

eBOUND
C A N A D A

EFFECTS OF KEYWORDS ON EBOOK SALES AND DISCOVERABILITY IN THE CANADIAN CONTEXT

Prepared by Amanda Lee
eBOUND Canada February 2019

Executive Summary	4
Introduction and Background	5
Keywords in Bibliographic Metadata	5
Objective	5
Introduction	6
Analysis Methodology	7
Methodology	7
Limitations	10
Findings	10
Overview	10
Effect on Sales	11
Effect on Discoverability	15
Conclusions	23
Recommendations for Publishers	23
Appendices	24
Appendix A: Development of Human-Created Keywords	24
Appendix B: Additional Analysis of the Effects of Keywords on Sales	26



Funding for this study was provided by Ontario Creates. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of Ontario Creates or the Government of Ontario. The Government of Ontario and its agencies are in no way bound by the recommendations contained in this document.

Funded by the Government of Canada



EXECUTIVE SUMMARY

eBOUND Canada, a not-for-profit agency dedicated to advancing the digital engagement of Canadian book publishers, commissioned this study to determine the effect of adding keywords to ebook metadata on the sales and discovery of those ebooks.

This report analyzes the change in sales and discovery of three different groups of ebooks: a control group with no keywords, a group with keywords created by human contractors, and a group with keywords created by Firebrand Technologies using machine learning-driven audience analysis. The keywords were distributed to Amazon Canada and Amazon US at the very end of September 2018, data was gathered from October 1 2018 to December 31 2018, and this report compares the study period to the previous three months (July, August and September 2018) as well as to the previous year (October, November and December 2017).

The goal of the study was to determine if there is value to Canadian publishers in adding keywords to their ebook metadata on a large scale and, if so, the most effective method of keyword creation.

Results of this analysis indicate that the addition of keywords, both machine-created and human-created, had no effect on sales regardless of the ebook's category, sales ranking or publication year. However, the ebooks with keywords saw a significant increase in discovery as measured by views on Amazon, and the books with human-created keywords saw the largest increase overall.

While further research into converting views to purchases is needed before clear recommendations can be made, the increase in discovery is significant enough that publishers looking to improve that metric should find value in adding keywords to their bibliographic metadata.

INTRODUCTION AND BACKGROUND

Keywords in Bibliographic Metadata

Keywords are a part of bibliographic metadata that is typically not displayed to the consumer. Few retailers are currently making use of keywords, but the ones that do use them as invisible but searchable—so that a consumer searching Amazon for “quirky Canadian cookbook” finds *Food That Really Schmecks* (Wilfrid Laurier University Press) even though that term isn’t used in the book’s descriptive copy or title. The consumer won’t see that phrase appear anywhere on the page.

Currently keyword standards are in a bit of flux: the current Book Industry Study Group’s (BISG) best practices document recommends a limit of 500 characters¹, which lines up with the ONline Information eXchange (ONIX) bibliographic metadata distribution guidelines. However, as of November 2018, Amazon will only accept 250 bytes in the keywords field. Generally, each character counts as one byte, but there are some special characters such as ‘é’, ‘ä’, and ‘&’ that count as two or more bytes. In contrast, punctuation (including commas, colons and semi-colons) do not count as bytes. Amazon is the only major retailer currently using keywords. Throughout this study, sales data and discoverability data are taken entirely from Amazon.

Keywords are distributed in ONIX using the Subject composite under the heading Keywords.

Objective

The primary objective of this research was to determine if adding keywords to bibliographic metadata had an effect on the discovery and sales of Canadian ebooks. The secondary objective was to determine whether keywords created by human readers or keywords created by machine learning would perform better in the Canadian context. The tertiary objective was to determine which types of books saw the greatest change after keywords were assigned.

The end goal was to provide Canadian publishers with guidance on the use of keywords in bibliographic metadata.

¹ Book Industry Study Group. “Revised Best Practices for Keywords in Metadata.” *BISG*, Feb. 2018, bisg.org/store/ViewProduct.aspx?id=6972954.

Introduction

eBOUND Canada

eBOUND Canada is an independent not-for-profit organization dedicated to advancing Canadian publishers' engagement in the digital marketplace regardless of their size or level of expertise. eBOUND provides digital asset and metadata management, marketing, conversion services, and professional development to its seventy member publishers.

Firebrand Technologies / Kadaxis

Firebrand Technologies is a software and services company that has served the publishing industry for over thirty years. Kadaxis is a technology company which uses artificial intelligence and machine learning to enhance bibliographic metadata. The two companies work together to create keywords by machine-analyzing audience reviews and other text about the book found on the internet.

Firebrand Technologies / Kadaxis will be referred to as “Firebrand” throughout this report.

Amanda Lee, Lead Consultant

Amanda Lee has worked in Canadian publishing for 8 years. She has held positions focused on metadata enhancement and digital distribution platforms at Dundurn Press, HarperCollins Canada and House of Anansi Press / Groundwood Books. She has worked closely with retailers, distributors and industry groups to understand and communicate bibliographic metadata standards, and has worked on metadata improvement projects for eBOUND, LitDistCo, and several independent authors.

ANALYSIS METHODOLOGY

Methodology

This research project had four primary milestones. First, selecting books from eBOUND client publishers' lists to add keywords to as well as books to use in the control group; second, creating and distributing those keywords to Amazon; third, gathering the sales and discoverability data and fourth, analyzing and presenting that data.

Title Selection

Firebrand Technologies was given a list of approximately 8,000 ebooks produced by eBOUND's client publishers. This list contained titles and international standard book numbers (ISBNs). They ran those titles through their predictive analysis model, which assigns each book a score based on the number of reviews that book has on Amazon and Goodreads, as well as its sales ranking and star ranking on Amazon. They test potential keywords against a variety of factors that they know about how Amazon utilizes keywords, including a database of keywords derived directly from Amazon. The score, which is presented as a number, indicates how well Firebrand's AI-driven keyword creation is likely to work as well as their predicted ability to generate high-quality keywords created for that book. The predictive analysis system filtered out approximately 5,000 ebooks which did not have enough reviews to allow the system to create keywords, and then Firebrand returned the remaining 3,000 ISBNs with scores.

After receiving the scores, lead consultant Amanda Lee chose 600 ebooks from the list of 3,000. She sorted out the books that already had keywords assigned, contacted publishers to attain permission to modify and distribute metadata, then selected books with scores that represented the books published by the majority of eBOUND's publishers. Then she further sorted them into categories by choosing from several Book Industry Standards and Communications (BISAC) code main subject headings.

The selected ebooks were split into three groups of 200 ISBNs each: one group to serve as a control group, one group to be given human-created keywords, and one group to be run through Firebrand Technologies / Kadaxis's keyword creation tool. Each group had the same approximate median score and the same approximate number of ISBNs from each BISAC main category.

Unfortunately, the selected group of titles had scores too low for Firebrand Technologies to work with, and many of the titles that scored high enough had already been assigned keywords by their publishers. We compromised by choosing higher-ranked titles for Firebrand, approximately fifty of which already had keywords, and asking those publishers to allow us to replace their keywords with Firebrand's machine-created keywords. Please see page 8 for more information on this limitation.

Keyword Creation and Distribution

Firebrand Technologies / Kadaxis

Firebrand received their 200 ISBNs and used their keyword creation tool to analyze the ebook's audience response and assign keywords using machine learning. Firebrand required one month to assign keywords to their group of 200 ISBNs.

Firebrand's ISBNs had an average of 71 individual keywords and an average length of 907 characters.

Human created

The 200 titles were split into three groups and assigned to three subcontractors trained by Amanda Lee. Keywords were chosen by reading the book's description, the author's biography, professional reviews and reader reviews, as well as by looking at keywords assigned to similar titles using BookNet's Biblio-o-Matic tool.² For more information on the methods used by the contractors, please see Appendix A.

It took the subcontractors approximately twenty minutes to create keywords for each ISBN. The set of human-created keywords was received in two weeks.

The ISBNs assigned to human contractors had an average of 42 individual keywords and an average length of 505 characters.

Distribution

All of the ebooks selected were stored in CoreSource's digital asset management system. The updated metadata of each ISBN, including the control titles, was uploaded to CoreSource and distributed to ebook retailers on September 26th, which allowed enough time for the keywords to be ingested into Amazon's backend. The keywords

² "BookNet Canada Biblio-o-Matic." *Chrome Web Store*, BookNet Canada, June 2018, chrome.google.com/webstore/detail/booknet-canada-biblio-o-m/gggkfkjokoepibflomjbjjffjoejplom.

were tested via search on the Amazon site in late September, and then periodically checked in Amazon’s Vendor Central platform throughout the period of the study to ensure they hadn’t been accidentally overwritten.

Metrics

The effect of keywords was measured using two metrics: sales and glance views.

Sales

The Canadian and US net sales by number of purchases, not income. Sales data came **only** from Amazon and **only** for the ebooks—not their print or audio versions. When we say “all sales”, we refer only to Amazon sales of ebooks.

Glance Views

The total number of glance views for an individual ebook across all Amazon markets. Glance views are a measure of discoverability—glance views show the number of times the Amazon Product Detail Page of the ebook is viewed by a consumer.

Gathering Data

The study took place over the course of three months: October, November and December 2018. Monthly Amazon sales reports were collected for those months, as well as the previous three months and October, November and December 2017.

Weekly Amazon glance view reports were gathered to measure discoverability. These were collected for fifteen weeks (the thirteen weeks of the study and the two weeks prior to the study) as well as for the three months leading up to the study (July, August and September of 2018). The glance view reports present a year-over-year increase, which allowed us to determine the effect on discoverability over the previous year.

Analyzing the Data

The data was analyzed to determine the increase in sales and in glance views. Sales data was analyzed month-over-month during the study period as well as against the previous three months and the same period in the previous year. Discoverability data was analyzed week-over-week during the study period as well as against the previous three months and the same period in the previous year.

Limitations

Title Selection

Because Firebrand Technologies was not able to work with the lower-ranking ebooks which make up the majority of the ISBNs published by eBOUND's publishers, we had to assign them a group of titles that see significantly higher sales than those assigned to the other two groups. In the analysis that follows, this is overcome by making a note of the significant difference or by looking at the change in sales by average percentage change rather than average number of sales. Number of sales are analyzed when looking at individual books.

Many of the high-ranked titles already had keywords assigned by the publishers. This means that approximately fifty of Firebrand's titles were assigned new keywords. It is possible that the effect on discoverability for Firebrand's keywords is smaller due to this discrepancy.

Collected Data

eBOUND Canada works with digital books only, and the scope of this study is such that only sales of ebooks were collected.

It is possible that there was an increase in sales of the print format of our keyworded titles. Consumers searching on Amazon may be clicking through to the digital version, thus increasing glance views, and then selecting and purchasing the print version. Further research collecting sales of all formats of a title may be of benefit here.

FINDINGS

Overview

The findings show that there is no statistically significant change in sales of ebooks with keywords assigned over ebooks without keywords assigned, regardless of the ebook's category or year of publication, and regardless of whether those keywords were assigned using machine learning or by human beings. However, the discoverability of ebooks with keywords assigned was increased over those without, and it was also increased over the same ebook last year and last quarter. This increase was apparent in books across all categories and publication year, and while the increase is more significant in those books with human-assigned keywords, the effect is still apparent in the ebooks with machine-assigned keywords.

Effect on Sales

Keyword-Assigned ebooks vs. Control ebooks

The following looks at the change in sales of the ebooks to which we assigned keywords (regardless of whether the keywords were created by machines or by humans) against the control group—those ebooks that did not have any keywords at all.

The most telling set of data is the average percentage change in sales month-over-month (**Fig. 1**). This chart looks at the average percentage change in Canadian Amazon sales in September 2018 over August 2018 (the month prior to the keywords being distributed), and then October over September (the month the keywords appeared), November over October and December over November. The next chart (**Fig. 2**) is the same but looking at US Amazon sales.

FIGURE 1: Average percentage change in Canadian sales of ebooks with keywords assigned (**Keyword**) versus ebooks without keywords assigned (**Control**), month over month. A **blue** dotted line is used to indicate the start of the study period.

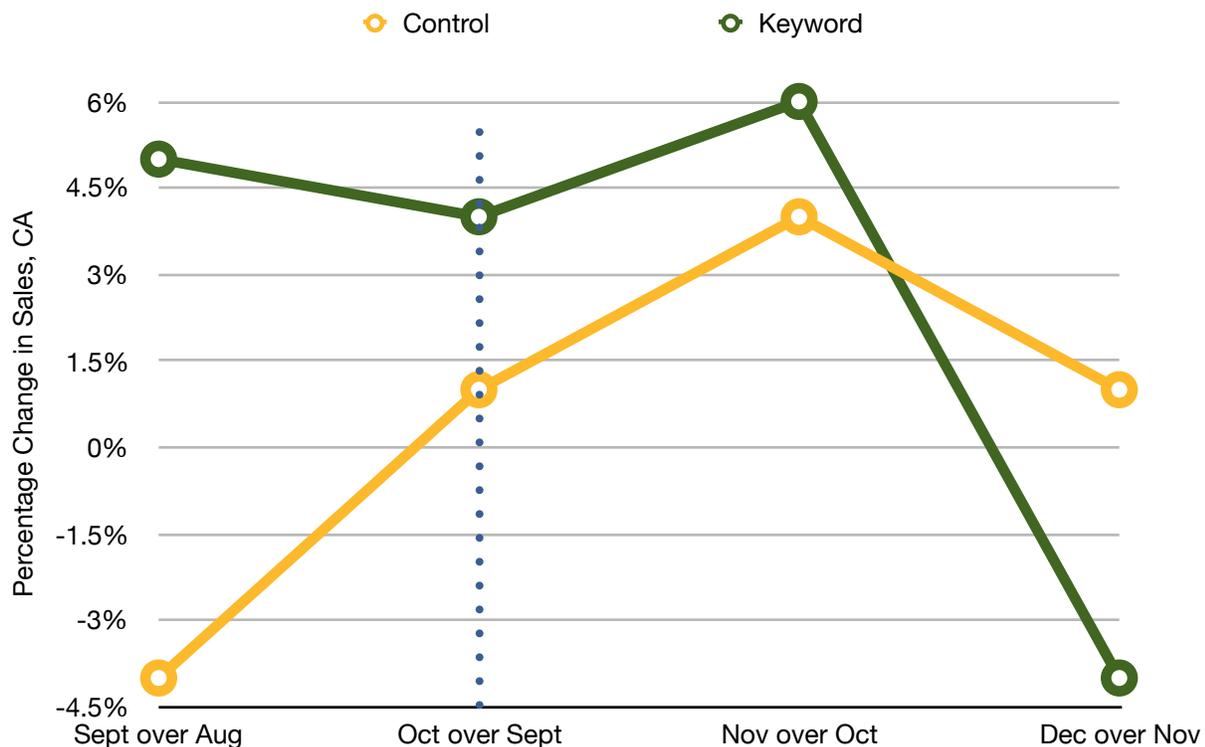


FIGURE 2: Average percentage change in US sales of ebooks with keywords assigned (Keyword) versus ebooks without keywords assigned (Control), month over month. A blue dotted line is used to indicate the start of the study period.

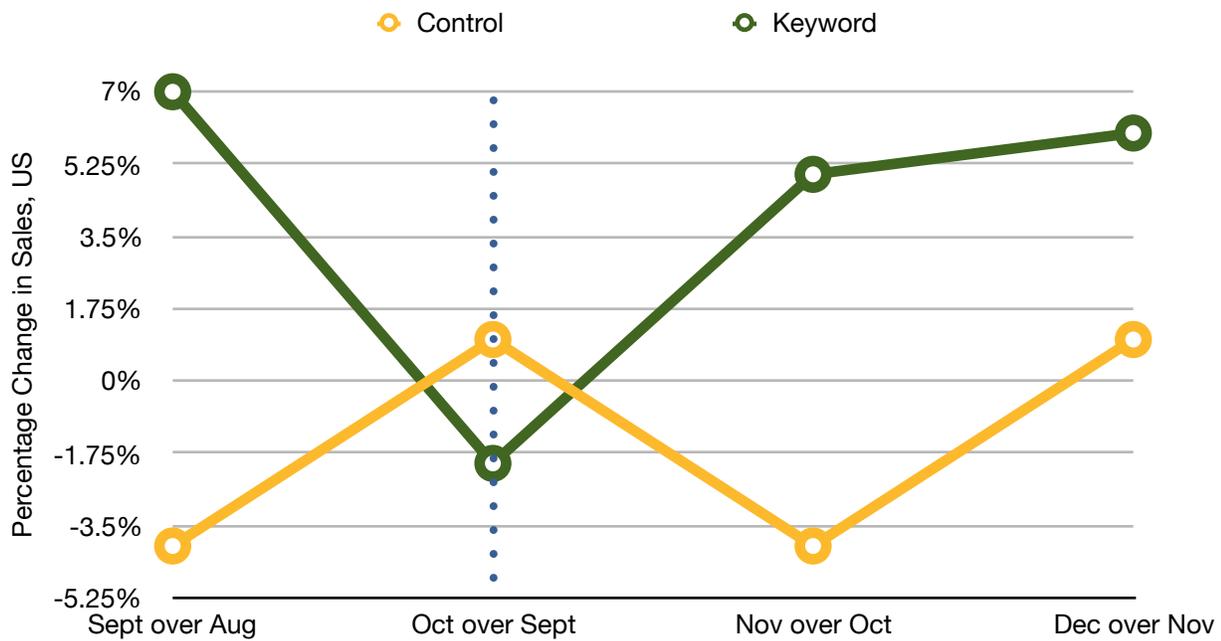
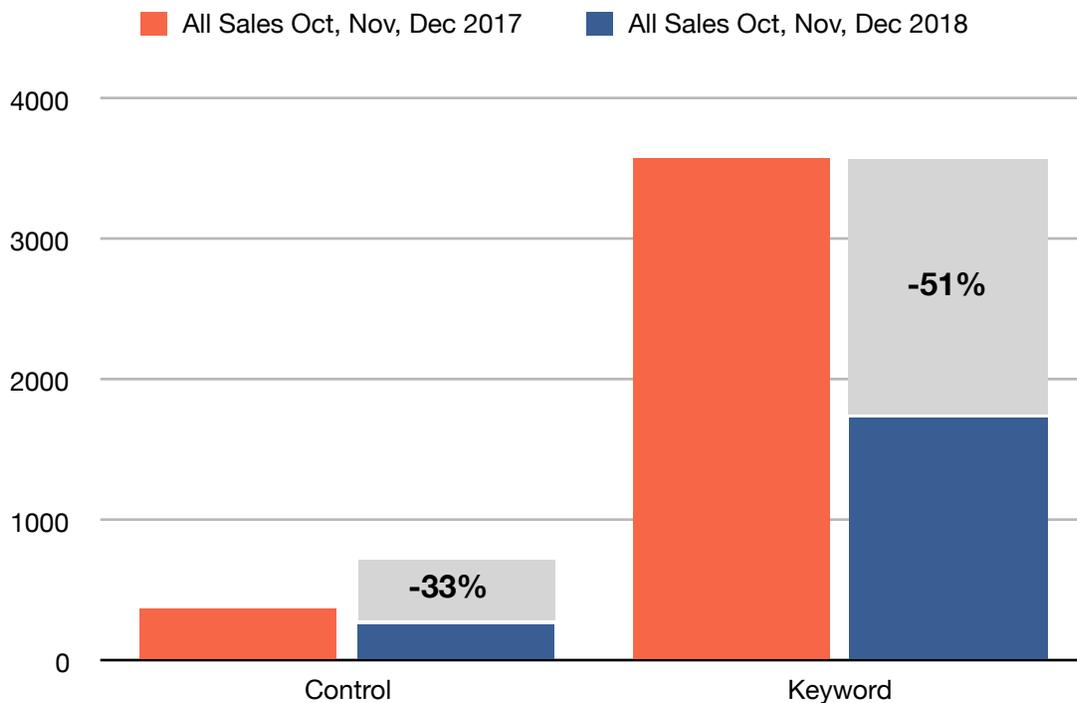


FIGURE 3: Total number of ebooks sold (US and CA) in each category for the study period vs the previous year.



We expected to see an increase in sales of titles with keywords assigned starting at the beginning of the study period that lasted into the holiday sales season. Unfortunately, the Canadian sales show a *decrease*, which was followed by an increase that mirrored the increase in sales of the control titles. The US sales show an even more marked decrease. The conclusion we came to is that the keywords had no effect on sales month-over-month.

The same conclusion can be reached when looking at total sales of the same period from 2017 (**Fig. 3**). It is important to note that while the keyworded titles had more sales than the control group due to the limitations discussed on page 8, the change in sales was not what we expected to see. The chart shows a decrease in sales in both the keyword group and the control group.

These sales results are consistent with current industry trends: Amazon sales of traditionally-published ebooks are declining across the board.³ Although we expected to see an improvement when we added keywords, after analyzing the data from multiple approaches we continued to see similar sales results between the two groups.

Human-Created Keywords vs. Machine-Created Keywords

The figures in this section look at the change in sales of the ebooks which had keywords created by humans against the ebooks which had keywords created by Firebrand’s machine-learning software.

The following charts show average percentage change in Amazon sales month-over-month (**Fig. 4, Fig. 5**) and total sales during the study period over the last year (**Fig. 6**).

Fig. 4 and **Fig. 5** indicate that Firebrand’s keywords perform slightly better than the human-created keywords when first assigned and distributed, but that evens out as time goes on. However, when looking at the change in total sales over last year (**Fig. 6**), we see that there is very little difference in year over year—both groups display about a 50% decrease in sales.

It appears that regardless of who created the keywords and how we explored the data (please refer to Appendix B for further analysis), there was no real sales impact of keywords on the selected ebooks.

³ Milliot, Jim. “E-Book Sales Fell 10% in 2017.” *PublishersWeekly.com*, 25 Apr. 2018, www.publishersweekly.com/pw/by-topic/digital/content-and-e-books/article/76706-e-book-sales-fell-10-in-2017.html.

FIGURE 4: Percentage change in CA sales of ebooks with human-created keywords (**Human**) versus ebooks machine-created keywords (**Machine**), month over month. A **blue** dotted line is used to indicate the start of the study period.

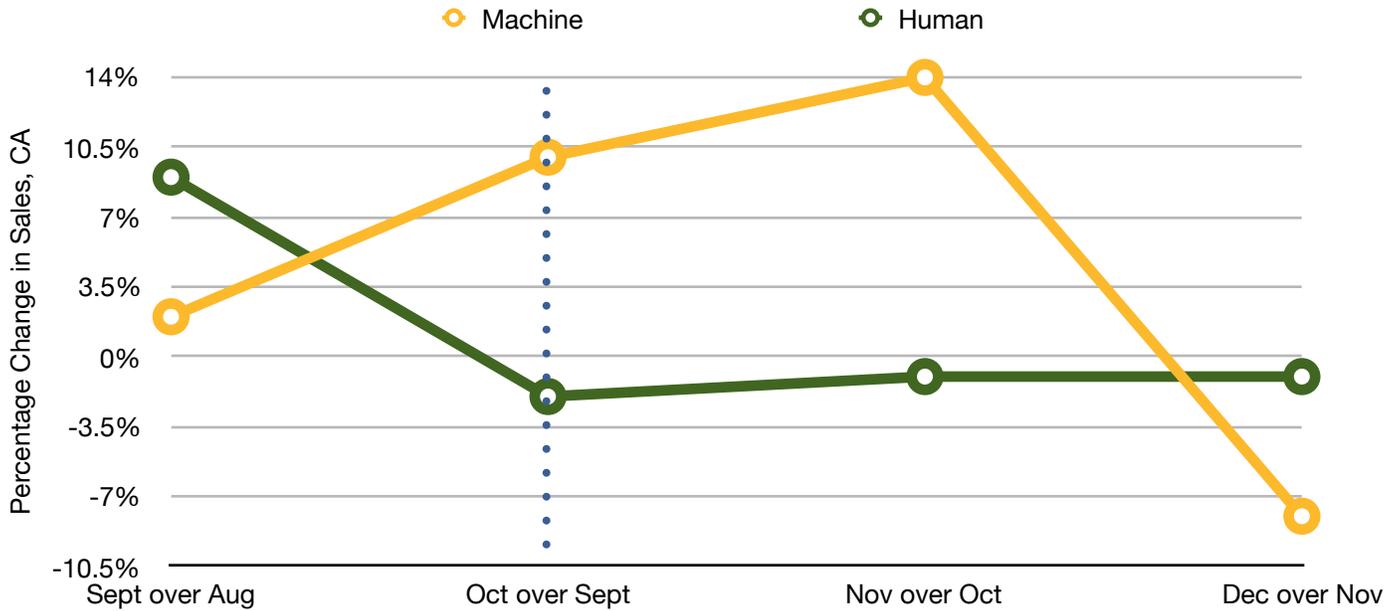


FIGURE 5: Average percentage change in US sales of ebooks with human-created keywords (**Human**) versus ebooks machine-created keywords (**Machine**), month over month. A **blue** dotted line is used to indicate the start of the study period.

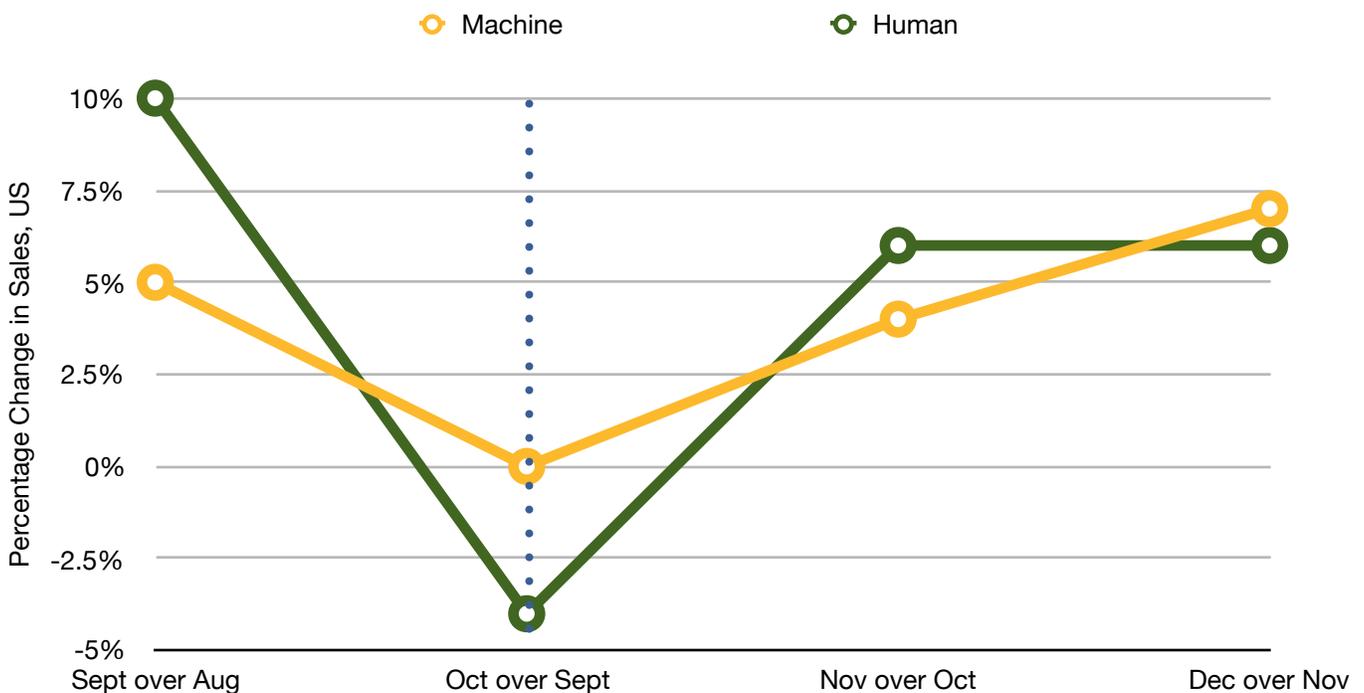
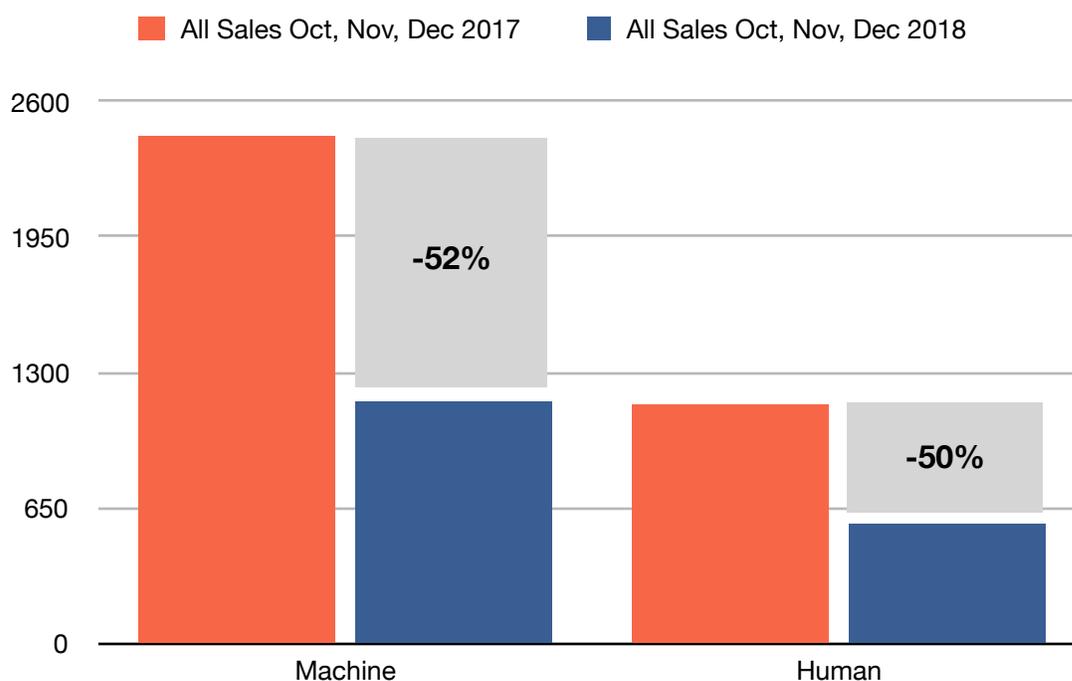


FIGURE 6: Total number of ebooks sold (US and CA) in each category for the study period vs the previous year.



Effect on Discoverability

Keyword-Assigned ebooks vs. Control ebooks

The following looks at change in Amazon glance views for the ebooks to which we assigned keywords (regardless of whether the keywords were created by machines or by humans) against the control group—those ebooks that did not have any keywords at all.

The most striking example of the increase in discoverability of keyworded titles is apparent when investigating the glance views of the study period (October, November and December 2018) over the same period in 2017. **Figure 7** shows a dramatic improvement in the glance views of those ebooks to which we assigned keywords over the ebooks in the control group.

The change in glance views over the previous three months (July, August and September 2018, **Fig. 8**) is also striking, displaying a similar marked improvement in glance views of keyworded ebooks over control ebooks.

FIGURE 7: Average percentage change in glance views between October, November and December 2017 and October, November and December 2018. The ebooks with keywords assigned (**Keyword**) show a 240% increase in glance views, and the control ebooks (**Control**) show a 48% increase.

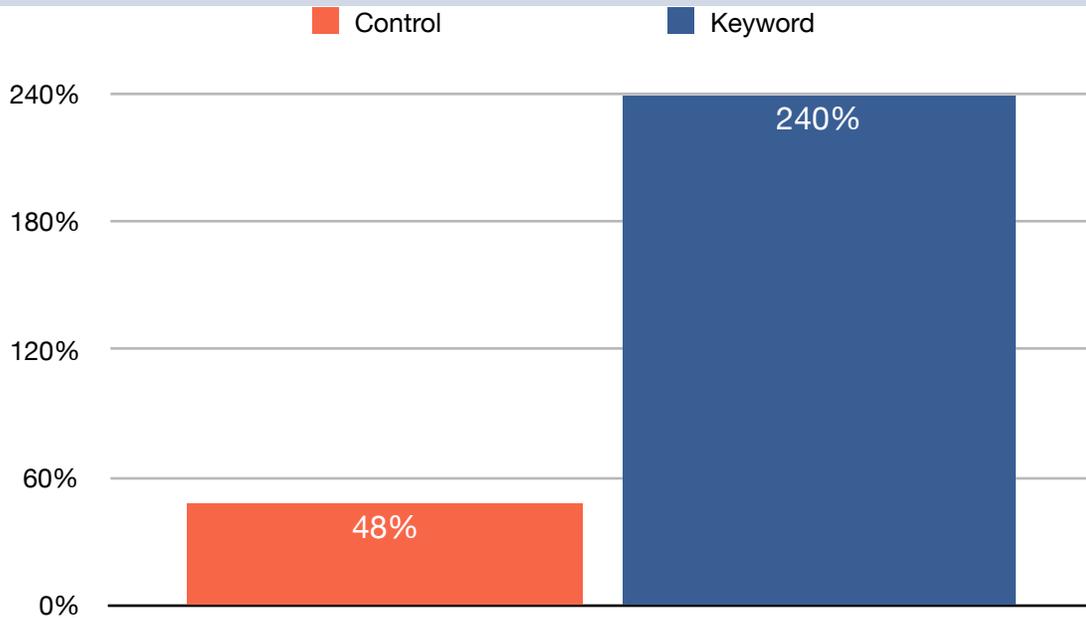
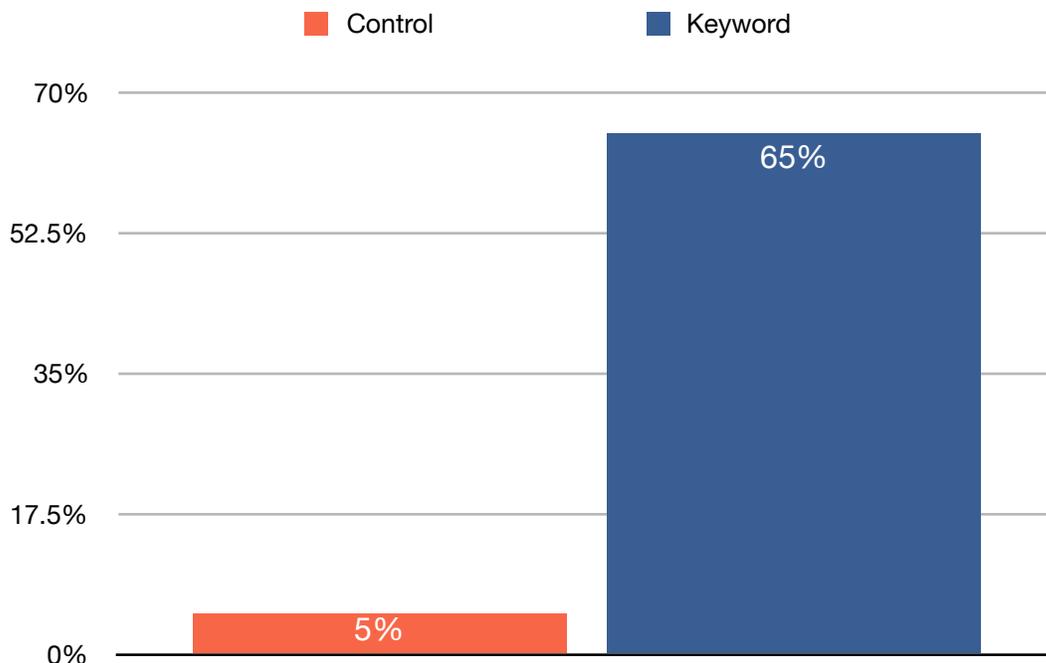


FIGURE 8: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018. The ebooks with keywords assigned (**Keyword**) show a 65% increase in glance views, and the control ebooks (**Control**) show a 5% increase.



Improvements in discoverability can be seen across all BISAC categories (**Fig. 9**) selected for the study, although the most dramatic increase is in the youth categories. And the same improvements are seen in backlist (publication years from 1989 to 2015) and frontlist (publication years in 2016 and 2017) titles (**Fig. 10**), although the keyworded frontlist titles saw the greatest improvement. For a more in-depth analysis of the effects of keywords on sales broken out across BISAC categories, publication years and ranking group, please refer to **Appendix B**.

FIGURE 9: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018, broken up by BISAC category. Ebooks with keywords assigned are in blue (**Keyword**) and the control ebooks (**Control**) are in orange.

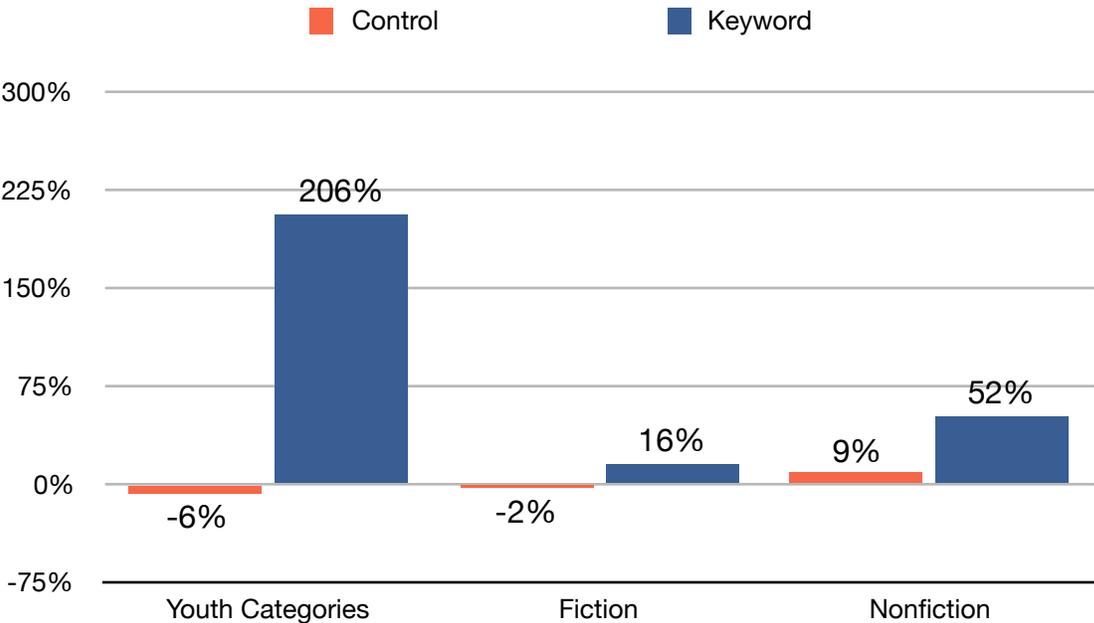
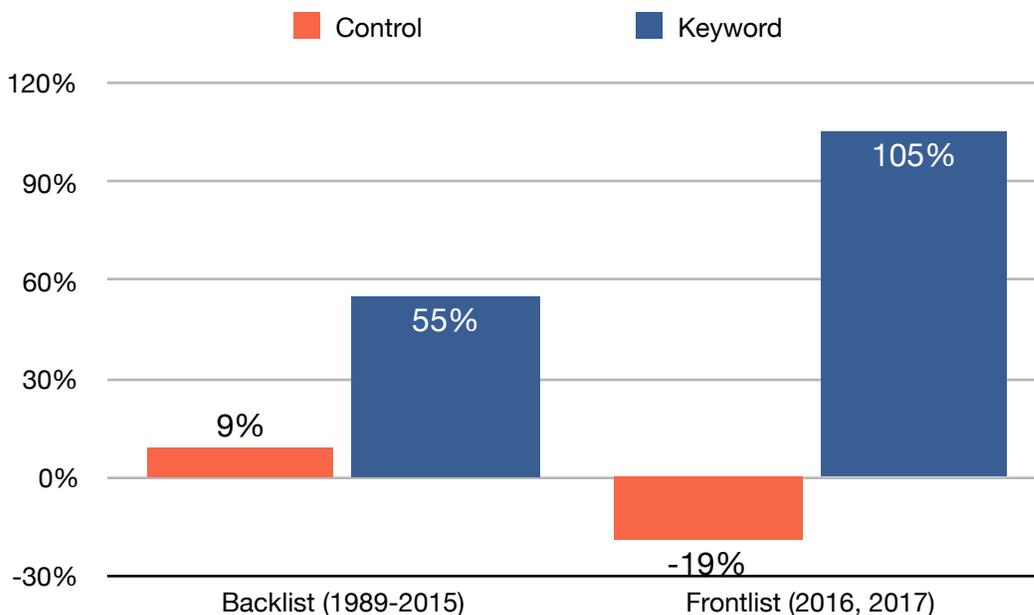


FIGURE 10: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018, broken up by publication year. Ebooks with keywords assigned are in blue (**Keyword**) and the control ebooks (**Control**) are in orange.



Human-Created Keywords vs. Machine-Created Keywords

The following looks at change in Amazon glance views for the ebooks with keywords created by Firebrand versus the change in glance views for ebooks with keywords created by humans.

While it is clear that keywords are effective regardless of whether they're human assigned or created by Firebrand's algorithms, the following data explores if one has a greater effect than the other.

Figure 11 shows glance views of the study period (October, November and December 2018) over the same period in 2017. When looking at the entire set of titles, the ebooks with human-created keywords performed significantly better than those with machine-created keywords. The same is true when looking at the change in glance views over the previous three months (July, August and September 2018, **Fig. 12**).

FIGURE 11: Average percentage change in glance views between October, November and December 2017 and October, November and December 2018. The ebooks with human-created keywords (**Human**) show a 381% increase in glance views, and the ebooks with machine-created keywords (**Machine**) show a 101% increase.

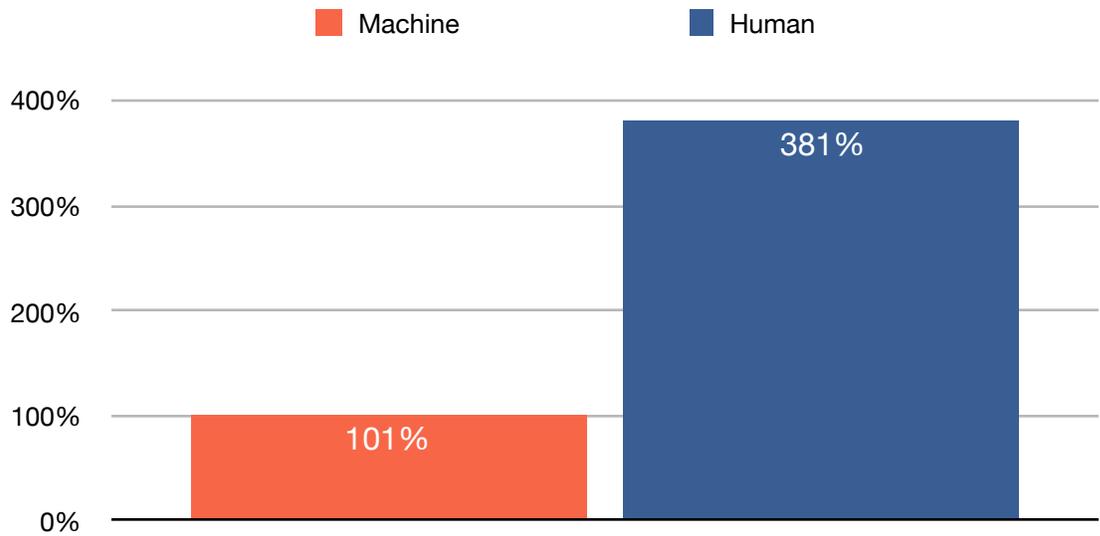
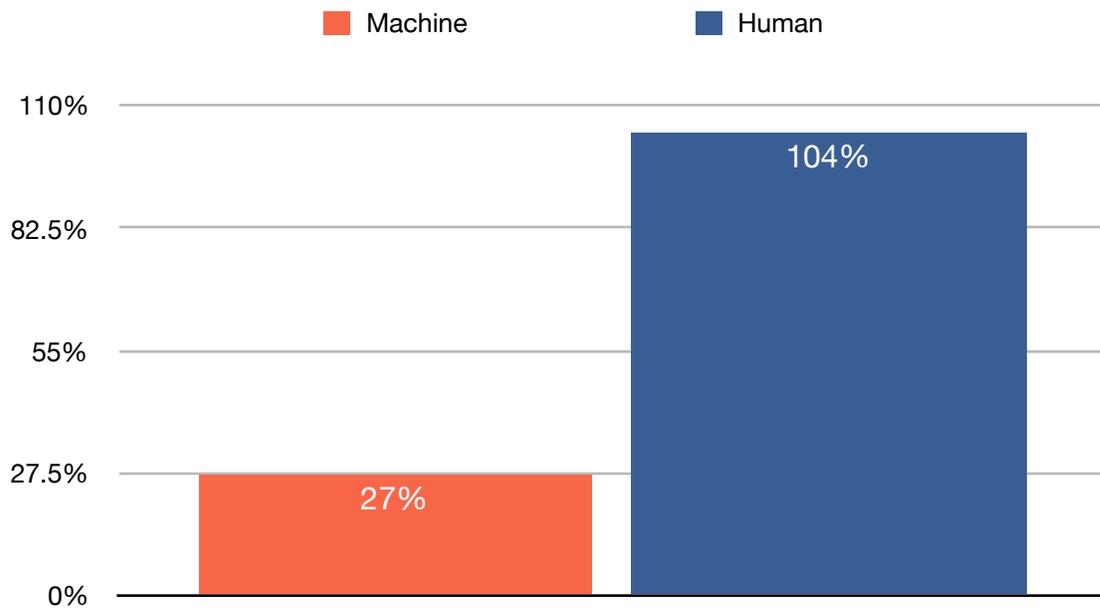


FIGURE 12: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018. The ebooks with human-created keywords (**Human**) show a 104% increase in glance views, and the ebooks with machine-created keywords (**Machine**) show a 27% increase.



When we break down the books by publication year (**Fig. 13**) or category (**Fig. 14**) and look at their glance views over the previous three months, we again see the human-created keywords performing better.

FIGURE 13: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018, broken up by publication year. Ebooks with human-created keywords are in blue (**Human**) and ebooks with machine-created keywords (**Machine**) are in orange.

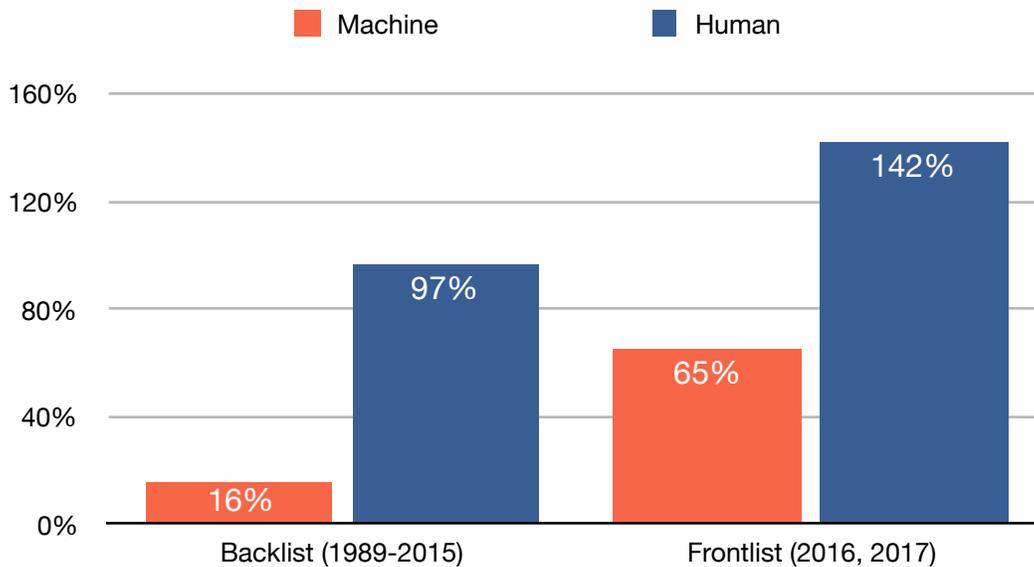
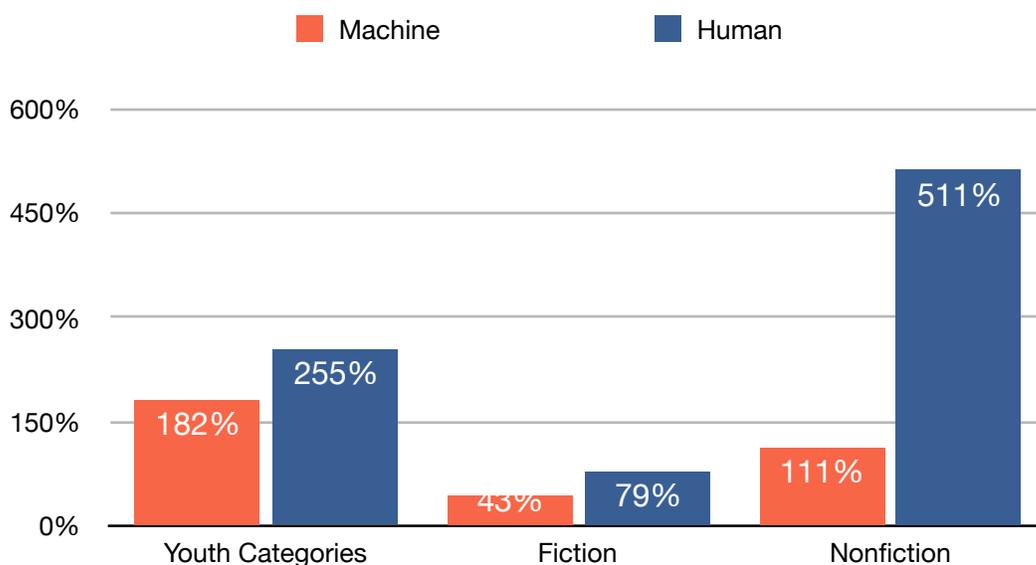


FIGURE 14: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018, broken up by BISAC category. Ebooks with human-created keywords are in blue (**Human**) and ebooks with machine-created keywords (**Machine**) are in orange.

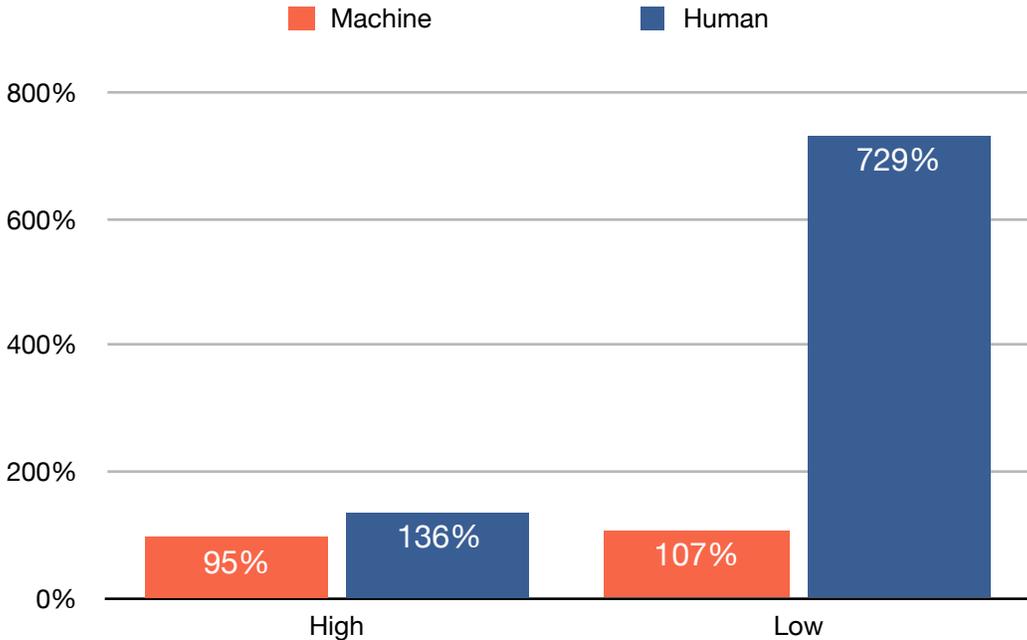


However, as mentioned on page 6, the ebooks were originally selected using Firebrand’s ranking system. Firebrand was assigned the highest-scoring books due to the constraints of their systems.

The ebooks were broken out based on their score. The first chart (**Fig. 15**) has them separated into a high-scoring group and a low-scoring group. The median score was 0.238, and the groups were arranged around that median.

Previous results show the human-created keywords outperforming the machine-created keywords significantly. But when the score group is taken into account, it’s clear that while human-created keywords are much more effective for the low-scoring ebooks, that difference is much less apparent for the high-scoring ebooks.

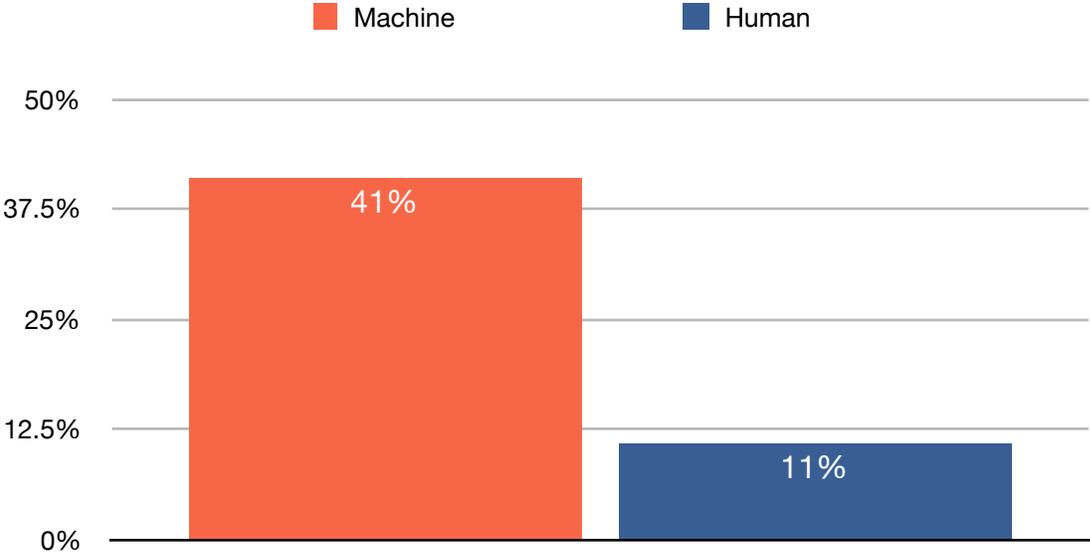
FIGURE 15: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018, broken up by score group (as determined by Firebrand). Ebooks with human-created keywords are in blue (Human) and ebooks with machine-created keywords (Machine) are in orange.



In **Figure 16**, we break out the twenty-five top-scoring groups into a very high score category, and here that effect is even more clear.

However, the majority of eBOUND’s client publishers have books that are not in that high scoring bracket. For most Canadian small-press ebooks, it appears that human-created keywords will result in a greater improvement in discoverability.

FIGURE 16: Average percentage change in glance views between July, August and September 2018 and October, November and December 2018 for the 25 top-scoring books. Ebooks with human-created keywords are in blue (**Human**) and ebooks with machine-created keywords (**Machine**) are in orange.



CONCLUSIONS

Based on the data collected and our analysis of that data, it appears that adding keywords to bibliographic metadata had no effect on the sales of ebooks published by independent Canadian presses. However, it did appear that adding keywords had a significant effect on the discovery of those ebooks, regardless of their category or when they were published. It also appears that for most ebooks, having keywords created by human readers is more effective for discovery overall.

Further Research

There is a lot of room for further research into the effect of keywords on the Canadian book market.

It may be useful to investigate the sales of all formats of a keyworded book, as mentioned on page 9. But for long-term success, it would be valuable to determine if there are other factors within the bibliographic metadata that can be used in tandem with keywords to increase conversion—that is, turn the extra glance views into higher sales. It is possible that with more robust, cleaner, and more up-to-date metadata, Canadian independent publishers would see increase ebook sales in line with the increased discoverability created by the addition of keywords.

RECOMMENDATIONS FOR PUBLISHERS

The goal of this research was to provide a guideline for Canadian independent publishers looking into adding keywords to their books. Based on the results, we cannot say that the addition of keywords will make a significant impact on sales, but we know that it will significantly improve discoverability.

Therefore, it is our recommendation that publishers add keywords to their bibliographic metadata where possible, but to do so in tandem with other updates to their metadata. We also recommend that keywords be created by human beings, as human-created keywords enhance discovery across the greatest number of books.

APPENDICES

Appendix A: Development of Human-Created Keywords

Current Best Practices

As mentioned in the main text, retailers have not yet settled on a standard use of keywords. However, it is strongly recommended that anyone planning to use keywords in their bibliographic metadata read the BISG’s “Revised Best Practices for Keywords in Metadata”, available for free on BISG’s website and referred to on page 4 of this document. It is revised often and is the best overview of the current understanding and usage of keywords.

Methods Used in This Study

In some places, the methods used in this study differed somewhat from the best practices outlined by the BISG. We’ve outlined the steps taken and resources used here and indicated if the methods used here differ from those in the BISG’s best-practices.

This study investigated the use of keywords, but it did not look at the quality of the keywords as a factor. The steps below are only a description of what the human contractors did to choose their keywords and are not intended as a guideline to create the best possible keywords.

Researching the Publisher-Created Metadata

The contractors were given a spreadsheet with their assigned books. They then went through each book individually and used the book’s metadata as revealed by BookNet’s Biblio-o-Matic tool as well as its Amazon page to get a sense of the book’s subject, content, tone and intended market. They created keywords based on this basic bibliographic metadata, including whether the book had Canadian content (“Canadian”, “Canada”) or was set in a specific place (“Quebec”, “Toronto”), or if it had an intended age range (“grade two”, “10”, “11”), or if the book was own-voices (“women”, “Japanese-Canadian”).

This is somewhat contrary to BISG’s best practices, as they recommend not creating keywords that can be found elsewhere in the metadata. In our circumstances, we wanted to ensure consumers were discovering the books by using more human-sounding keywords. For example, the book’s metadata may list “ages 10-14”, but a consumer is more likely to search for “books for grade 8”.

The contractors also used tone words (“funny”, “sad”) and other content gleaned from the publisher’s descriptive text.

Investigating Reader-Created Content

Once the contractors had exhausted the publisher-generated content, they moved on to reader-created content such as reviews or categories. These are an excellent way to determine how readers are talking about the book—is it a “good book for slow readers”, or perhaps a “scary mystery”? These are things that a reader will tell us in their reviews on Amazon or Goodreads but that might not be evident in the publisher’s metadata, and by using the reader’s language we’re more likely to capture the way potential readers are searching for books.

Using Comparative Titles

In some cases, the books did not have enough reader-created content to create good keywords. When this happened, the contractors took as much as they could from the metadata and from the book’s table of contents, and then they moved on to comparative titles. They looked at the reviews of those books to find the way readers spoke about similar titles.

Organizing the Keywords

The contractors were asked to stay under 600 characters. BISG’s recommendation is 500, and at the time of this study Amazon had not indicated their character limit. We based our limit on the information reported by Erica Leeman of MIT Press.⁴

It is important to note that as of November 2018, Amazon has changed their keyword character limit to 250 bytes.

The contractors organized their keywords so that the most valuable ones were at the beginning. While there’s no evidence that Amazon prioritizes keywords by location, it’s wise to keep the best at the top in case a different vendor has a lower character count—or, as mentioned above, in case Amazon lowers their limit.

⁴ Leeman, Erica. “Demystifying the Inner Workings of Amazon Keywords.” *LinkedIn SlideShare*, 24 Mar. 2017, www.slideshare.net/booknetcanada/demystifying-the-inner-workings-of-amazon-keywords.

Appendix B: Additional Analysis of the Effects of Keywords on Sales

There was no conclusive evidence that the addition of keywords to the bibliographic metadata of the ebooks studied had any effect on sales. In the body of this report we explored overall sales. In this appendix, we'll look at the sales of the control group against sales of the keyword books broken out across BISAC category (**Fig. 17**), by publication year (**Fig. 18**) and by ranking group (**Fig. 19**).

All charts display the change in sales represented by a percentage, to avoid the limitations discussed on page 8 and for easier visualization.

FIGURE 17: Percentage change in number of ebooks sold (US and CA) in each category for the study period (Oct., Nov., Dec., 2018) vs the same period in 2017, broken up by main BISAC category. Ebooks with keywords are in blue (Keywords) and ebooks in the control group (Control) are in orange.

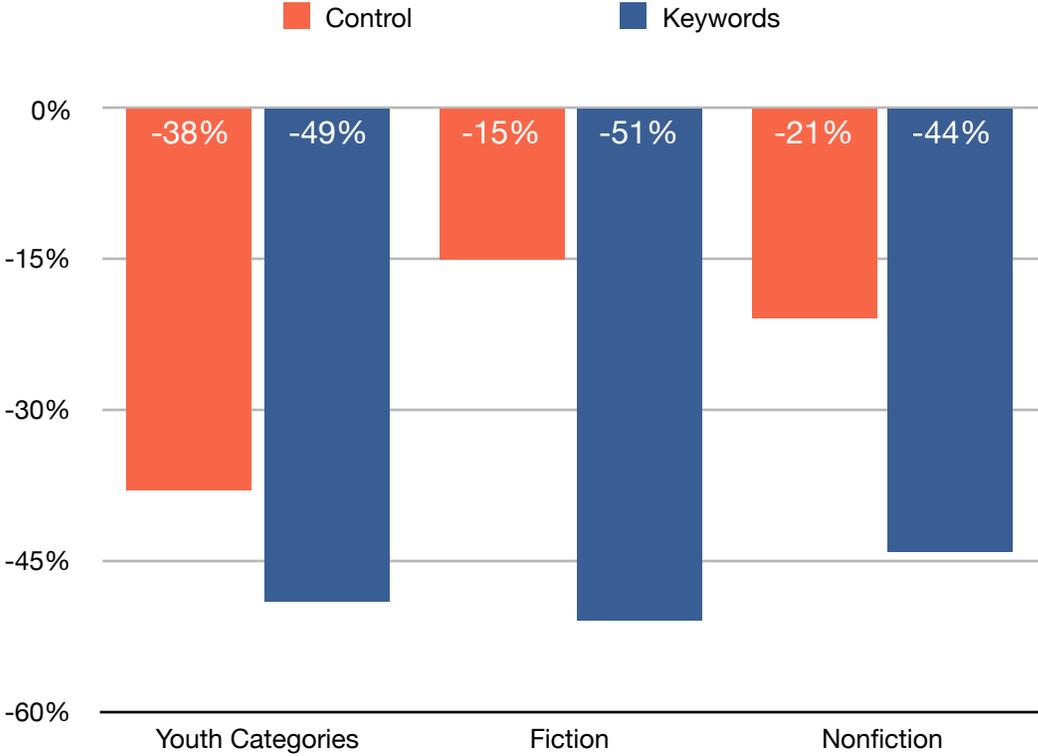


FIGURE 18: Percentage change in number of ebooks sold (US and CA) in each category for the study period (Oct., Nov., Dec., 2018) vs the same period in 2017, broken up by publication year. Ebooks with keywords are in blue (**Keywords**) and ebooks in the control group (**Control**) are in orange.

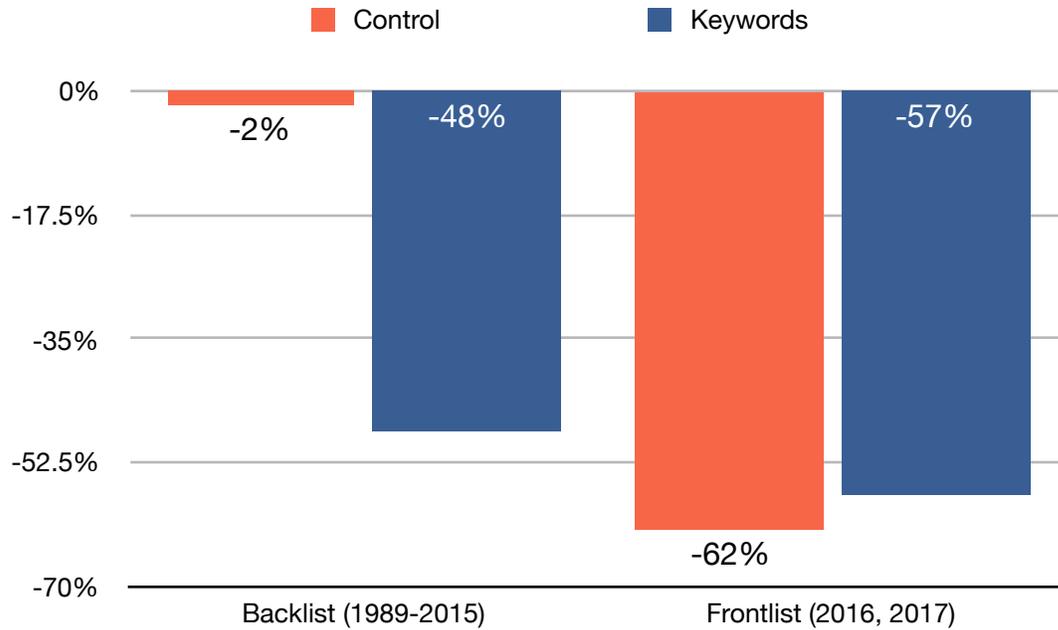


FIGURE 19: Percentage change in number of ebooks sold (US and CA) in each category for the study period (Oct., Nov., Dec., 2018) vs the same period in 2017, broken up by score group. Ebooks with keywords are in blue (**Keywords**) and ebooks in the control group (**Control**) are in orange.

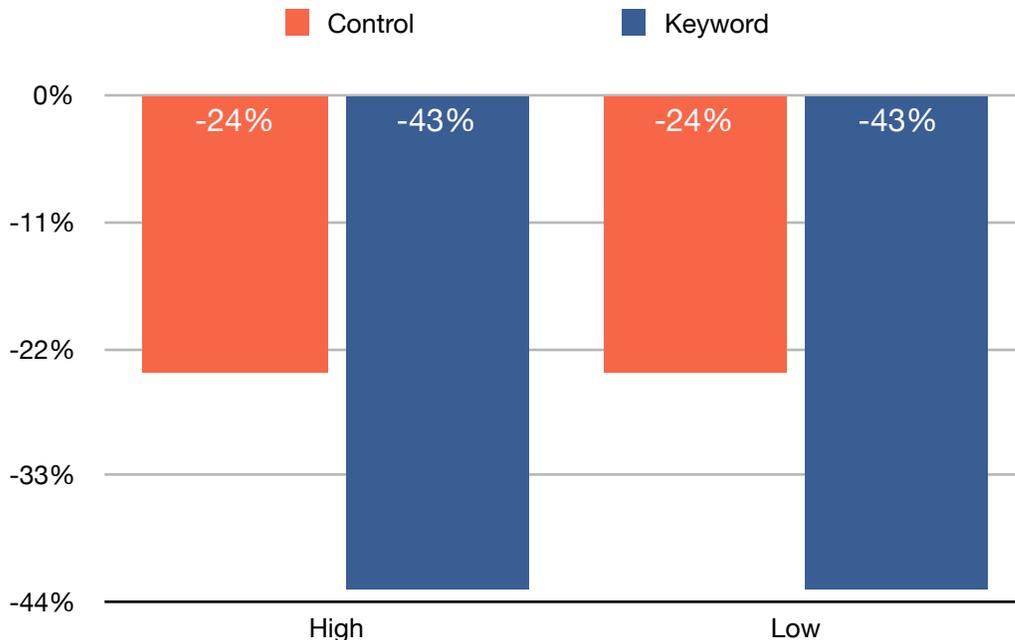


Figure 18 does indicate that keywords, overall, perform slightly better on frontlist titles, which tracks with the change in glance views as presented in figure 10. However, the difference is so small as to be a curiosity rather than a piece of significant data upon which to base a recommendation.

It's clear from this data that sales were down regardless of category, publication year, group or whether or not the ebook had been assigned keywords. In fact, it appears that the keyword books performed *worse* than the control books overall. However, when examined together with the significant increase in discoverability it's hard to suggest that the keywords were the cause of the decrease.