Machine-Readable or Marked-Up?: A Comparison of MARC and HTML

Amelia Burlingame

IS_LT 7302: Organization of Information

March 4, 2023

In this essay I will compare the similarities and differences between encoding formats Machine-Readable Cataloging (MARC) and HyperText Markup Language (HTML). Both encoding formats provide a framework for computers to read and decipher multiple types of information, which can then be displayed for users online. Although MARC was developed much earlier than HTML, and is specialized for libraries, both have evolved since their creation and continue to be relevant despite the rapid changes that the internet and widespread computer use have brought to the modern world. Initially created in the 1980s, HTML has continued to be the backbone of web page structure. According to Joudrey & Taylor (2018), HTML5 was released in 2014 to allow more seamless integration of multimedia on web pages. Included as a specification of HTML5 is *microdata*, which "can be used to nest metadata," previously provided limited support, "...within the content of web resources" (Joudrey & Taylor, 2018, p. 267). MARC format, on the other hand, was developed by the Library of Congress in the 1960s "as a way to share bibliographic data among libraries" (Joudrey & Taylor, 2018, p. 253). It has continued to evolve as the industry standard, and in the year 2000, MARC 21 was developed as a way to standardize library metadata in the US and Canada, and has since been adopted by several other countries (Joudrey & Taylor, 2018). As the internet has increased global interconnectivity, this type of international standardization can support interoperability among libraries.

The similarities between MARC and HTML fall far short of their differences, however. While both are computer languages, HTML is a basic language for web page structure, whereas MARC is specific to library catalog creation. MARC has become the standard for libraries particularly because it offers unlimited fields and field lengths for the librarian to utilize when creating a catalog entry (Library of Congress, 2009). MARC's usefulness is up for debate in more recent years as its integration with HTML and other, newer, encoding languages has proven to be limited (Joudrey & Taylor, 2018). The tags used by each for computer readability are another area of notable difference. MARC "[t]ags are the three-digit numbers (from 001 to 999) that designate the kind of content that will be entered into the field" (Joudrey & Taylor, 2018, p. 256). These tags designate certain criteria of the bibliographic metadata, and tags starting with the same number are categorized as a group. These tags would hardly be known or understood outside of library metadata creation, however. Tags used in HTML focus on how the information is structured on a webpage, and are widely used to create a variety of web content. Tags can indicate headers, page breaks, lists, or how text is aligned on the page. They can also be used to insert links, images, video, or other media and interactive tools onto a webpage, or change the typeface displayed. While any novice to both MARC and HTML may see both as an unfamiliar and mysterious language, they are used for purposes as unique as any web page or bibliographic data.

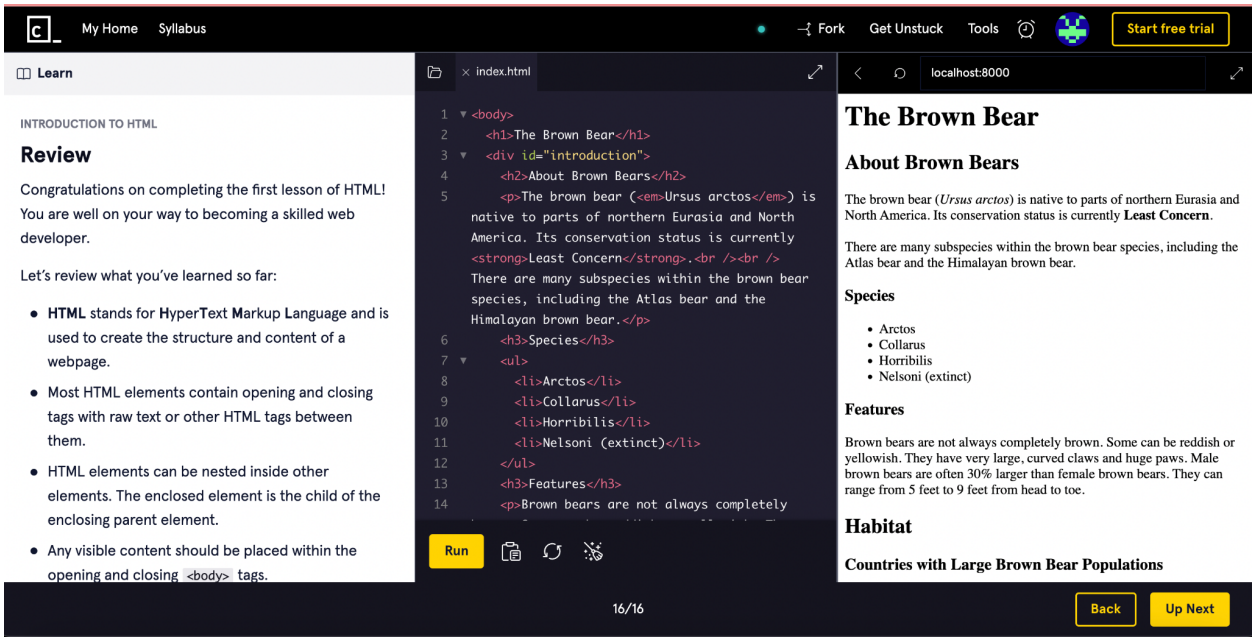
References

Joudrey, D. N, & Taylor, A. G. (with Wisser, K. M.). (2018). *The organization of information* (4th ed.). Santa Barbara, CA: Libraries Unlimited.

Library of Congress. (2009). *Understanding MARC Bibliographic: Parts 1 to 6*. Loc.gov. https://www.loc.gov/marc/umb/um01to06.html

Appendix

Figure 1



Note: The final page of my lesson in CodeAcademy's Learn HTML *course.*