Book Publishing Final Research Paper

2019

Coding LGBTQ Content: BISACs, Fanfiction, and Searchability in the Digital Age

Hanna Ziegler
Portland State University

Let us know how access to this document benefits you.

Recommended Citation

Coding LGBTQ Content

BISACs, Fanfiction, and Searchability in the Digital Age

By Hanna Ziegler

Introduction

After making observations about the fanfiction I was reading, I noticed that fanfiction authors were making room in their fics for queer characters and queer themes. In most, if not all of these fics, the author had used tags to alert potential readers of the represented queerness and other relevant information that did not appear in the description. These tags served as a kind of abstract for non-spoiler events that can be potentially triggering for readers, and are almost always updated by the author at the request of a reader, should the need arise. These tags, in addition to the description, are instrumental in a readers choice to read one fic over another, and can be used to help readers find specific types of diversity or representation. This same pattern of openness, diversity, and support, especially for the LGBTQ community, is something I was not seeing as often in traditionally published work.

As long as fanfiction has been around it's been a safe space for queer content, and that is still true as more LGBTQ content makes its way into popular media. Certain tropes remain in these society approved works however, like the trend towards tragic lesbian love affairs and the continued invisibility of bisexual protagonists, which has been a constant as recently as 2015.[1] Some of these tropes persist in fanfiction as well, but overall the trends in fanfiction are towards a general queering of work and the validation and exploration of queer identity through creative expressions in the context of familiar characters and settings. At its root, fanfiction is a reaction to the straightness of popular culture and frequently serves as a correction or recasting of the source media. Because readers cannot find queer material in traditional publishing, they are making their own, and in doing so have created a new mode of searching for that content. Publishers are limited to BISAC classifications, keywords in metadata, and back cover copy to get their books in the hands of potential readers, but fanfiction published on the website Archive of Our Own (AO3) has its own elaborate search functions that allow readers to see the tags authors are using to make their works more searchable. The purpose of this research is to see how BISAC codes and keyword searches can be used to find traditionally published LGBTQ-related content and compare them to the content available in AO3. Because this kind of study has never been done before, a lot of it will be setting groundwork with potentials for future studies.

---

1   Laura M. Jiménez, "Representations in Award-Winning LGBTQ Young Adult Literature from 2000–2013," Journal of Lesbian Studies 19, no. 4 (October 2015): 406–22, https://doi.org/10.1080/10894160.2015.1057795.

Methods

In order to analyze how different tagging systems are working in different communities, I pulled data from NPD Decision Key, the Barnes & Noble (B&N) website, Goodreads, and the popular fanfiction website AO3. I pulled a complete list of LGBTQ-related BISAC codes from the Book Industry Study Group (BISG) website (see Appendix A). The list contains forty codes that I used to create reports in Decision Key. I used a standard "Query by BISAC" search and changed the dates to include everything from January 1, 2017 to the present.[2] I chose this date because it is within the two-year range that most publishing houses build their comparative titles around when acquiring a new book or doing marketing for an existing book. I built a report for each BISAC code, searching for titles that listed that code as the primary BISAC only. I removed all titles published in physical audio, ebook, and mass market formats and removed all duplicate titles so that each book was only counted once.[3] These included special or large print editions and omnibuses or collections.[4] If the collection was listed but not the original than I kept it. I also removed titles published after April 30, 2019 because in analyzing the data I broke each year into three four month sections, where April 30 marked the end of a section.

Next I collected a series of LGBTQ-related keywords and phrases to run through B&N, Goodreads, and AO3. These keywords were pulled from LGBTQ lists on Goodreads, LGBTQ focused websites, and my own knowledge of the community, and were collected over a period of two weeks as new keywords emerged in my secondary research. When the list was compiled I ran it through B&N and Goodreads, ending with AO3. I wanted to use a bookstore website instead of searching on Amazon to better understand how a brick and mortar store would organize its content. Barnes & Noble is the last remaining nationally recognized bookstore chain, but it also has a large online retail component, which I thought may give them a fairly searchable website. Goodreads, as a reader community, should have a large scope of content since there are other readers interested in reading and sharing content. AO3 is a sprawling

---

2 May 4, 2019 was the most recent data pull.
3 Duplicate titles were removed by sorting the titles alphabetically and removing the lowest ranked duplicates (ones that have sold fewer copies between January 1, 2017 and present). If the ranking was the same, I removed the newest edition to minimize the risk of skewing my data with reprints.
4 There were several titles that changed between editions, either hardcover to paperback, or just updated reprintings. I have chosen to only count these titles once.

mass of fan-related content that has intricate and advanced search options that allow users to find content in many different ways (see analysis). One of these is to search by tags, which searches the entire Archive rather than a specific fandom, and this is the tool I used to run my keywords and phrases.

Each platform has its own purpose, and therefore its own organizational system. For example, BISAC codes are industry-generated forms of metadata made to inform bookstores and libraries about where to shelve books, while Goodreads uses book descriptions and reader-generated keywords and lists to help like-minded readers find similar books and reader communities. Author-generated fanfiction tags on AO3 operate similarly in that they can be used to narrow the search field and help a reader find specific content. My analysis will include comparing search results for LGBTQ-related content across these platforms to see how easily readers interested in this content can find it with the goal of seeing how these systems are working and whether they are working for or against readers interested in LGBTQ-related books.

The Industry Now

According to a study done by Malinda Lo in 2012, only 1.6% of all young adult books published that year featured an LGBTQ representative protagonist.[5] Of those, she was able to categorize character gender with only four genders: female, male, trans, and multiple. Lo provides no data on the percentage of LGBTQ YA books published in later years, but in her 2016 study she was able to broaden those categories to include the following: cisgender male, cisgender female, transgender male, transgender female, gender-destabilizing,[6] intersex, nonbinary, genderfluid, multiple, and problem novels.[7] These expanded categories show a changing attitude towards including these underrepresented queer identities.[8] Though there is clearly a move towards greater inclusivity, there is also still strong bias in favor of cisgender male protagonists. A study in 2015 shows that only 24% of protagonists in award-winning LGBTQ young

5   Malinda Lo, "YA Pride: 2012 LGBT YA by the Numbers," Malinda Lo, accessed February 11, 2019, https://www.malindalo.com/blog/2012/06/ya-pride-2012-lgbt-ya-by-the-numbers.

6   This is a category invented by Lo to include characters in speculative fiction who change gender but are not trans.

7   These are books that feature straight protagonists dealing with queer themes, topics, or stories.

8   Malinda Lo, "LGBTQ YA by the Numbers: 2015-16," Malinda Lo, accessed February 11, 2019, https://www.malindalo.com/blog/2017/10/12/lgbtq-ya-by-the-numbers-2015-16.

adult novels from 2011–2013 were lesbian, and of those none of the characters have a long-lasting or healthy relationship.[9]

Publishing is a notoriously slow changing industry, and one of the ways to measure that change is by looking at BISAC code updates. As recently as 2015 the only BISAC codes for publishers wanting to categorize LGBTQ titles were "homosexual" and "LGBT."[10] The former is considered offensive by many members of the queer community and the latter is simply out of date. The initialism "LGBTQ" has been in use since 1996,[11] and was cited as the the most preferred and inclusive acronym in 2016 when the BISAC codes were updated.[12] Since 2015, many BISAC codes have been added for a total of forty codes (see Appendix A). In 2017, Booknet Canada found that only "75 publishers were using LGBT BISAC codes on their titles," with that number jumping to 435 publishers in 2018.[13] New codes were added between 2017 and 2018 for Bisexual and Transgender titles under LGBT[14] Fiction and LGBT Romance. Since 2018, more codes have been added for LGBT Erotica, Religion (Sexuality & Gender Studies), JUV Fiction, Young Adult, and Social Sciences. These changes are both reflective of changing attitudes about the LGBTQ community and also of the slowness of change in publishing, since all of these codes, though recently added, are still being tagged as "LGBT." That said, the BISAC codes do allow for additional content with an "LGBT General" code in Fiction, Romance, Erotica, and Social Sciences, but even this broad pool could become easily convoluted in years to come with the increase in LGBTQ content and the continued expansion of the queer spectrum and the limitless nature of identity.

Despite the obvious problems with these codes it is still unquestionably a positive move towards inclusivity, and it is heartening to see that these codes are already being

9   Jiménez, "Representations in Award-Winning LGBTQ Young Adult Literature from 2000–2013."

10   Kira Harkonen, "Taking Pride in New BISAC Codes," BookNet Canada, accessed May 2, 2019, https://www.booknetcanada.ca/blog/2018/6/21/taking-pride-in-new-bisac-codes.

11   Steven Petrow, "Civilities: What Does the Acronym 'LGBTQ' Stand For?," Washington Post, accessed May 11, 2019, https://www.washingtonpost.com/news/arts-and-entertainment/wp/2014/05/23/civilities-what-does-the-acronym-lgbtq-stand-for/.

12   Trudy Ring, "Expanding the Acronym: GLAAD Adds the Q to LGBT," Advocate, October 26, 2016, http://www.advocate.com/media/2016/10/26/expanding-acronym-glaad-adds-q-lgbt.

13   Harkonen, "Taking Pride in New BISAC Codes."

14   When discussing BISAC codes, the initialism LGBT is used, as that is the language used by BISG and throughout the industry generally.

used, even if the number of books using them as primary BISAC codes is small. The purpose of BISAC codes is to identify at a simultaneously general and specific level what a book is about, and so having more options to choose from could make it easier for publishers to find markets for their books. However, the structure of BISAC codes is such that publishers will always have to make a choice between what comes first in the classification of a book and what comes after. There is also the danger that what the publisher deems worthy of primary order BISAC categorization will be something that a reader will find of secondary importance. It is perhaps a reason why BISAC codes are doomed to fail. Categorizing a book as one thing limits the potential it has to a reader, and is perhaps why keyword optimization has grown in importance across the industry.

BISAC codes are meant to help bookstores and even libraries organize titles in ways that make it easier for readers to browse, but a reader looking for their next favorite book is not searching for BISAC codes. They are searching keywords on Google or in their local bookstore or library databases and finding lists on Goodreads made by other readers with the same interests. Goodreads is a key database for readers because it allows a high level of community involvement through reviews—which have a limit of 20,000 characters—and the creation and curation of reading lists. Goodreads can be an honest look into an individual's book buying and reading habits. Tracking what books were read and when, how long it took to read, how many friends have read the same book, it all paints what Pandell calls an "intimate inventory of [a reader's] experiences."[15] It is also one of the key draws for readers because it allows them to find other like-minded individuals in places all over the country and world. It has turned into a combination of platforms, equal parts reporting on books read and digital networking opportunities. In some ways, Goodreads is the closest thing to a fanfiction community most traditionally published books will ever see. Readers come together to talk about their favorite books and find a community of peers with the same reading interests, and one of the ways they can do that is through the review process. Readers can tag books with keywords they thought were relevant to the book, and those keywords will come up when other readers search for them. This is a great way to expand on the metadata created by the publisher and to fill in the gaps left by potentially snappy and vague back cover copy.

Contrary to industry standards, the world of fanfiction has long been a digital mecca for queer content. It is a platform for "essentially anyone who doesn't identify

---

15   Lexi Pandell, "Goodreads Is Finally Cashing in on Its Devoted Community," Wired, May 19, 2016, https://www.wired.com/2016/05/goodreads-selling-books/.

as a straight, white, cis man" to create and explore the kinds of non-heteronormative stories and relationships that are unavailable in mainstream media.[16] Frequently this manifests as a fanfiction writer taking a plot of a popular TV show, movie, or book and rewriting elements of it to create a new queer narrative. Examples of this include the fancast relationship of Sherlock Holmes and John Watson in BBC's *Sherlock* as a romantic one (affectionately called "Johnlock" in the fandom), writing Sirius Black as a trans demisexual man, or writing the popular all-female group Little Mix into a polyamorous relationship. Fanfiction is a place where a broader demographic of viewers, readers, artists, and fans have the opportunity to "subvert the mainstream perspective [by fracturing] a story and recasting it in their own way."[17] In this way it is possible for fans to take their own experiences, interests, and identities and express them through their favorite characters and since a fan base is built in, those experiences are guaranteed an audience of readers who are also looking for narratives beyond the heteronormative ones perpetuated by mainstream media.

AO3 was founded in 2007 as a branch of the nonprofit Organization for Transformative Works. The site is in its simplest form, a fanfiction hosting site operated by a small army of over 700 volunteers serving as everything from coders to lawyers.[18] The site currently contains over 4.5 million written works representing over 312,000 fandoms and 1.8 million registered users with roughly 3,000 new users per week, all using a system of tags to help readers find the content they want, and avoid the content they do not.[19] Because it is a nonprofit, it has been allowed to grow at its own pace, and has been able to maintain an open-door policy on the fanfiction that it allows users to post, namely whatever they want. Rather than policing the site for taboo works, AO3 trusts its readers to be their own gatekeepers and decide for themselves what content they wish to engage with. One way users can do this is through the elaborate tagging system that allows authors to create

---

16   Sarah Manavis, "Internet Fanfiction Is Becoming Mainstream – but Not Necessarily More Representative," NewStatesmanAmerica, accessed March 13, 2019, https://www.newstatesman.com/culture/fiction/2018/11/internet-fanfiction-becoming-mainstream-after-movie-harry-styles-potter-one-direction.

17   Manavis.

18   Caitlin Busch, "An Archive of Our Own: How AO3 Built a Nonprofit Fanfiction Empire and Safe Haven | SYFY WIRE," SYFY WIRE, February 19, 2019, https://web.archive.org/web/20190219085837/https://www.syfy.com/syfywire/an-archive-of-our-own-how-ao3-built-a-nonprofit-fanfiction-empire-and-safe-haven.

19   Busch.

tags, while volunteers for the archive "wrangle" the tags to keep the overall library organized. These author-generated tags can be as vague or detailed as an author chooses, and can be highly personalized and inventive since AO3 has roughly 300 volunteers working behind the scenes to "corral the community's enthusiasm" with much stricter tagging guidelines.[20] For example, there are 30,060 unique author-generated tags in the archive that use the keyword "gay" at the time of this writing. This includes everything from "Canon Gay Character" to "yes i love my gay flowerboys" and "no gay panic." Each of these unique tags can be associated with one or more fanfictions—"Canon Gay Character" has been used by 1,753 fics while "yes i love my gay flowerboys" has only been used by one—but all of them are searchable for users interested in searching for content based on tagged queer-related keywords and can be changed at any time. This is a significant factor in fanfiction as some fanfic authors change tags at the request of readers to mark things like trigger warnings and past or background character relationships. This searchability gives fanfiction an edge that traditionally published books lack because they are limited to BISAC codes, back cover copy, and the insertion of keywords into title metadata which may or may not include relevant keywords.

Searchability

BISAC Codes

The BISG currently has forty BISAC codes that include the keyword "LGBT." These codes appear on a total of 17,409 unique titles that have sold at least one copy since January 01, 2017. Of these titles 73% are coded as fiction and 60% are coded as "Gay." To better understand the break down of these titles, I organized them according to the publication date and isolated the books published between January 1, 2017–April 30, 2019. I further separated each year into three four month terms in order to best incorporate the limited data. I analyzed each BISAC individually and then combined them according to expressed sexuality. For example, the data for the five BISAC codes with the word "gay" in the code description were combined,[21] and the same for the four codes each for "lesbian," "bisexual," and "transgender." Codes marked "general"

---

20   Busch.
21   The CCN004210 code used for Manga/ Yaoi was combined with the gay-specific codes because 'Yaoi' translates to "boy love," which is a genre of manga with a M/M pairing.

or left unspecified as simply "LGBT" were also grouped together. I expected to see patterns in when content was being published, for example, a spike in releases during the second third of every year (May 1–Aug 31) since June is Pride Month. However, I only saw a noticeable increase in sales during this period in the combined gay-specific data, and when observing the combined fiction data. All other groups had no noticeable correlation in their spikes during this time.

I also grouped the fiction codes together, since they make up a substantial majority of the data. Adult fiction makes up 94% of all fiction coded titles. Within adult fiction, 53% have no further separation other than sexuality-specific subheadings. Forty-two percent are subcategorized as Fiction / Romance, with the remaining 5% subcategorized as Fiction / Erotica. This is interesting because it does not support some of the stigma about queer literature that I have encountered in and outside of my own queer communities, namely that it is dominated by hardcore and borderline pornographic content. That is not to say that those stories cannot still be found in titles that have been categorized under Romance, but usually erotic titles are self-proclamatory, and using the provided Erotica categorization would increase the visibility of that proclamation in bookstores that provided those sections. (B&N

LGBT Manga

Gay Erotica

Gay Romance
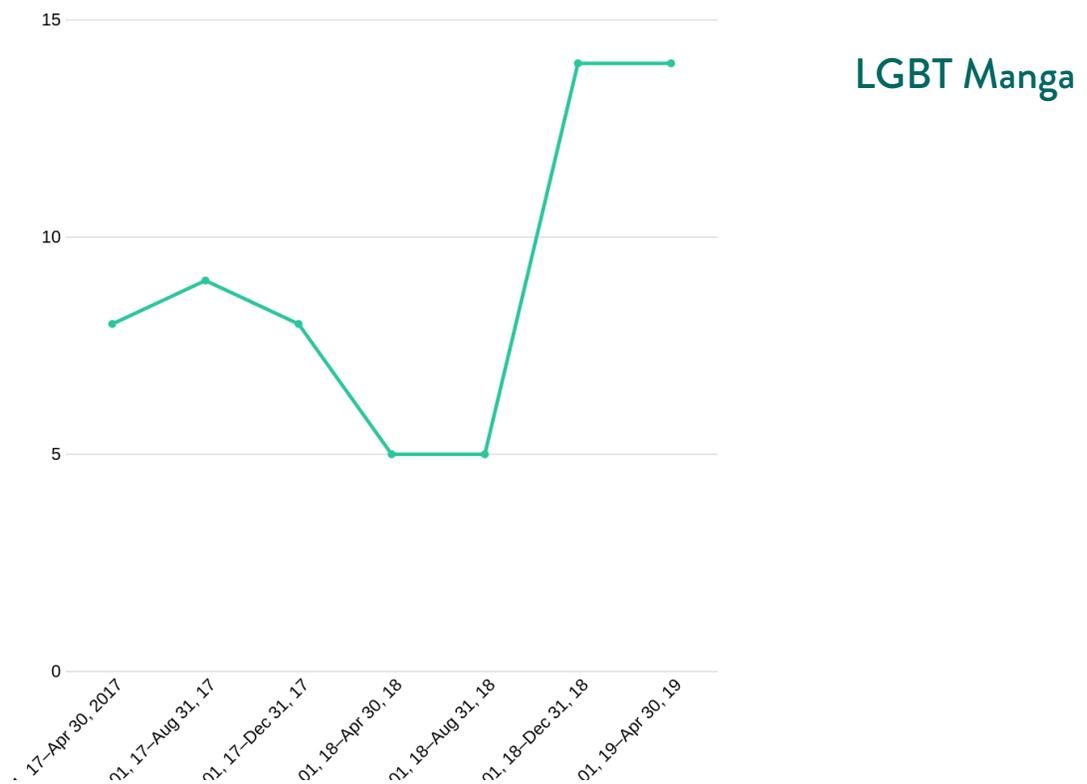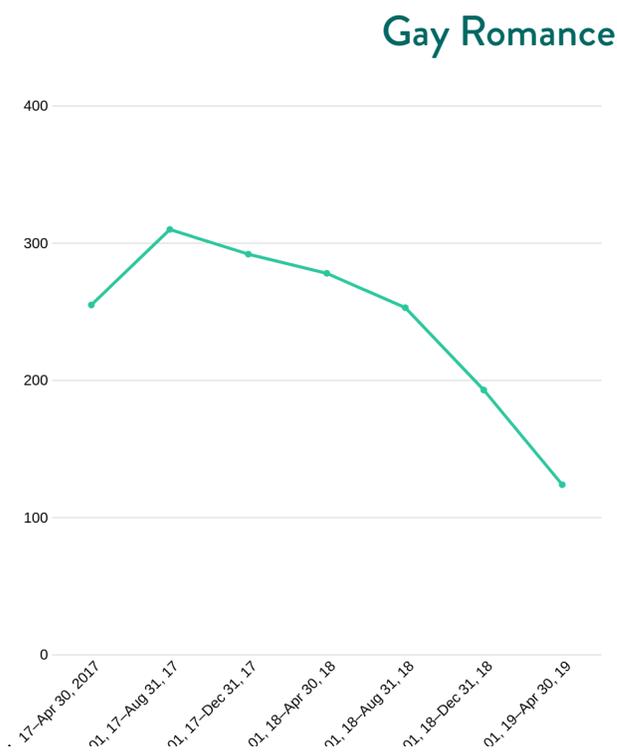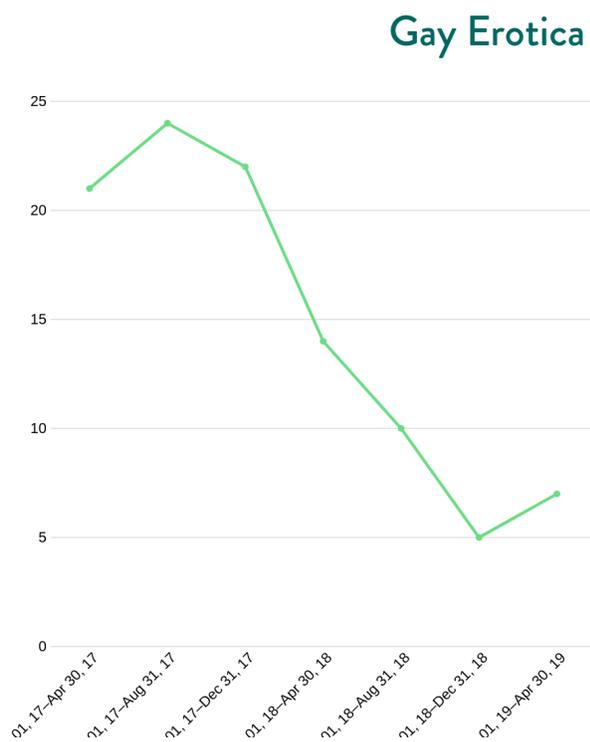
combines erotica and romance into one category.) There is currently no data available for general LGBT erotica.

The codes for LGBT manga, JUV fiction, literary collections, and gay studies in social sciences all showed an increase in the use of their codes over the two year period. I was surprised that there were only four codes with this pattern as I had expected to see an increase in more of them as the attitudes in western society began to shift away from stigmatizing queer identities and relationships. However, publishing is a slow changing industry, and its possible that greater increases in the number of titles using these codes will take much longer to see. It is also possible that there has been a shift in the industry away from labeling titles with LGBT codes in order to increase the chance of other readers who are not specifically interested in these titles to find them.

Another trend in the data was the decrease in use over time of the codes for gay erotica and gay romance. This surprised me because gay-specific codes make up 55% of the total data for LGBTQ-related BISAC codes. Two explanations I have for this are 1) publishers are broadening the range of queer representation in the books they publish, thereby publishing fewer gay-specific books, and 2) that this is

representative of a decrease in all LGBTQ-related codes. The first is optimistic and unsupported by the very limited data on non-gay LGBTQ codes, and the second is true to my data, which shows an average of 904 queer titles per four month period until the last data set. However, this data could be skewed by the proximity to that time. It's possible that there are books published in the last period that have not sold a book yet, and therefore would not appear in Decision Key. It is also problematic that M/M content makes up more than half of the available LGBTQ content. This shows that even in a minority population, the male perspective is still the most normalized. In contrast, lesbian-specific content makes up 18% while general LGBTQ content is 26% and the remaining 1% is divided between bisexual and transgender-specific titles.

The remaining data is very inconclusive. Lesbian erotica and poetry both show erratic spikes and falls, while Yaoi manga and drama have spikes in Sep 2018–Dec 2018 followed by a fall in the final period. There is also the problem of other sexuality-specific BISAC codes, such as bisexual and transgender. These codes were just added by the BISG in 2017, but there doesn't seem to be a high demand for them as evidenced by their single digits data. I was not surprised by the bisexual codes since the invisibility of bisexuality as an identity in media and queer culture is a much discussed topic, but I assumed that the numbers for transgender books would be higher because there has been so much media attention in the last five or so years to transgender topics.

In engaging with these codes, I noticed some problems. First, the identifier is "LGBT" and not the more commonly accepted LGBTQ, or the more recently introduced initialism LGBTQIA+. This would not be a difficult change to make, and it is disheartening to see this outdated initialism persisting. I also noticed that there are no LGBT codes in sections like History or Self-Help. Instead, Self-Help has "Gender & Identity," and "Sexuality."[22] There are also no codes related to LGBTQ families, and no fiction subcategories beyond romance and erotica (i.e. mystery, historical, fantasy, science fiction). Another discrepancy I found was that there is a yaoi manga code but not one for yuri manga, the lesbian equivalent. This goes back to the conversation on the normalization of the male perspective, and while interesting, is a different research topic.

---

22   These are clearly related BISAC categories with LGBT overlap, however, I chose not to include them in my data set because they are not LGBT-specific. Religion is also technically not LGBT-specific, but it includes the specific LGBT code in the expanded description (see Appendix A).

Keyword Searches
B&N and Goodreads

When curating my list of keywords, I tried to consider everything a casual reader may search for when looking for a queer book. This included a range of sexuality identifiers such as aromantic, demisexual, and nonbinary, as well as queer themes like coming out, crossdressing, and sexual experimentation. I thought that running these terms through the B&N website and Goodreads would reveal more than just books with these keywords included in the title. Especially in the case of B&N, where they as the bookseller are trying to accommodate as many search functions as possible in order to promote sales. They also have a more direct link with the publisher than the general consumer, meaning they should have access to title metadata, which should include searchable keywords. This did not seem to be the case. Running these key-words through B&N came up with only items that had the keyword in the title. This still uncovered a fairly large data set, but some of them are skewed. For example, the keyword "Gay" pulled 30,245 items from B&N and 28,066 titles for Goodreads, but this does not account for duplicate titles in different formats or for author names like Roxane Gay and John Gay. Unfortunately for B&N, there is no way to narrow that search to a simple keyword search, or to exclude items that were not books like movies, toys, or games.[23]

The B&N website does have several browsable categories dedicated to LGBT fiction (11,959), gay and lesbian manga (129), gay, lesbian, and LGBT romance (1,446), and lesbian and gay teen fiction (596). Though these categories exist, they are difficult to find. Readers intentionally searching out this content may be more inclined to put in the work, but a casual browser may miss it if it lives in such a hidden place. Additionally, I noticed that even among these categories there were some books that I had read or knew about that were missing or incorrect. For example, *A Lady's Guide to Petticoats and Piracy* by Mackenzie Lee features an asexual protagonist but is for some reason shelved in YA Romance. Perhaps even more upsetting is the fact that this book is the second of a series, and that the first, which features her gay older brother Percy, is shelved in regular YA Fiction.

Goodreads had the same problem pulling books that only had the keyword in the title or as part of an author name. However, unlike B&N, Goodreads allows you to choose from several search options: all, title, author, or genre. I chose to run the

---

23   I did not scroll through every page of every keyword search, so I do not know if these keywords pulled any non-book materials, but they could have.

AO3 Search Options

Genre: Format of the source material, i.e. anime and manga, books and literature, movies, music and bands, tv shows, etc

Fandoms: i.e. *Harry Potter*, *Sherlock*, BTS, and Marvel

Ratings: General, Teen, PG-13, Mature, and Explicit

Warnings: e.g. graphic depictions of violence

Categories: Used to describe the nature of the relationships, i.e. M/M, F/F, F/M, or Multi

Additional Fandoms: generally used for crossover fanfiction with multiple fandoms

Relationships: describes the kind of relationships depicted in the work. Can be both romantic (denoted by a '/') and platonic (denoted by a '&').

Additional Tags: AO3 pulls a list of the top ten additional tags, usually describing themes like romance or drama (called "fluff" and "angst"). Readers can also type their own tags into a provided search bar.

Language: Some fics have been translated into other languages, but there are also fics written in languages other than English.

Hits: Number of times it has been opened or read.

Bookmarks: Number of times a work has been saved by a user.

Kudos: Another word for likes, marked with a heart icon.

keywords through "all" in the hope that it would pull more content, which it did, so I also ran a few using the genre search. These pulled fewer titles but used metadata instead of title words, so they may be more accurate. However, I don't believe the average reader would search for content that way, and so I did not include those figures in my data set.

AO3

Fanfiction websites like AO3 have a very extensive and easily searchable database of content. Readers can browse broadly by genre or narrowly by a fandom. They can narrow the search further by choosing to either include or exclude specific tags based on categories such as rating, warnings, categories, fandoms, characters, relationships, additional tags, completion status, word count, date updated, and language. Readers can also decide how these search results appear, either by date added, date updated, author name, title, word count, hits, bookmarks, kudos, and comments.

This is a vast difference from the search options in B&N and Goodreads. The volume of searchable choices can be overwhelming to a casual reader, but to an avid fan with clearly defined tastes it is the perfect browsing experience. Performing a keyword search in the archive tags was astonishing. Where "polyamory" pulled up 62 items in B&N and 154 in Goodreads, it pulled up 36,214 works in AO3, which is more than 580 times the amount of works in B&N. This extremely disproportionate difference between traditional coding systems and the author-generated free for all of AO3 tags continues throughout the entire list. It is most noticeable in the gender expression, identity, and sexuality keywords like agender,

demisexual, genderfluid, and nonbinary. Overall, compound keywords like "gay romance" and "lesbian fiction" did better in B&N and Goodreads than AO3. Other keywords that performed better outside of AO3 include: bisexual, drag, gay, gay fiction, gay manga, gay romance, gender, GLBTQ, lesbian, lesbian fiction, lesbian manga, lesbian romance, LGBT, LGBTQ, queer, questioning, and sexual orientation. However, this does not include the full data set. When a keyword is searched in AO3's "search by tag" search field, the archive will pull up every single tag that uses that keyword and then show the number of works that use that tag in parentheses. For this data, I found the exact tag that matched each keyword search and used the number of works that use that tag in order to accurately compare the number of works found in B&N and Goodreads to the number of works in AO3.

However, that means that my searches are excluding all of the other unique and potentially relevant tags that fanfic authors have created and used on their works, which is another obscenely large data set. Just to show the scale, let's say that every unique tag other than the exact keyword match have only been used once. Then the total number of exact keyword matches can be added to the number of unique tags. Using the keyword "gay" as an example, that puts the total potential number of works

Keyword Comparisons
B&N
Goodreads
AO3
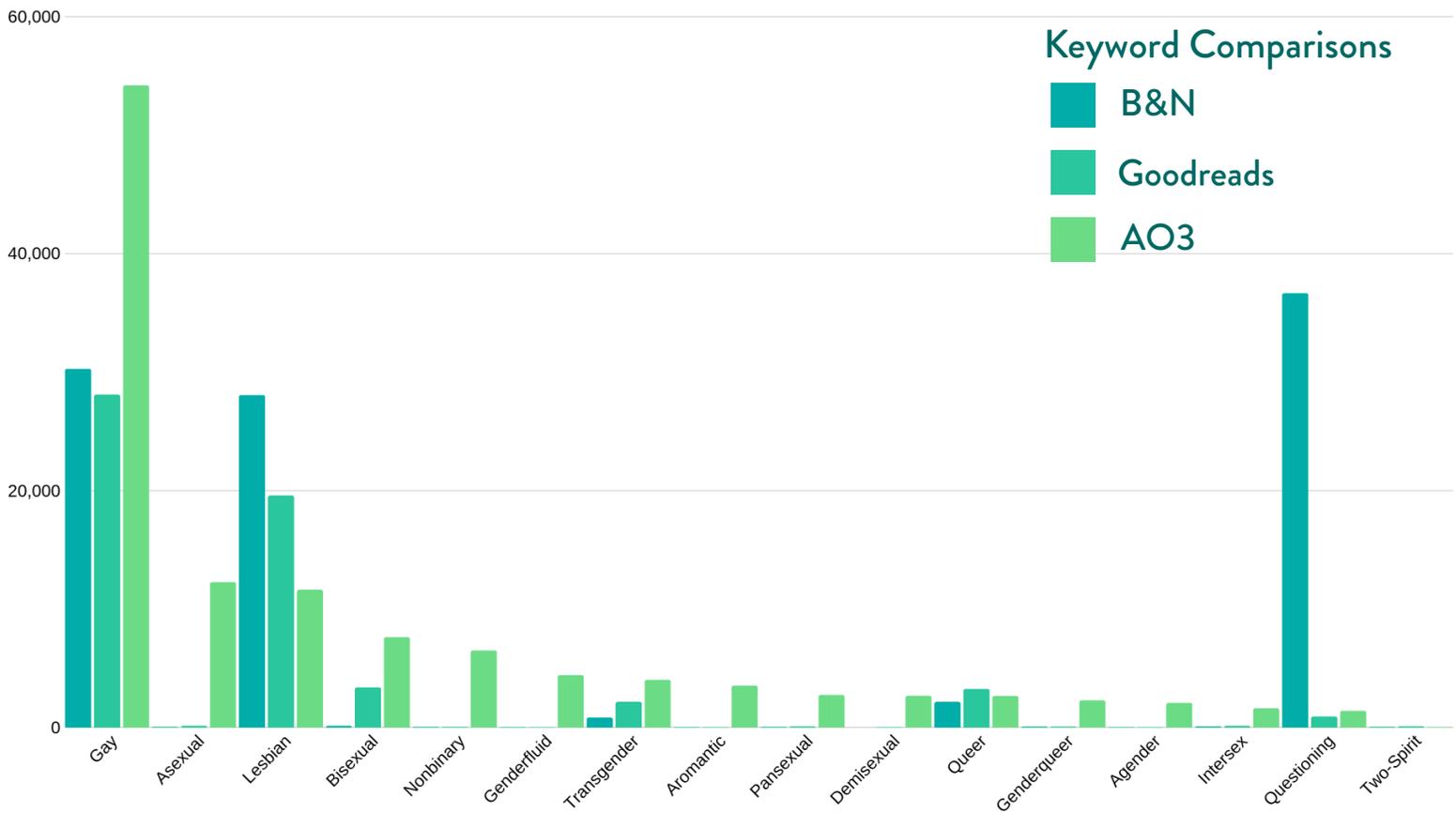
in AO3 tagged as "gay" at 54,176 works which is just over 4,000 works shy of being the total items in B&N and Goodreads combined. Even adding these figures together doesn't tip the scales for any keyword other than "gay." Lesbian content is shockingly even more underrepresented on AO3 than it is in traditional publishing losing second place to asexual content. Bisexual stories are ranked fourth and transgender coming in seventh. Despite the extensive search capabilities of the archive, I found that there was in fact something B&N and Goodreads did better, which was to include extensions of a keyword in the search results. For example, in AO3 if I search "demisexual" it will only pull up tags that use the word "demisexual" with no other variations. B&N and Goodreads both include additional results, such as "demisexuals" and "demisexuality." This is problematic to readers searching additional tags in a specific fandom because it has the potential to limit the results too much, or not enough.

Limitations

The first limitation I found deals with BISAC codes which, as industry standards, are not something a casual consumer can type into a search field on a website like that of B&N. Publishers almost never keep lists of the BISAC codes they use for their books in a place that is readily accessible to the general public. This means that readers have no way of knowing whether a book has been coded with an LGBTQ-related BISAC unless it is actually shelved according to those codes. This leads to the second limitation, which is the shelving guidelines of bookstores. These guidelines undoubtedly change from store to store, but companies like B&N that have all of their shelving decisions made on a corporate level, may never shelve a book in its coded location, as the corporate body may decide to code it themselves as something different.

Additionally, BISAC codes are updated every year. On the one hand, this means that outdated codes like "homosexuality" can be retired and replaced with more appropriate, less offensive, and more representative variations. On the other hand, not every book is recategorized after its initial printing, so some books will not be updated even if a better suited code is created. The standard amount of BISAC codes a book can have is four, but most publishers stick to one or two, especially as BISACs become increasingly specific. Another limitation is the functionality of Decision Key, which is most useful for tracking sales data, which

I was not interested in during this preliminary evaluation. I was not able to run a BISAC query that would pull a complete list of every book ever published using the forty LGBTQ-related BISAC codes.

The keyword searches also had limitations. For example, the keyword searches performed in B&N and Goodreads did not account for duplicate title formats (hardcover vs. paperback vs. ebook). They also did not search metadata keywords, but used the keyword as part of a book title. Some keywords have meaning outside of the LGBTQ community, such as "gay," which can also lead a reader to any book written by Roxane Gay, and "questioning," "sexual experimentation," and "coming out," which pulled the book of the same name by Danielle Steel. I chose not to remove these instances from the data, as they imitated a natural search that may have been performed by a casual reader.

Conclusion

There are clearly a lot of barriers to performing accurate and comparable research in this area of publishing. On the one hand, the industry has Decision Key for tracking sales data, which is important to the industry at large, but this database is limited in many ways, including searching info beyond sales, despite tracking the data anyway (i.e. searching by pub date regardless of sales data). Even the BISAC queries are limited to sales history, which isn't relevant to the purpose of this study. Similarly, retailers like B&N and reader community websites like Goodreads are not as keyword-friendly as they should be, meaning that readers could have trouble finding the representation they are looking for in a traditionally published work. Alternatively, AO3 has an extensive search category that incorporates everything from maturity ratings to searching customized tags. These tags can be highly specific while remaining spoiler-free and allowing the reader to have useful additional information about the story before they begin reading. This is a search feature publishers should be mirroring because it allows readers to see the kinds of metadata tags publishers are using and claim a sense of ownership by choosing a book that suits their interests on a more customized level than simply reading the back cover copy. Moving to a taggable model would allow publishers to include more relevant information about the book without relying on metadata and keyword search habits of the average reader. This would benefit the entire sales chain

from publisher to consumer. Bookstores would have greater flexibility in creating unique store lists and promos, to a broader range of searchable categorizations that aren't limited to BISAC codes. It would also benefit the consumer as it could give them much more information about the title and encourage them to make more informed choices about their reading habits. At the end of the day, publishing is about putting a book in the hands of someone who is going to enjoy it, and I believe that moving towards better searchability through improved keyword searches is an important first step towards that goal.

Process and Acknowledgments

My interest in this topic came about as a result of a convergence of my two great loves: fanfiction and all things queer. These interests set me up for several research papers during my time in the Book Publishing master's program at Portland State University, including a paper on fanfiction and copyright law. Originally, I had wanted to continue that research for my final project, but realized that I was more interested in analyzing queer representation in fanfiction compared to queer representation in traditional publishing. It was only through my preliminary research that I considered focusing on the searchability of content across platforms, and whether one is better than another. However, because this work has never been done before, a lot of what I have discovered is intended as groundwork for future research.

I could not have performed this research without the wisdom and guidance of my core faculty and faculty advisors—Dr. Rachel Noorda, Kathi Inman Berens, and Abbey Gaterud. Each of them supported me with their advice, good humor, and enthusiasm, and the research would not be what it is without them. I would also like to thank my peers in the program for their solidarity and encouragement to me and one another, and to Kelley Dodd, Kent Watson and Brittney Finato for providing additional feedback on this research.

Bibliography

Busch, Caitlin. "An Archive of Our Own: How AO3 Built a Nonprofit Fanfiction Empire and Safe Haven | SYFY WIRE." SYFY WIRE, February 19, 2019. https://web.archive.org/web/20190219085837/https://www.syfy.com/syfywire/an-archive-of-our-own-how-ao3-built-a-nonprofit-fanfiction-empire-and-safe-haven.

Harkonen, Kira. "Taking Pride in New BISAC Codes." BookNet Canada. Accessed May 2, 2019. https://www.booknetcanada.ca/blog/2018/6/21/taking-pride-in-new-bisac-codes.

Jiménez, Laura M. "Representations in Award-Winning LGBTQ Young Adult Literature from 2000–2013." Journal of Lesbian Studies 19, no. 4 (October 2015): 406–22. https://doi.org/10.1080/10894160.2015.1057795.

Lo, Malinda. "LGBTQ YA by the Numbers: 2015-16." Malinda Lo. Accessed February 11, 2019. https://www.malindalo.com/blog/2017/10/12/lgbtq-ya-by-the-numbers-2015-16.

———. "YA Pride: 2012 LGBT YA by the Numbers." Malinda Lo. Accessed February 11, 2019. https://www.malindalo.com/blog/2012/06/ya-pride-2012-lgbt-ya-by-the-numbers.

Manavis, Sarah. "Internet Fanfiction Is Becoming Mainstream – but Not Necessarily More Representative." NewStatesmanAmerica. Accessed March 13, 2019. https://www.newstatesman.com/culture/fiction/2018/11/internet-fanfiction-becoming-mainstream-after-movie-harry-styles-potter-one-direction.

Pandell, Lexi. "Goodreads Is Finally Cashing in on Its Devoted Community." Wired, May 19, 2016. https://www.wired.com/2016/05/goodreads-selling-books/.

Petrow, Steven. "Civilities: What Does the Acronym 'LGBTQ' Stand For?" Washington Post. Accessed May 11, 2019. https://www.washingtonpost.com/news/arts-and-entertainment/wp/2014/05/23/civilities-what-does-the-acronym-lgbtq-stand-for/.

Ring, Trudy. "Expanding the Acronym: GLAAD Adds the Q to LGBT." Advocate, October 26, 2016. http://www.advocate.com/media/2016/10/26/expanding-acronym-glaad-adds-q-lgbt.

Appendix A: BISAC Codes

BIO031000 BIOGRAPHY & AUTOBIOGRAPHY / LGBT

CGN009000 COMICS & GRAPHIC NOVELS / LGBT

CGN004130 COMICS & GRAPHIC NOVELS / Manga / LGBT

CGN004210 COMICS & GRAPHIC NOVELS / Manga / Yaoi

DRA017000 DRAMA / LGBT

FAM056000 FAMILY & RELATIONSHIPS / LGBT

FIC005070 FICTION / Erotica / LGBT / General *
- NO DATA FOR THIS CODE

FIC005080 FICTION / Erotica / LGBT / Bisexual

FIC005030 FICTION / Erotica / LGBT / Gay

FIC005040 FICTION / Erotica / LGBT / Lesbian

FIC005090 FICTION / Erotica / LGBT / Transgender

FIC068000 FICTION / LGBT / General

FIC072000 FICTION / LGBT / Bisexual

FIC011000 FICTION / LGBT / Gay

FIC018000 FICTION / LGBT / Lesbian

FIC073000 FICTION / LGBT / Transgender

FIC027300 FICTION / Romance / LGBT / General

FIC027390 FICTION / Romance / LGBT / Bisexual

FIC027190 FICTION / Romance / LGBT / Gay

FIC027210 FICTION / Romance / LGBT / Lesbian

FIC027400 FICTION / Romance / LGBT / Transgender

HUM024000 HUMOR / Topic / LGBT / General

JUV060000 JUVENILE FICTION / LGBT
- JUVENILE FICTION / Social Themes / Homosexuality see LGBT
- JUVENILE FICTION / Social Themes / Sexuality see Social Themes / Dating & Relationships

JNF007150 JUVENILE NONFICTION / Biography & Autobiography / LGBT

JNF053080 JUVENILE NONFICTION / LGBT

LCO016000 LITERARY COLLECTIONS / LGBT

LIT004160 LITERARY CRITICISM / LGBT

POE021000 POETRY / LGBT

REL105000 RELIGION / Sexuality & Gender Studies
   • RELIGION / LGBT see Sexuality & Gender Studies

SOC064000 SOCIAL SCIENCE / LGBT Studies / General

SOC064010 SOCIAL SCIENCE / LGBT Studies / Bisexual Studies

SOC012000 SOCIAL SCIENCE / LGBT Studies / Gay Studies

SOC017000 SOCIAL SCIENCE / LGBT Studies / Lesbian Studies

SOC064020 SOCIAL SCIENCE / LGBT Studies / Transgender Studies

TRV026070 TRAVEL / Special Interest / LGBT

YAF010140 YOUNG ADULT FICTION / Comics & Graphic Novels / LGBT

YAF031000 YOUNG ADULT FICTION / LGBT

YAF052040 YOUNG ADULT FICTION / Romance / LGBT

YAN006150 YOUNG ADULT NONFICTION / Biography & Autobiography / LGBT

YAN032000 YOUNG ADULT NONFICTION / LGBT

Appendix B: Keywords

Agender
Aromantic
Asexual
Bisexual
Coming of age
Coming out
Crossdress
Demisexual
Drag
Gay
Gay Fiction
Gay Manga
Gay Romance
Gender
Genderfluid
Genderqueer
GLBT
GLBTQ
Homosexual
Intersex
Lesbian

Lesbian Fiction
Lesbian Manga
Lesbian Romance
LGBT
LGBTQ
LGBTQ Characters
LGBTQ Themes
Nonbinary
Pansexual
Polyamory
Queer
Queer Characters
Queer Themes
Queerplatonic
Questioning
Sexual Experimentation
Sexual Orientation
Sexuality
Transgender
Two-Spirit
Yaoi