Metadata Standards in Digital Audio

M

Kimmy Szeto
Baruch College, City University of New York, USA

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.

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

DOI: 10.4018/978-1-5225-2255-3.ch560

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.

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 the advent of storing and transmitting audio as computer files, however, audio consumption was no long tied to specific media or carriers. Personal computers allowed manipulation of audio and the associated metadata, and distribution of audio files over data networks. Widespread integration between personal computers, data networks, and personal portable players spurred consumer demands for describing and organizing digital audio. Descriptive metadata standards evolved and expanded in scope alongside techniques to generate audio semantics and analyze consumer preferences and behaviors, resulting in the ability to create personalized and customizable audio collections.

EMBEDDED METADATA IN DIGITAL AUDIO FILE FORMATS

Pulse-Code Modulation

Pulse-code modulation (PCM) is a method of representing analog signals in digital binaries. It was first developed in 1937 by British engineer Alec H. Reeves for audio transmissions over telegraph. Reeves was granted patents in France in 1938, in Britain in 1939, and in the United States in 1942, but lacked the equipment to produce the digital audio efficiently (Deloraine & Reeves, 1965; Robertson, 2005). The digital coding of analog signals is accomplished in three steps. Sound signals are

15 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/184341

Related Content

Food Security Policy Analysis Using System Dynamics: The Case of Uganda
Isdore Paterson Guma, Agnes Semwanga Rwashanaand Benedict Oyo (2018). *International Journal of Information Technologies and Systems Approach (pp. 72-90).*
www.irma-international.org/article/food-security-policy-analysis-using-system-dynamics/193593

An Evidence-Based Health Information System Theory
Daniel Carbone (2009). *Handbook of Research on Contemporary Theoretical Models in Information Systems (pp. 95-111).*
www.irma-international.org/chapter/evidence-based-health-information-system/35826

ICT Standardization
Kai Jakobs (2018). *Encyclopedia of Information Science and Technology, Fourth Edition (pp. 4679-4691).*
www.irma-international.org/chapter/ict-standardization/184174

Management Model for University-Industry Linkage Based on the Cybernetic Paradigm: Case of a Mexican University
Yamilet Nayeli Reyes Moralesand Javier Suárez-Rocha (2022). *International Journal of Information Technologies and Systems Approach (pp. 1-18).*
www.irma-international.org/article/management-model-for-university-industry-linkage-based-on-the-cybernetic-paradigm/304812

The Analysis of Instrument Automatic Monitoring and Control Systems Under Artificial Intelligence
Qinmei Wang (2024). *International Journal of Information Technologies and Systems Approach (pp. 1-13).*
www.irma-international.org/article/the-analysis-of-instrument-automatic-monitoring-and-control-systems-under-artificial-intelligence/336844