 50 artists or more. Levy and Sandler (2009) found that about a third of the 5,265 artists in their sample had never received a tag for any of their tracks and that amongst the artists who had tagged tracks, a third had no more than five unique tags on average per song. Laplante (2015) examined the social tagging of songs during the first three months following their release and found that a majority of songs do not get tagged during the first week and that the number of tags assigned to a song was positively related to its popularity.

Blogging platforms

While there is extensive literature on blogs (Scale & Quan-Haase, 2014), blog users (Boyd, 2006), and motivations for blogging (Fullwood *et al.*, 2014), much less academic work looks directly at tagging practices in blogging platforms such as WordPress, Blogger, or Joomla.

Aharony (2009) examined the distribution by content and form of tags assigned by bloggers of 30 library and information science blogs in the now-defunct Technorati platform. Chopin (2008) examined the usefulness of searching Technorati tags, compared to using a general search engine, and noted that people did not tag much, with the consequence that, in most cases, tag search did not allow users to access a variety of viewpoints. She stressed nonetheless the usefulness of tags for discovery, as they allowed users to navigate between various topics in an undirected fashion.

In light of the above studies, many potential avenues exist for the study of social tagging across platforms. The differences in user practices are linked to a wide range of factors: The nature of the resource or entity being tagged, the design of the platform, the culture of the online community, and, of course, the individual user.

Overarching methods

The study seeks to address the following four research questions:

1. What kinds of tags do users contribute to *Casino Royale* or its derivatives in the different platforms?
2. How do tagging practices compare across platforms?
3. What categories of tags do users provide for the items being tagged in the various platforms?
 - 3.1. What does this reveal about the platforms' cultures?
4. How do the research approaches utilized to study each of the four case studies compare?

The proposed overarching approach was to collect data from various platforms in order to analyze and compare tagging practices, taking into consideration the differences inherent to the platforms studied. Taking a page from qualitative fieldwork methods, the researchers adopted a layered and nested case study approach (Patton, 2015). The top-level case study is the project as a whole and uses the book *Casino Royale*, by Ian Fleming, as the focus of inquiry. Individual case studies are nested within this top-level case study and were performed on Goodreads, Last.fm, WordPress, as well as in public library social discovery platforms. Given the varied nature of the content (books, music, blogs), 'Casino Royale' was operationalized as a cultural product (the book, in all editions), which could also be studied through its derived products (movies, theme songs).

Due to the exploratory and comparative nature of the project, a qualitative, inductive, and situational research framework was chosen (Schreier, 2012). This allowed the researchers to adapt their sampling strategies and analysis to the individual platforms studied whilst adhering to common overarching methodological parameters. This was particularly relevant, and indeed necessary, for the multi-case comparison at hand, and made possible through the established flexibility of certain methods, such as content analysis (White and Marsh, 2006; Hsieh and Shannon, 2005).

Within this layered framework, the specific qualitative analyses performed are presented in the individual case study reports, along with their accompanying measures of ongoing verification (Morse *et al.*, 2002). The intent here was to make sure that no "flattening" of the data or results would occur due to the fact that one method was chosen over another. As stated in RQ4, one of the purposes of the study was to compare approaches and methods as well as results. Since no similar cross-platform, cross-product qualitative study had been identified in the review of the

literature, a “top-down” rigid approach seemed counterproductive and in direct opposition to the study’s objectives. While quantitative studies aim for generalizability, qualitative studies such as this one aim instead to point out where *not* to generalize too hastily, thereby complementing and informing large studies. The choice was therefore made to let each researcher or team design and conduct their own case study independently within the overarching parameters of the broader study, with all the challenges this entails, but with all of the true inductive quality it brought to the research as well. The overarching research questions and qualitative outlook ensured that the individual case studies yielded “intersecting and overlapping units of analysis” (Patton, 2015, p. 384) in the form of individual tags (and, in one case, subject headings) suitable for a common discussion aimed at “recognizing meanings” (Krippendorff, 2013, p. 27) from the textual content studied.

Platform-specific qualitative analyses were thus performed from the perspective that “[i]f meaning depends on context, context in fact becomes part of the data” (Schreier, 2012, p. 22). Each case study therefore maintained its own integrity, and discussions amongst all involved ensured that the results were presented in light of both the limitations and possibilities afforded by this approach. This was an intrinsic part of the study and its effects are therefore discussed throughout the report. Further efforts were made to streamline the reporting style by including both descriptive statistics, best suited to qualitative results (White and Marsh, 2006), and rich description, deemed essential for contextualizing the findings (Patton, 2015). The use of examples was also perceived as important in order to fully illustrate the scope of the tagging practices revealed by the case studies (Bradley, 1993). This two-pronged reporting style, combining counts and narration, helps paint a complete portrait of tags as the online manifestations studied (Krippendorff, 2013; White and Marsh, 2006).

Case study 1: Goodreads

The platform

Goodreads, a social cataloguing platform owned by Amazon (Herther, 2013; Olanoff, 2013), has a popularity ranking of 307th in the world and 122nd in the United States, and places 2nd overall in the Reference/Libraries subcategory, and first in the Reference/Libraries/Library and Information Science subcategory (Alexa, n.d. a; n.d. b) as of May 2016. Goodreads has 50 million members, 1.5 billion book titles, and 50 million reviews (Goodreads Inc., n.d.). The NoveList database integrates links to Goodreads reviews (Kenney, 2011) and these links are visible in many library catalogues.

Casino Royale content studied

The data were collected on August 31, 2013 through a title search for the book *Casino Royale*. Amongst the entries retrieved, which included translations and compilations, “Casino Royale (James Bond (Original Series) #1)” was chosen as the most relevant.

Case-specific methods

In a throwback to physical libraries, tagging, in Goodreads, is known as shelving. Using the Goodreads ranking, the first 100 English-language users who had “shelved” *Casino Royale* were identified. Language was restricted to English since this was the language of the study and of comparison with the other platforms.

These users provided a population of 198 tag uses. After the removal of duplicates (identical tags provided by more than one user), 117 unique tags were identified. The three default shelves suggested by the platform (“Read”, “Currently Reading”, “Want to Read”) were not considered.

Using the QDA Miner software, a computer-assisted, directed qualitative analysis (Hsieh and Shannon, 2005) was conducted, yielding an initial list of 63 codes based in part on the RA appeal framework. Two researchers coded the 117 unique tags, achieving an intercoder reliability of 99.8%. From these original codes, 10 categories were derived and a second round of coding was performed by both coders on the entire set. Any remaining disparities were resolved through reconciliation, a method used to remove any human error in coding, to ensure that any variations in interpretation of the codes are taken into consideration, and to facilitate the presentation and visualization of the final results as a coherent whole (Campbell *et al.*, 2013; Hruschka *et al.*, 2004; Schreier, 2012). After the analysis was completed, it was deemed that the sample provided the coders with enough saturation in the context of this exploratory study.

Case-specific findings

Table 1 presents an overview of the distribution of the tags. Some of the tags were composite tags were coded in more than one category, which explains why the total number of occurrences coded reaches 131.

Table 1. Distribution of tags across categories on Goodreads

Category	Number of shelves/occurrences in each category*	% in terms of representation of each category	Examples
Genre	36	27.5	mystery-thriller; political-fiction; fiction; classic-hardboiled-noir
Reading experience	28	21.4	summer-reading; books-read-2012-2013; ultimate-reading-list; adult-recreational
Appeal	18	13.7	england; spy-vs-spy; james-bond
Reading lists	15	11.5	1001-core
Format	12	9.2	ebook; kindle; audiobook
Reaction	8	6.1	favorites; crap-tastic; ugh; not-worth-the-money; good; what-is-the-hype-all-about
Movie reference	6	4.6	book-to-movie; film-and-tv
Ownership	4	3.1	read-unowned;

Cryptic	3	2.3	fill-in-the-pages
Author reference	1	0.8	ian-fleming
Total	131	100	

* The initial number of unique tags was 117, but some of the tags were composite and therefore coded in more than one category, which explains why the total reaches 131 coded occurrences.

Although the most popular tags in the sample were linked to genre, shared uses of the exact same term by multiple users was rare (Table 2).

Table 2. Uses of the most popular Genre tags for *Casino Royale* on Goodreads

Category	Tag (Shelf)	Number of occurrences	% (total number of tags used = 198)
Genre	fiction	13	6,6
Genre	espionage	7	3,5
Genre	thriller	7	3,5
Genre	mystery	6	3,0

As mentioned above, composite tags were found and therefore coded in more than 1 category to reflect their complex nature. For example, 10.3 % of unique tags in the Genre category were considered “double label” tags, meaning they could nuance the genre alluded to, specify a sub-genre, or allude to the central character or other contextual references, thereby creating multifaceted shelves. Table 3 presents reconciled partial results from the first round of coding, providing insight into the level of complexity that can be achieved in a single shelf and a sense of how individual these combinations can be.

Table 3. Occurrences of single and composite Genre codes on Goodreads

Code within the “Genre” Category	Unique tags (N=117)	Total population of tags (N=198)	Examples
Genre (single label)	22 (18.8%)	64 (32.3%)	fiction; thriller; mystery
Genre (double label)	12 (10.3%)	20 (10.1%)	action-adventure; crime-fiction; crime-horror
Genre (triple label)	3 (2.6%)	3 (1.5%)	classic-hardboiled-noir; crime-detective-thriller; mystery-spy-novels
Genre with other	1 (0.9%)	1 (0.5%)	war-espionage-fiction

reference (triple label); also coded as Appeal (“war”)			
Genre with other reference (quadruple label); also coded as Appeal (“mi6”; “spy”)	1 (0,9%)	1 (0.5%)	mi6-mystery-spy-thriller

The recurrence of the term “fiction” further shows the importance of combinations: While it was assigned alone 13 times, its use in combination with other terms brought its total number of occurrences to 23. There were also 29 uses of 13 tags that contained the terms “spy”, “spies”, or “espionage” (again, alone or in combination with other terms). Variations or combinations of such classic genres as “mystery”, “thriller”, “detective”, and “crime”, were also popular. Rather surprisingly, few tags (8) depicted an appreciation or reaction, whether positive (“favorite-books-ever”), negative (“let-it-gather-dust”), or ambivalent (“crap-tastic”) to the book.

There were few direct allusions to other types of products from the James Bond franchise. The most prevalent were the films, as there have been, in fact, three adaptations of the book to the small (*Casino Royale*, 1954) and big (*Casino Royale*, 1967; 2006) screens. Six single tags were found, with only 2 clearly favouring the film. There were no references to music or to any visual collection or gallery. The tag “curious-endeavourances,” however, led to the blog of a Goodreads reading group that bears the same name and uses the WordPress platform (see Case study 4).

Case study 2: Public library catalogues

The platform

This case study examines the contribution of social discovery platforms and user-generated content in RA services in Canadian public libraries. Social discovery platforms provide an enhanced search and discovery experience for the users. These platforms allow users to enhance the content of catalogue records by adding their own tags, ratings, and reviews. The BiblioCommons and Encore platforms were used in this study, as an examination of the Canadian Public Libraries Gateway (<http://www.collectionscanada.gc.ca/gateway/s22-200-e.html>) revealed them to be the most commonly used social platforms in Canadian public libraries.

Casino Royale content studied

Attention was focused on the catalogue records of all print editions of *Casino Royale* found in Canadian public libraries that use the BiblioCommons and Encore social discovery platforms.

Case-specific methods

The methodology used in this case study is based on a similar, but much larger study conducted by the author on a selection of 22 award-winning fiction titles drawn from Canadian public libraries that use BiblioCommons (BC) and Encore (Pecoskie, Tarulli, & Spiteri, 2014). For the sake of consistency, only the records for print editions of monographs were selected, since these records are structured using the standard Anglo-American Cataloguing Rules (AACR) and encoded in the Machine-Readable Cataloging (MARC) framework. Records for e-books are often provided by OverDrive, which does not use the same AACR and MARC formats.

All qualifying libraries that had the *Casino Royale* monograph in their collections were examined: 28 BC and 5 Encore (n=33). Grounded Theory (Hollan *et al.*, 2000; Strauss and Corbin, 1990; Walker and Myrick, 2006) was used to derive categories of access represented by Library of Congress subject or genre headings, user tags, and user reviews in the *Casino Royale* bibliographic records. Data were collected in June 2013. From the final set of library-located bibliographic records (n=33), the unique user-generated content in the form of tags and reviews/comments was extracted to determine what type of content users contributed to the records. In the case of tags, spelling variations (e.g., labor/labour), and single or plural variants of terms (e.g., dog/dogs) were considered as non-unique terms. Two researchers independently derived categories from the tags and subject headings from each record using the Grounded Theory method. Each researcher coded independently and inductively, allowing categories to emerge from the dataset. Similarly, two other researchers worked independently from each other to derive categories from the user reviews and comments. In both cases, a third researcher, who was not involved in the first round of analysis, independently coded for categories for the tags, Library of Congress subject or genre headings, and reviews. This researcher subsequently assessed the three sets of categorical analyses of the tags, subject headings, and review data and examined them for overlap, clarity, exclusivity, and relevance. These three sets of independent categories were assessed for similarity and subsequently grouped into one finalized set of categories.

Case-specific findings

In total, 97 Library of Congress Subject/Genre Headings were assigned to the 31 bibliographic records, with an average of 3.13 headings per record. None of the Encore records had any user-assigned tags; this same pattern was observed in the larger study of the 22 fiction titles. Only one tag was assigned, wcls2013, by 3 BC catalogues; the meaning of this tag could not be determined. In order to focus on user-generated content provided specifically by public library users, imported reviews from private, non-library entities, such as Goodreads or LibraryThing were not considered. The Encore libraries all contained links to 1482 user reviews exclusively from Goodreads, while the BC libraries (n=25) contained a total of 7 unique reviews, with an average of 4.38 reviews per record. BC libraries share the same bibliographic record, so these 7 unique reviews are repeated across the 25 BC libraries.

The Grounded Theory analysis revealed that subject headings represented the following categories: Genre (28.4%), Protagonists (27.3%), Location (17%), Topic (13.6%), and People (13.6%). User reviews represented the following categories: Recommendation (33.3%), Protagonists (23.8%), Tone (19.1%), Author (9.5%), Historical event (4.8%), and Pace (9.5%). Examples for each category can be found in Table 4.

Table 4. Categories of subject headings and user reviews in library catalogues

Categories	Subject Headings	User Reviews	Examples
Genre	28.4%	N/A	spy stories; spy thrillers
Protagonists	27.3%	23.8%	Bond, James (fictitious character) - LCSH; James Bond - Reviews
Location	17%	N/A	France; Great Britain
Topic	13.6%	N/A	casinos; gambling
People	13.6%	N/A	intelligence officers; gamblers
Recommendation	N/A	33.3%	fun story to read
Tone	N/A	19.1%	gritty; raw
Author	N/A	9.5%	Ian Fleming
Historical event	N/A	4.8%	Cold War
Pace	N/A	9.5%	well-paced book; very slow

Library-contributed access points focus on the more tangible aspects of the work, namely Genre, Protagonists (James Bond), and Location; the Topic of a work does not feature as prominently. User reviews provide more affective access points, namely Recommendations and Tone. User-contributed content can provide valuable access points for topic and affect that are not always reflected well in library-assigned subject headings; this content can thus serve as an important supplement to Machine-Readable Catalog (MARC) records.

Case study 3: Last.fm

The platform

Last.fm is a music recommender system based on collaborative filtering, and provides personalized recommendations by comparing a user's taste profile with other taste profiles in order to identify like-minded users. Last.fm is ranked 1,643rd in the world in popularity, 2,048th in the United States, and 1,754th rank in Canada as of June 2016 (Alexa, n.d. c). Last.fm provides an application programming interface (API) that allows other systems to extract various types of data from Last.fm (including social tags). All users can assign tags to albums, tracks, and artists. For each resource, Last.fm displays the most popular tags (up to 60) in a cloud.

Casino Royale content studied

Since Last.fm is a music platform, the focus was on the theme song of the *Casino Royale* movie, “You know my name”, performed by Chris Cornell. To collect the data, a request was sent to the Last.fm API method “track.getTopTags” on September 4, 2014. This API method is used to retrieve the top tags (up to 100) for a given track and their relative weight based on the number of times they have been applied to the track.

Case-specific methods

For this case study, the codebook was developed based on a larger dataset, not related to *Casino Royale* theme song. More specifically, the tags applied to 630 songs in Last.fm were analyzed. In order to ensure that the sample would reflect a variety of social tagging practices, a non-proportional quota sample was used, containing 30 songs from 21 different genres. The list of distinct tags that had been applied to each song was retrieved, for a total of 23,052 tags. The grounded theory approach was used to analyze the tags (Corbin and Strauss, 2008; Creswell, 2009), which involves a process that is both inductive and iterative. The categories were compared and refined until they were distinct and their properties could be defined clearly. The final codebook contained 12 categories. This codebook was then used to code the top 100 tags that had been applied to *Casino Royale* theme song. The coding was done independently by two researchers, with an intercoder reliability of 94%. As in the Goodreads case study, the two coders discussed the disparities to achieve reconciliation. A few composite tags were assigned two or more categories.

Case-specific findings

The number of distinct tags in each category is presented in Table 5. Although Lamere (2008) and Bischoff and colleagues (2008) found that the most prevalent category of tags in Last.fm is Genre, this category ranks only third for the theme song. The category with the most tags is the Link to cultural products or events, or to another artist category. This is not surprising, considering that the song being tagged is part of a movie soundtrack. Thirty of the 33 tags in this category provide a direct reference to the movie (e.g., “bond theme,” “Casino Royale”), or to the fact that it is a theme song (e.g., “theme,” “soundtrack”). Another tag refers to an actor in the movie (i.e., “Daniel Craig”). The remaining two tags in this category refer to music groups of which Chris Cornell is or was a member (e.g., “Audioslave,” “soundgarden”). Interestingly, except for the tag “Casino Royale” that could refer to both the book’s and the movie’s title, there was no direct reference to Ian Fleming’s book.

Table 5. Distribution of distinct tags by category for the song *You know my name* on Last.fm

Category	Number of distinct tags	% of distinct tags	Examples
Link to cultural product or event, or to another artist	33	29.7	casino royale theme, cinema, soundgarden
Opinion	31	27.9	love at first listen, 10 of 10 stars
Genre/Style	12	10.8	indie, hard rock, post-grunge
Instrumentation	9	8.1	male vocalist, guitar
Other/indecipherable	6	5.4	power, SICmuSIC
Bibliographic information	4	3.6	Chris Cornell, you know my name
Place of origin/Language	3	2.7	american, seattle
Mood/Emotion	3	2.7	angry, energetic
Period	3	2.7	2006, 2000s
Theme/Topic	3	2.7	spy, secret agent
Usage context	2	1.8	my morning, workout
Self-reference	2	1.8	songs I know by heart

Total	111	100
--------------	------------	------------

The second most popular tag category is Opinion, which accounts for more than a quarter of the distinct tags. The large number of opinion tags appears to confirm what previous research revealed, which is that there is a performance aspect to tagging in Last.fm (Laplante, 2015; Zollers, 2007), as there is in the way the users of other social networks list their cultural interests in their profile (Liu, 2007). An examination of the relative weight of each tag reveals that opinion tags are rarely reapplied. The fact that many users choose their own tags and not simply reuse existing tags suggests that, at least for certain users, applying tags involves much more than assigning indexing terms for future retrieval: Tags can be used as a statement of good taste.

In the 10 most frequently applied tags for the theme song, five are genre-related, four refer to the movie, and one falls under the bibliographic information category (see Table 6). The placement of five different genres in the top 10 tags shows that users do not always agree on the genre of a song (“alternative” vs. “grunge”), nor the level at which it should be classified (“rock” vs. “alternative rock”).

Table 6. Most popular tags for the song *You know my name* on Last.fm

rock	100	Genre/Style
Soundtrack	91	Link to cultural product or event, or to another artist
james bond	68	Link to cultural product or event, or to another artist
alternative rock	48	Genre/Style
Casino Royale	39	Link to cultural product or event, or to another artist
007	28	Link to cultural product or event, or to another artist
alternative	23	Genre/Style
Chris Cornell	20	Bibliographic information
Grunge	15	Genre/Style
hard rock	13	Genre/Style

Case study 4: WordPress

The platform

WordPress is an online publishing platform that allows users to create and manage personal and professional blogs and web sites. WordPress has become particularly popular for the purpose of blogging. Blogs are often compared to web pages, but differ in that the content is chronologically arranged and updated more regularly (Clyde, 2004a; 2004b). Rankings by Alexa (n.d. d) from May 2016 put the popularity of WordPress at 41th and 41th worldwide and in the US, respectively. WordPress (2016) statistics from May 2016 show 59.3 million new posts and 48.4 million new comments each month. WordPress allows only the creator of the blog post to tag and it also limits the number of tags per blog post to 15, but recommends a much lower number. Tags can also be renamed as categories, or filed under a category heading, both of which are also user-defined.

Casino Royale content studied

One of the key ways to make information more accessible on WordPress is the use of tags, which are posted directly at the bottom of each blog post. The tags are often user generated, but some tags are suggested by the system. Bloggers have the option to accept or reject the suggestions made by Wordpress prior to publishing their posts. Data were collected by searching for the term “Casino Royale”, of which the 20 most recent WordPress blog posts as of October 2013 were selected.

Case-specific methods

The 20 selected posts were analyzed through a two-stage process. First, all the blog posts were read in full, and memos were made regarding the topic of each post. Based on this preliminary analysis, a list of central codes was generated (see Table 7). This list was then employed to classify the tags based on the content of the blog post to which they were linked. For example, a post on travel that mentions Ian Fleming’s birthplace and tagged with the name (“Ian Fleming”) would be classified under the Places code, while an identically tagged post about a James Bond novel would be classified under Books. This helped in the identification of tags that were rarely employed in the data set and which were coded as “single-use”. These outliers, it was hypothesized, would be the links in which users might find related, surprising, or unintended information. In the second stage of the analysis, the tags were imported into Gephi, an interactive visualization and exploration platform commonly utilized for the analysis of networked data. The nodes in the visualization represent the tags and the edges represent co-usage either within or across a blog post. If they were used in the same blog post, the edge and node are of the identical colour.

Case-specific findings

There were 179 tags used in the 20 blog posts analyzed. The total number of tags used per blog post ranged from 1 to 15, with a mean of 8.4. Of these, the three main codes were: “movies”, “books”, and “places”, which accounted for 61.2% of the total tags in the corpus. Table 7 shows that the tags related to movies accounted for the largest number of tags, often referring directly to the movie title “Kill Bill 3” (*Casino Royale* was a movie Tarantino wanted to direct, but it was given to someone else, whereas *Kill Bill 3* was a possible movie for him to direct, but there is uncertainty as to whether that will ever happen). Many of the tags also referred to the original book by either listing “Ian Fleming” or the title itself “Casino Royale”.

Table 7. Coding chart for WordPress tags (N=179)

Code	Total	%	Example of tag
Movies	63	35	“Kill Bill 3”
Single use	52	29	“nail polish”
Books	24	13.4	“Ian Fleming”
Places	23	12.8	“Belarus”
Topic of Blog	12	6.7	“Travel destinations”
Bond-associated	5	2.7	“007”

In addition to tags referring to the actual cultural products (either movies or books), tags often also referred to places associated with these products (12 tags in total). For instance, a blog would be tagged “Belarus” and “Casino Royale” to indicate that the physical place was referenced in relation to the cultural product.

The network visualization (Figure 1) shows all the tags used to describe the 20 posts; the more times a tag was used, the greater the size of the word and node used to represent it. In the visualization we see that Bond related topics (“james bond”, “daniel craig”, etc.) are most popular, while the smallest words represent the outliers; topics that we termed “single use” in the coding and can lead the farthest away from the original topic. The network topology indicates that there is strong clustering within each blog post, but few connections that bridge across them. This would suggest that there are few popular closely related topics that overlap; otherwise, the blogs are disconnected.

The network graph visualization is centered on the node 'casinoroyale' (James Bond), which is highlighted in a large green circle. The graph consists of numerous nodes and edges, representing relationships between various entities. The nodes are color-coded and connected by lines of varying thickness, indicating the strength or type of relationship.

Central Node: casinoroyale (James Bond)

Key Clusters and Nodes:

- Actors and Characters:** Daniel Craig, James Bond, Pierce Brosnan, Michael Madsen, John Travolta, Vincent Vega, Double V Vega, Reservoir Dogs, Uma Thurman, Quentin Tarantino, Roger Moore, Len Deighton, Frederick Forsyth, Song of Reason, Nathi Polish, Matte, Noe Crazy, United, Noe, Fitt the Gap, Penit Talk, Monster University, Zombieland, Movie Rotations, World War Z, Kolor, Montenegro, Backpacking, Destinations, Travel Blog, Travel, Perast, Sveti Stefan, Europe, Budva, Eastern Europe, Karlo Vary, Sabirshakeel Photography, Imperia Hotel, Lazne 1, Karlobad, The Grand Hotel Pupp, Girls with Guns, Movie, Skullfall, Cold War, Barbara Bouchet, Bone Gin, Travel Destination, Czech Republic, Karlbad Plaza Spa & Wellness Hotel, Proof of Concept, Iran, North Korea, Wordpress Risk, Jason Sacks, Book Review, Alfred Hitchcock, Charles Goldman, Ben Buchgrego, Rattoff, James Bond in Film, Michael Madsen, Vincent Vega, Pulp Fiction, John Travolta, Reservoir Dogs, Uma Thurman, Quentin Tarantino, Roger Moore, Len Deighton, Frederick Forsyth, Song of Reason, Nathi Polish, Matte, Noe Crazy, United, Noe, Fitt the Gap, Penit Talk, Monster University, Zombieland, Movie Rotations, World War Z, Kolor, Montenegro, Backpacking, Destinations, Travel Blog, Travel, Perast, Sveti Stefan, Europe, Budva, Eastern Europe, Karlo Vary, Sabirshakeel Photography, Imperia Hotel, Lazne 1, Karlobad, The Grand Hotel Pupp, Girls with Guns, Movie, Skullfall, Cold War, Barbara Bouchet, Bone Gin, Travel Destination, Czech Republic, Karlbad Plaza Spa & Wellness Hotel.
- Locations and Settings:** London, Las Vegas, Clean Bathroom, Interior Design, Vegas, From Russia with Love, John Barry, Pierce Brosnan, Michael Madsen, Vincent Vega, Pulp Fiction, John Travolta, Reservoir Dogs, Uma Thurman, Quentin Tarantino, Roger Moore, Len Deighton, Frederick Forsyth, Song of Reason, Nathi Polish, Matte, Noe Crazy, United, Noe, Fitt the Gap, Penit Talk, Monster University, Zombieland, Movie Rotations, World War Z, Kolor, Montenegro, Backpacking, Destinations, Travel Blog, Travel, Perast, Sveti Stefan, Europe, Budva, Eastern Europe, Karlo Vary, Sabirshakeel Photography, Imperia Hotel, Lazne 1, Karlobad, The Grand Hotel Pupp, Girls with Guns, Movie, Skullfall, Cold War, Barbara Bouchet, Bone Gin, Travel Destination, Czech Republic, Karlbad Plaza Spa & Wellness Hotel.
- Media and Entertainment:** The Man with the Golden Eye, Goldfinger, The Man from U.N.C.L.E., The Return of the Man from U.N.C.L.E., The Man from U.N.C.L.E. Movie Starts Filming, The Other Spies, Daniel Craig, Ian Fleming, James Bond Films, Henry Cavill, Norman Felton, Sam Rolfe, A Movie Version of the Man from U.N.C.L.E., Ford Motor Co., Armie Hammer, Ian Fleming Foundation, Ian Fleming's Collection of 007 Vehicles, TV Spy Shows, Tomhawk Press, Albert R. Broccoli Maker, Peter Lamont, Michael Lamont's New Book, Neil Lamont, The Spy Who Loved Me, Michael Caine, Sean Connery, Terence Young, Robert Shaw, Las Vegas, Clean Bathroom, Interior Design, Vegas, From Russia with Love, John Barry, Pierce Brosnan, Michael Madsen, Vincent Vega, Pulp Fiction, John Travolta, Reservoir Dogs, Uma Thurman, Quentin Tarantino, Roger Moore, Len Deighton, Frederick Forsyth, Song of Reason, Nathi Polish, Matte, Noe Crazy, United, Noe, Fitt the Gap, Penit Talk, Monster University, Zombieland, Movie Rotations, World War Z, Kolor, Montenegro, Backpacking, Destinations, Travel Blog, Travel, Perast, Sveti Stefan, Europe, Budva, Eastern Europe, Karlo Vary, Sabirshakeel Photography, Imperia Hotel, Lazne 1, Karlobad, The Grand Hotel Pupp, Girls with Guns, Movie, Skullfall, Cold War, Barbara Bouchet, Bone Gin, Travel Destination, Czech Republic, Karlbad Plaza Spa & Wellness Hotel.
- Other Entities:** Ian Fleming Foundation, Ian Fleming's Collection of 007 Vehicles, TV Spy Shows, Tomhawk Press, Albert R. Broccoli Maker, Peter Lamont, Michael Lamont's New Book, Neil Lamont, The Spy Who Loved Me, Michael Caine, Sean Connery, Terence Young, Robert Shaw, Las Vegas, Clean Bathroom, Interior Design, Vegas, From Russia with Love, John Barry, Pierce Brosnan, Michael Madsen, Vincent Vega, Pulp Fiction, John Travolta, Reservoir Dogs, Uma Thurman, Quentin Tarantino, Roger Moore, Len Deighton, Frederick Forsyth, Song of Reason, Nathi Polish, Matte, Noe Crazy, United, Noe, Fitt the Gap, Penit Talk, Monster University, Zombieland, Movie Rotations, World War Z, Kolor, Montenegro, Backpacking, Destinations, Travel Blog, Travel, Perast, Sveti Stefan, Europe, Budva, Eastern Europe, Karlo Vary, Sabirshakeel Photography, Imperia Hotel, Lazne 1, Karlobad, The Grand Hotel Pupp, Girls with Guns, Movie, Skullfall, Cold War, Barbara Bouchet, Bone Gin, Travel Destination, Czech Republic, Karlbad Plaza Spa & Wellness Hotel.

18

In sum, the “Casino Royale” tag often served to disambiguate reference to other products, such as other movies and books, and to indicate that a place, person, or event was being discussed in relation to *Casino Royale*.

Discussion

The four case studies examined in this paper focus on tagging systems that differ in several respects including the nature of the resource or entity being tagged, the platform design, and the community’s culture. Referring back to the overarching research questions of the study, they inform what the different platforms reveal, but also how they can be studied to take their contextual differences into consideration.

Table 8. Comparison of rankings of tag categories across platforms

Category	Rank (percentage)				
	Goodreads	Catalogues, LCSH	Catalogues, Reviews	Last.fm	WordPress
Genre	1 (27.5%)	1 (28.4%)		3 (10.8%)	
Reading experience	2 (21.4%)				
Reading lists	4 (11.5%)			8 (1.8%)	
Usage context				8 (1.8%)	
Self-reference					3 (13.4%)
Books					
Appeal	3 (13.7%)				
Protagonists		2 (27.3%)	2 (23.8%)		
Location		3 (17%)			
Topic		4 (13.6%)			
People		4 (13.6%)			
Tone			3 (19.1%)		
Historical event			5 (4.8%)		
Pace			4 (9.5%)		
Mood/emotion				7 (2.7%)	
Period				7 (2.7%)	
Theme/topic				7 (2.7%)	
Places					4 (12.8)
Topic of blog					5 (6.7%)
Bond-associated					6 (2.7%)
Format	5 (12%)				
Reaction	8 (6.1%)				
Recommendation			1 (33.3%)		
Opinion				2 (27.9%)	
Movie Reference	6 (4.6%)				
Link to another product or event, or to another artist				1 (29.7%)	
Movies					1 (35%)
Ownership	7 (3.1%)				

Cryptic Other/indecipherable	8 (2.3%)			5 (5.4%)	
Author reference Bibliographic information	9 (0.8%)		4 (9.5%)	6 (3.6%)	
Instrumentation				4 (8.1%)	
Place (of origin)/Language				7 (2.7%)	
Single use					2 (29%)

RQ1. What kinds of tags do users contribute to Casino Royale or its derivatives in the different platforms?

Communities in Goodreads and Last.fm use composite tags, and so it was often necessary to place these tags into more than one category to reflect their complexity. Authors in WordPress create tags where multiple terms are used to fully represent one concept. Interestingly, these practices, while slightly different, all mirror the construction of controlled vocabularies in traditional library cataloguing, such as LCSH, in which headings, free-floating subdivisions, and genre or form terms can be associated to form a more complete representation of the object being described, as seen in these examples:

- Bond, James (fictitious character), LCSH
- war-espionage-fiction, Goodreads
- great male voice, Last.fm
- ian fleming foundation's collection of 007 vehicles, WordPress

Unlike traditional subject headings, however, the folksonomies studied here also presented sentence-like tags that either described the cultural product in relationship to the user (“songs i know by heart”, Last.fm), or in relationship with other cultural products. The latter could either be purely descriptive (“a movie version of the man from u.n.c.l.e”, WordPress), or opinion-laden (“the-movie-is-better”, Goodreads).

The form of the tags is influenced if not dictated at least in part by the tagging system design, which is discussed below.

RQ 2. How do tagging practices compare across platforms?

The platforms provide access to different types of resources (i.e., books, music recordings, and blog posts) and this influences the way they are designed. All platforms, except for WordPress, adopt a broad folksonomy approach whereby “many people [are] tagging the same object and every person can tag the object with their own tags in their own vocabulary” (Vander Wal, 2005). At the moment of tagging, Last.fm suggests tags that have been applied the most frequently to the resource being tagged, as well as the tags most applied by the user. Only the top 60 tags assigned to an item are displayed. Goodreads suggests its three default shelves (“Read”, “Currently Reading”, and “Want to Read”) in addition to the shelves most applied by the user; suggestions are therefore linked directly to the reader and individual timelines or practices.

RQ3. What categories of tags do users provide for the items being tagged in the various platforms?

It is in the comparison of the results of the various analyses that the layered and nested approach revealed the full contextual importance of each platform in relationship to the products tagged.

Table 8 establishes comparisons between the categories found on each platform, keeping in mind that no initial grid or list was imposed. These categories are therefore indigenous to the analysis of the platforms studied and a reflection of what the team responsible for each case study found through the qualitative analysis of the data. Groupings of categories were not established based on the content of the tags, but rather on the categories. This highlights the importance of respecting and reporting anomalies (Bradley, 1993) rather than overly flattening results. A telling example of this is the “Instrumentation” category unique to the Last.fm platform; another would be the format, which indicates if a resource is an audiobook and ebook, of importance to the Goodreads community, but irrelevant to WordPress. It is also normal not to find recommendations or opinions in subject headings.

The comparison of the results reveals interesting differences and commonalities between tagging practices across platforms. Compared with traditional access points provided in library catalogues, users of social media platforms tended to provide more subjective access points, such as opinion or experience-related tags, which parallels Šauperl’s (2012) finding that UGC goes beyond the bibliographic description provided in traditional library catalogues. For instance, in Last.fm, opinion tags accounted for 27.9% of tags; in Goodreads, although the reaction category only accounted for 6.1% of the tags, the reading-experience tags accounted for 21.4% of tags, and the reading-list tags for 11.5%; and the recommendation category ranked first in the library catalogue reviews. This shows why knowing how people tag items in social media platforms is increasingly important for libraries, especially since many library catalogues now import Goodreads reviews and LibraryThing tags. The use of tags referring to traditional appeal factors to describe the characteristics of books in readers’ advisory services was relatively low (13.7%), whereas they made up 71.5% of the LCSH in the library catalogue. Goodreads users and cataloguers remain faithful to the quintessential notion of genre to describe fiction books as reflected by 27.5% of tags in the social platform and 28.4% of subject headings in the catalogues. Interestingly, although previous studies show that genre is the most popular tag category at the platform level, only 10.8% of the tags assigned to the theme song of the movie were genre related; in this case, the reference to the movie took precedence over genre and represented 29.7% of the tags. A similar trend was found in WordPress where the reference to the movie also ranked first and accounted for 35% of the tags. This seems to suggest that the relationship to the root product is perceived by users as necessary when tagging derivative products.

RQ 3.1 What does this reveal about the site culture?

The platforms bring together communities of users with differing cultures and practices. As the review of the literature on social tagging revealed (see Background section), these differences can affect tagging practices (e.g., Bischoff *et al.*, 2008; Iofciu *et al.*, 2011; Peters, 2009; Strohmaier *et al.*, 2012).

It is in the relationship between the tags, the original cultural product (*Casino Royale*), and its derivatives that the most revealing differences across platforms occur. Cultural products do not exist in a vacuum; this is perhaps evident, but too easy to forget when one sees “james bond” as a tag in all platforms. A place, such as “England”, can be the setting of a book, a place of origin for its creator, or a travel destination. References to the movie were prevalent on Last.fm, indicating the strong importance of linking a theme song to the movie for which it was created, rather than the life the song has known outside of it. Indeed, this goes against findings from other research which shows that genre is usually the favoured tag category on Last.fm (Lamere, 2008; Laplante,

2015), as it was here for Goodreads, where references to the movie had only a weak representation. This suggests that the reading act far outweighs the fact that this book was turned into a movie, and supports a very book-centric, almost book-exclusive nature to the Goodreads platform, isolating the book from other cultural products and the reading experience from a broader cultural context. On WordPress, the results suggest the opposite, since the relationship to movies was also the strongest (63 cases), seemingly placing the movie *Casino Royale* at the core of the tagging and establishing more relationships with other movie references, and then with books, which ranked third with 24 cases (the outliers taking the second position). Unsurprisingly perhaps due to the far-reaching nature of the blogging experience, the tags on WordPress offer the farthest-reaching network in terms of the cultural context.

RQ 4. How do the research approaches utilized to study each of the four cases compare?

As stated in this research question, it was an intrinsic part of this study to explore how inductive methods would compare in a cross-platform study. This approach allowed the researchers to use their expertise in information behaviour, organization of information, library systems, readers' advisory (RA), social media, serendipitous encountering of information, music, and literature, to lay some exploratory groundwork for larger-scale studies. By exploring different qualitative, inductive analyses, the case studies revealed where the parallels lie across the platforms, but also the pitfalls linked to underestimating not only the importance of context in terms of the meaning associated with each semantic unit in tagging, but also the strength of the relationship between the object tagged and said semantic unit.

Qualitative studies do not aim for generalizability; rather, their strength lies in revealing the detail and anomalies (Bradley, 1993) in the reduction of data. Given the results presented here, the approaches of content analysis, directed or non-, and grounded theory are all relevant to this type of content, since the categories are created inductively and therefore truly reflect site practices. The network analysis performed further revealed how tags were related to one another and how co-usage of tags allows readers of blogs to encounter new content that is only peripherally related to the tag of *Casino Royale*. The comparison of the results provided by each case study is also important because it tells us something about what inductive methods can bring: contextualized codebooks, unique to each platform, unique to the relationships of the tags to the objects, and reflective of particular communities—even when the terms used for tagging are the same. The final contributions of this paper are therefore to state:

- Design methods for the analysis of this type of content should always take the following differences into consideration: the actual product being tagged and its relationship to any root product, user motivation (experience, retrieval, etc.), and tag style or composition, as well as, quite obviously, the design of the tagging system;
- Researchers might be weary of establishing single, cross-platform codebooks; rather, multi-pronged and inductive approaches built on nested and layered case studies should be explored and pursued further.

Limitations and further studies

There are limitations to this study, some pertaining to the choice of focusing on one cultural product. The use of *Casino Royale* as our cultural product meant that a large and diverse set of tags was encountered, which may not be the case with other cultural products, which are more focused to a single medium (e.g., only available in book or audio format). Additionally, *Casino Royale* has a loyal following, often with a strong emotional connection to the movies and books;

this is reflected in the way that tags are assigned, often with a positive valence, as seen in the travel tags found in WordPress blog posts.

The adoption of a nested and layered case study approach also entailed some limitations. Firstly, there was a time interval in the data collection. For three case studies, the data were collected in the summer or fall of 2013 while the Last.fm data were collected in the fall of 2014. Nonetheless, there are two reasons why this would not influence the results drastically: the *Casino Royale* movie was released in 2006, so much of the initial increase in tags around cultural products would have occurred shortly thereafter; and all four case studies were cross-sectional, snapshots and as such reflect data for that point, without much consideration of changes over time. To illustrate this, the data for Last.fm were actually also collected and analyzed summarily in the fall 2013 (Desrochers, Laplante, Martin, Quan-Haase, Rasmussen Pennington, and Spiteri, 2013) and very few differences were found between the two datasets.

Secondly, the inductive and qualitative paradigms, which imposed no pre-determined set of categories, also had both advantages and disadvantages. On the one hand, this approach limited the comparisons to categories and format of tags across the different platforms, rather than the tags themselves; on the other hand, it reflected the reality of the different platforms, as established by the literature. Further and larger studies should be conducted to compare the data at the collection phase and to perform parallel comparisons of semantic units and categories across the platforms studied; however, the caveats and pitfalls mentioned here remain of crucial importance in order for context to be taken into consideration. Given the exploration performed in this study, the suggested next step would be to take a population of tags from each platform and perform again inductive analyses in order to establish not simply a common codebook yet, but parallel codebooks. Then, a certain merging of categories could be performed, on two fronts, creating two codebooks: one for categories which are common to some or all platforms, and one for those who are not. This way, the similarities would be studied alongside the differences without running the risk of reducing the analysis to an illusory convergence of meaning, an illusory equivalence between cultural products, or an illusory ‘bond’ between communities and their practices.

Conclusion

The present study allows us to draw some interesting conclusions. User-generated content on social media necessitates the development of innovative research methodologies that allow for cross-platform comparisons to identify differences and similarities in user behaviour, resulting from the unique cultural norms, practices, and technical and social affordances inherent to each platform. Our study presents a layered and nested case study approach as one way of meeting the challenge of examining multiple, diverse platforms in a single study by focusing on one franchise that provided multiple cultural products. We were also able to identify differences across platforms in how users tag content related to our key topic “Casino Royale” that demonstrate a fundamental shift: Digital platforms with a focus on the material book *Casino Royale*, like Goodreads and library catalogues, focus primarily on the book itself; by contrast, digital platforms with a focus on music and content, namely Last.fm and WordPress, are more likely to cross-reference the root product and other cultural products, thereby expanding the domain in multiple and sometimes surprising directions. Finally, it is interesting to note that the incentives to tag still reflect those presented in early studies of tagging systems such as Marlow, Namaan, boyd, and Davis (2006). While Goodreads and Last.fm users tag mainly for future retrieval and

self-representation (through the expression of their reading experience and opinion, respectively), WordPress users seek to attract attention of readers to their posts. All of this shows that tagging practices on social platforms should always be studied in relationship to both the nature of the product being tagged and the design of the tagging system. This is especially true when comparing platforms, for removing context could lead to hasty conclusions where user practices are concerned.

Acknowledgements

This research received financial support from the Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Development Grant #430-2011-0396 and the Social Sciences and Humanities Research Council of Canada (SSHRC) Insight Grant #R3603A13. Further thanks are given to Catherine Saint-Arnaud-Babin, research assistant for both the Goodreads and Last.fm study.

References

1. Abel, F., Araújo, S., Gao, Q. and Houben, G.-J. (2011), "Analyzing cross-system user modeling on the social Web", in Auer, S., Díaz, O. and Papadopoulos, G. (Eds), *Web Engineering: 11th International Conference, ICWE 2011, Paphos, Cyprus, June 20-24, 2011*, Vol. 6757, Springer Berlin Heidelberg, pp. 28-43.
2. Aharony, N. (2009), "Librarians and information scientists in the blogosphere: an exploratory analysis", *Library & Information Science Research*, Vol. 31 No. 3, pp. 174-181.
3. Alexa. (n.d. a), "How popular is goodreads.com?", available at: <http://www.alexa.com/siteinfo/goodreads.com> (accessed: 2 September 2015).
4. Alexa. (n.d. b), "Top sites in: all categories > reference > libraries > library and information science", available at: http://www.alexa.com/topsites/category/Top/Reference/Libraries/Library_and_Information_Science (accessed: 31 May 2016)
5. Alexa. (n.d. c), "How popular is last.fm?", available at: <http://www.alexa.com/siteinfo/last.fm> (accessed: 3 June 2016).
6. Alexa. (n.d. d), "How popular is wordpress.fm?", available at: <http://www.alexa.com/siteinfo/wordpress.com> (accessed: 6 June 2016).
7. Bates, J. and Rowley, J. (2011), "Social reproduction and exclusion in subject indexing", *Journal of Documentation*, Vol. 67 No. 3, pp. 431-448.
8. Bertin-Mahieux, T., Eck, D., Mailliet, F. and Lamere, P. (2008), "Autotagger: a model for predicting social tags from acoustic features on large music databases", *Journal of New Music Research*, Vol. 37 No. 2, pp. 115-135.
9. Bischoff, K., Firan, C. S., Nejd, W. and Paiu, R. (2008), "Can all tags be used for search?", in *Proceeding of the 17th ACM conference on Information and knowledge management*, ACM Press, New York, pp. 193-202.
10. Boyd, D. (2006), "A blogger's blog: exploring the definition of a medium", *Reconstruction*, Vol. 6 No. 4.
11. Bradley, J. (1993), "Methodological issues and practices in qualitative research", *The Library Quarterly: Information, Community, Policy*, Vol. 63 No. 4, pp. 431-449.
12. Braun, L. W. (2013), "Can it be done? Readers' advisory in a digital world", *Voice of Youth Advocates*, Vol. 36 No. 5, pp. 44-45.

13. Campbell, J. L., Quincy, C., Osserman, J. and Pedersen, O. K. (2013), "Coding in-depth semistructured interviews: problems of unitization and intercoder reliability and agreement", *Sociological Methods & Research*, Vol. 42 No. 3, pp. 294-320.
14. *Casino Royale* (2006), motion picture, Columbia Pictures, United Kingdom.
15. *Casino Royale* (1967), motion picture, Metro Goldwyn Mayer (MGM) British Studio, United Kingdom.
16. *Casino Royale* (1954), television series episode, Columbia Broadcasting System (CBS), Los Angeles, CA.
17. Chang, H.-C. 'Emotion barometer of reading: user interface design of a social cataloging website', in *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, ACM Press, Boston, MA, pp. 3371-3376.
18. Chopin, K. (2008), "Finding communities: alternative viewpoints through weblogs and tagging", *Journal of Documentation*, Vol. 64 No. 4, pp. 552-575.
19. Clyde, L. A. (2004a), "Library weblogs", *Library Management*, Vol. 25 No. 4/5, pp. 183-189.
20. Clyde, L. A. (2004b), "Weblogs – are you serious?", *The Electronic Library*, Vol. 22 No. 5, pp. 390-392.
21. Corbin, J. M. and Strauss, A. L. (2008), *Basics of qualitative research techniques and procedures for developing grounded theory*, 3rd ed., Sage, Los Angeles; London.
22. Creswell, J. W. (2009), *Research design: qualitative, quantitative, and mixed methods approaches*, 3rd ed., Sage, Thousand Oaks.
23. Desrochers, N., Laplante, A., Martin, K., Quan-Haase, A., Rasmussen Pennington, D. and Spiteri, L. (2013). "Beyond the playlist: Looking at user-generated collocation of cultural products through social tagging", in *Proceedings of the American Society for Information Science and Technology*, Vol. 50, No. 1, pp. 1-4.
24. Fullwood, C., Nicholls, W. and Makichi, R. (2014), "We've got something for everyone: how individual differences predict different blogging motivations", *New Media & Society*.
25. Geleijnse, G., Schedl, M. and Knees, P. (2007), "The quest for ground truth in musical artist tagging in the social Web era", in Dixon, S., Bainbridge, D. and Typke, R. (Eds), *ISMIR 2007: Proceedings of the 8th International Conference on Music Information Retrieval, Vienna, Austria, September 23-27*, Österreichische Computer Gesellschaft, Vienna, pp. 525-530.
26. Gerolimos, M. (2013), "Tagging for libraries: a review of the effectiveness of tagging systems for library catalogs", *Journal of Library Metadata*, Vol. 13 No. 1, pp. 36-58.
27. Golder, S. and Huberman, B. (2006), "Usage patterns of collaborative tagging systems", *Journal of Information Science*, Vol. 32 No. 2, pp. 198-208.
28. Goodreads Inc. (n.d.), "About Goodreads", available at: <https://www.goodreads.com/about/us> (accessed: 2 September 2015).
29. Heckner, M., Heilemann, M. and Wolff, C. (2009), "Personal information management vs. resource sharing: towards a model of information behavior in social tagging systems", in *Proceeding of the Third International Conference on Weblogs and Social Media*, AAAI Press, Menlo Park, CA, pp. 42-49.
30. Herther, N. K. (2013), "Goodreads: social media meets readers advisory", *Online searcher*, Vol. 37 No. 4, pp. 38-41.
31. Hollan, J., Hutchins, E. and Kirsh, D. (2000), "Distributed cognition: toward a new foundation for human-computer interaction research", *ACM Transactions on Computer-Human Interaction*, Vol. 7 No. 2, pp. 174-196.
32. Hruschka, D. J., Schwartz, D., St.John, D. C., Picone-Decaro, E., Jenkins, R. A. and Carey, J. W. (2004), "Reliability in coding open-ended data: lessons learned from HIV behavioral research",

- Field Methods*, Vol. 16 No. 3, pp. 307-331.
33. Hsieh, H.-F. and Shannon, S. E. (2005), "Three approaches to qualitative content analysis", *Qualitative Health Research*, Vol. 15 No. 9, pp. 1277-1288.
 34. Huang, J., Thornton, K. M. and Efthimiadis, E. N. 'Conversational tagging in twitter', in *Proceedings of the 21st ACM conference on Hypertext and hypermedia*, ACM Press, New York, pp. 173-178.
 35. Hvass, A. (2008), "Cataloguing with LibraryThing: as easy as 1,2,3!", *Library Hi Tech News*, Vol. 25 No. 10, pp. 5-7.
 36. Iofciu, T., Fankhauser, P., Abel, F. and Bischoff, K. (2011), "Identifying users across social tagging systems", in *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*, AAAI Press, Menlo Park, CA, pp. 522-525.
 37. Jeffries, S. (2008), "Social cataloging tools: a comparison and application for librarians", *Library Hi Tech News*, Vol. 25 No. 10, pp. 1-4.
 38. Jonker, T. (2013), "Social media's best-kept secret: Goodreads is a fabulous site to revolutionize your literary life", *School Library Journal*, Vol. 59 No. 1, pp. 32-34.
 39. Kathuria, S. (2011), "Content analysis of social tags on intersectionality for works on asian women: an exploratory study of LibraryThing", Master's thesis, University of Tennessee, Knoxville, TN, available at: http://trace.tennessee.edu/utk_gradthes/988 (accessed: 2 September 2015).
 40. Kenney, B. (2011), "EBSCO's NoveList and Goodreads strike content deal", *Library Journal*, available at: <http://lj.libraryjournal.com/2011/06/industry-news/ebscos-novelist-and-goodreads-strike-content-deal/> (accessed: 2 September 2015).
 41. Krippendorff, K. (2013), *Content analysis: an introduction to its methodology*, 3rd ed., Sage, Thousand Oaks, CA.
 42. Kroski, E. (2005), "The hive mind: folksonomies and user-based tagging", *InfoTangle Blog*, available at: <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/> (accessed: 18 August 2014).
 43. Laplante, A. (2008), *Everyday life music information-seeking behaviour of young adults: An exploratory study*, Ph.D. Thesis, McGill University, available at: <http://digitool.library.mcgill.ca/thesisfile22017.pdf> (accessed: June 3rd, 2016)
 44. Laplante, A. (2015), "Tagged at first listen: an examination of social tagging practices in a music recommender system", *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, Vol. 20 No. 1, pp. 33-54.
 45. Levy, M. and Sandler, M. (2009), "Music information retrieval using social tags and audio", *IEEE Transactions on Multimedia*, Vol. 11 No. 3, pp. 383-395.
 46. Liu, H. (2007), "Social network profiles as taste performances", *Journal of Computer-Mediated Communication*, Vol. 13 No. 1, pp. 252-275.
 47. Lu, C., Park, J.-R. and Hu, X. (2010), "User tags versus expert-assigned subject terms: a comparison of LibraryThing tags and Library of Congress Subject Headings", *Journal of Information Science*, Vol. 36 No. 6, pp. 763-779.
 48. Marcotte, A. (2011), "Partager le plaisir de lire avec www.babelio.com!", *Québec français*, Vol. No. 160, pp. 60-61.
 49. Marlow, C., Naaman, M., boyd, d. and Davis, M. (2006), "HT06, tagging paper, taxonomy, Flickr, academic article, to read", in *Proceedings of the seventeenth conference on Hypertext and hypermedia*, ACM Press, New York, pp. 31-40.
 50. Mendes, L. H., Quiñonez - Skinner, J. and Skaggs, D. (2009), "Subjecting the catalog to tagging", *Library Hi Tech*, Vol. 27 No. 1, pp. 30-41.

51. Merholz, P. (2004), "Ideas: metadata for the masses", *Adaptive Path*, available at: <http://www.adaptivepath.com/ideas/e000361/> (accessed: 2 September 2015).
52. Morse, J. M., Barrett, M., Mayan, M., Olson, K. and Spiers, J. (2002), "Verification strategies for establishing reliability and validity in qualitative research", *International journal of qualitative methods*, Vol. 1 No. 2, pp. 13-22.
53. Moulaison, H. L. (2008), "Social tagging in the Web 2.0 environment: author vs. user tagging", *Journal of Library Metadata*, Vol. 8 No. 2, pp. 101-111.
54. Naik, Y. (2012), "Finding good reads on Goodreads", *Reference & User Services Quarterly*, Vol. 51 No. 4, pp. 319-323.
55. O'Leary, M. (2012), "Reading dead? No way! See Goodreads", *Information Today*, Vol. 29 No. 1, pp. 22-23.
56. Olanoff, D. (2013), "Amazon acquires social reading site Goodreads, which gives the company a social advantage over Apple", *TechCrunch*, available at: <http://techcrunch.com/2013/03/28/amazon-acquires-social-reading-site-goodreads/> (accessed: 2 September 2015).
57. Patton, M. Q. (2015), *Qualitative research & evaluation methods*, 4th ed., Sage, Thousand Oaks, CA.
58. Pecoskie, J., Tarulli, L., & Spiteri, L. F. (2014), "OPACs, users, and readers' advisory: exploring the implication of user-generated content for readers' advisory in Canadian public libraries", *Cataloging & Classification Quarterly*, Vol. 52 No. 4, pp. 431-453.
59. Peters, I. (2009), *Folksonomies. Indexing and retrieval in Web 2.0*, de Gruyter, Berlin.
60. Pew Research Center (2014), "Internet & tech: Internet user demographics", available at: <http://www.pewinternet.org/data-trend/internet-use/latest-stats/> (accessed: 2 September 2015).
61. Quan-Haase, A., Martin, K., & McCay-Peet, L. (2015). Networks of digital humanities scholars: The informational and social uses and gratifications of Twitter. *Big Data & Society*, 2(1). <http://doi.org/10.1177/2053951715589417>
62. Saricks, J. G. (2009), *The readers' advisory guide to genre fiction*, 2nd ed., American Library Association, Chicago.
63. Šauperl, A. (2012), "Pinning down a novel: characteristics of literary works as perceived by readers", *Library Review*, Vol. 61 No. 4, pp. 286-303.
64. Scale, M. S., & Quan-Haase, A. (2014). Categorizing blogs as information sources: Implications for libraries and information science. In M. Khosrow-Pour (Ed.), *Encyclopedia of information science and technology*, 3rd e. Hershey, NJ: IGI Global.
65. Schreier, M. (2012), *Qualitative content analysis in practice*, Sage, London.
66. Smith, G. (2004), "Folksonomy: social classification", *Atomic: a weblog by Gene Smith*, available at: http://atomiq.org/archives/2004/08/folksonomy_social_classification.html (accessed: 21 August 2014).
67. Starr, J. (2007), "LibraryThing.com: the Holy Grail of book recommendation engines", *Searcher*, Vol. 15 No. 7, pp. 25.
68. Starr, J. (2008), "Social networks: another avenue to new books and friends", *Searcher*, Vol. 16 No. 6, pp. 42.
69. Stover, K. M. (2009), "Stalking the wild appeal factor", *Reference & User Services Quarterly*, Vol. 48 No. 3, pp. 243-246, 269.
70. Strauss, A. and Corbin, J. (1990), *Basics of qualitative research: grounded theory procedures and techniques*, Sage, Newbury Park, CA.
71. Strohmaier, M., Körner, C. and Kern, R. (2012), "Understanding why users tag: A survey of tagging motivation literature and results from an empirical study", *Web Semantics: Science, Services and*

- Agents on the World Wide Web*, Vol. 17 No., pp. 1-11.
72. Thomas, M., Caudle, D. M. and Schmitz, C. (2010), "Trashy tags: problematic tags in LibraryThing", *New Library World*, Vol. 111 No. 5/6, pp. 223-235.
 73. Thompson, A. E. (2008), "Playing tag: an analysis of vocabulary patterns and relationships within a popular music folksonomy", Master's thesis, University of North Carolina at Chapel Hill, available at: <http://www.ils.unc.edu/MSpapers/3422.pdf> (accessed: 2 September 2015).
 74. Vander Wal, T. (2005), "Off the top: explaining and showing broad and narrow folksonomies", *Vanderwal.net*, available at: <http://www.vanderwal.net/random/entrysel.php?blog=1635> (accessed: 2 September 2015).
 75. Walker, D. and Myrick, F. (2006), "Grounded theory: an exploration of process and procedure", *Qualitative Health Research*, Vol. 16 No. 4, pp. 547-559.
 76. White, M. D. and Marsh, E. E. (2006), "Content analysis: a flexible methodology", *Library trends*, Vol. 55 No. 1, pp. 22-45.
 77. WordPress (2015), "A live look at activity across WordPress.com", available at: <http://wordpress.com/activity/> (accessed: 2 September 2015).
 78. Wyatt, N. (2007), "2.0 for readers", *Library Journal*, Vol. 132 No. 18, pp. 30-33.
 79. Yang, S., Quan-Haase, A., & Rannenber, K. (2016). The changing public sphere on Twitter: Network structure, elites, and topics of the #righttobeforgotten. *New Media & Society*. doi : 10.1177/1461444816651409
 80. Zollers, A. (2007), "Emerging motivations for tagging: expression, performance, and activism", in *Proceedings of the WWW 2007 Workshop Tagging and Metadata for Social Information Organization*, paper 55.
 81. Zuccala, A. A., Verleysen, F. T., Cornacchia, R. and Engels, T. C. E. (2015), "Altmetrics for the humanities: comparing Goodreads reader ratings with citations to history books", *Aslib Journal of Information Management*, Vol. 67 No. 3, pp. 320-336.