

# Chapter 20

## Metadata Standards in Digital Audio

**Kimmy Szeto**

*Baruch College, City University of New York, USA*

### ABSTRACT

*Audio metadata is an essential tool that supports control and management of systems that create, transmit, describe, manage, and store digital audio. Technical and structural metadata enable audio in devices and software applications; descriptive metadata provide context. Metadata also relate and incorporate audio in a multimedia environment resulting in profound effects on conception, reception, and consumption. This chapter explores metadata standards in digital audio.*

### INTRODUCTION

Audio metadata are an essential tool that supports control and management of systems that create, transmit, describe, manage, and store digital audio. Throughout the lifecycle of digital audio objects—pre-production, acquisition and production, post-production, distribution, storage, transmission, and archiving—metadata “describe the attributes of a resource, characterize resource relationships, and support resource discovery, management, and effective use” (Vellucci, 1999). Technical and structural metadata enable audio in devices and software applications; descriptive metadata provide context. Metadata also relate and incorporate audio in a multimedia environment resulting in profound effects on conception, reception, and consumption.

Issues of audio metadata are intertwined with audio equipment, music production, and information technology. Holmes (2006) gives a broad overview of the technology; Gilmer (2004) provides the background for usage in multimedia production; Lu & Hanjalic (2009) and Kriechbaum (2009) detail the technical aspects in database systems; Casey & Gordon (2007) offer the information professional perspective. International audio metadata standards are chiefly issued by the European Broadcasting Union (EBU), the Society of Motion Picture and Television Engineers (SMPTE), and the Audio Engineering Society (AES). Guidelines for usage and implementation are routinely published by these three organizations, as well as by manufacturers, libraries, and archives.

DOI: 10.4018/978-1-5225-7601-3.ch020

## **BACKGROUND**

Analog sound recording technologies had been in use for over a century when the Compact Disc Digital Audio (CD-DA) technology emerged in the late 1970s following breakthroughs in signal processing and the optical medium. Audio continued to be produced in a mixture of analog and digital systems consisting of single-purpose devices designed for recording and producing audio to be transmitted in either analog or digital form. The production system was largely determined by the transmission medium—wire, wax cylinder, shellac disc, vinyl disc, magnetic tape, broadcast, optical disc—each requiring an assortment of specialty equipment, typically supplied by a single manufacturer. In such vertically integrated systems, controls and switches that supplied parameters and characteristics (metadata) to the audio were generally integrated in the hardware. Where metadata were specified and shared, the necessity for interoperability was limited to the devices immediately before and after in the production chain.

As digital signal processing technologies evolved, the proliferation of recording formats, compression methods, and media formats made it no longer sensible to build a separate integrated system for each device combination. In such a mixed production environment, devices need to be able to pass along and recognize standardized metadata throughout the production process. For several decades since the 1970s, the broadcast industry attempted to standardize metadata within device categories to allow for interchangeable equipment for each step of the production. But demands for higher flexibility and broader scope mounted as increasing prevalence of computer software and data networks shifted production workflow and transmission toward a distributed model. The industry was compelled to re-conceive use of metadata in not only audio, but also video production in a digital environment where data and content would be stored and transmitted in the form of computer files. These files, ideally, would encapsulate all necessary metadata, interoperable through the production chain, regardless of media format or equipment.

EBU and SMPTE established the Task Force for Harmonized Standards for the Exchange of Program Material as Bit Streams in 1995 to study the long-term interoperability and stability of production systems. In its 1997 report, the Task Force elevated metadata to be equal in significance to the signals, casting the formulas “Content = Essence + Metadata” and “Asset = Content + Rights.” (Essence includes video, audio, and data; metadata represents information about the essence, including rights.) And, among other technical details, the Task Force, anticipating the need to store and transmit digital multimedia as computer files, specified the generalized file format, or “wrapper.” This format consists of three parts—the preamble, consisting solely of metadata, the body, containing the essence and additional metadata, and the end-of-file marker. The wrapper provided the means to link metadata and essence logically and physically. This structure allowed audio files to contain an ever-increasing variety of metadata. SMPTE published the Metadata Element Dictionary Structure (SMPTE 335:2012) and established the SMPTE Registration Authority (2010), a public metadata registry, to serve as a basis and the administrative body for the development of future metadata standards. Open standards will guide the development of audio and multimedia metadata standards to handle increasingly complex and horizontal system design. Global adoption of well-defined international standards will facilitate data exchange and re-use, as well as guarantee long-term preservation and combat format obsolescence.

Efforts to standardize audio metadata also dramatically shaped the consumer experience and conception of audio. With analog audio, consumers’ ability to produce and manipulate audio was limited to using audiocassette systems. As the Sony Corporation introduced the Walkman, a portable cassette player, in 1979, listening became personalized. Later, personal portable players were extended to the Compact Disc medium, even though recording devices did not reach consumers until much later. With

19 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/metadata-standards-in-digital-audio/214557](http://www.igi-global.com/chapter/metadata-standards-in-digital-audio/214557)

## Related Content

---

### Social Networks and Cultural Differences: Adidas's Case on Twitter and Sina Weibo

José Duarte Santos and Steffen Mayer (2021). *Analyzing Global Social Media Consumption* (pp. 121-137).

[www.irma-international.org/chapter/social-networks-and-cultural-differences/264935](http://www.irma-international.org/chapter/social-networks-and-cultural-differences/264935)

### Asia-Pacific Students' Awareness and Behaviour Regarding Social Networking in the Education Sector

Tomayess Issa, Sulaiman Alqahtani, Theodora Issa, Noorminshah A. Iahad, Peldon Peldon, Sooyoung Kim, Samant Saurabh, Sumaiya Pervaiz and Sun Joo Yoo (2023). *Research Anthology on Applying Social Networking Strategies to Classrooms and Libraries* (pp. 1652-1680).

[www.irma-international.org/chapter/asia-pacific-students-awareness-and-behaviour-regarding-social-networking-in-the-education-sector/313003](http://www.irma-international.org/chapter/asia-pacific-students-awareness-and-behaviour-regarding-social-networking-in-the-education-sector/313003)

### Computational Trust in SocialWeb: Concepts, Elements, and Implications

Kiyana Zolfaghar and Abdollah Aghaie (2010). *International Journal of Virtual Communities and Social Networking* (pp. 60-74).

[www.irma-international.org/article/computational-trust-socialweb/45779](http://www.irma-international.org/article/computational-trust-socialweb/45779)

### How Consumption Provides a Relational Understanding of Place

Mitrajit Biswas (2023). *Social Media and Online Consumer Decision Making in the Fashion Industry* (pp. 108-113).

[www.irma-international.org/chapter/how-consumption-provides-a-relational-understanding-of-place/327689](http://www.irma-international.org/chapter/how-consumption-provides-a-relational-understanding-of-place/327689)

### Expanding Opportunities in a Shrinking World: A Conceptual Model Explicating the Role of Social Networks and Internet-Based Virtual Environments in Social Entrepreneurship

Avimanyu Datta and Len Jessup (2009). *International Journal of Virtual Communities and Social Networking* (pp. 33-49).

[www.irma-international.org/article/expanding-opportunities-shrinking-world/37562](http://www.irma-international.org/article/expanding-opportunities-shrinking-world/37562)