

Metadata is becoming more and more important for publishers in this digital age, because that's how customers discover their books. Here, Shivangi Ramachandran discusses the imperative role of metadata for publishers selling their books online. The publishing industry has been changing rapidly for the past few decades. From the advent of the superstores like Barnes and Noble, to now Amazon selling both ebooks and physical books online, internet and mobile technology have changed our business models. Easy access to both the retail of books as well as information about old, new and upcoming books, have allowed customers to shop for the books they want online. In 2012, 44% of all sales were on Amazon (Digital Book World, 2013). A source at a Big Five publisher in the United States confirmed that 60% of their sales now happen online. Therefore, this has changed how and where customers discover their books. No longer do book readers find the next book that they will read through a serendipitous browse through bookstore; instead they're shopping on Amazon, relying on search results, browsing reviews on online portals like Goodreads to find their next book. And how does all of that work? It works through invisible code, algorithms, and the publishing buzz word of our times – "metadata."

What is metadata?



Metadata can be defined as a “collection of attributes.” Metadata is the information about each book that is channeled appropriately to categorize it and place it in relevant spaces on the Internet. Metadata helps consumers find their books when they want it—whether that be through a direct search, or a search for related books and topics. The most common pieces of metadata are a book’s title, subtitle, and author. Other data like ISBN, format, and pub date are also part of its metadata.

Metadata: then and now

The concept of digital metadata began in the late 1970s. This is when “bar coding on books was introduced and electronic transactions between retailers and publishers began.” While scholarly publishing was quite detailed in its metadata even back then, trade publishing only listed the bare minimum – the ISBN, availability, and price. This was because transactions were still based on the physical book bought in a physical bookstore. Big chain retailers like Barnes & Noble and Borders, however, recognized the value in book metadata. They were the first to start using metadata efficiently as computer transactions and scanners allowed them to save on time and helped with logistics. It was an important part of the success of the superstore. A detailed inventory that listed the ISBN, author, title, price, and other core metadata allowed bookstore personnel to know where each book was shelved.

The use of metadata again changed and grew with the rise of Amazon. Computer software and interfaces like Microsoft Windows made it easier in the 1990s to display and use information about books and support innovation. Amazon took full advantage of this and launched in 1990,

using metadata to grow its eRetail. For the first time, metadata was not just used for librarian, retailer or wholesaler's reference; instead, it was available to the customer, and lo and behold, the customer wanted to know more about the book than just its price and ISBN. They were looking for descriptions of a book, cover images, excerpts, and reviews. They wanted to know the publication date, and sometimes, they wanted to order in advance of the publication date. As a response, publishers scrambled to supply their warehouse data to Amazon. The warehouse data was not meant to be viewed by consumers, and it came disorganized, with misspelled titles and author names in all caps. That is, until Amazon hired its own staff of editors to reorganize the metadata and make it seamless. Libraries soon followed this example, and a host of new companies arose from this need for metadata. The value of metadata can be seen via the longevity of the firms that arose to help publishers and libraries with their metadata. Companies like Syndetics (founded in 1998), Muze/Rovi (founded in 2009, and now a part of TiVo Corporation) and Firebrand's Eloquence (founded in 1987) are all still in business.

Importance of metadata

Metadata is not only important for book publishers, but for all industries that are trying to sell things online. Therefore, it is imperative and interesting to look at how different industries are tackling the problem of eRetail and Discovery. A good juxtaposition to books is music – and how the music industry coaxes listeners to discover new music that they will actually like. In that respect, Pandora is a great example of metadata harnessed well, along with human intelligence.

The core of Pandora's metadata is the Music Genome project, which uses up to 450 characteristics to describe a piece of music. Jeffrey Pomerantz, in his book, *Metadata*, covers how Pandora uses metadata: "These characteristics [...] run the gamut of relatively simple (for example, key, tempo, beats per minute, gender of the vocalist) to the highly subjective (for example, vocal characteristics, degree of distortion of instruments). Pandora employs a team of musicians whose job it is to listen to every song Pandora licenses, and to describe each song according to as many of these hundreds of characteristics as are relevant." Pandora exists as a platform to help users discover new music, and therefore it allows users to find music similar to what they like based on the metadata of the music they are already listening to. In a similar way, metadata for books can help generate recommendations and minimize the number of books for a consumer to dig through. The key thing here to note is that both objective and subjective pieces of metadata used together made for perfect recommendations.

Similarly, metadata has come to be an irreplaceable part of the marketing strategy for publishers. Metadata will not function alone to shoot up the sales of a book. A source at an independent publisher in New York City agrees that "while metadata doesn't necessarily create the buzz needed to sell a book, it does allow the people who could create that buzz to find it easily." However, even with that understanding of the importance of metadata as a tool, it is still something most publishers are struggling to harness. Simon & Schuster became the first company to hire a team of people to look solely at their metadata and create tools to make it more seamless. We are now coming to understand how different things like BISAC codes and competitive titles allow a book to be discovered and sold. An independent publisher in New York said, "Because of BISAC codes, I'm going to know where to show [the books]. And if your

competitive titles in your metadata aren't accurate and aren't representative of the title in question, bookstores won't be able to contextualize the work. If you don't make your manuscripts available early through Edelweiss or Netgalley, the likelihood of the bookseller or librarian being able to find your book is low. It's always about doing all you can to make your book as accessible as possible to tastemakers and influencers who have a tangible impact on the life of book. It's an essential part of the business that most book publishers take it very seriously. It's one of those things in the publishing business that is very tedious, has much to do with the work itself, but something that must be done."

The science of metadata

Metadata is not a dry science. It's something everyone collaborates on it— from the editor, the marketing team, the publicity team, the sales team, and sometimes even reviewers. It's really a collaboration of people who know their content, think strategically, and are using new technology to help their product. In his book about metadata in music, Jeffrey Pomerantz wrote, "The sound of the music is something automation can't determine by itself, so a human team had to painstakingly map these attributes to the bands and maintain the database. This is what makes the system work: a human filter applying a custom taxonomy. In other words, the human classification team allows the technology to work." For book publishing as well, it is that combination of human intelligence that will allow us to harness the tool that is metadata to allow it to work properly and to the best of its ability.



*Shivangi Ramachandran has recently completed her Master's in Publishing from Emerson College. This is an excerpt from her thesis, *Book Discoverability in the Age of eRetail*, which won the Best Thesis Prize in the graduate program. She has worked for companies like Pearson Education in New Jersey, Oxford University Press in India, Harvard Education Publishing Group in Boston. She is currently based in Boston.*