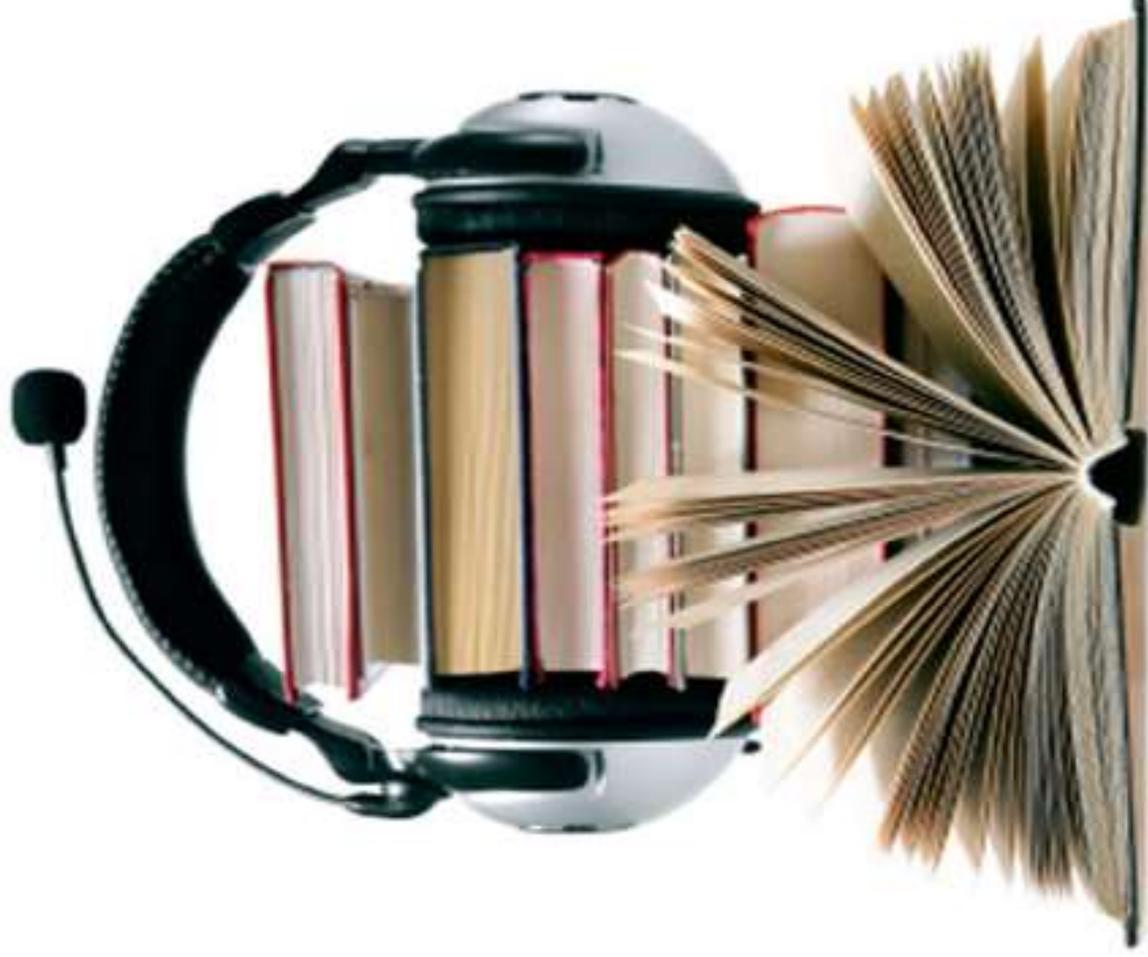


Know All About

AudioBooks



Barney Everhart

First Edition, 2011

ISBN 978-93-81157-04-6

WWT

© All rights reserved.

Published by:

The English Press

4735/22 Prakashdeep Bldg,

Ansari Road, Darya Ganj,

Delhi - 110002

Email: info@wtbooks.com

WORLD TECHNOLOGIES

Table of Contents

Chapter 1- Introduction to Audiobook

Chapter 2 - Windows Media Audio

Chapter 3 - Advanced Audio Coding

Chapter 4 - DAISY Digital Talking Book

Chapter 5 - Audio Codecs

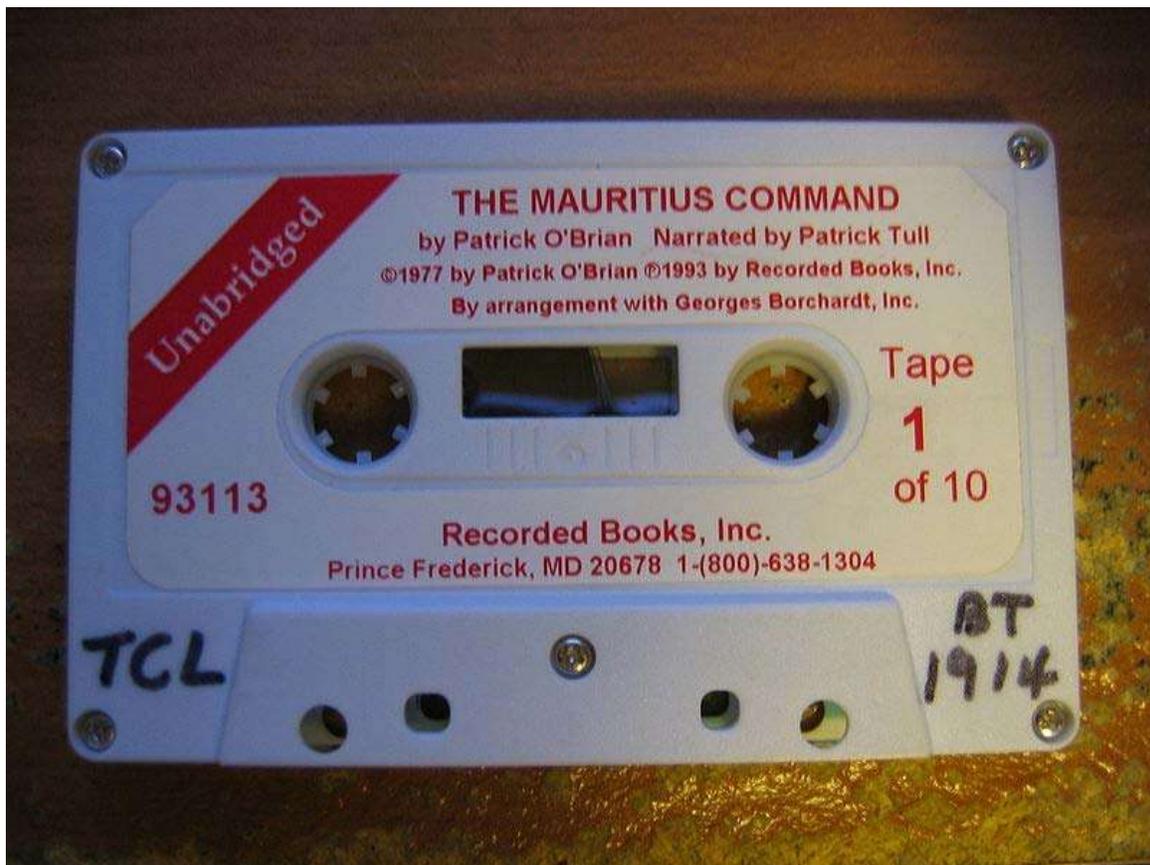
Chapter 6 - Digital Rights Management

Chapter 7 - Fair use

WWT

Chapter- 1

Introduction to Audiobook



An audio cassette recording

An **audiobook** is a recording that is primarily spoken word. It is often based on a recording of commercial printed material. It is not necessarily an exact audio version of a book.

Spoken audio was available in school and public libraries and to a lesser extent in music shops. It was not until the 1980s that there began a concerted effort to attract book retailers. As publishers entered the field of spoken-word publishing, the transition to book

retailers carrying audiobooks became commonplace on bookshelves rather than in separate displays.

Formats

Audiobooks are distributed on CDs, cassette tapes, downloadable digital formats (e.g., MP3 (.mp3), Windows Media Audio (.wma), and Advanced Audio Coding (.aac)). A benefit to the CD format is that CDs can be ripped and then put onto an MP3 player; a 14 CD "book" can be put onto a tiny & cheap (\$10USD) MP3 player.

In 2005 cassette-tape sales were 16% of the audiobook market, with CD sales accounting for 74% of the market and downloadable audio books accounting for approximately 9%. In the United States, a sales survey (performed by the Audio Publishers' Association in the summer of 2006 for the year 2005) estimated the industry to be worth 871 million US dollars. Current industry estimates are around two billion US dollars at retail value per year. In recent years, the Internet has introduced another powerful means of delivery for audiobooks and many titles are now available on-line, as downloads and as audio streams.

Sometimes audiobook format is available simultaneously with book publication. There are 50,000 titles on cassette, CD or digital format.

Unabridged audiobooks are word for word readings of a book, while abridged audiobooks have text removed by the abridger. Abridgements may be wanted to reduce the cost or for other reasons, but are also criticized for being incomplete versions of the original work. The increasing use of digital formats for audiobook recording has led to much less abridgment of texts, the most common reason cited for abridgment was the cost of production and with digital recordings being stored as data rather than on discs or tapes, the differential in cost of producing the full text was minimised. Audiobooks may come as fully dramatized versions of the printed book, sometimes calling upon a complete cast, music, and sound effects, though many consumers have indicated a preference for less music and fewer multiple voices and sound effects. Each spring, the Audie Awards are given to the top nominees for performance and production in several genre categories. Relatedly, a dramatized audio adaptation of a book is one form of an audio drama.

Occasionally there are radio programs serializing books, sometimes read by the author or sometimes by an actor, with most of them on the BBC.

History

United States

In 1931, Congress established the talking-book program, which was intended to help blind adults who couldn't read print. This program was called "Books for the Adult Blind

Project." The American Foundation for the Blind developed the first talking books in 1932. One year later the first reproduction machine began the process of mass publishing. In 1933 anthropologist J.P. Harrington drove the length of North America to record oral histories of Native American tribes on aluminum discs using a car battery-powered turntable. Audiobooks preserve the oral tradition of storytelling that J.P. Harrington pursued many years ago. By 1935, after Congress approved free mailings of audio books to blind citizens, the Books for the Adult Blind Project was in full operation. In 1992 the National Library Service for the Blind and Physically Challenged (NLS) network circulated millions of recorded books to more than 700,000 Physically challenged listeners. All NLS recordings were created by professionals.

Though spoken recordings were popular in 33-1/3 vinyl record format for schools and libraries into the early 1970s, the beginning of the trade acceptance of this medium can be traced to the introduction of the audio cassette and, most importantly, to the prevalence of these cassette players as standard equipment (rather than as options which older drivers did not choose) in imported (Japanese) automobiles, which became very popular during the 1979 energy crisis. Thereafter, consumers and authors slowly accepted the medium. Into the early 1980s there were still many authors who refused to have their books created as audiobooks, so a good many of the audiobooks were original productions not based upon printed books.

With the development of portable cassette recorders, audiotapes had become very popular and by the late 1960s libraries became a source of free audiobooks, primarily on vinyl records but also on cassettes. Instructional and educational recordings came first, followed by self-help tapes and then by literature. In 1975, Olympic gold medalist, Duvall Hecht founded Books on Tape, Inc. as a direct to consumer mail order rental service for unabridged audiobooks and expanded their services selling their products to libraries and audiobooks gaining popularity with commuters and travelers. By the middle of 1980s the audio publishing business grew to several billion dollars a year in retail value.

Caedmon was the first to work with integrated production teams and professional actors, while Nightingale Conant featured business and self-help authors reading their own works first on vinyl records and then on cassettes.

The Audio Publishers Association was established in 1986 by six competitive companies who joined together to promote the consumer awareness of spoken word audio. In 1996 the Audio Publishers Association established the Audie Awards for audio books, which is equivalent to the Oscar for the talking books industry. The nominees are announced each year in January. The winners are announced at a gala banquet in the spring, usually in conjunction with BookExpo America.

While music fans rapidly accepted CDs, audiobook listeners were slower. One reason may have been that a cassette tape by nature retains its position when the player is turned off, but many early CD players did not retain the playing position of CDs when turned

off. Also, it was not until cassette players were replaced by CD players in most automobiles that this format eventually took hold.

With the advent of the Internet, broadband technologies, new compressed audio formats and portable media players, the popularity of audiobooks has increased significantly. This growth was reflected with the advent of audiobook download subscription services.

Use, distribution and popularity

Recent technology has encouraged the proliferation of free audiobooks that take works from the public domain and enlist volunteers to read them. Audiobooks also can be created with text to speech computer software, although the quality of synthesised speech may suffer by comparison to recordings by a human voice. On the other hand, computer-voiced reading enables the proliferation of more works faster through automation, than if read by humans.

Audiobooks in the private domain are also distributed online by for-profit companies. It is believed that most major audiobook publishers insist that their works, when sold as downloads, be protected by Digital Rights Management (DRM), but this is not the case. Companies such as Apple Inc. have licensed their proprietary FairPlay DRM system (for DRM protection of iPod files in the .aa file format) to only one company. . Because of the major publishers' insistence on DRM, this has effectively created a monopoly in the sale of the works of major publishers to iPod users, who make up the majority of the portable audio market. However, with the release of new sites such as We Read 4 You, which releases all of its content as DRM-free .mp3 files, there may be a slow conversion by other companies, such as Apple and Amazon to a DRM-free format.

Audiobooks on cassette or CD are typically more expensive than hardbacks because of the added expense of recording and the lack of the economy of scale in high "print" runs that are available in the publishing of printed books. Preloaded digital formats are similar in price to their CD counterparts. The audio content is preloaded on a small and simple player, which removes the need for a separate piece of technology such as a CD player or an MP3 player.

Downloadable audiobooks tend to cost slightly less than hardbacks but more than their paperback equivalents. For this reason, market penetration of audiobooks is substantially lower than for their printed counterparts despite the high market penetration of the hardware (MP3 and WMA players) and despite the massive market penetration achieved by audio music products. Given the elasticity of demand for audiobooks and the availability of cheaper alternatives, slow and steady growth in sales seems more likely than a mass market explosion.

However, economics are on the side of downloadable audiobooks in the long run. They do not carry mass production costs, do not require storage of a large inventory, do not require physical packaging or transportation and do not face the problem of returns that add to the cost of printed books. Received wisdom of market forces suggests that

significant price reductions to customers, while cutting into per unit profit margins, will be offset by increased volumes of sales. This will increase absolute profits to the industry while bringing audiobooks to a wider public.

One of the factors holding back price competition is the fear that low-price audiobooks might simply take business away from more traditional forms of publishing. This is especially significant in the case of publishers who have interests in both print and audiobook publishing. However, most major book publishers now actively participate in audiobook publishing and see it as a complement to their publishing operations.

Resellers of audiobooks that acquire much of their content from major publishers, must price their content at such a level as to take account of their cost of goods as well as operating costs. On the other hand, audiobook sellers that sell their own content or publish lesser known authors have lower operating costs and can therefore sell at lower prices using a "lower-margin-higher-sales" business model. However, they still have to meet the costs of writer's royalties, performers fees and production facility costs. The shift from CDs and cassettes to downloadable audiobooks, whilst doing nothing to reduce initial recording and editing costs, creates further downward pressure on price, by removing some of the other costs, such as production, packaging and physical distribution.

Audiobooks have been used to teach children to read and to increase reading comprehension. They are also useful for the blind. The National Library of Congress in the U.S. and the CNIB Library in Canada provide free audiobook library services to the visually impaired; requested books are mailed out (at no cost) to clients.

About forty percent of all audiobook consumption occurs through public libraries, with the remainder served primarily through retail book stores. Library download programs are currently experiencing rapid growth (more than 5,000 public libraries offer free downloadable audio books). Libraries are also popular places to check out audio books in the CD format. According to the National Endowment for the Arts' recent study, "Reading at Risk", audio book listening is one of very few "types" of reading that is increasing general literacy.

Listening practices

Audio books are considered a valuable learning tool because of their format. Unlike traditional books or a video program, one can learn from an audiobook while doing other tasks, although it should be noted that this can detract from the primary task, assuming the learning is not the main activity. Such multitasking is feasible when doing mechanical tasks that do not require much thought and have only little or no chance of an emergency arising. Such tasks include doing the laundry and exercising indoors, among others. The most popular general use of audiobooks by adults is when driving an automobile or traveling with public transport, as an alternative to radio. Many people listen as well just to relax or as they drift off into sleep.

Common practices include:

- **Replaying:** Depending upon one's degree of attention and interest, it is often necessary to listen to segments of an audio book more than once to allow the material to be understood and retained satisfactorily. Replaying may be done immediately or after extended periods of time.
- **Learning:** People may listen to an audio book (usually an unabridged one) while following along in an actual book. This helps them to learn words that they may not learn correctly if they were only to read the book. This can also be a very effective way to learn a new language.

Audiobook charities in the UK

Listening Books is an audiobook charity in the UK providing an internet streaming and postal service to anyone who has a disability or illness which makes it difficult to hold a book, turn its pages, or read in the usual way. They have audiobooks for both leisure and learning and a library of over 4,000 titles which are recorded in their own digital studios or commercially sourced.

The Royal National Institute of Blind People (RNIB) is a UK charity which offers a Talking Books library service. The audio books are provided in DAISY format and delivered to the reader's house by post. There are over 18,000 audio books available to borrow, paid for by annual subscription. RNIB subsidises the Talking Books service by around £4 million a year.

Spoken word album

A **spoken word album** was a record album that did not consist mainly of music or songs, but of spoken material. It could be said to be the ancestor of today's audiobook format. Spoken word albums ranged from such items as recordings of actual political speeches and/or dramatic readings of historical documents, to dialogue from the soundtrack of a film, to condensed dramatized versions of literary classics, to complete performances of plays by Shakespeare or other great authors, to stories for children, or to standup comedy routines recorded live in nightclubs.

Early beginnings

Spoken word albums have been made since the early days of recording; examples include the popular Ronald Colman 1941 version of Charles Dickens's *A Christmas Carol* on American Decca Records. However, a true milestone was reached when Columbia Masterworks, which had previously released an album of excerpts from Shakespeare's *Richard II* with Maurice Evans, recorded most of Margaret Webster's famed (and never filmed) 1943 Broadway production of *Othello*, starring Paul Robeson, José Ferrer, and Uta Hagen, on an eighteen-record 78-RPM set that was later transferred to LP. It was the longest spoken word album made up to that time. The album gave millions of listeners

who otherwise were unable to attend a theatrical performance a chance to hear Robeson as Othello and Ferrer as Iago. Sales of the album, however, were affected after Robeson was blacklisted by the U.S. government in 1950 for his alleged Communist sympathies.

LP influence and educational value

After the advent of LPs, spoken word albums became much more common. The Ronald Colman *A Christmas Carol* was transferred to LP, as were many other 78-RPM spoken word albums made by Decca, such as *Moby Dick*, with Charles Laughton as Captain Ahab. An uncredited actor provided the voice of Ishmael, the narrator. Disneyland Records issued many spoken word albums for children, including narrated adaptations of the Disney classic films in their "Disneyland Storyteller" series. Notable Broadway productions, such as the 1950 *Don Juan in Hell* (with Charles Boyer, Charles Laughton, Cedric Hardwicke and Agnes Moorehead), the 1953 dramatized reading of the poem *John Brown's Body* (with Tyrone Power, Judith Anderson and Raymond Massey), the original 1962 Broadway version of *Who's Afraid of Virginia Woolf?* (with Uta Hagen, Arthur Hill, George Grizzard and Melinda Dillon), and Richard Burton's Broadway production of *Hamlet* (1964), were all recorded complete with their original casts by Columbia Masterworks. José Ferrer was heard in an album of excerpts from *Cyrano de Bergerac*, with members of the cast of the 1946 Broadway revival in which Ferrer first played Cyrano, on Capitol Records. Caedmon Records recorded the complete plays of Shakespeare, as well as recordings of other plays such as *Death of a Salesman* with original stars Lee J. Cobb and Mildred Dunnock, *The Glass Menagerie* with Jessica Tandy, and a nearly complete *Cyrano de Bergerac* with Ralph Richardson, who had triumphed in the role in London in 1946. Many of these recorded dramas were played in high school literature classes to enable students to hear the play and follow along in their textbooks at the same time.

After the assassination of President John F. Kennedy in 1963, memorial collections of his speeches began to appear on LP. Most of the soundtrack of the commemorative 1966 documentary *John F. Kennedy: Years of Lightning, Day of Drums*, narrated by Gregory Peck, was released on a Capitol Records LP.

Dialogue excerpts were also released of the film soundtracks of Franco Zeffirelli's *The Taming of the Shrew* (1967) (on RCA Victor), and *Romeo and Juliet* (1968) (on Capitol). RCA Victor also released a virtually complete 2-LP album of the film soundtrack of *A Man for All Seasons*. RCA Victor was also responsible for complete 4-LP album sets of the Laurence Olivier *Othello* and Zeffirelli's National Theatre of Great Britain production of *Much Ado About Nothing*, starring Maggie Smith, Robert Stephens, and Albert Finney. Comedy routines by such celebrities as Bill Cosby became extremely popular on Warner Bros. Records. Emlyn Williams recorded an edition of his one-man performance as Charles Dickens (for Argo Records), and Hal Holbrook recorded excerpts from his one-man *Mark Twain Tonight!* for Columbia Masterworks.

Decline

With the advent of videocassettes and compact discs, however, original cast albums of non-musical plays, as well as spoken word albums of film soundtracks, went into a serious decline from which they have never completely recovered. CDs usually place more emphasis on music than on the spoken word, and there was little interest in only listening to a play or dialogue excerpts from a film when one could now buy plays and films on video and watch them at home whenever they wished. While the Cosby albums have resurfaced on CD, most of the other albums mentioned above have not. (Some of the Caedmon albums have been released on CD by Harper Audio, a division of Harper Collins, which now owns Caedmon.) The 1968 album of *Romeo and Juliet* excerpts has also appeared on CD, and Pearl has issued the Robeson *Othello* in that medium, but the CD edition of the *Othello* has, unfortunately, attracted little attention in comparison to the history-making vinyl record release of the 1940s, and now that *Cyrano de Bergerac*, *A Man for All Seasons*, the Olivier *Othello*, the Zeffirelli versions of *Romeo and Juliet* and *The Taming of the Shrew*, the television version of *Mark Twain Tonight*, and *Richard Burton's Hamlet* are all available on DVD, this has become for most a more preferred way to experience these productions.

Although Naxos Records is a major producer of audiobooks, many famous spoken word recordings of the past, such as Columbia Masterworks' *John Brown's Body* and *Don Juan in Hell* have yet to be released on CD, although *Don Juan in Hell* has become available as an mp3 download. Whether or not it will appear in CD form is still unknown. Also online (but not yet on CD) is Capitol Records' *The Story Teller: A Session with Charles Laughton*, a Grammy-winning one-man stage performance by the actor, featuring dramatic readings from the Bible, Shakespeare, George Bernard Shaw and Jack Kerouac, as well as autobiographical reminiscences.

Today's spoken word albums

There have been some spoken word albums over the past fifteen years or so that have been recorded specifically for compact disc; these have often been combined with classical music. Among them are the Naxos audiobooks, as well as a Chandos Records series of albums which combine the music that William Walton wrote for several Shakespeare productions (including the Olivier film adaptations) with readings from the author performed by such actors as John Gielgud and Christopher Plummer. There is also a Hyperion Records stereophonic re-creation of Ralph Vaughan Williams's 1942 radio play adaptation of *The Pilgrim's Progress*, again with Gielgud.

Today, such websites as BBC, ZBS, L.A. Theatre Works and The Hollywood Theater of the Ear offer full-length recordings on CD of their dramatic productions. These recordings are possibly the closest that modern day discs have come to the spoken word albums of the 1960's.

Chapter- 2

Windows Media Audio

Windows Media Audio



Filename extension	.wma
Internet media type	audio/x-ms-wma
Uniform Type Identifier	com.microsoft.windows-?media-wma
Developed by	Microsoft

Windows Media Audio (WMA) is an audio data compression technology developed by Microsoft. The name can be used to refer to its audio file format or its audio codecs. It is a proprietary technology that forms part of the Windows Media framework. WMA consists of four distinct codecs. The original WMA codec, known simply as *WMA*, was conceived as a competitor to the popular MP3 and RealAudio codecs. *WMA Pro*, a newer and more advanced codec, supports multichannel and high resolution audio. A lossless codec, *WMA Lossless*, compresses audio data without loss of audio fidelity (the regular WMA format is not lossless). And *WMA Voice*, targeted at voice content, applies compression using a range of low bit rates.

Development history

The first WMA codec was based on earlier work by Henrique Malvar and his team which was transferred to the Windows Media team at Microsoft. Malvar was a senior researcher and manager of the Signal Processing Group at Microsoft Research, whose team worked on the *MSAudio* project. The first finalized codec was initially referred to as *MSAudio 4.0*. It was later officially released as *Windows Media Audio*, as part of Windows Media Technologies 4.0. Microsoft claimed that WMA could produce files that were half the size of equivalent-quality MP3 files; Microsoft also claimed that WMA delivered "near CD-quality" audio at 64 kbit/s. The former claim however was rejected by some

audiophiles. RealNetworks also challenged Microsoft's claims regarding WMA's superior audio quality compared to RealAudio.

Newer versions of WMA became available: *Windows Media Audio 2* in 1999, *Windows Media Audio 7* in 2000, *Windows Media Audio 8* in 2001, and *Windows Media Audio 9* in 2003. Microsoft first announced its plans to license WMA technology to third-parties in 1999. Although earlier versions of Windows Media Player played WMA files, support for WMA file creation was not added until the seventh version. In 2003, Microsoft released new audio codecs that were not compatible with the original WMA codec. These codecs were *Windows Media Audio 9 Professional*, *Windows Media Audio 9 Lossless*, and *Windows Media Audio 9 Voice*.

Container format

A WMA file is in most circumstances encapsulated, or contained, in the Advanced Systems Format (ASF) container format, featuring a single audio track in one of following codecs: WMA, WMA Pro, WMA Lossless, or WMA Voice. These codecs are technically distinct and mutually incompatible. The ASF container format specifies how metadata about the file is to be encoded, similar to the ID3 tags used by MP3 files. Metadata may include song name, track number, artist name, and also audio normalization values.

This container can optionally support digital rights management (DRM) using a combination of elliptic curve cryptography key exchange, DES block cipher, a custom block cipher, RC4 stream cipher and the SHA-1 hashing function.

Codecs

Windows Media Audio

Windows Media Audio (WMA) is the most common codec of the four WMA codecs. Colloquial usage of the term *WMA*, especially in marketing materials and device specifications, usually refers to this codec only. The first version of the codec released in 1999 is regarded as WMA 1. In the same year, the bit stream syntax, or compression algorithm, was altered in minor ways and became WMA 2. Since then, newer versions of the codec were released, but the decoding process remained the same, ensuring compatibility between codec versions. WMA is a lossy audio codec based on the study of psychoacoustics. Audio signals that are deemed to be imperceptible to the human ear are encoded with reduced resolution during the compression process.

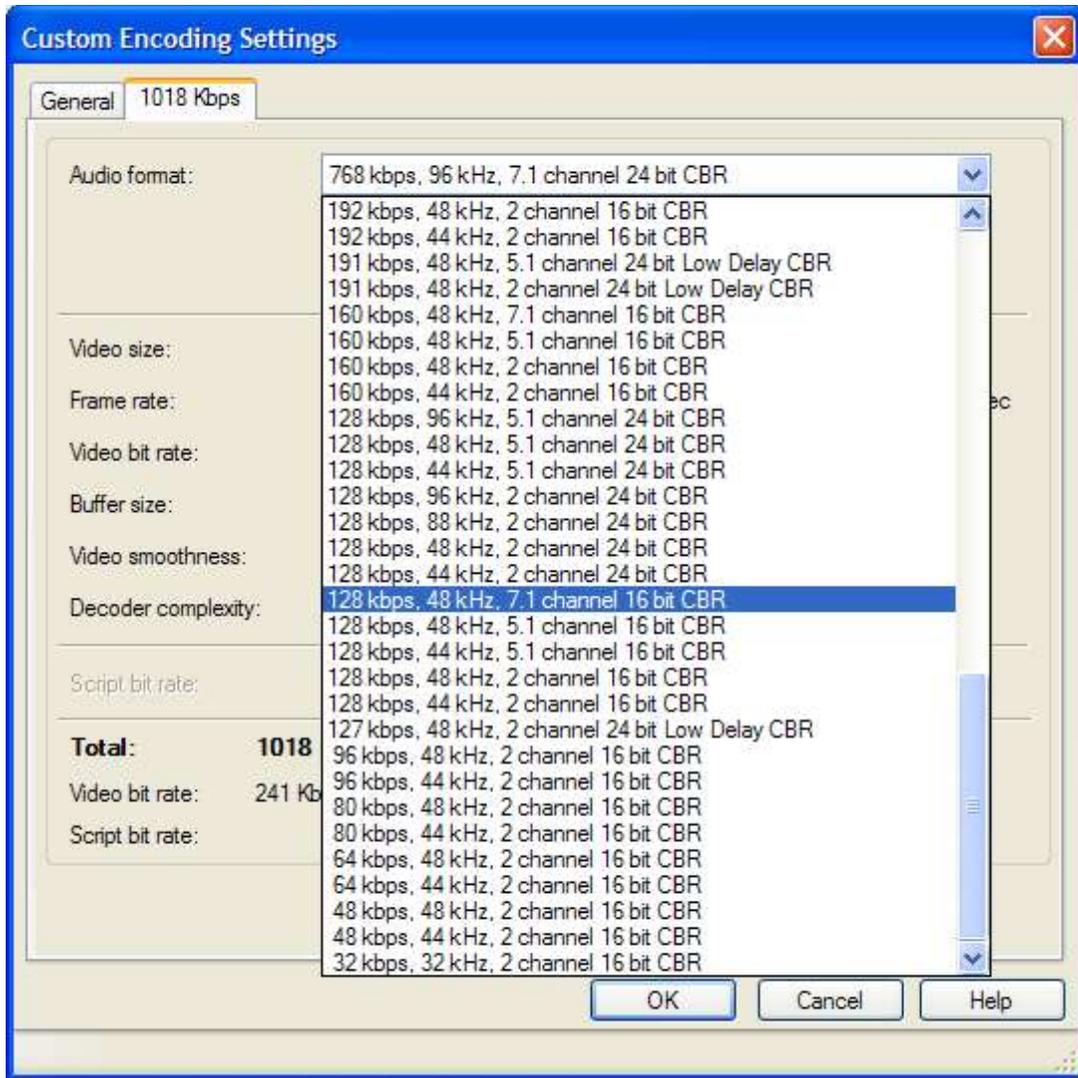
WMA can encode audio signals sampled at up to 48 kHz with up to two discrete channels (stereo). WMA 9 introduced variable bit rate (VBR) and average bit rate (ABR) coding techniques into the MS encoder although both were technically supported by the original format. WMA 9.1 also added support for low-delay audio, which reduces latency for encoding and decoding.

Fundamentally, WMA is a transform coder based on modified discrete cosine transform (MDCT), somewhat similar to AAC, Cook and Vorbis. The bit stream of WMA is composed of superframes, each containing 1 or more frames of 2048 samples. If the bit reservoir is not used, a frame is equal to a superframe. Each frame contains a number of blocks, which are 128, 256, 512, 1024, or 2048 samples long after being transformed into the frequency domain via the MDCT. In the frequency domain, masking for the transformed samples is determined, and then used to requantize the samples. Finally, the floating point samples are decomposed into coefficient and exponent parts and independently huffman coded. Stereo information is typically mid/side coded. At low bit rates, line spectral pairs (typically less than 17 kbit/s) and a form of noise coding (typically less than 33 kbit/s) can also be used to improve quality.

Like AAC and Ogg Vorbis, WMA was intended to address perceived deficiencies in the MP3 standard. Given their common design goals, it's not surprising that the three formats ended up making similar design choices. All three are pure transform codecs. Furthermore the MDCT implementation used in WMA is essentially a superset of those used in Ogg and AAC such that WMA iMDCT and windowing routines can be used to decode AAC and Ogg Vorbis almost unmodified. However, quantization and stereo coding is handled differently in each codec. The primary distinguishing trait of the WMA Standard format is its unique use of 5 different block sizes, compared to MP3, AAC, and Ogg Vorbis which each restrict files to just two sizes. WMA Pro extends this by adding a 6th block size used at 88.1/96 kHz sampling rate.

Certified PlaysForSure devices, as well as a large number of uncertified devices, ranging from portable hand-held music players to set-top DVD players, support the playback of WMA files. Most PlaysForSure-certified online stores distribute content using this codec only. In 2005, Nokia announced its plans to support WMA playback in future Nokia handsets. In the same year, an update was made available for the PlayStation Portable (version 2.60) which allowed WMA files to be played on the device for the first time.

Windows Media Audio Professional



Screenshot of Windows Media Encoder 9 Series, displaying new encoding options for Windows Media Audio 10 Professional.

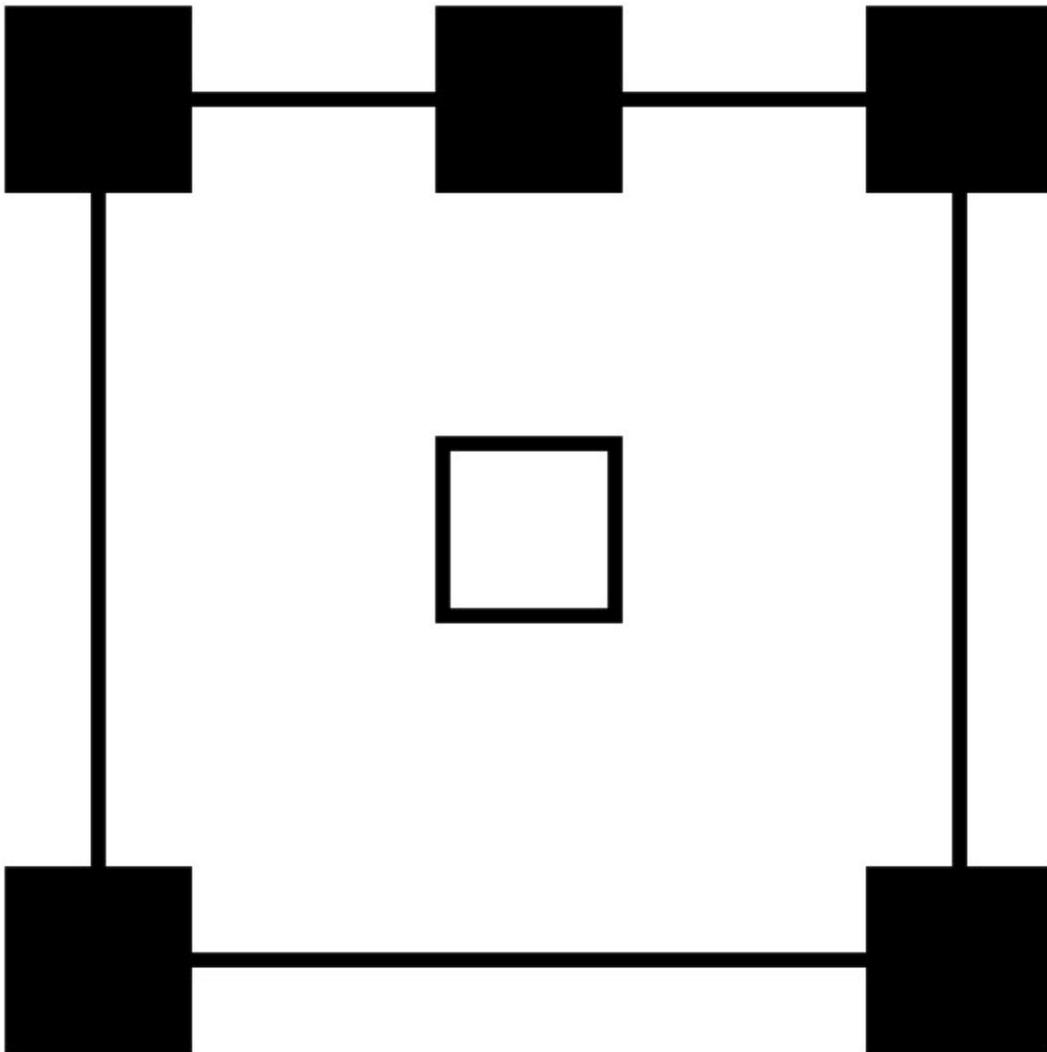
Windows Media Audio Professional (WMA Pro) is an improved lossy codec closely related to WMA standard. It retains most of the same general coding features, but also features improved entropy coding and quantization strategies as well as more efficient stereo coding. Notably, many of the WMA standard's low bitrate features have been removed, as the core codec is designed for efficient coding at most bitrates. Its main competitors include AAC, HE-AAC, Vorbis, Dolby Digital, and DTS. It can support audio resolutions of up to 96 kHz and up to eight discrete channels (7.1 channel surround). WMA Pro also supports dynamic range compression, which reduces the volume difference between the loudest and quietest sounds in the audio track. According to Microsoft's Amir Majidimehr, WMA Pro can technically go beyond 7.1 surround sound and support "an unlimited number of channels."

The codec's bit stream syntax was frozen at the first version, WMA 9 Pro. Later versions of WMA Pro introduced low-bit rate encoding, low-delay audio, frequency interpolation mode, and an expanded range of sampling rate and bit-depth encoding options. A WMA 10 Pro file compressed with frequency interpolation mode comprises a WMA 9 Pro track encoded at half the original sampling rate, which is then restored using a new compression algorithm. In this situation, WMA 9 Pro players which have not been updated to the WMA 10 Pro codec can only decode the lower quality WMA 9 Pro stream. Starting with WMA 10 Pro, eight channel encoding starts at 128 kbit/s, and tracks can be encoded at the native audio CD resolution (44.1 kHz, 16-bit), previously the domain of WMA Standard.

Despite a growing number of supported devices and its superiority over WMA, WMA Pro still has little hardware and software support. Some notable exceptions to this are the Microsoft Zune (limited to stereo), Xbox 360, Windows Mobile-powered devices with Windows Media Player 10 Mobile, newer Toshiba Gigabeat and Motorola devices, and devices running recent versions of the Rockbox alternative firmware. In addition, WMA Pro is a requirement for the WMV HD certification program. On the software side, Verizon utilizes WMA 10 Pro for its V CAST Music Service, and Windows Media Player 11 has promoted the codec as an alternative to WMA for copying audio CD tracks. WMA Pro is supported in Silverlight as of version 2 (though only in stereo mode). In the absence of the appropriate audio hardware, WMA Pro can automatically downmix multichannel audio to stereo or mono, and 24-bit resolution to 16-bit during playback.

A notable example of WMA Pro being used instead of WMA Standard is the NBC Olympics website which uses WMA 10 Pro in its low-bitrate mode at 48 kbit/s.

Windows Media Audio Lossless



Label for 5.1 surround sound, the maximum channel configuration for Windows Media Audio Lossless.

Windows Media Audio Lossless (WMA Lossless) is a lossless audio codec that competes with ATRAC Advanced Lossless, Dolby TrueHD, DTS-HD Master Audio, Apple Lossless, Shorten, Monkey's Audio, FLAC, and WavPack (the last two have the advantage of being open source software and available for nearly any operating system). Designed for archival purposes, it compresses audio signals without loss of quality from the original using VBR. When decompressed, the audio signal is an exact replica of the original. The first version of the codec, WMA 9 Lossless, and its revisions support up to 96 kHz, 24-bit audio for up to 6 discrete channels (5.1 channel surround) with dynamic range compression control. The typical compression ratio for music varies between 1.7:1 and 3:1.

Hardware support for the codec is available on the Cowon A3 , Bang & Olufsen Serenata, Sony Walkman NWZ-A and NWZ-S series, Zune 4, 8, 80 30, Zune 120 (with firmware version 2.2 or later) and the new Zune HD, Xbox 360, Windows Mobile-powered devices with Windows Media Player 10 Mobile, Toshiba Gigabeat S and V models, Toshiba T-400, the Meizu M3, and Best Buy's Insignia NS-DV, Pilot, and Sport music players. Contrary to some claims, the Archos make of media devices do not support WMA Lossless, nor does the SONOS system . Like WMA Standard, WMA Lossless is being used by a few online stores to distribute music online. Similar to WMA Pro, the WMA Lossless decoder can perform downmixing when capable audio hardware is not present .

Windows Media Audio Voice

Windows Media Audio Voice (WMA Voice) is a lossy audio codec that competes with Speex (used in Microsoft's own Xbox Live online service), ACELP, and other codecs. Designed for low-bandwidth, voice playback applications, it employs low-pass and high-pass filtering of sound outside the human speech frequency range to achieve higher compression efficiency than WMA. It can automatically detect sections of an audio track containing both voice and music and use the standard WMA compression algorithm instead. WMA Voice supports up to 22.05 kHz for a single channel (mono) only. Encoding is limited to constant bit rate (CBR) and up to 20 kbit/s. The first and only version of the codec is WMA 9 Voice.

Windows Mobile-powered devices with Windows Media Player 10 Mobile have native support for WMA 9 Voice playback. In addition, BBC World Service has employed WMA Voice for its Internet radio streaming service.

Sound quality

Microsoft claims that audio encoded with WMA sounds better than MP3 at the same bit rate; Microsoft also claims that audio encoded with WMA at lower bit rates sound better than MP3 at higher bit rates. Double blind listening tests with other lossy audio codecs have shown varying results, from failure to support Microsoft's claims about its superior quality to supremacy over other codecs. One independent test conducted in May 2004 at 128 kbit/s showed that WMA was roughly equivalent to LAME MP3; inferior to AAC and Vorbis; and superior to ATRAC3 (software version).

Some conclusions made by recent studies:

- At 32 kbit/s, WMA Standard was noticeably better than LAME MP3, but not better than other modern codecs in a collective, independent test in July 2004.
- At 48 kbit/s, WMA 10 Pro was ranked second after Nero HE-AAC and better than WMA 9.2 in an independent listening test organized and supported by Sebastian Mares and Hydrogenaudio Forums in December 2006. This test, however, used CBR for WMA 10 Pro and VBR for the other codecs.

- At 64 kbit/s, WMA Pro outperformed Nero HE-AAC in a commissioned, independent listening test performed by the National Software Testing Labs in 2005. Out of 300 participants, "71% of all listeners indicated that WMA Pro was equal to or better than HE AAC."
- At 80 kbit/s and 96 kbit/s, WMA had lower quality than HE-AAC, AAC-LC, and Vorbis; near-equivalent quality to MP3, and better quality than MPC in individual tests done in 2005.
- At 128 kbit/s, there was a four-way tie between aoTuV Vorbis, LAME MP3, WMA 9 Pro and AAC in a large scale test in January 2006, with each codec sounding close to the uncompressed music file for most listeners.
- At 768 kbit/s, WMA 9 Pro delivered full-spectrum response at half the bit rate required for DTS in a comparative test done by EDN in October 2003. The test sample was a 48 kHz, 5.1 channel surround audio track.

Criticism of claimed quality

Microsoft's claims of WMA sound quality have frequently drawn complaints. "Some audiophiles challenge Microsoft's claims regarding WMA's quality," according to a published article from EDN. Another article from MP3 Developments wrote that Microsoft's claim about CD-quality audio at 64 kbit/s with WMA was "very far from the truth." At the early stages of WMA's development, a representative from RealNetworks claimed that WMA was a "clear and futile effort by Microsoft to catch up with RealAudio 8."

Microsoft has sometimes claimed that the sound quality of WMA at 64 kbit/s equals or exceeds that of MP3 at 128 kbit/s (both WMA and MP3 are considered near-transparent at 192 kbit/s by most listeners). In a 1999 study funded by Microsoft, National Software Testing Laboratories (NSTL) found that listeners preferred WMA at 64 kbit/s to MP3 at 128 kbit/s (as encoded by MusicMatch Jukebox). However, a September 2003 public listening test conducted by Roberto Amorim found that listeners preferred 128 kbit/s MP3 to 64 kbit/s WMA audio with greater than 99% confidence. This conclusion applied equally to other codecs at the same bitrate, leading him to conclude that:

“ No codec delivers the marketing plot of same quality as MP3 at half the bitrates. ”

It is important to note that both MP3 and WMA encoders have undergone active development and improvement for many years, so their relative quality may change over time.

A July 2007 public listening test by Sebastian Mares found that 64 kbit/s HE-AAC audio (encoded by Nero Digital) was statistically tied with 64 kbit/s WMA Pro audio, in terms of listener preference.

Players

Apart from Windows Media Player, most of the WMA compression formats can be played using ALLPlayer, VLC media player, MPlayer, RealPlayer, Winamp, Zune Software (with certain limitations—DSP plugin support and DirectSound output is disabled using the default WMA plugin), and many other software media players. The Microsoft Zune media management software supports most WMA codecs, but uses a variation of Windows Media DRM which is used by PlaysForSure.

The FFmpeg project has reverse-engineered and re-implemented the WMA codecs (except WMA Lossless) to allow their use on POSIX-compliant operating systems such as Linux. The rockbox project further extended this codec to be suitable for embedded cores, allowing playback on portable MP3 players and cell phones running open source software. RealNetworks has announced plans to support playback of DRM-unprotected WMA files in RealPlayer for Linux. On the Macintosh platform, Microsoft released a PowerPC version of Windows Media Player for Mac OS X in 2003, but further development of the software has ceased. Microsoft currently endorses the third-party Flip4Mac WMA, a QuickTime component that allows Macintosh users to play WMA files in any player that uses the QuickTime framework. Flip4Mac, however, does not currently support the Windows Media Audio Voice codec.

Encoders

Software that can export audio in WMA format include Windows Media Player, Windows Movie Maker, KM Player, Microsoft Expression Encoder, Sony Sound Forge, GOM Player, RealPlayer, Adobe Premiere Pro, Adobe Audition, and Adobe Soundbooth. Microsoft Office OneNote supports encoding in all WMA codecs, Music Morpher. and Windows Media Encoder supports all available bit rate and resolution options as well.

Digital rights management

The WMA codecs are most often used with the ASF container format, which has an optional DRM facility. Windows Media DRM, which can be used in conjunction with WMA, supports time-limited music subscription services such as those offered by unlimited download services, including MTV's URGE, Napster, Rhapsody, Yahoo! Music Unlimited, and Virgin Digital. Windows Media DRM, a component of PlaysForSure and Windows Media Connect, is supported on many modern portable audio devices and streaming media clients such as Roku, SoundBridge, Xbox 360, and Wii. Players that support the WMA format but not Windows Media DRM list protected titles as unplayable.

Chapter- 3

Advanced Audio Coding

Advanced Audio Codings



iTunes DRM-restricted AAC file icon

Filename extension	.m4a, .m4b, .m4p, .m4v, .m4r, .3gp, .mp4, .aac
Internet media type	audio/aac, audio/aacp, audio/3gpp, audio/3gpp2, audio/mp4, audio/MP4A-LATM, audio/mpeg4-generic
Initial release	1997
Type of format	Audio compression format, Lossy compression
Contained by	MPEG-4 Part 14, 3GP and 3G2, ISO base media file format and Audio Data Interchange Format (ADIF)
Standard(s)	ISO/IEC 13818-7, ISO/IEC 14496-3

Advanced Audio Coding (AAC) is a standardized, lossy compression and encoding scheme for digital audio. Designed to be the successor of the MP3 format, AAC generally achieves better sound quality than MP3 at similar bit rates.

AAC has been standardized by ISO and IEC, as part of the MPEG-2 and MPEG-4 specifications. The MPEG-2 standard contains several audio coding methods, including the MP3 coding scheme. AAC is able to include 48 full-bandwidth (up to 96 kHz) audio

channels in one stream plus 16 low frequency effects (LFE, limited to 120 Hz) channels, up to 16 "coupling" or dialog channels, and up to 16 data streams. The quality for stereo is satisfactory to modest requirements at 96 kbit/s in joint stereo mode; however, hi-fi transparency demands data rates of at least 128 kbit/s (VBR). The MPEG-2 audio tests showed that AAC meets the requirements referred to as "transparent" for the ITU at 128 kbit/s for stereo, and 320 kbit/s for 5.1 audio.

AAC is also the default or standard audio format for: Apple's iPhone, iPod, iPad, Nintendo DSi, iTunes, DivX Plus Web Player, Sony's PlayStation 3 and is supported by Sony's PlayStation Portable, latest generation of Sony Walkman, phones from Sony Ericsson, the latest S40 and S60 models from Nokia, Android based phones, Nintendo's Wii (with the Photo Channel 1.1 update installed for Wii consoles purchased before late 2007), and the MPEG-4 video standard.

"*High-Efficiency AAC*" is part of digital radio standards like DAB+ and Digital Radio Mondiale, and mobile television standards DVB-H and ATSC-M/H.

History

AAC was developed with the cooperation and contributions of companies including AT&T Bell Laboratories, Fraunhofer IIS, Dolby, Sony Corporation and Nokia, and was officially declared an international standard by the Moving Picture Experts Group in April 1997. MPEG-2 AAC-LC profile consists of a base format very much like AT&T's Perceptual Audio Coding (PAC) coding format, with the addition of temporal noise shaping (TNS), the Dolby Kaiser Window described below, a nonuniform quantizer, and a reworking of the bitstream format to handle up to 16 stereo, 16 mono, 16 LFE, and 16 commentary channels in one bitstream. The Main profile adds a set of recursive predictors that are calculated on each tap of the filterbank. The SSR uses a 4-band PQMF filterbank, with four shorter filterbanks following, in order to allow for scalable sampling rates.

Standardization

It is specified both as *Part 7 of the MPEG-2 standard*, and *Subpart 4 in Part 3 of the MPEG-4 standard*. As such, it can be referred to as **MPEG-2 Part 7** and **MPEG-4 Part 3 AAC** depending on its implementation, however it is most often referred to as **MPEG-2 AAC**, **MPEG-4 AAC**, or **AAC** for short.

AAC was first specified in the standard MPEG-2 Part 7 (known formally as ISO/IEC 13818-7:1997) in 1997 as a new "part" (distinct from ISO/IEC 13818-3, a.k.a. MPEG-2 BC - backwards compatible) in the MPEG-2 family of international standards. It is known as MPEG-2 NBC (Non-Backward Compatible), because it is not compatible with the MPEG-1 Audio formats (MP3, MP2, MP1). It defined three profiles: Low complexity profile (AAC LC), Main profile (AAC Main) and Scalable sampling rate profile (AAC SSR).

It was updated in MPEG-4 Part 3 (MPEG-4 Audio) (known formally as ISO/IEC 14496-3:1999) in 1999. The MPEG-4 Part 3 standard also defined usage of other audio compression formats (a.k.a. Audio Object Types), such as TwinVQ, CELP, HVXC, Text-To-Speech Interface, Structured Audio and others. A notable addition in this version of the AAC standard is Perceptual Noise Substitution (PNS). MPEG-4 Audio is defined in a way that it remains backwards compatible to MPEG-2 AAC. The MPEG-2 Part 7 profiles - AAC LC profile, AAC Main profile and AAC SSR profile are combined with Perceptual Noise Substitution and defined in the MPEG-4 Audio standard as Audio Object Types (using the same names AAC LC, AAC Main and AAC SSR). MPEG-4 Audio Object Types are combined in four MPEG-4 Audio profiles: Main (which includes most of the MPEG-4 Audio Object Types), Scalable (AAC LC, AAC LTP, CELP, HVXC, TwinVQ, Wavetable Synthesis, TTSI), Speech (CELP, HVXC, TTSI) and Low rate Synthesis (Wavetable Synthesis, TTSI). The reference software for MPEG-4 Part 3 is specified in MPEG-4 Part 4 and the conformance bit-streams are specified in MPEG-4 Part 5.

The MPEG-4 Audio Version 2 (ISO/IEC 14496-3:1999/Amd 1:2000) defined new Audio Object Types - the Low Delay AAC (AAC-LD) object type, Bit-Sliced Arithmetic Coding (BSAC) object type, Parametric audio coding using Harmonic and Individual Line plus Noise and Error Resilient (ER) versions of object types. It also defined four new audio profiles: High Quality Audio Profile, Low Delay Audio Profile, Natural Audio Profile and Mobile Audio Internetworking Profile.

The HE-AAC Profile (AAC LC with SBR) and AAC Profile (AAC LC) were first standardized in ISO/IEC 14496-3:2001/Amd 1:2003. The HE-AAC v2 Profile (AAC LC with SBR and Parametric Stereo) was first specified in ISO/IEC 14496-3:2005/Amd 2:2006. The Parametric Stereo audio object type used in HE-AAC v2 was first defined in ISO/IEC 14496-3:2001/Amd 2:2004.

The current version of the AAC standard is defined in ISO/IEC 14496-3:2009.

AAC+ v2 is also standardized by ETSI (European Telecommunications Standards Institute) as TS 102005.

The MPEG-4 Part 3 standard also contains other ways of compressing sound. These include lossless compression formats, synthetic audio and low bit-rate compression formats generally used for speech.

AAC's improvements over MP3

Advanced Audio Coding is designed to be the successor of the *MPEG-1 Audio Layer 3*, known as MP3 format, which was specified by ISO/IEC in 11172-3 (MPEG-1 Audio) and 13818-3 (MPEG-2 Audio).

Blind tests show that AAC demonstrates greater sound quality and transparency than MP3 for files coded at the same bit rate.

Improvements include:

- More sample frequencies (from 8 to 96 kHz) than MP3 (16 to 48 kHz)
- Up to 48 channels (MP3 supports up to two channels in MPEG-1 mode and up to 5.1 channels in MPEG-2 mode)
- Arbitrary bit-rates and variable frame length. Standardized constant bit rate with bit reservoir.
- Higher efficiency and simpler filterbank (rather than MP3's hybrid coding, AAC uses a pure MDCT)
- Higher coding efficiency for stationary signals (AAC uses a blocksize of 1024 or 960 samples, allowing more efficient coding than MP3's 576 sample blocks)
- Higher coding accuracy for transient signals (AAC uses a blocksize of 128 or 120 samples, allowing more accurate coding than MP3's 192 sample blocks)
- Can use Kaiser-Bessel derived window function to eliminate spectral leakage at the expense of widening the main lobe
- Much better handling of audio frequencies above 16 kHz
- More flexible joint stereo (different methods can be used in different frequency ranges)
- Adds additional modules (tools) to increase compression efficiency: TNS, Backwards Prediction, PNS etc... These modules can be combined to constitute different encoding profiles.

Overall, the AAC format allows developers more flexibility to design codecs than MP3 does, and corrects many of the design choices made in the original MPEG-1 audio specification. This increased flexibility often leads to more concurrent encoding strategies and, as a result, to more efficient compression. However, in terms of whether AAC is better than MP3, the advantages of AAC are not entirely decisive, and the MP3 specification, although antiquated, has proven surprisingly robust in spite of considerable flaws. AAC and HE-AAC are better than MP3 at low bit rates (typically less than 128 kilobits per second). This is especially true at very low bit rates where the superior stereo coding, pure MDCT, and more optimal transform window sizes leave MP3 unable to compete. However, as bit rate increases, the efficiency of an audio format becomes less important relative to the efficiency of the encoder's implementation, and the intrinsic advantage AAC holds over MP3 no longer dominates audio quality.

How AAC works

AAC is a wideband audio coding algorithm that exploits two primary coding strategies to dramatically reduce the amount of data needed to represent high-quality digital audio.

1. Signal components that are perceptually irrelevant are discarded;
2. Redundancies in the coded audio signal are eliminated.

The actual encoding process consists of the following steps:

- The signal is converted from time-domain to frequency-domain using forward modified discrete cosine transform (MDCT). This is done by using filter banks that take an appropriate number of time samples and convert them to frequency samples.
- The frequency domain signal is quantized based on a psychoacoustic model and encoded.
- Internal error correction codes are added;
- The signal is stored or transmitted.
- In order to prevent corrupt samples, a modern implementation of the Luhn mod N algorithm is applied to each frame

The MPEG-4 audio standard does not define a single or small set of highly efficient compression schemes but rather a complex toolbox to perform a wide range of operations from low bitrate speech coding to high-quality audio coding and music synthesis.

- The MPEG-4 audio coding algorithm family spans the range from low bitrate speech encoding (down to 2 kbit/s) to high-quality audio coding (at 64 kbit/s per channel and higher).
- AAC offers sampling frequencies between 8 kHz and 96 kHz and any number of channels between 1 and 48.
- In contrast to MP3's hybrid filter bank, AAC uses the modified discrete cosine transform (MDCT) together with the increased window lengths of 1024 or 960 points.

AAC encoders can switch dynamically between a single MDCT block of length 1024 points or 8 blocks of 128 points (or between 960 points and 120 points, respectively).

- If a signal change or a transient occurs, 8 shorter windows of 128/120 points each are chosen for their better temporal resolution.
- By default, the longer 1024-point/960-point window is otherwise used because the increased frequency resolution allows for a more sophisticated psychoacoustic model, resulting in improved coding efficiency.

Modular encoding

AAC takes a modular approach to encoding. Depending on the complexity of the bitstream to be encoded, the desired performance and the acceptable output, implementers may create profiles to define which of a specific set of tools they want to use for a particular application.

The MPEG-2 Part 7 standard (Advanced Audio Coding) was first published in 1997 and offers three default profiles :

- **Low Complexity (LC)** – the simplest and most widely used and supported;
- **Main Profile (Main)** – like the LC profile, with the addition of backwards prediction;

- **Scalable Sample Rate (SSR)** (MPEG-4 AAC-SSR) – a.k.a. *Sample-Rate Scalable (SRS)*;

The MPEG-4 Part 3 standard (MPEG-4 Audio) defined various new compression tools (a.k.a. Audio Object Types) and their usage in brand new profiles. AAC is not used in some of the MPEG-4 Audio profiles. The MPEG-2 Part 7 AAC LC profile, AAC Main profile and AAC SSR profile are combined with Perceptual Noise Substitution and defined in the MPEG-4 Audio standard as Audio Object Types (under the name AAC LC, AAC Main and AAC SSR). These are combined with other Object Types in MPEG-4 Audio profiles. Here is a list of some audio profiles defined in the MPEG-4 standard:

- **Main Audio Profile** – defined in 1999, uses most of the MPEG-4 Audio Object Types (AAC Main, AAC LC, AAC SSR, AAC LTP, AAC Scalable, TwinVQ, CELP, HVXC, TTSI, Main synthesis)
- **Scalable Audio Profile** – defined in 1999, uses AAC LC, AAC LTP, AAC Scalable, TwinVQ, CELP, HVXC, TTSI
- **Speech Audio Profile** – defined in 1999, uses CELP, HVXC, TTSI
- **Synthetic Audio Profile** – defined in 1999, TTSI, Main synthesis
- **High Quality Audio Profile** – defined in 2000, uses AAC LC, AAC LTP, AAC Scalable, CELP, ER AAC LC, ER AAC LTP, ER AAC Scalable, ER CELP
- **Low Delay Audio Profile** – defined in 2000, uses CELP, HVXC, TTSI, ER AAC LD, ER CELP, ER HVXC
- **Mobile Audio Networking Profile** – defined in 2000, uses ER AAC LC, ER AAC Scalable, ER TwinVQ, ER BSAC, ER AAC LD
- **AAC Profile** – defined in 2003, uses AAC LC
- **High Efficiency AAC Profile** – defined in 2003, uses AAC LC, SBR
- **HE-AAC v2 Profile** – defined in 2006, uses AAC LC, SBR, PS

(One of many improvements in MPEG-4 Audio is the Object Type - Long Term Prediction (LTP), which is an improvement of the Main profile using a forward predictor with lower computational complexity.)

Depending on the AAC profile and the MP3 encoder, 96 kbit/s AAC can give nearly the same or better perceptual quality as 128 kbit/s MP3.

AAC error protection toolkit

Applying error protection enables error correction up to a certain extent. Error correcting codes are usually applied equally to the whole payload. However, since different parts of an AAC payload show different sensitivity to transmission errors, this would not be a very efficient approach.

The AAC payload can be subdivided into parts with different error sensitivities.

- Independent error correcting codes can be applied to any of these parts using the Error Protection (EP) tool defined in MPEG-4 Audio standard.

- This toolkit provides the error correcting capability to the most sensitive parts of the payload in order to keep the additional overhead low.
- The toolkit is backwardly compatible with simpler and pre-existing AAC decoders. A great deal of the tool kit's error correction functions are based around spreading information about the audio signal more evenly in the datastream.

Error Resilient (ER) AAC

Error Resilience (ER) techniques can be used to make the coding scheme itself more robust against errors.

For AAC, three custom-tailored methods were developed and defined in MPEG-4 Audio

- **Huffman Codeword Reordering (HCR)** to avoid error propagation within spectral data;
- **Virtual Codebooks (VCB11)** to detect serious errors within spectral data;
- **Reversible Variable Length Code (RVLC)** to reduce error propagation within scale factor data.

AAC Low Delay

The **MPEG-4 Low Delay Audio Coder (AAC-LD)** is designed to combine the advantages of perceptual audio coding with the low delay necessary for two-way communication. It is closely derived from the MPEG-2 Advanced Audio Coding (AAC) format.

Licensing and patents

No licenses or payments are required to be able to stream or distribute content in AAC format. This reason alone makes AAC a much more attractive format to distribute content than MP3, particularly for streaming content (such as Internet radio).

However, a patent license is required for all manufacturers or developers of AAC codecs. It is for this reason FOSS implementations such as FFmpeg and FAAC are distributed in source form only, in order to avoid patent infringement.

Extensions and improvements

Some extensions have been added to the first AAC standard (defined in MPEG-2 Part 7 in 1997):

- **Perceptual Noise Substitution (PNS)**, added in MPEG-4 in 1999. It allows the coding of noise as pseudorandom data;
- **Long Term Predictor (LTP)**, added in MPEG-4 in 1999. It is a forward predictor with lower computational complexity.

- **Error Resilience (ER)**, added in MPEG-4 Audio version 2 in 2000, used for transport over error prone channels;
- **AAC-LD (Low Delay)**, defined in 2000, used for real-time conversation applications;
- **High Efficiency AAC (HE-AAC)**, a.k.a. aacPlus v1 or AAC+, the combination of SBR (Spectral Band Replication) and AAC LC; used for low bitrates; defined in 2003;
- **HE-AAC v2**, a.k.a. aacPlus v2 or eAAC+, the combination of Parametric Stereo (PS) and HE-AAC; used for even lower bitrates; defined in 2004 and 2006;
- **MPEG-4 Scalable To Lossless (SLS)**, defined in 2006, can supplement an AAC stream to provide a lossless decoding option, such as in Fraunhofer IIS's "HD-AAC" product;

Container formats

In addition to the MP4, 3GP and other ISO base media file format-based container formats for storage, AAC audio data may be packaged in a more basic format called Audio Data Interchange Format (ADIF), consisting of a single header followed by the raw AAC audio data blocks. Alternatively, it may be packaged in a streaming format called Audio Data Transport Stream (ADTS), consisting of a series of frames, each frame having a header followed by the AAC audio data. Both formats are defined in MPEG-2 Part 7, but are only considered informative by MPEG-4, so an MPEG-4 decoder does not need to support either format. Two more formats are defined in MPEG-4 Part 3: Low-overhead MPEG-4 Audio Transport Multiplex (LATM), which provides a way to combine separate audio payloads, and Low Overhead Audio Stream (LOAS), a self-synchronizing streaming format.

Products that support AAC

HDTV Standards

Japanese ISDB-T

In December 2003, Japan started broadcasting terrestrial DTV ISDB-T standard that implements MPEG-2 video and MPEG-2 AAC audio. In April 2006 Japan started broadcasting the ISDB-T mobile sub-program, called 1seg, that was the first implementation of video H.264/AVC with audio HE-AAC in Terrestrial HDTV broadcasting service on the planet.

International ISDB-Tb

In December 2007, Brazil started broadcasting terrestrial DTV standard called International ISDB-Tb that implements video coding H.264/AVC with audio AAC-LC on main program(single or multi) and video H.264/AVC with audio HE-AACv2 in the 1seg mobile sub-program.

DVB

The ETSI, the standards governing body for the DVB suite, supports AAC, HE-AAC and HE-AAC v2 audio coding in DVB applications since at least 2004. DVB broadcasts which use the h.264 codec for video normally use the HE-AAC codec for audio.

Hardware

iTunes and iPod

In April 2003, Apple Computer brought mainstream attention to AAC by announcing that its iTunes and iPod products would support songs in MPEG-4 AAC format (via a firmware update for older iPods). Customers could download music in a closed-source Digital Rights Management (DRM)-restricted form of AAC via the iTunes Store or create files without DRM from their own CDs using iTunes. In later years, Apple began offering music videos and movies, which also use AAC for audio encoding.

On May 29, 2007, Apple began selling songs and music videos free of DRM from participating record labels. These files mostly adhere to the AAC standard and are playable on many non-Apple products but they do include custom iTunes information such as album artwork and a purchase receipt, so as to identify the customer in case the file is leaked out onto peer-to-peer networks. It is possible, however, to remove these custom tags to restore interoperability with players that conform strictly to the AAC specification. As of January 6, 2009, nearly all music on the iTunes Store became DRM-free, with the remainder becoming DRM-free by the end of March 2009.

iTunes supports a "Variable bit rate" (VBR) encoding option which encodes AAC tracks in an "Average bit rate" (ABR) scheme. As of September 2009, Apple has added support for HE-AAC (which is fully part of the MP4 standard) but iTunes still lacks support for true VBR encoding. The underlying QuickTime API does offer a true VBR encoding profile however.

Other portable players

- **Archos**
- **Creative Zen Portable**
- **Microsoft Zune**
- **SanDisk Sansa** (some models)
- **Sony PlayStation Portable (PSP)** with firmware 2.0 or greater
- **Sony Walkman**
- **Nintendo DSi**
- **Any portable player that fully supports the Rockbox third party firmware**

Mobile phones

For a number of years, many mobile phones from manufacturers such as Nokia, Motorola, Samsung, Sony Ericsson, BenQ-Siemens and Philips have supported AAC playback. The first such phone was the Nokia 5510 released in 2002 which also plays MP3s. However, this phone was a commercial failure and such phones with integrated music players did not gain mainstream popularity until 2005 when the trend of having AAC as well as MP3 support continued. Most new smartphones and music-themed phones support playback of these formats.

- **Sony Ericsson** phones support various AAC formats in MP4 container. AAC-LC is supported in all phones beginning with K700, phones beginning with W550 have support of HE-AAC. The latest devices such as the P990, K610, W890i and later support HE-AAC v2.
- **Nokia XpressMusic** and other new generation Nokia multimedia phones like N- and E-Series: also support AAC format in LC, HE, M4A and HEv2 profiles
- **BlackBerry**: RIM's latest series of Smartphones such as the 8100 ("Pearl"), 9500 ("Storm") and 8800 support AAC.
- **Apple's iPhone** supports AAC and FairPlay protected AAC files formerly used as the default encoding format in the iTunes store until the removal of DRM restrictions in March 2009.
- The **Motorola Droid** Family supports AAC along with several other audio codecs.
- The HTC Dream (Also known as the T-Mobile G1) is described as supporting certain subset of the full AAC format. As of 2009-04-13 at least several forms of AAC files played while others did not play.

Other devices

- **Apple's iPad**: Supports AAC and FairPlay protected AAC files used as the default encoding format in the iTunes store.
- **Palm OS PDAs**: Many Palm OS based PDAs and smartphones can play AAC and HE-AAC with the 3rd party software Pocket Tunes. Version 4.0, released in December 2006, added support for native AAC and HE-AAC files. The AAC codec for TCPMP, a popular video player, was withdrawn after version 0.66 due to patent issues, but can still be downloaded from sites other than corecodec.org. CorePlayer, the commercial follow-on to TCPMP, includes AAC support. Other PalmOS programs supporting AAC include Kinoma Player and AeroPlayer.
- **Microsoft Windows Mobile** platforms support AAC either by the native Windows Media Player or by third-party products (TCPMP, CorePlayer)
- **Epson** supports AAC playback in the P-2000 and P-4000 Multimedia/Photo Storage Viewers. This support is not available with their older models, however.
- The **Sony Reader** portable eBook plays M4A files containing AAC, and displays metadata created by iTunes. Other Sony products, including the A and E series Network Walkmans, support AAC with firmware updates (released May 2006) while the S series supports it out of the box.

- Nearly every major car stereo manufacturer offers models that will play back .m4a files recorded onto CD in a data format. This includes Pioneer, Sony, Alpine, Kenwood, Clarion, Panasonic, and JVC.
- The **Sonos Digital Media Player** supports playback of AAC files.
- The **Roku SoundBridge** network audio player supports playback of AAC encoded files.
- The **Squeezebox** network audio player (made by Slim Devices, a Logitech company) supports playback of AAC files.
- The *PlayStation 3* supports encoding and decoding of AAC files.
- The *Xbox 360* supports streaming of AAC through the Zune software, and off supported iPods connected through the USB port
- The **Wii** video game console supports AAC files through version 1.1 of the Photo Channel as of December 11, 2007. All AAC profiles and bitrates are supported as long as it is in the.m4a file extension. This update removed MP3 compatibility, but users who have installed this may freely downgrade to the old version if they wish.
- The **Livescribe Pulse and Echo Smartpens** record and store audio in AAC format. The audio files can be replayed using the pen's integrated speaker, attached headphones, or on a computer using the Livescribe Desktop software. The AAC files are stored in the user's "My Documents" folder of the Windows OS and can be distributed and played without specialized hardware or software from Livescribe.

Software

Flash Player

A very common program supporting AAC playback is Flash player, version 9, update 3 and above. Since flash player is also a browser plugin, it can play AAC files through a browser as well.

The Rockbox open source firmware (available for multiple portable players) also offers support for AAC to varying degrees, depending on the model of player and the AAC profile.

Optional iPod support (playback of unprotected AAC files) for the Xbox 360 is available as a free download from Xbox Live.

Other software media players

Almost all current computer media players include built-in decoders for AAC, or can utilize a library to decode it. On Microsoft Windows, DirectShow can be utilized this way with the corresponding filters to enable AAC playback in any DirectShow based player. Mac OS X supports AAC via the QuickTime libraries. Software player applications of particular note include:

- 3ivx MPEG-4 - A suite of DirectShow and QuickTime plugins which support AAC encoding and/or AAC/ HE-AAC decoding in any DirectShow application
- Audio Transcoder - CD Ripper, audio converter, tag editor for Windows, allows to convert from/to AAC.
- CorePlayer also supports LC and HE AAC.
- Easy CD-DA Extractor for Windows, CD Ripper and audio converter, which includes an AAC encoder that supports LC and HE AAC.
- ffdshow is a free open source DirectShow filter for Microsoft Windows operating systems that uses FAAD2 to support AAC decoding.
- foobar2000 is a freeware audio player for Windows that supports LC and HE AAC.
- K-Multimedia Player also supports AAC.
- KSP Sound Player also supports AAC.
- Media Player Classic Home Cinema
- MPlayer or xine are often used as AAC decoders on Linux or Apple Macintosh.
- RealPlayer includes RealNetworks' RealAudio 10 AAC encoder.
- Songbird for Windows, Linux and Mac OS X supports AAC, including the DRM rights management encoding used for purchased music from the iTunes Store, with a plug-in.
- Sony SonicStage also support AAC.
- VLC media player supports playback and encoding of MP4 and AAC files.
- Winamp for Windows, which includes an AAC encoder that supports LC and HE AAC;
- Windows Media Player 12, released with Windows 7, supports playback of AAC files natively.
- Another Real product, Rhapsody supports the RealAudio AAC codec, in addition to offering subscription tracks encoded with AAC.
- XBMC (**X**Box **M**edia **C**enter) supports both AAC (LC and HE) on modified Xbox game-consoles.
- XMMS supports mp4 playback using a plugin provided by the faad2 library.

Some of these players (e.g., foobar2000, Winamp, and VLC) also support the decoding of ADTS (Audio Data Transport Stream) or MP4-contained AAC streamed over HTTP using the SHOUTcast protocol. Plug-ins for Winamp and foobar2000 enable the creation of such streams.

Nero Digital Audio

In May 2006, Nero AG released an AAC encoding tool free of charge, *Nero Digital Audio* (Nero AAC Codec), which is capable of encoding LC-AAC, HE-AAC and HE-AAC v2 streams. The tool is a Command Line Interface tool only, and a separate utility is included to decode to PCM WAV.

Various tools including the foobar2000 audio player, MediaCoder, MeGUI encoding front end and dBpoweramp can provide a GUI for the encoder.

FAAC and FAAD2

FAAC and FAAD2 stand for Freeware Advanced Audio Coder and Decoder 2 respectively. FAAC supports audio object types LC, Main and LTP. FAAD2 supports audio object types LC, Main, LTP, SBR and PS. Although FAAD2 is free software, FAAC is not free software.

FFmpeg

FFmpeg's libavcodec library contains free software codecs for both encoding and decoding AAC (decoding is experimental).

WWT

Chapter- 4

DAISY Digital Talking Book

Digital Accessible Information System, or DAISY, is a means of creating digital talking books for people who wish to hear—and navigate—written material presented in an audible format; many such listeners have "print disabilities," including blindness, impaired vision, dyslexia or other issues.

Using DAISY, a talking book format is presented with enabled navigation within a sequential and hierarchical structure consisting of (marked-up) text synchronized with audio.

DAISY 2 was a standard based on XHTML and SMIL . DAISY 3 is also based on XML and is standardized as ANSI/NISO Z39.86-2005.

DAISY assists people who, for different reasons, have problems using regular printed media. DAISY books have the benefits of regular audiobooks, but they are superior because DAISY 2.02 provides up to six embedded "navigation levels" for content (i.e. other objects such as images, graphics, MathML etc.) and for displaying synchronized text to speech. DAISY Multimedia can be a talking book, computerised text or a synchronised presentation of text and audio.

As a result, DAISY books allow the blind listener to navigate an encyclopedia; this is impossible using conventional audio recordings because they lack search and navigation features and they require linear listening . While reading a DAISY book, a reader can go to the next or previous page, chapter or sentence. DAISY is for everyone who needs accessible information and for everyone who loves to read.

The DAISY Consortium has been selected by the National Information Standards Organization (NISO) as the official maintenance agency for the DAISY/NISO Standard, officially, the ANSI/NISO Z39.86, Specifications for the Digital Talking Book, known as DAISY 3.

The DAISY Consortium was founded in 1996 and consists of a growing membership of organizations around the world committed to developing equitable access to information for people who have a print disability.

Specification

A DTB is a collection of electronic files arranged to present information to the target population via alternative media, namely, human or synthetic speech, refreshable Braille, or visual display, e.g., large print. The files comprising the Daisy formats are:

- **Package File:** Drawn from the Open eBook Publication Structure 1.2. It is a XML1.0 file with a set of metadata describing the DTB, a list of files that make up the DTB (the manifest) and a spine that defines the default reading order of the document. Standard: OEBF * Publication Structure 1.2 (the file extension is “.opf”).
- **Textual content file:** this document contains the text of the document as an XML1.0 according to a specific DTD (dtbook.dtd).
- **Audio Files:** human or synthetic speech recordings.
- **Image files:** for visual displays
- **Synchronization files:** To synchronize the different media files of a DTB during playback, this standard specifies the use of the SMIL2.0.
- **Navigation control file:** use the NCX (Navigation Control File for XML applications) to view the document’s hierarchical structure, allowing user to move through the book’s larger division or progressively smaller steps (footnotes, for example).
- **Bookmark/Highlight file:** support to user-set highlights to which text and audio notes can be applied.
- **Resource file:**
- **Distribution Information File:** describes how to map each SMIL file to a specific media unit.

Distribution and playback

DAISY books can be heard on standalone DAISY players, computers using DAISY playback software, mobile phones, and MP3 players (with limited navigation). DAISY books can be distributed on a CD/DVD, memory card or through the Internet.

A computerized text DAISY book can be read using refreshable Braille display or screen-reading software, printed as Braille book on paper, converted to a talking book using synthesised voice or a human narration, and also printed on paper as large print book. In addition, it can be read as large print text on computer screen.

Software players

Software-based players include:

- Read:OutLoud 6
- Read:OutLoud Bookshare Edition
- gh PLAYER, a commercial DAISY player
- AMIS - Adaptive Multimedia Information System - an open-source self-voicing player for Windows that works with several screen readers; developed by the DAISY Consortium
- FSReaderDAISY Player Software for PAC Mate and Desktop
- emerson-reader, an open-source and cross-platform Epub and DAISY player
- AnyDaisy Firefox Extension, by Benetech
- Dorina DAISY Reader (DDReader), a Firefox add-on for DAISY 3.0
- ButtercupReader, a web-based application for DAISY 3 books
- CUCAT Olearia, an open-source DAISY reader for Mac OS X
- Daisy Delight (DAISY 2.02, for Mac OS X and Unix-based systems)
- DAISY Book Reader, GTK Daisy Talking Book reader application
- DAISYPlayer (for Windows; only available in Spanish)
- Daisy Player, an OSS/GPL CLI Linux/Ubuntu Daisy Talking Book reader application
- DaisyWorm (for iPhone, iPod touch and iPad; iOS 3.1 or higher)
- Dolphin EasyReader and EasyReader Express, for Windows

Other relevant software includes:

- Daisy Uppsala Archive Project, server-side system for managing DAISY (Digital Talking Books) files
- Online Daisy Delivery Technology, open-source software to deliver DAISY books online

Hardware players

There are also a wide range of hardware products available that can play DAISY content, usually in a portable form factor. Some of these devices are dedicated to playback of books, while others focus on other functionality, such as PDA or mobile Internet access, and offer DAISY playback as either a feature of the unit or as a software add-on.

A short (incomplete) list of products that have built-in support for DAISY playback includes:

- Victor Reader Stream, a hand-held portable DAISY player for the blind, visually handicapped and print impaired, produced by HumanWare
- Victor Reader Wave, also by HumanWare, is a portable CD player that can play DAISY content from CD media
- BookSense, a similar, smaller unit produced by GW Micro; the advanced XT model features built-in flash memory and Bluetooth headset support for playback, as well as an FM radio
- The National Library Service for the Blind and Physically Handicapped (NLS) in the United States has developed a proprietary DAISY player designed for use by

its print-disabled patrons. The player will replace the aging cassette-based distribution system.

Access to materials by the disabled

Since DAISY is often used by people with disabilities, many of the existing organizations which produce accessible versions of copyrighted content are moving to the DAISY standard, and slowly moving away from more traditional methods of distribution such as cassette tape.

In the United States, Recording for the Blind and Dyslexic (RFB&D), BookShare and the National Library Service for the Blind and Physically Handicapped (NLS), among others, are offering content to blind and visually impaired individuals. RFB&D also allows access by those with dyslexia or other disabilities which impair the person's ability to read print. The NLS uses a library methodology, on the basis that the books are loaned (as they traditionally have been, on physical cassette), hence they are able to offer content free of charge, just as any public library can. RFB&D and Bookshare both are subscription-based services, however RFB&D has made their membership free of charge to all qualified individuals in the United States

Content from both the NLS and the RFB&D organizations utilizes proprietary encryption extensions to the DAISY standard. The basic structure of the DAISY definition files remains the same, however, the audio itself, and in some cases certain information tags in the DAISY SMIL files, are encrypted and must be decrypted in order to be read/played back. This is done to prevent unauthorized individuals, such as those who do not have a qualifying disability, from accessing the materials. The organization which offers the content provides a decryption key to the user, which can be installed into a DAISY player to allow decryption. As the encryption schemes are not part of the core DAISY standard, only players which specifically implement the necessary algorithms and key management will be able to access these titles. Bookshare does not use such encryption; the data as downloaded from the server is encrypted using only the user's password, which is used to decrypt the data to an open format, and thus content from Bookshare generally is readable on any compliant DAISY player.

Production

Add-ins or extensions to create DAISY files from office software are also available:

- Microsoft and Sonata Software created a Save as DAISY add-in for Microsoft Word to convert Office Open XML text documents to DAISY XML.
- Odt2DAISY is an extension for OpenOffice.org that exports OpenDocument Text to DAISY XML or to Full DAISY (both XML and audio).

Other tools for DAISY production include:

- the DAISY Pipeline, a cross-platform "open source framework for document- and DTB-related pipelined transformations", developed by the DAISY Consortium ,
- the DAISY Pipeline GUI ,
- PipeOnline, a web interface for the DAISY Pipeline ,
- Daisy Producer, an integrated production management system for Digital Talking Books based on the DAISY Pipeline and liblouis
- Z39.86 DTB Validator, "Zedval": "a Java-based conformance validator for ANSI/NISO Z39.86 Digital Talking Books"
- Dolphin Publisher and Dolphin EasyProducer (commercial products)
- Obi, a free authoring tool

WWT

Chapter- 5

Audio Codecs

G.719

G.719 is an ITU-T standard audio codec providing high quality, moderate bit rate (32 to 128 kbit/s) wideband (20 Hz - 20 kHz audio bandwidth, 48 kHz audio sample rate) audio coding at low computational load. It was produced through a collaboration between Polycom and Ericsson.

G.719 incorporates elements of Polycom's Siren22 codec (22 kHz) and Ericsson codec technology, as well as Polycom's Siren7 and Siren14 codecs (G.722.1 and G.722.1 Annex C), which have been used in videoconferencing systems for many years. As ITU-T Recommendation G.719, it was approved on June 13, 2008.

G.719 is optimized for both speech and music. It is based on transform coding with adaptive time-resolution, adaptive bit-allocation and low complexity lattice vector quantization. The computational complexity is quite low (18 floating-point MIPS) for an efficient high-quality compressor. The codec operates on 20 ms frames, and the algorithmic delay end-to-end is 40 ms. The encoder input and decoder output are sampled at 48 kHz.

In addition to the nominal bit rates of 32, 48 and 64 kbit/s, the G.719 codec has an inherent feature of flexible rate selection. In fact, it is possible to accommodate any rate between 32 kbit/s and 64 kbit/s by steps of 4 kbit/s. Moreover, the codec can also provide higher rates than 64 kbit/s and up to 128 kbit/s.

Amendment 1 of the ITU-T G.719 specification defined the use of the ISO base media file format (ISO/IEC 14496-12 a.k.a. MPEG-4 Part 12) as container for the G.719 bitstream. It also defined stereo and multichannel use of G.719 bitstreams in the ISO base media file format. It addresses non-conversational use cases of the codec (e.g. call waiting music playback and recording of teleconferencing sessions, voice mail messages). Thus, media file formats such as MP4 (audio/mp4 or video/mp4) and 3GP (audio/3GPP and video/3GPP) can contain G.719-encoded audio.

RFC 5404 defined media type audio/G719.

Licensing

G.719 is licensed by Polycom, Inc. and by Ericsson; both licenses are necessary for use. Polycom licensees also receive the right to use G.722.1 (Siren7, Polycom's 7 kHz codec), and G.722.1 Annex C (Siren14, the 14 kHz equivalent).

G.722

G.722 is a ITU-T standard 7 kHz wideband speech codec operating at 48, 56 and 64 kbit/s. It was approved by ITU-T in November 1988. Technology of the codec is based on sub-band ADPCM (SB-ADPCM).

G.722 sample audio data at a rate of 16 kHz (using 14 bits), double that of traditional telephony interfaces, which results in superior audio quality and clarity.

Other ITU-T 7 kHz wideband codecs include G.722.1 and G.722.2. These codecs are not variants of G.722 and they use different patented compression technologies. G.722.1 is based on Siren codecs and offers lower bit-rate compressions. A more recent G.722.2, also known as AMR-WB ("Adaptive Multirate Wideband") is based on ACELP and offers even lower bit-rate compressions, as well as the ability to quickly adapt to varying compressions as the network topography mutates. In the latter case, bandwidth is automatically conserved when network congestion is high. When congestion returns to a normal level, a lower-compression, higher-quality bitrate is restored.

Applications

G.722 is an ITU standard codec that provides 7 kHz wideband audio at data rates from 48, 56 and 64 kbit/s. This is useful for voice over IP applications, such as on a local area network where network bandwidth is readily available, and offers a significant improvement in speech quality over older narrowband codecs such as G.711, without an excessive increase in implementation complexity. Environments where bandwidth is more constrained may prefer one of the more bit-efficient codecs, such as G.722.1 (Siren7) or G.722.2 (AMR-WB).

G.722 is also widely used by broadcasters for sending commentary grade audio over a single 64kbps ISDN B channel.

RTP Encapsulation

G.722 VoIP is typically carried in RTP payload type 9. Note that IANA records the clock rate for type 9 G.722 as 8 kHz (instead of 16 kHz), RFC3551 clarifies that this is due to a historical error and is retained in order to maintain backward compatibility. Consequently

correct implementations represent the value 8,000 where required but encode and decode audio at 16 kHz.

Whilst G.722 allows for bitrates of 64, 56 and 48 kbit/s, in practice, data is encoded at 64 kbit/s, with bits from the lower sub-band being used to encode auxiliary data. The greater the number of bits allocated to aux data, the lower the bit rate.

G.722.1

G.722.1 is a licensed royalty-free ITU-T standard audio codec providing high quality, moderate bit rate (24 and 32 kbit/s) wideband (50 Hz - 7 kHz audio bandwidth, 16 ksamples per second) audio coding. It is a partial implementation of Siren 7 codec (which offers bit rates 16, 24, 32 kbit/s) developed by PictureTel Corp. (now Polycom, Inc.). Its official name is *Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss*.

G.722.1 Annex C (or G.722.1C) is a low-complexity extension mode to G.722.1, which doubles the algorithm to permit 14 kHz audio bandwidth using a 32 kHz audio sample rate, at 24, 32, and 48 kbit/s. It is included in the official ITU-T Recommendation G.722.1. The name of this annex is *Annex C - 14 kHz mode at 24, 32, and 48 kbit/s*. It is an implementation of the mono version of Polycom's Siren 14 codec.

G.722.1 is the successor to PT716plus developed by PictureTel Corp. (now Polycom, Inc.), which has been used in videoconferencing systems for many years. As ITU-T Recommendation G.722.1, it was approved on September 30, 1999 after a four-year selection process involving extensive testing. G.722.1/Annex C was approved by ITU-T on May 14, 2005.

G.722.1 is a transform-based compressor that is optimized for both speech and music. The G.722.1 algorithm is based on transform technology, using a Modulated Lapped Transform (MLT). The computational complexity is quite low (5.5 floating-point MIPS) for an efficient high-quality compressor, and the algorithmic delay end-to-end is 40 ms. A 14 kHz (32 ksamples) extension, G.722.1/Annex C, was approved by ITU-T on May 14, 2005. Also known as the mono version of Siren 14, this extension is also available from Polycom as a royalty-free license.

The numbering of the wideband ITU audio codecs is sometimes confusing. There are three principal codecs, which are unrelated, but all carrying the G.722 label. G.722 is the original 7 kHz codec, using ADPCM and operating at 48 – 64 kbit/s. G.722.1, another 7 kHz codec, operates at half the data rate while delivering comparable or better quality as G.722, but is a transform-based codec. G.722.1 Annex C is very similar to G.722.1, but provides twice the audio bandwidth, 14kHz. And G.722.2, which operates on wideband speech and delivers very low bitrates, is an ACELP-based algorithm.

G.723

G.723 is a ITU-T standard speech codec using extensions of G.721 providing voice quality covering 300 Hz to 3400 Hz using Adaptive Differential Pulse Code Modulation (ADPCM) to 24 and 40 kbit/s for digital circuit multiplication equipment (DCME) applications. The standard G.723 is **obsolete** and has been superseded by G.726.

Note that this is a completely different codec from G.723.1.

G.723.1

G.723.1 is an audio codec for voice that compresses voice audio in 30 ms frames. An algorithmic look-ahead of 7.5 ms duration means that total algorithmic delay is 37.5 ms. Its official name is *Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s*.

Note that this is a completely different codec from G.723.

There are two bit rates at which G.723.1 can operate:

- 6.3 kbit/s (using 24 byte frames) using a MPC-MLQ algorithm (MOS 3.9)
- 5.3 kbit/s (using 20 byte frames) using an ACELP algorithm (MOS 3.62)

Use

G.723.1 is mostly used in Voice over IP (VoIP) applications due to its low bandwidth requirement. Music or tones such as DTMF or fax tones cannot be transported reliably with this codec, and thus some other method such as G.711 or out-of-band methods should be used to transport these signals. The complexity of the algorithm is below 16 MIPS. 2.2 kilobytes of RAM is needed for codebooks.

G.723.1 is a required audio codec in the H.324 ITU-T recommendation for H.324 terminals offering audio communication. In 3GPP 3G-324M specification support for G.723.1 is not mandatory, but recommended.

Features

- Sampling frequency 8 kHz/16-bit (240 samples for 30 ms frames)
- Fixed bit rate (5.3 kbit/s with 20 byte 30 ms frames, 6.3 kbit/s with 24 byte 30 ms frames)
- Fixed frame size for each rate (5.3 kbit/s with 20 byte 30 ms frames, 6.3 kbit/s with 24 byte 30 ms frames)
- Algorithmic delay is 37.5 ms per frame, with 7,5 ms look-ahead delay

- G.723 is a hybrid speech coder, with high bit rate using Multi-Pulse Maximum Likelihood Quantization (MP-MLQ) and low bit rate using Algebraic Code Excited Linear Prediction (ACELP)
- The complexity of the algorithm is rated at 25, using a relative scale where G.711 is 1 and G.729a is 15.
- G.723.1 Annex A defines 4 byte Silence Insertion Descriptor (SID) frame for Comfort Noise Generation
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.08 for G.723.1 (6.3 kbit/s), compared to 4.45 for G.711 (μ -law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.57 for G.723.1 (6.3 kbit/s), compared to 4.13 for G.711 (μ -law)

G.726

G.726 is an ITU-T ADPCM speech codec standard covering the transmission of voice at rates of 16, 24, 32, and 40 kbit/s. It was introduced to supersede both G.721, which covered ADPCM at 32 kbit/s, and G.723, which described ADPCM for 24 and 40 kbit/s. G.726 also introduced a new 16 kbit/s rate. The four bit rates associated with G.726 are often referred to by the bit size of a sample, which are 2-bits, 3-bits, 4-bits, and 5-bits respectively.

The most commonly used mode is 32 kbit/s, which doubles the usable network capacity by using half the rate of G.711. It is primarily used on international trunks in the phone network. The principal application of 24 and 16 kbit/s channels is for overload channels carrying voice in digital circuit multiplication equipment (DCME). The principal application of 40 kbit/s channels is to carry data modem signals in DCME, especially for modems operating at greater than 4800 kbit/s.

It also is the standard codec used in DECT wireless phone systems and is used on some Canon cameras.

History

G.721 was introduced in 1984, while G.723 was introduced in 1988. They were folded into G.726 in 1990.

G.727 was introduced at the same time as G.726, and includes the same bit rates, but is optimized for packet circuit multiplex equipment (PCME) environment. This is achieved by embedding 2-bit quantizer to 3-bit quantizer and same for the higher modes. This allows dropping of the least significant bit from the bit stream without adverse effects on speech signal.

Features

- Sampling frequency 8 kHz
- 16 kbit/s, 24 kbit/s, 32 kbit/s, 40 kbit/s bit rates available
- Generates a bitstream, therefore frame length is determined by packetization (typically 80 samples for 10 ms frame size)
- Typical algorithmic delay is 0.125 ms, with no look-ahead delay
- G.726 is a waveform speech coder which uses Adaptive Differential Pulse Code Modulation (ADPCM)
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.30 for G.726 (32 kbit/s), compared to 4.45 for G.711 (μ -law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.79 for G.726 (32 kbit/s), compared to 4.13 for G.711 (μ -law)
- 40 kbit/s G.726 can carry 12000 bit/s and slower modem signals, while 32 kbit/s G.726 can carry 2400 bit/s and slower modem signals well and 4800 bit/s with some more degradation than clear channel codecs.

G.728

G.728 is an ITU-T standard for speech coding operating at 16 kbit/s. It is officially described as *Coding of speech at 16 kbit/s using low-delay code excited linear prediction*.

Technology used is LD-CELP, low-delay code excited linear prediction. Delay of the codec is only 5 samples (0.625 ms). The linear prediction is calculated backwards with a 50th order linear predictive coding filter. The excitation is generated with gain scaled VQ. The standard was finished in 1992 in the form of algorithm exact floating point code. In 1994 a bit exact fixed point codec was released. G.728 passes low bit rate modem signals up to 2400 bit/s. Also network signaling goes through. The complexity of the codec is 30 MIPS. 2 kilobytes of RAM is needed for codebooks. Mean opinion score for G.728 is 3.61.

The essence of CELP techniques, which is an analysis-by-synthesis approach to codebook search, is retained in LD-CELP. The LD-CELP however, uses backward adaptation of predictors and gain to achieve an algorithmic delay of 0.625 ms.

RealAudio 28_8 is a reduced-bitrate variant of this standard, using 15.2 kbit/s.

G.729

G.729 is an audio data compression algorithm for voice that compresses digital voice in packets of 10 milliseconds duration. It is officially described as *Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)*.

Because of its low bandwidth requirements, G.729 is mostly used in Voice over Internet Protocol (VoIP) applications (such as Skype) where bandwidth must be conserved. Standard G.729 operates at a bit rate of 8 kbit/s, but there are extensions, which provide

rates of 6.4 kbit/s (Annex D, F, H, I, C+) and 11.8 kbit/s (Annex E, G, H, I, C+) for worse and better speech quality, respectively.

G.729 has been extended with various features, commonly designated as G.729a and G.729b.

DTMF tones, Fax transmissions, and high-quality audio cannot be transported reliably with this codec. DTMF requires the use of the RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals as specified in RFC 2833.

G.729 Annexes

Functionality	G.729 Annexes											
	-	A	B	C	D	E	F	G	H	I	C+	J
Low complexity		X	X									
Fixed-point	X	X	X		X	X	X	X	X	X		X
Floating-point				X							X	
8 kbit/s	X	X	X	X	X	X	X	X	X	X	X	X
6.4 kbit/s					X		X		X	X	X	
11.8 kbit/s						X		X	X	X	X	
DTX			X				X	X		X	X	
Embedded variable bit rate, wideband												X

G.729 Annex A

G.729a is a compatible extension of G.729, but requires less computational power. This lower complexity, however, bears the cost of marginally reduced speech quality.

G.729a was developed by a consortium of organizations: France Telecom, Mitsubishi Electric Corporation, Nippon Telegraph and Telephone Corporation (NTT), and Université de Sherbrooke.

The features of G.729a are:

- Sampling frequency 8 kHz/16-bit (80 samples for 10 ms frames)
- Fixed bit rate (8 kbit/s 10 ms frames)
- Fixed frame size (10 bytes for 10 ms frame)
- Algorithmic delay is 15 ms per frame, with 5 ms look-ahead delay
- G.729a is a hybrid speech coder which uses Algebraic Code Excited Linear Prediction (ACELP)

- The complexity of the algorithm is rated at 15, using a relative scale where G.711 is 1 and G.723.1 is 25.
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.04 for G.729a, compared to 4.45 for G.711 (u-law)
- PSQM testing under network stress yields Mean Opinion Scores of 3.51 for G.729a, compared to 4.13 for G.711 (u-law)

G.729 Annex B

G.729 has been extended in Annex B (G.729b) which provides a silence compression method that enables a voice activity detection (VAD) module. It is used to detect voice activity in the signal. It also includes a discontinuous transmission (DTX) module which decides on updating the background noise parameters for non speech (noisy frames). It uses 2-byte Silence Insertion Descriptor (SID) frames transmitted to initiate comfort noise generation (CNG). If transmission is stopped, and the link goes quiet because of no speech, the receiving side may assume that the link has been cut. By inserting comfort noise, analog hiss is simulated digitally during silence to assure the receiver that the link is active and operational.

Other extensions

Recently, G.729 has been extended (with Annex J) to provide support for wideband speech and audio coding, i.e., the transmitted acoustic frequency range is extended to 50 Hz - 7 kHz. The respective extension to G.729 is referred to as G.729.1. The G.729.1 coder is hierarchically organized: Its bit rate and the obtained quality are adjustable by simple bitstream truncation.

G.729.1

G.729.1 is an 8-32 kbit/s embedded speech and audio codec providing bitstream interoperability with G.729, G.729 Annex A and G.729 Annex B. Its official name is *G.729-based embedded variable bit rate codec: An 8-32 kbit/s scalable wideband coder bitstream interoperable with G.729.*

This codec has been designed to provide better quality and more flexibility than the existing ITU-T G.729 speech coding standard. G.729.1 is scalable in bit rate, acoustic bandwidth and complexity. In addition it offers various encoder and decoder modes, including the support of both 8 and 16 kHz input/output sampling frequency, compatibility with G.729B, and reduced algorithmic delay. The bitstream of G.729.1 is structured into 12 hierarchical layers. The first layer (or core layer) at 8 kbit/s follows the G.729 format. The second layer (adds 4 kbit/s for a total of 12 kbit/s) is a narrowband enhancement layer. The third layer (2 kbit/s for a total of 14 kbit/s) is a bandwidth extension layer. Further layers (in 2 kbit/s steps) are wideband enhancement layers. The

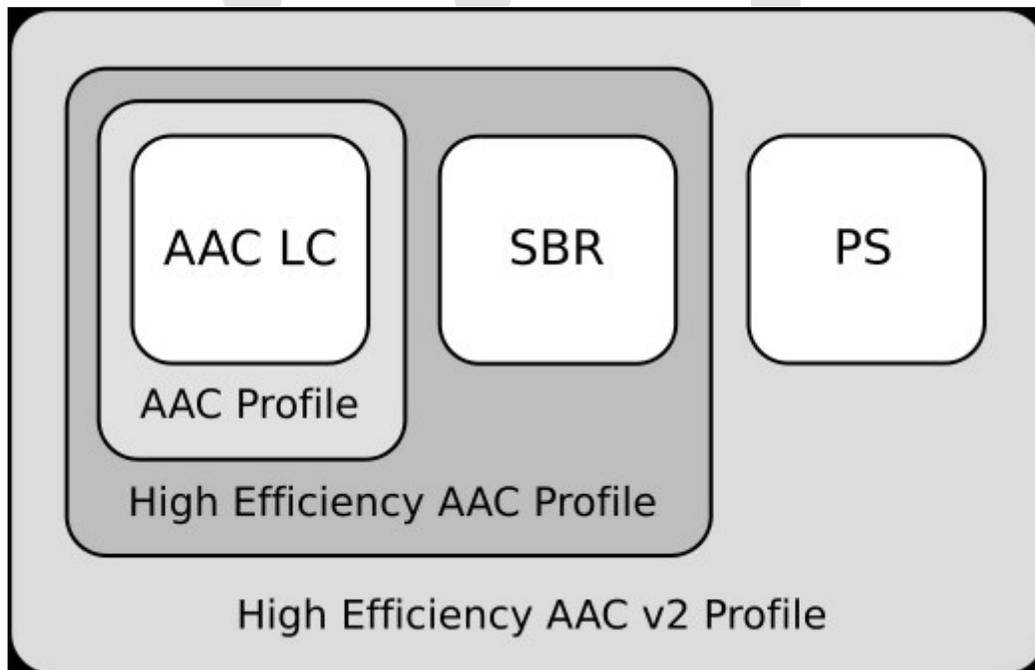
G.729.1 output bandwidth is 50-4000 Hz at 8 and 12 kbit/s, and 50-7000 Hz from 14 to 32 kbit/s. G.729.1 is also known as G.729 Annex J and G.729EV where EV stands for Embedded Variable (bit rate).

The G.729.1 algorithm is based on a three-stage coding structure: embedded Code-excited linear prediction (CELP) coding of the lower band (50-4000 Hz), parametric coding of the higher band (4000-7000 Hz) by Time-Domain Bandwidth Extension (TDBWE), and enhancement of the full band (50-7000 Hz) by a predictive transform coding technique referred to as Time-Domain Aliasing Cancellation (TDAC).

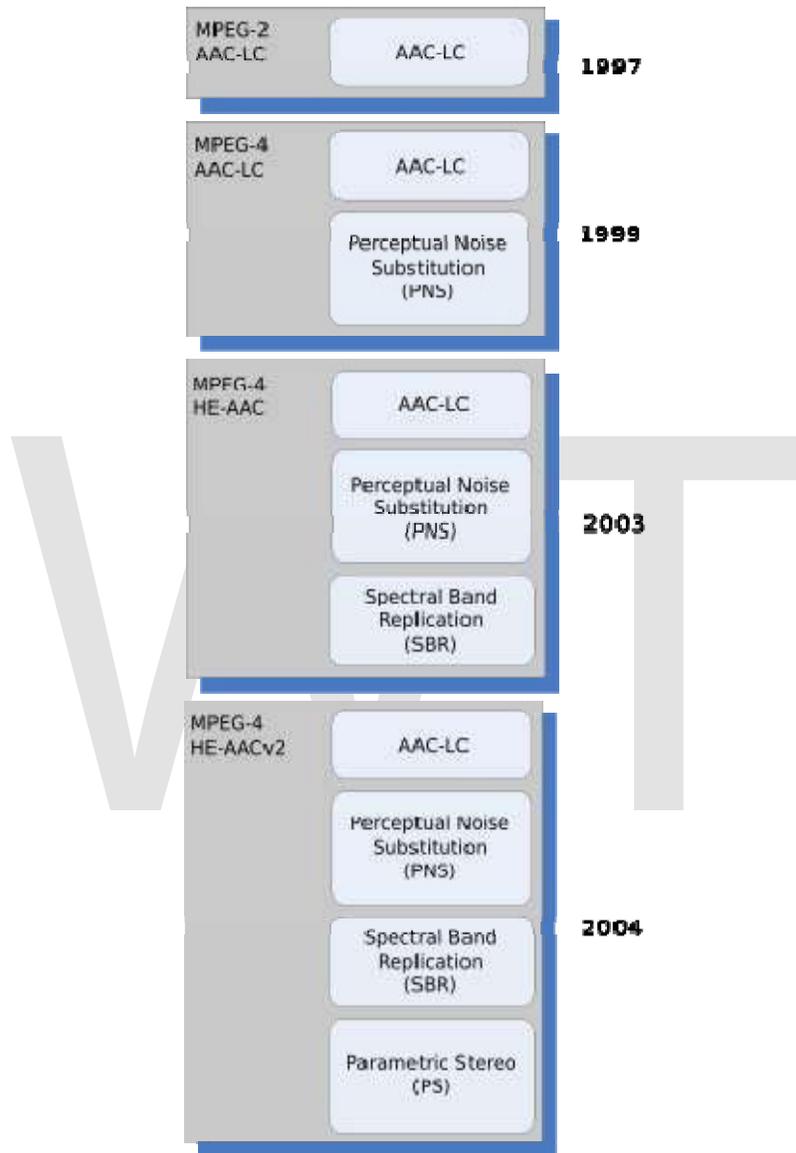
High-Efficiency Advanced Audio Coding

High-Efficiency Advanced Audio Coding

Internet media type	audio/aacp, audio/3gpp, audio/3gpp2
Developed by	ISO
Type of format	Audio compression format
Contained by	3GP, MP4, .dvb, etc.
Extended from	AAC
Standard(s)	ISO/IEC 14496-3



Hierarchical structure of AAC Profile, HE-AAC Profile and HE-AAC v2 Profile, and compatibility between them. The HE-AAC Profile decoder is fully capable of decoding any AAC Profile stream. Similarly the HE-AAC v2 decoder can handle all HE-AAC Profile streams as well as all AAC Profile streams. Based on the MPEG-4 Part 3 technical specification.



Evolution from MPEG-2 AAC LC (Low Complexity) Profile and MPEG-4 AAC LC Object Type to HE-AACv2 Profile.

High-Efficiency Advanced Audio Coding (HE-AAC) is a lossy data compression scheme for digital audio defined as a MPEG-4 Audio profile in ISO/IEC 14496-3. It is an extension of Low Complexity AAC (AAC LC) optimized for low-bitrate applications such as streaming audio. HE-AAC version 1 profile (HE-AAC v1) uses spectral band replication (SBR) to enhance the compression efficiency in the frequency domain. HE-

AAC version 2 profile (HE-AAC v2) couples SBR with Parametric Stereo (PS) to enhance the compression efficiency of stereo signals. It is a standardized and improved version of the AACplus codec.

HE-AAC is used in digital radio standards like DAB+ and Digital Radio Mondiale.

History

The HE-AAC version 1 was standardized as a profile of MPEG-4 Audio in 2003 by MPEG and published as part of MPEG-4 in document ISO/IEC 14496-3:2001/Amd 1:2003. The HE-AAC version 2 profile was standardized in 2006 as ISO/IEC 14496-3:2005/Amd 2:2006. A coding tool (Parametric Stereo) used in HE-AAC v2 was standardized in 2004 by MPEG and published in document ISO/IEC 14496-3:2001/Amd 2:2004.

HE-AAC version 2 was also standardized under the name Enhanced aacPlus by 3GPP for 3G UMTS multimedia services in September 2004 (3GPP TS 26.401). It is based on the AAC LC, SBR and Parametric Stereo coding tools defined in the MPEG-4 Audio standard (and it refers to ISO/IEC 14496-3:2001/Amd.1:2003 - Bandwidth Extension, ISO/IEC 14496-3:2001/Amd.2:2004 - Parametric Coding for High Quality Audio, ISO/IEC 14496-3:2001 and ISO/IEC 14496-3:2001/Amd.1:2003/DCOR1). In addition it includes further tools such as error concealment, spline resampler, and stereo-to-mono downmix. It was also published by ETSI as TS 126 401 V6.1.0 in December 2004.

The progenitor of HE-AAC was developed by Coding Technologies under the trade name CT-aacPlus. CT-aacPlus combined MPEG-2 AAC LC with the Coding Technologies invented spectral band replication (SBR). CT-aacPlus is the codec used by XM Radio for their satellite radio service. Subsequently, Coding Technologies submitted their SBR to MPEG as a basis of HE-AAC.

Later, but prior to the standardization of HE AAC v2 by MPEG, Coding Technologies submitted the combination of HE-AAC v1 coupled with Parametric Stereo to 3GPP under the name Enhanced AAC+. As a result, **aacPlus v2** and **eAAC+** are now common trade names that refer to HE-AAC v2. **AAC+** and **aacPlus** are common trade names that refer to HE-AAC v1.

Perceived quality

Scientific testing by the European Broadcasting Union has indicated that HE-AAC at 48 kbit/s was ranked as "Excellent" quality using the MUSHRA scale. MP3 in the same testing received a score less than half that of HE-AAC and was ranked "Poor" using the MUSHRA scale. Data from this testing also indicated that some individuals confused 48 kbit/s encoded material with an uncompressed original.

Testing indicates that material decoded from 64 kbit/s HE-AAC does not yet have similar audio quality to material decoded from MP3 at 128 kbit/s using high quality encoders. The test, taking bitrate distribution and RMSD into account, is a tie between MP3pro, HE AAC and Ogg Vorbis.

Further controlled testing by 3GPP during their revision 6 specification process indicates that HE-AAC and its derivative MPEG-4 HE-AAC v2 provide "Good" audio quality for music at low bit rates (e.g. 24 kbit/s).

MPEG-2 and MPEG-4 AAC LC decoders without SBR support will decode the AAC LC part of the audio, resulting in audio output with only half the sampling frequency, thereby reducing the audio bandwidth.

Support

Encoding

Sony supports HE-AAC encoding since SonicStage version 4. iTunes 9 supports HE-AAC encoding and playback. Nero has released a free-of-charge command line HE-AAC encoder, and also supports HE-AAC inside the Nero software suite. Sorenson Media's Squeeze Compression Suite includes an HE-AACv1 encoder and is available for Mac OS X as well as Windows. The 3GPP consortium released source code of a reference HE-AACv2 encoder that appears to offer competitive quality. Winamp Pro also supports ripping music to HE-AAC. Using a transcoding plugin for Winamp's media library, any file can be transcoded to HE-AAC. XLD, an OS X audio encoding program, offers encoding from any of its supported formats to HE-AAC. Nokia PC Suite may encode audiofiles to eAAC+ format before transmitting them to mobile phone.

Decoding

HE-AAC is supported in the open source FAAD/FAAD2 decoding library (and all players incorporating it): VLC media player, Winamp, foobar2000, Audacious Media Player, and Sony's latest SonicStage version 4. HE-AAC is also used by AOL Radio clients to deliver high-fidelity music at low bitrates.

Adobe's Flash Player 9 supports HE-AACv2.

iTunes 9.2 and iOS4 include full decoding of HE-AAC v2 parametric stereo streams. iTunes 9 thru 9.1, iPhone OS 3.1 and Fall 2009 iPods have support for HE-AAC playback for version 1 with no parametric stereo. Other third-party software featuring support for HE-AAC include FStream, VLC (without metadata/title streaming), Songbird, Snowtape, and recent development versions of Audacious Media Player.

Older versions of Apple iTunes, iPod Touch, and iPhone will play HE-AAC files at reduced fidelity because they ignore the spectral-band replication and parametric stereo information, instead playing them as though they were standard AAC-LC files without

the high-frequency, or "treble," information that is only present in the SBR part of the signal. These will report the track length as twice its actual length.

Dolby released Dolby Pulse decoders and encoders in September 2008. HE-AAC v2 is the core of Dolby Pulse so files and streams encoded in Dolby Pulse will playback on AAC, HE-AAC v1 and v2 decoders. Conversely files and streams encoded in AAC, HE-AAC v1 or v2 will playback on Dolby Pulse decoders.

Dolby Pulse provides the following additional capabilities beyond HE-AAC v2:

- Ability to intelligently generate and insert reversible loudness normalization and dynamic range metadata into the encoded file/stream, this metadata can then be used to optimize the playback experience based on application and/or device.
- Ability to insert custom metadata into the encoded file, and extract this metadata on playback

Dolby has additionally released a PC decoder as an SDK suitable for integration into PC applications requiring Dolby Pulse, HE-AAC or AAC playback capabilities.

Promotion aspects

Commercial trademarks and labeling

HE-AAC is marketed under the trademark aacPlus(tm) by Coding Technologies and under the trademark Nero Digital(tm) by Nero AG. Sony Ericsson, Nokia and Samsung use AAC+ to label support for HE AAC v1 and eAAC+ to label support for HE-AAC v2 on their phones. Motorola uses AAC+ to indicate HE AAC v1 and "AAC+ Enhanced" to indicate HE AAC v2.

Licensing and patents

Companies holding patents for HE AAC have formed a patent pool administered by Via Licensing Corporation to provide a single point of license for product makers.

Patent licenses are required for end-product companies that make hardware or software products that include HE AAC encoders and/or decoders. Unlike the MP3 format, content owners are not required to pay license fees to distribute content in HE AAC.

Standards

HE-AAC profile was first standardized in ISO/IEC 14496-3:2001/Amd 1:2003. HE-AAC v2 profile (HE-AAC with Parametric Stereo) was first specified in ISO/IEC 14496-3:2005/Amd 2:2006. The Parametric Stereo coding tool used by HE-AAC v2 was standardized in 2004 and published as ISO/IEC 14496-3:2001/Amd 2:2004.

The current version of the MPEG-4 Audio (including HE-AAC standards) is published in ISO/IEC 14496-3:2009.

Enhanced aacPlus is required audio compression format in 3GPP technical specifications for 3G UMTS multimedia services and should be supported in IP Multimedia Subsystem (IMS), Multimedia Messaging Service (MMS), Multimedia Broadcast/Multicast Service (MBMS) and Transparent end-to-end Packet-switched Streaming Service (PSS). HE-AAC version 2 was standardized under the name Enhanced aacPlus by 3GPP for 3G UMTS multimedia services in September 2004 (3GPP TS 26.401).

HE-AAC and HE-AAC v2 audio coding for DVB applications is standardized by TS 101 154. AacPlus v2 by Coding Technologies is also standardized by the ETSI as TS 102 005 for Satellite services to Handheld devices (DVB-SH) below 3 GHz.

In December 2007, Brazil started broadcasting terrestrial DTV standard called International ISDB-Tb that implements video coding H.264 with audio AAC LC on main program (single or multi) and video H.264 with audio HE-AACv2 in the 1Seg mobile sub-program.

Versions

The following is the summary of the different versions of HE-AAC:

Version	Common trade names	Codec feature	Standards
HE-AAC v1	aacPlus v1, eAAC, AAC+, CT-aacPlus	AAC LC + SBR	ISO/IEC 14496-3:2001/Amd 1:2003
HE-AAC v2	aacPlus v2, eAAC+, AAC++, Enhanced AAC+	AAC LC + SBR + PS	ISO/IEC 14496-3:2005/Amd 2:2006

G.711

G.711 is an ITU-T standard for audio companding. It is primarily used in telephony. The standard was released for usage in 1972. Its formal name is *Pulse code modulation (PCM) of voice frequencies*. It is required standard in many technologies, for example in H.320 and H.323 specifications. It can be also used in one of methods for fax communication over IP networks (as defined in T.38 specification).

G.711 represents logarithmic pulse-code modulation (PCM) samples for signals of voice frequencies, sampled at the rate of 8000 samples/second.

G.711.0 (G.711 LLC) - *Lossless compression of G.711 pulse code modulation* was approved by ITU-T in September 2009. It gives as much as 50 percent reduction in bandwidth use.

G.711.1 is an extension to G.711, published as ITU-T Recommendation G.711.1 in March 2008. Its formal name is *Wideband embedded extension for G.711 pulse code modulation*.

G.711, also known as Pulse Code Modulation (PCM), is a very commonly used waveform codec. G.711 uses a sampling rate of 8,000 samples per second, with the tolerance on that rate 50 parts per million (ppm). Non-uniform quantization (logarithmic) with 8 bits is used to represent each sample, resulting in a 64 kbit/s bit rate. There are two slightly different versions; μ -law, which is used primarily in North America, and A-law, which is in use in most other countries outside North America.

Types

G.711 defines two main compression algorithms, the μ -law algorithm (used in North America & Japan) and A-law algorithm (used in Europe and the rest of the world). Both are logarithmic, but A-law was specifically designed to be simpler for a computer to process. The standard also defines a sequence of repeating code values which defines the power level of 0 dB.

The μ -law and A-law algorithms encode 14-bit and 13-bit signed linear PCM samples (respectively) to logarithmic 8-bit samples. Thus, the G.711 encoder will create a 64 kbit/s bitstream for a signal sampled at 8 kHz.

G.711 μ -law tends to give more resolution to higher range signals while G.711 A-law provides more quantization levels at lower signal levels. When using μ -law G.711 in networks where suppression of the all 0 character signal is required, the character signal corresponding to negative input values between decision values numbers 127 and 128 should be 0000010 and the value at the decoder output is -7519. The corresponding decoder output value number is 125.....

G.711 A-Law

A-law encoding thus takes a 13-bit signed linear audio sample as input and converts it to an 8 bit value as follows:

Linear input code Compressed code

```
s0000000wxyz...  s000wxyz  
s0000001wxyz...  s001wxyz  
s000001wxyzab... s010wxyz  
s00001wxyzabc... s011wxyz
```

s0001wxyzabcd... s100wxyz
s001wxyzabcde... s101wxyz
s01wxyzabcdef... s110wxyz
s1wxyzabcdefg... s111wxyz

Where *s* is the sign bit, and the ellipsis represents additional low-order bits that are not encoded. So for example, 1000'0000'1010'1111 maps to 1000'1010 (according to the first row of the table), and 0000'0001'1010'1111 maps to 0001'1010 (according to the second).

This can be seen as a floating point number with 4 bits of mantissa and 3 bits of exponent.

In addition, the standard specifies that all resulting even bits are inverted before the octet is transmitted. This is to provide plenty of 0/1 transitions to facilitate the clock recovery process in the PCM receivers. Thus, a silent A-law encoded PCM channel has the 8 bit samples coded 0x55 instead of 0x00 in the octets (or 0xD5 if the sign bit happens to be set).

Note that the ITU define bit 1 to have the value 128 and bit 8 to have the value 1.

The more widely accepted convention has bit 7 = 128 and bit 0 = 1.

Note that when data is sent over E0 (G.703), MSB (signbit) is sent first and LSB is sent last.

G.711 μ -Law

μ -law encoding takes a 14-bit signed linear audio sample as input, increases the magnitude by 32 (binary 100000), and converts it to an 8 bit value as follows:

Linear input code Compressed code

s00000001wxyza... s000wxyz
s0000001wxyzab... s001wxyz
s000001wxyzabc... s010wxyz
s00001wxyzabcd... s011wxyz
s0001wxyzabcde... s100wxyz
s001wxyzabcdef... s101wxyz
s01wxyzabcdefg... s110wxyz
s1wxyzabcdefgh... s111wxyz

Where *s* is the sign bit, and the ellipsis represents additional low-order bits that are not encoded.

In addition, the standard specifies that all result bits are inverted before the octet is transmitted. Thus, a silent μ -law encoded PCM channel has the 8 bit samples coded 0xFF instead of 0x00 in the octets.

Also the "trick" of adding 32 means μ -law does not encode all 14-bit values; inputs must be within ± 8159 .

Features

- Sampling frequency 8 kHz
- 64 kbit/s bitrate (8 kHz sampling frequency x 8 bits per sample)
- Typical algorithmic delay is 0.125 ms, with no look-ahead delay
- G.711 is a waveform speech coder
- G.711 Appendix I defines a Packet Loss Concealment (PLC) algorithm to help hide transmission losses in a packetized network
- G.711 Appendix II defines a Discontinuous Transmission (DTX) algorithm which uses Voice Activity Detection (VAD) and Comfort Noise Generation (CNG) to reduce bandwidth usage during silence periods
- PSQM testing under ideal conditions yields Mean Opinion Scores of 4.45 for G.711 u-law, 4.45 for G.711 a-law
- PSQM testing under network stress yields Mean Opinion Scores of 4.13 for G.711 u-law, 4.11 for G.711 a-law

G.711.1

A recent extension to G.711, G.711.1, allows the addition of narrowband and/or wideband (16000 samples/s) enhancements, each at 25 % of the bitrate of the (included) base G.711 bitstream, leading to data rates of 64, 80 or 96 kbit/s.

G.711.1 is compatible with G.711 at 64 kbit/s, hence an efficient deployment in existing G.711-based voice over IP (VoIP) infrastructures is foreseen. The G.711.1 coder can encode signals at 16 kHz with a bandwidth of 50–7000 Hz at 80 and 96 kbit/s, and for 8-kHz sampling the output may produce signals with a bandwidth ranging from 50 up to 4000 Hz, operating at 64 and 80 kbit/s.

The G.711.1 encoder creates embedded bitstream structured in three layers corresponding to three available bit rates: 64, 80 and 96 kbit/s. The bitstream does not contain any information on which layers are contained, an implementation would require outband signalling on which layers are available. The three G.711.1 layers are: log companded pulse code modulation (PCM) of the lower band including noise feedback, embedded PCM extension with adaptive bit allocation for enhancing the quality of the base layer in the lower band, and weighted vector quantization coding of the higher band based on modified discrete cosine transformation (MDCT).

Two extensions for G.711.1 are planned in 2010: superwideband extension (bandwidth to 14000 Hz) and lossless bitstream compression.

WWT

Chapter- 6

Digital Rights Management

Digital rights management (DRM) is a term for access control technologies that can be used by hardware manufacturers, publishers, copyright holders and individuals to limit the usage of digital content and devices. The term is used to describe any technology that inhibits uses of digital content not desired or intended by the content provider. The term does not generally refer to other forms of copy protection which can be circumvented without modifying the file or device, such as serial numbers or keyfiles. It can also refer to restrictions associated with specific instances of digital works or devices. Digital rights management is used by companies such as Sony, Amazon, Apple Inc., Microsoft, AOL and the BBC.

The use of digital rights management is controversial. Proponents argue it is needed by copyright holders to prevent unauthorized duplication of their work, either to maintain artistic integrity or to ensure continued revenue streams. Some opponents, such as the Free Software Foundation (through its Defective By Design campaign), maintain that the use of the word "rights" is misleading and suggest that people instead use the term **digital restrictions management**. Their position is essentially that copyright holders are restricting the use of material in ways that are beyond the scope of existing copyright laws, and should not be covered by future laws. The Electronic Frontier Foundation, and other opponents, also consider DRM systems to be anti-competitive practices. This position holds that the rights of the user need legal protection.

Introduction

DRM technologies attempt to control use of digital media by preventing access, copying or conversion to other formats by end users. Long before the arrival of digital or even electronic media, copyright holders, content producers, or other financially or artistically interested parties had business and legal objections to copying technologies. Examples include: player piano rolls early in the 20th century, audio tape recording, and video tape recording (e.g., the "Betamax case" in the U.S.). Copying technology thus exemplifies a disruptive technology.

The advent of digital media and analog/digital conversion technologies, especially those that are usable on mass-market general-purpose personal computers, has vastly increased the concerns of copyright-dependent individuals and organizations, especially within the music and movie industries, because these individuals and organizations are partly or wholly dependent on the revenue generated from such works. While analog media inevitably loses quality with each copy generation, and in some cases even during normal use, digital media files may be duplicated an unlimited number of times with no degradation in the quality of subsequent copies. The advent of personal computers as household appliances has made it convenient for consumers to convert media (which may or may not be copyrighted) originally in a physical/analog form or a broadcast form into a universal, digital form (this process is called ripping) for location- or timeshifting. This, combined with the Internet and popular file sharing tools, has made unauthorized distribution of copies of copyrighted digital media (digital piracy) much easier.

DRM technologies have enabled publishers to enforce access policies that not only disallow copyright infringements, but also prevent lawful fair use of copyrighted works, or even implement use constraints on non-copyrighted works that they distribute; examples include the placement of DRM on certain public-domain or open-licensed e-books, or DRM included in consumer electronic devices that time-shift (and apply DRM to) both copyrighted and non-copyrighted works.

DRM is most commonly used by the entertainment industry (e.g., film and recording). Many online music stores, such as Apple Inc.'s iTunes Store, as well as many e-book publishers have implemented DRM. In recent years, a number of television producers have implemented DRM on consumer electronic devices to control access to the freely-broadcast content of their shows, in response to the rising popularity of time-shifting digital video recorder systems such as TiVo.

Technologies

DRM and film

An early example of a DRM system was the Content Scrambling System (CSS) employed by the DVD Forum on film DVDs since ca. 1996. CSS used a simple encryption algorithm, and required device manufacturers to sign license agreements that restricted the inclusion of features, such as digital outputs that could be used to extract high-quality digital copies of the film, in their players. Thus, the only consumer hardware capable of decoding DVD films was controlled, albeit indirectly, by the DVD Forum, restricting the use of DVD media on other systems until the release of DeCSS by Jon Lech Johansen in 1999, which allowed a CSS-encrypted DVD to play properly on a computer using Linux, for which the Alliance had not arranged a licensed version of the CSS playing software.

Microsoft's Windows Vista contains a DRM system called the Protected Media Path, which contains the Protected Video Path (PVP). PVP tries to stop DRM-restricted content from playing while unsigned software is running in order to prevent the unsigned

software from accessing the content. Additionally, PVP can encrypt information during transmission to the monitor or the graphics card, which makes it more difficult to make unauthorized recordings.

Advanced Access Content System (AACSS) is a DRM system for HD DVD and Blu-ray Discs developed by the AACSS Licensing Administrator, LLC (AACSS LA), a consortium that includes Disney, Intel, Microsoft, Matsushita (Panasonic), Warner Brothers, IBM, Toshiba and Sony. In December 2006 a process key was published on the internet by hackers, enabling unrestricted access to AACSS-restricted HD DVD content. After the cracked keys were revoked, further cracked keys were released.

DRM and television

The CableCard standard is used by cable television providers in the United States to restrict content to services to which the customer has subscribed.

The broadcast flag concept was developed by Fox Broadcasting in 2001 and was supported by the MPAA and the U.S. Federal Communications Commission (FCC). A ruling in May 2005 by a US Court of Appeals held that the FCC lacked authority to impose it on the TV industry in the US. It required that all HDTVs obey a stream specification determining whether or not a stream can be recorded. This could block instances of fair use, such as time-shifting. It achieved more success elsewhere when it was adopted by the Digital Video Broadcasting Project (DVB), a consortium of about 250 broadcasters, manufactures, network operators, software developers, and regulatory bodies from about 35 countries involved in attempting to develop new digital TV standards.

An updated variant of the broadcast flag has been developed in the Content Protection and Copy Management (DVB-CPCM). It was developed in private, and the technical specification was submitted to European governments in March 2007. As with much DRM, the CPCM system is intended to control use of copyrighted material by the end-user, at the direction of the copyright holder. According to Ren Bucholz of the EFF, which paid to be a member of the consortium, "You won't even know ahead of time whether and how you will be able to record and make use of particular programs or devices". The DVB supports the system as it will harmonize copyright holders' control across different technologies and so make things easier for end users. The normative sections have now all been approved for publication by the DVB Steering Board, and will be published by ETSI as a formal European Standard as ETSI TS 102 825-X where X refers to the Part number of specification. Nobody has yet stepped forward to provide a Compliance and Robustness regime for the standard (though several are rumoured to be in development), so it is not presently possible to fully implement a system, as there is nowhere to obtain the necessary device certificates.

DRM and music

Audio CDs

Discs with digital rights management schemes are not legitimately standards-compliant Compact Discs (CDs) but are rather CD-ROM media. Therefore they all lack the CD logotype found on discs which follow the standard (known as Red Book). Therefore these CDs could not be played on all CD players. Many consumers could also no longer play purchased CDs on their computers. Personal computers running Microsoft Windows would sometimes even crash when attempting to play the CDs.

In 2005, Sony BMG introduced new DRM technology which installed DRM software on users' computers without clearly notifying the user or requiring confirmation. Among other things, the installed software included a rootkit, which created a severe security vulnerability others could exploit. When the nature of the DRM involved was made public much later, Sony initially minimized the significance of the vulnerabilities its software had created, but was eventually compelled to recall millions of CDs, and released several attempts to patch the surreptitiously included software to at least remove the rootkit. Several class action lawsuits were filed, which were ultimately settled by agreements to provide affected consumers with a cash payout or album downloads free of DRM.

Sony's DRM software actually had only a limited ability to prevent copying, as it affected only playback on Windows computers, not on other equipment. Even on the Windows platform, users regularly bypassed the restrictions. And, while the Sony DRM technology created fundamental vulnerabilities in customers' computers, parts of it could be trivially bypassed by holding down the "shift" key while inserting the CD, or by disabling the autorun feature. In addition, audio tracks could simply be played and re-recorded, thus completely bypassing all of the DRM (this is known as the analog hole). Sony's first two attempts at releasing a patch which would remove the DRM software from users' computers failed.

In January 2007, EMI stopped publishing audio CDs with DRM, stating that "the costs of DRM do not measure up to the results." Following EMI, Sony BMG was the last publisher to abolish DRM completely, and audio CDs containing DRM are no longer released by the four record labels.

Internet music

Many online music stores employ DRM to restrict usage of music purchased and downloaded online. There are many options for consumers wishing to purchase digital music over the internet:

- The iTunes Store (run by Apple), allows users to purchase a track online. The tracks purchased use Apple's FairPlay DRM system. Apple later launched *iTunes Plus*, which offered higher quality DRM-free tracks for a higher price. On October 17, 2007, iTunes Plus became available at the usual US\$0.99 price, replacing the non-Plus tracks. On January 6, 2009 Apple announced at its Macworld Expo keynote that iTunes music would be available completely DRM free by the end of the month. Videos sold and rented through iTunes, as well as

mobile software sold through the iTunes App Store for the iPhone and iPod touch, continue to use Apple's FairPlay DRM to inhibit casual copying. Apple increased the price of many hit singles and selected classic tracks to \$1.29 on April 7, 2009.

- Napster music store offers a subscription-based approach to DRM alongside permanent purchases. Users of the subscription service can download and stream an unlimited amount of music transcoded to Windows Media Audio (WMA) while subscribed to the service. But when the subscription period lapses, all of the downloaded music is unplayable until the user renews his or her subscription. Napster also charges users who wish to use the music on their portable device an additional \$5 per month. In addition, Napster gives users the option of paying an additional \$0.99 per track to burn it to CD or for the song to never expire. Music bought through Napster can be played on players carrying the Microsoft PlaysForSure logo (which, notably, do not include iPods or even Microsoft's own Zune). As of June 2009 Napster is giving DRM free MP3 music, which can be played on iPhones and iPods.
- Wal-Mart Music Downloads, another online music download store, charges \$0.94 per track for all non-sale downloads. All Wal-Mart, Music Downloads are able to be played on any Windows PlaysForSure marked product. The music does play on the SanDisk's Sansa mp3 player, for example, but must be copied to the player's internal memory. It cannot be played through the player's microSD card slot, which is a problem that many users of the mp3 player experience.
- Sony operated an online music download service called "Connect" which used Sony's proprietary OpenMG DRM technology. Music downloaded from this store (usually via Sony's SonicStage software) was only playable on computers running Microsoft Windows and Sony hardware (including the PSP and some Sony Ericsson phones).
- Kazaa is one of a few services offering a subscription-based pricing model. However, music downloads from the Kazaa website are DRM-protected, and can only be played on computers or portable devices running Windows Media Player, and only as long as the customer remains subscribed to Kazaa.

The various services are currently not interoperable, though those that use the same DRM system (for instance the several Windows Media DRM format stores, including Napster, Kazaa and Yahoo Music) all provide songs that can be played side-by-side through the same player program. Almost all stores require client software of some sort to be downloaded, and some also need plug-ins. Several colleges and universities, such as Rensselaer Polytechnic Institute, have made arrangements with assorted Internet music suppliers to provide access (typically DRM-restricted) to music files for their students, to less than universal popularity, sometimes making payments from student activity fee funds. One of the problems is that the music becomes unplayable after leaving school unless the student continues to pay individually. Another is that few of these vendors are compatible with the most common portable music player, the Apple iPod. The Gowers

Review of Intellectual Property (to HMG in the UK; 141 pages, 40+ specific recommendations) has taken note of the incompatibilities, and suggests (Recommendations 8—12) that there be explicit fair dealing exceptions to copyright allowing libraries to copy and format-shift between DRM schemes, and further allowing end users to do the same privately. If adopted, some of the acrimony may decrease.

Although DRM is prevalent for Internet music, some online music stores such as eMusic, Dogmastic, Amazon, and Beatport, do not use DRM despite encouraging users to avoid sharing music. Another online retailer, Xiie.net, which sells only unsigned artists, encourages people to share the music they buy from the site, to increase exposure for the artists themselves. Major labels have begun releasing more online music without DRM. Eric Bangeman suggests in *Ars Technica* that this is because the record labels are "slowly beginning to realize that they can't have DRMed music and complete control over the online music market at the same time... One way to break the cycle is to sell music that is playable on any digital audio player. eMusic does exactly that, and their surprisingly extensive catalog of non-DRMed music has vaulted it into the number two online music store position behind the iTunes Store." Apple's Steve Jobs has called on the music industry to eliminate DRM in an open letter titled *Thoughts on Music*. Apple's iTunes store will start to sell DRM-free 256 kbit/s (up from 128 kbit/s) AAC encoded music from EMI for a premium price (this has since reverted to the standard price). In March 2007, Musicload.de, one of Europe's largest online music retailers, announced their position strongly against DRM. In an open letter, Musicload stated that three out of every four calls to their customer support phone service are as a result of consumer frustration with DRM.

Computer games

Computer games sometimes use DRM technologies to limit the number of systems the game can be installed on by requiring authentication with an online server. Most games with this restriction allow three or five installs, although some allow an installation to be 'recovered' when the game is uninstalled. This not only limits users who have more than three or five computers in their homes (seeing as the rights of the software developers allow them to limit the number of installations), but can also prove to be a problem if the user has to unexpectedly perform certain tasks like upgrading operating systems or reformatting the computer's hard drive, tasks which, depending on how the DRM is implemented, count a game's subsequent reinstall as a new installation, making the game potentially unusable after a certain period even if it is only used on a single computer.

In mid-2008, the publication of *Mass Effect* marked the start of a wave of titles primarily making use of SecuROM and Steam for DRM and requiring authentication via an online server. The use of DRM scheme in 2008's *Spore* backfired and there were protests, resulting in a considerable number of users seeking a pirated version instead. This backlash against 3 activation limit was a significant factor in *Spore* becoming the most pirated game in 2008.

Many mainstream publishers continued to rely on online-based DRM throughout the later half of 2008 and early 2009, including Electronic Arts, Ubisoft and Atari. Ubisoft broke with the tendency to use online DRM in late 2008 with the release of *Prince of Persia* as an experiment to "see how truthful people really are" regarding the claim that DRM was inciting people to use pirated copies. Although Ubisoft has not commented on the results of the 'experiment', the majority of their subsequent titles in 2009 contained no online-based DRM since the release of *Prince of Persia* - notable examples being *Anno 1404* and *James Cameron's Avatar: The Game* making use of the online version of the TAGES copy protection system. An official patch has since been released stripping *Anno 1404* of the DRM. Electronic Arts followed suit in June 2009 with *The Sims 3*, with subsequent EA and EA Sports titles also being devoid of online DRM.

Ubisoft formally announced a return to on-line authentication on 9 February 2010 through its Uplay on-line gaming platform, starting with *Silent Hunter 5*, *The Settlers 7* and *Assassin's Creed 2*. *Silent Hunter V* was first reported to have been compromised within 24 hours of release, but users of the cracked version soon found out that only early parts of the game were playable. The Uplay system works by having the installed game on the local PCs incomplete and then continuously downloading parts of the game-code from Ubisoft's servers as the game progresses, making cracking games using the system a daunting task. It was only more than a month after the PC release in the first week of April that software was released that could bypass Ubisoft's DRM in *Assassin's Creed 2*, demonstrating its strength. The software did this by emulating a Ubisoft server for the game. Later that month, a real crack was released that was able to remove the connection requirement altogether. No fully working crack for *Silent Hunter V* has been confirmed.

In early March, 2010, Uplay servers suffered a period of inaccessibility due to a large scale DDoS attack, causing around 5% of game owners to become locked out of playing their game. The company later credited owners of the affected games with a free download, and there has been no further downtime.

Some most prominent cases making use of online DRM technology SecuROM include *Spore*, *BioShock*, *Mass Effect* and *Gears Of War*.

E-books

Electronic books read on a personal computer or an e-book reader typically use DRM restrictions to limit copying, printing, and sharing of e-books. E-books are usually limited to a certain number of reading devices and some e-publishers prevent any copying or printing. Some commentators believe that DRM is something that makes E-book publishing complex.

There are four main ebook formats at present. Mobipocket, Topaz, ePub and PDF. The Amazon Kindle uses both Mobipocket and Topaz format ebooks. Other ebook readers mostly use ePub format ebooks, but with differing DRM schemes.

There are four main ebook DRM schemes at present, one each from Adobe, Apple, Barnes & Noble and Amazon. Adobe's Adept DRM is applied to ePubs and PDFs, and can be read by several third-party ebook readers, as well as Adobe's Adobe Digital Editions software. Apple's Fairplay DRM is applied to ePubs, and can currently only be read by Apple's iBooks app on iOS devices. Barnes & Noble's DRM scheme is implemented by Adobe, and is applied to ePubs and the older Palm format ebooks. Amazon's DRM is an adaption of the original Mobipocket encryption, and is applied to Amazon's Mobipocket and Topaz format ebooks

Two software programs to view e-books are Adobe Reader and Microsoft Reader. Each program uses a slightly different approach to DRM. The first version of Adobe Acrobat e-book Reader to have encryption technologies was version 5.05. In the later version 6.0, the technologies of the PDF reader and the e-book reader were combined, allowing it to read both DRM-restricted and unrestricted files. After opening the file, the user is able to view the rights statement, which outlines actions available for the specific document. For example, for a freely transferred PDF, printing, copying to the clipboard, and other basic functions are available to the user. However, when viewing a more highly restricted e-book, the user is unable to print the book, copy or paste selections. The level of restriction is specified by the publisher or distribution agency.

Microsoft Reader, which exclusively reads e-books in a .lit format, contains its own DRM software. In Microsoft Reader there are three different levels of access control depending on the e-book: sealed e-books, inscribed e-books and owner exclusive e-books. Sealed e-books have the least amount of restriction and only prevents the document from being modified. Therefore, the reader cannot alter the content of the book to change the ending, for instance. Inscribed e-books are the next level of restriction. After purchasing and downloading the e-book, Microsoft Reader puts a digital ID tag to identify the owner of the e-book. Therefore, this discourages distribution of the e-book because it is inscribed with the owner's name making it possible to trace it back to the original copy that was distributed. Other e-book software uses similar DRM schemes. For example, Palm Digital Media, now known as Ereader, links the credit card information of the purchaser to the e-book copy in order to discourage distribution of the books.

The most stringent form of security that Microsoft Reader offers is called owner exclusive e-books, which uses traditional DRM technologies. To buy the e-book the consumer must first open Microsoft Reader, which ensures that when the book is downloaded it becomes linked to the computer's Microsoft Passport account. Thus the e-book can only be opened with the computer with which it was downloaded, preventing copying and distribution of the text.

Amazon.com has remotely deleted purchased copies of George Orwell's *1984* and *Animal Farm* from customer's Amazon Kindles. Commenters have widely described these actions as Orwellian, and have alluded to Big Brother from Orwell's *1984*. After an apology from Amazon CEO Jeff Bezos, the Free Software Foundation has written that this is just one more example of the excessive power Amazon has to remotely censor

what people read through its software, and called upon Amazon to free its e-book reader and drop DRM.

DRM and documents

Enterprise digital rights management (E-DRM or ERM) is the application of DRM technology to the control of access to corporate documents such as Microsoft Word, PDF, and AutoCAD files, emails, and intranet web pages rather than to the control of consumer media. E-DRM, now more commonly referenced as IRM (Information Rights Management), is generally intended to prevent the unauthorized use (such as industrial or corporate espionage or inadvertent release) of proprietary documents. IRM typically integrates with content management system software.

DRM has been used by organizations such as the British Library in its secure electronic delivery service to permit worldwide access to substantial numbers of rare (and in many cases unique) documents which, for legal reasons, were previously only available to authorized individuals actually visiting the Library's document centre at Boston Spa in England.

Watermarks

Digital watermarks are features of media that are added during production or distribution. Digital watermarks involve data that is arguably steganographically embedded within the audio or video data.

Watermarks can be used for different purposes that may include:

- recording the copyright owner
- recording the distributor
- recording the distribution chain
- identifying the purchaser of the music

Watermarks are not complete DRM mechanisms in their own right, but are used as part of a system for Digital Rights Management, such as helping provide prosecution evidence for purely legal avenues of rights management, rather than direct technological restriction. Some programs used to edit video and/or audio may distort, delete, or otherwise interfere with watermarks. Signal/modulator-carrier chromatography may also separate watermarks from original audio or detect them as glitches. Use of third party media players and other advanced programs render watermarking useless. Additionally, comparison of two separately obtained copies of audio using simple, home-grown algorithms can often reveal watermarks. New methods of detection are currently under investigation by both industry and non-industry researchers.

Metadata

Sometimes, metadata is included in purchased music which records information such as the purchaser's name, account information, or email address. This information is not embedded in the played audio or video data, like a watermark, but is kept separate, but within the file or stream.

As an example, metadata is used in media purchased from Apple's iTunes Store for DRM-free as well as DRM-restricted versions of their music or videos. This information is included as MPEG standard metadata.

Laws regarding DRM

Digital rights management systems have received some international legal backing by implementation of the 1996 WIPO Copyright Treaty (WCT). Article 11 of the Treaty requires nations party to the treaties to enact laws against DRM circumvention.

The WCT has been implemented in most member states of the World Intellectual Property Organization. The American implementation is the Digital Millennium Copyright Act (DMCA), while in Europe the treaty has been implemented by the 2001 European directive on copyright, which requires member states of the European Union to implement legal protections for technological prevention measures. In 2006, the lower house of the French parliament adopted such legislation as part of the controversial DADVSI law, but added that protected DRM techniques should be made interoperable, a move which caused widespread controversy in the United States.

Digital Millennium Copyright Act

The Digital Millennium Copyright Act (DMCA) is an amendment to United States copyright law passed unanimously on May 14, 1998, which criminalizes the production and dissemination of technology that allows users to circumvent technical copy-restriction methods. Under the Act, circumvention of a technological measure that effectively controls access to a work is illegal if done with the primary intent of violating the rights of copyright holders.

Reverse engineering of existing systems is expressly permitted under the Act under specific conditions. Under the reverse engineering safe harbor, circumvention necessary to achieve interoperability with other software is specifically authorized. Open-source software to decrypt content scrambled with the Content Scrambling System and other encryption techniques presents an intractable problem with the application of the Act. Much depends on the intent of the actor. If the decryption is done for the purpose of achieving interoperability of open source operating systems with proprietary operating systems, the circumvention would be protected by Section 1201(f) the Act. Cf., *Universal City Studios, Inc. v. Corley*, 273 F.3d 429 (2d Cir. 2001) at notes 5 and 16. However, dissemination of such software for the purpose of violating or encouraging others to violate copyrights has been held illegal.

On 22 May 2001, the European Union passed the EU Copyright Directive, an implementation of the 1996 WIPO Copyright Treaty that addressed many of the same issues as the DMCA.

The DMCA has been largely ineffective in protecting DRM systems, <http://boingboing.net/2007/03/24/dmcas_author_says_th.html> as software allowing users to circumvent DRM remains widely available. However, those who wish to preserve the DRM systems have attempted to use the Act to restrict the distribution and development of such software, as in the case of DeCSS.

Although the Act contains an exception for research, the exception is subject to vague qualifiers that do little to reassure researchers. Cf., 17 U.S.C. Sec. 1201(g). The DMCA has had an impact on cryptography, because many fear that cryptanalytic research may violate the DMCA. The arrest of Russian programmer Dmitry Sklyarov in 2001, for alleged infringement of the DMCA, was a highly publicized example of the law's use to prevent or penalize development of anti-DRM measures. Sklyarov was arrested in the United States after a presentation at DEF CON, and subsequently spent several months in jail. The DMCA has also been cited as chilling to non-criminal inclined users, such as students of cryptanalysis (including, in a well-known instance, Professor Felten and students at Princeton), and security consultants such as the Netherlands based Niels Ferguson, who has declined to publish information about vulnerabilities he discovered in an Intel secure-computing scheme because of his concern about being arrested under the DMCA when he travels to the US.

On 25 April 2007 the European Parliament supported the first directive of EU, which aims to harmonize criminal law in the member states. It adopted a first reading report on harmonizing the national measures for fighting copyright abuse. If the European Parliament and the Council approve the legislation, the submitted directive will oblige the member states to consider a crime a violation of international copyright committed with commercial purposes. The text suggests numerous measures: from fines to imprisonment, depending on the gravity of the offense.

The EP members supported the Commission motion, changing some of the texts. They excluded patent rights from the range of the directive and decided that the sanctions should apply only to offenses with commercial purposes. Copying for personal, non-commercial purposes was also excluded from the range of the directive.

International issues

In Europe, there are several ongoing dialog activities that are characterized by their consensus-building intention:

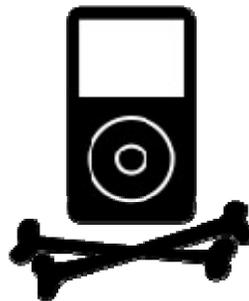
- Workshop on Digital Rights Management of the World Wide Web Consortium (W3C), January 2001.

- Participative preparation of the European Committee for Standardization/Information Society Standardisation System (CEN/ISSS) DRM Report, 2003 (finished).
- DRM Workshops of Directorate-General for Information Society and Media (European Commission) (finished), and the work of the DRM working groups (finished), as well as the work of the High Level Group on DRM (ongoing).
- Consultation process of the European Commission, DG Internal Market, on the Communication COM(2004)261 by the European Commission on "Management of Copyright and Related Rights" (closed).
- The INDICARE project is an ongoing dialogue on consumer acceptability of DRM solutions in Europe. It is an open and neutral platform for exchange of facts and opinions, mainly based on articles by authors from science and practice.
- The AXMEDIS project is a European Commission Integrated Project of the FP6. The main goal of AXMEDIS is automating the content production, copy protection and distribution, reducing the related costs and supporting DRM at both B2B and B2C areas harmonising them.
- The Gowers Review of Intellectual Property is the result of a commission by the British Government from Andrew Gowers, undertaken in December 2005 and published in 2006, with recommendations regarding copyright term, exceptions, orphaned works, and copyright enforcement.

Controversy

DRM opposition

**DRM IS
KILLING MUSIC**



AND IT'S A HIP OFF!

A parody on the Home Taping Is Killing Music logo.

Many organizations, prominent individuals, and computer scientists are opposed to DRM. Two notable DRM critics are John Walker, as expressed for instance, in his article *The Digital Imprimatur: How big brother and big media can put the Internet genie back in the bottle*, and Richard Stallman in his article *The Right to Read* and in other public statements: "DRM is an example of a malicious feature - a feature designed to hurt the user of the software, and therefore, it's something for which there can never be

toleration". Professor Ross Anderson of Cambridge University heads a British organization which opposes DRM and similar efforts in the UK and elsewhere. Cory Doctorow, a prominent writer and technology blogger, spoke on the Microsoft campus criticizing the technology, the morality, and the marketing of DRM.

There have been numerous others who see DRM at a more fundamental level. TechMediums.com argues that DRM-free music allows for viral marketing, arguing that independent artists benefit from "free marketing" and can then focus on revenues from higher margin products like merchandise and concert ticket sales. This is similar to some of the ideas in Michael H. Goldhaber's presentation about "The Attention Economy and the Net" at a 1997 conference on the "Economics of Digital Information." (sample quote from the "Advice for the Transition" section of that presentation: "If you can't figure out how to afford it without charging, you may be doing something wrong.")

The Electronic Frontier Foundation and similar organizations such as FreeCulture.org also hold positions which are characterized as opposed to DRM.

The Foundation for a Free Information Infrastructure has criticized DRM's impact as a trade barrier from a free market perspective.

The final version of the GNU General Public License version 3, as released by the Free Software Foundation, has a provision that 'strips' DRM of its legal value, so people can break the DRM on GPL software without breaking laws like the DMCA. Also, in May 2006, the FSF launched a "Defective by Design" campaign against DRM.

Creative Commons provides licensing options encouraging the expansion of and building upon creative work without the use of DRM. In addition, the use of a Creative Commons-licensed work on a device which incorporates DRM is a breach of the Baseline Rights asserted by each license.

Bill Gates spoke about DRM at CES in 2006. According to him, DRM is not where it should be, and causes problems for legitimate consumers while trying to distinguish between legitimate and illegitimate users.

According to Steve Jobs, Apple opposes DRM music after a public letter calling its music labels to stop requiring DRM on its iTunes Store. As of January 6, 2009, the iTunes Store is DRM-free for songs. However, Apple considers DRM on video content as a separate issue and has not removed DRM from all of its video catalog.



Defective by Design member protesting DRM on May 25, 2007.

As already noted, many DRM opponents consider "digital rights management" to be a misnomer. They argue that DRM manages rights (or access) the same way prison manages freedom and often refer to it as "digital restrictions management". Alternatively, ZDNet Executive Editor David Berlind suggests the term "Content Restriction, Annulment and Protection" or "CRAP" for short.

The Norwegian Consumer rights organization "Forbrukerrådet" complained to Apple Inc. in 2007 about the company's use of DRM in, and in conjunction with, its iPod and iTunes products. Apple was accused of restricting users' access to their music and videos in an unlawful way, and of using EULAs which conflict with Norwegian consumer legislation. The complaint was supported by consumers' ombudsmen in Sweden and Denmark, and is currently being reviewed in the EU. Similarly, the United States Federal Trade Commission held hearings in March 2009 to review disclosure of DRM limitations to customers' use of media products.

The use of DRM may also be a barrier to future historians, since technologies designed to permit data to be read only on particular machines, or with particular keys, or for certain periods, may well make future data recovery impossible.

DRM opponents argue that the presence of DRM violates existing private property rights and restricts a range of heretofore normal and legal user activities. A DRM component would control a device a user owns (such as a Digital audio player) by restricting how it may act with regards to certain content, overriding some of the user's wishes (for example, preventing the user from burning a copyrighted song to CD as part of a compilation or a review). An example of this effect may be seen in Microsoft's Windows Vista operating system in which content using a Protected Media Path is disabled or degraded depending on the DRM scheme's evaluation of whether the hardware and its use are 'secure'. All forms of DRM depend on the DRM enabled device (e.g., computer, DVD player, TV) imposing restrictions that (at least by intent) cannot be disabled or modified by the user. Key issues around digital rights management such as the right to make personal copies, provisions for persons to lend copies to friends, provisions for service discontinuance, hardware agnosticism, software and operating system agnosticism, contracts for public libraries, and customers' protection against one-side amendments of the contract by the publisher have not been fully addressed. It has also been pointed out that it is entirely unclear whether owners of content with DRM are legally permitted to pass on their property as inheritance to another person.

Tools like FairUse4WM have been created to strip Windows Media of DRM restrictions.

Valve Corporation President Gabe Newell also stated "most DRM strategies are just dumb" because they only decrease the value of a game in the consumer's eyes. Newell's suggests combating piracy by "[creating] greater value for customers through service value".

"DRM-Free"

Due to the strong opposition that exists to DRM, many companies and artists have begun advertising their products as "DRM-Free".

Apple began selling "DRM-Free" music through their iTunes store in April 2007. It was later revealed that the DRM-Free iTunes files were still embedded with each user's account information, a technique called Digital watermarking generally not regarded as DRM. In January 2009, iTunes began marketing all of their songs as "DRM-Free", however iTunes continues to use DRM on movies, TV shows, ringtones, and audiobooks.

Impossible task

Bruce Schneier has written about the futility of digital copy prevention and says it's an impossible task. He says "What the entertainment industry is trying to do is to use technology to contradict that natural law. They want a practical way to make copying hard enough to save their existing business. But they are doomed to fail." He has also described trying to make digital files uncopyable as being like "trying to make water not wet". The creators of StarForce also take this stance, stating that "The purpose of copy protection is not making the game uncrackable - it is impossible."

Both the Association for Computing Machinery and the Institute of Electrical and Electronics Engineers have historically opposed DRM, even going so far as to name AAC3 as a technology "most likely to fail" in an issue of IEEE Spectrum.

Shortcomings

Methods to bypass DRM

There are many methods to bypass DRM control on audio and video content.

One simple method to bypass DRM on audio files is to burn the content to an audio CD and then rip it into DRM-free files. This is only possible when the software that plays these DRM-restricted audio files allows CD-burning. Some software products simplify and automate this burn-rip process by allowing the user to burn music to a CD-RW disc or to a Virtual CD-R drive, then automatically ripping and encoding the music, and automatically repeating this process until all selected music has been converted, rather than forcing the user to do this one CD (72–80 minutes worth of music) at a time.

Many software programs have been developed that intercept the data stream as it is decrypted out of the DRM-restricted file, and then use this data to construct a DRM-free file. These programs require a decryption key. Programs that do this for DVDs, HD DVDs, and Blu-ray Discs include universal decryption keys in the software itself. Programs that do this for TiVo ToGo recordings, iTunes audio, and PlaysForSure songs, however, rely on the user's own key — that is, they can only process content that the user has legally acquired under his or her own account.

Another method is to use software to record the signals being sent through the audio or video cards, or to plug analog recording devices into the analog outputs of the media player. These techniques utilize the so-called "analog hole" (see below).

Analog hole

All forms of DRM for audio and visual material (excluding interactive materials, e.g. videogames) are subject to the *analog hole*, namely that in order for a viewer to play the material, the digital signal must be turned into an analog signal containing light and/or sound for the viewer, and so available to be copied as no DRM is capable of controlling content in this form. In other words, a user could play a purchased audio file while using a separate program to record the sound back into the computer into a DRM-free file format.

All DRM to date can therefore be bypassed by recording this signal and digitally storing and distributing it in a non DRM limited form, by anyone who has the technical means of recording the analog stream. However the conversion from digital to analog and back is likely to force a loss of quality, particularly when using lossy digital formats. HDCP is an attempt to restrict the analog hole, although it is largely ineffective.

Asus released a soundcard which features a function called "Analog Loopback Transformation" to bypass the restrictions of DRM. This feature allows the user to record DRM-restricted audio via the soundcard's built-in analog I/O connection.

DRM on general computing platforms

Many of the DRM systems in use are designed to work on general purpose computing hardware, such as desktop PCs apparently because this equipment is felt to be a major contributor to revenue loss from disallowed copying. Large commercial copyright infringers ("pirates") avoid consumer equipment, so losses from such infringers will not be covered by such provisions.

Such schemes, especially software based ones, can never be wholly secure since the software must include all the information necessary to decrypt the content, such as the decryption keys. An attacker will be able to extract this information, directly decrypt and copy the content, which bypasses the restrictions imposed by a DRM system.

DRM on purpose-built hardware

Many DRM schemes use encrypted media which requires purpose-built hardware to hear or see the content. This appears to ensure that only licensed users (those with the hardware) can access the content. It additionally tries to protect a secret decryption key from the users of the system.

While this in principle can work, it is extremely difficult to build the hardware to protect the secret key against a sufficiently determined adversary. Many such systems have failed in the field. Once the secret key is known, building a version of the hardware that performs no checks is often relatively straightforward. In addition user verification provisions are frequently subject to attack, pirate decryption being among the most frequented ones.

A common real-world example can be found in commercial direct broadcast satellite television systems such as DirecTV and Malaysia's Astro. The company uses tamper-resistant smart cards to store decryption keys so that they are hidden from the user and the satellite receiver. However, the system has been compromised in the past, and DirecTV has been forced to roll out periodic updates and replacements for its smart cards.

Watermarks

Watermarks can very typically be removed, although degradation of video or audio can occur.

Mass piracy failure

Mass piracy of hard copies does not necessarily need DRM to be decrypted or removed, as it can be achieved by bit-perfect copying of a legally obtained medium without accessing the decrypted content. Additionally, still-encrypted disk images can be distributed over the Internet and played on legitimately licensed players.

Obsolescence

When standards and formats change, it may be difficult to transfer DRM-restricted content to new media. Additionally, any system that requires contact with an authentication server is vulnerable to that server becoming unavailable, as happened in 2007 when videos purchased from Major League Baseball (mlb.com) prior to 2006 became unplayable due to a change to the servers that validate the licences.

Microsoft Zune - When Microsoft introduced their Zune media player in 2006, it did not support content that uses Microsoft's own PlaysForSure DRM scheme they had previously been selling. The EFF calls this "a raw deal".

MSN Music - In April 2008, Microsoft sent an email to former customers of the now-defunct MSN Music store: "As of August 31, 2008, we will no longer be able to support the retrieval of license keys for the songs you purchased from MSN Music or the authorization of additional computers. You will need to obtain a license key for each of your songs downloaded from MSN Music on any new computer, and you must do so before August 31, 2008. If you attempt to transfer your songs to additional computers after August 31, 2008, those songs will not successfully play."

However, to avoid a public relations disaster, Microsoft re-issued MSN Music shutdown statement on June 19th and allowed the users to use their licenses until the end of 2011: "After careful consideration, Microsoft has decided to continue to support the authorization of new computers and devices and delivery of new license keys for MSN Music customers through at least the end of 2011, after which we will evaluate how much this functionality is still being used and what steps should be taken next to support our customers. This means you will continue to be able to listen to your purchased music and transfer your music to new PCs and devices beyond the previously announced August 31, 2008 date."

Yahoo! Music Store - On July 23, 2008, the Yahoo! Music Store emailed its customers to tell them it will be shutting down effective September 30, 2008 and the DRM license key servers will be taken offline.

Walmart - In August 2007, Walmart's online music division started offering (DRM-free) MP3s as an option. Starting in February 2008, they made all sales DRM-free. On September 26, 2008, the Walmart Music Team notified its customers via email they will be shutting down their DRM servers October 9, 2008 and any DRM-encumbered music acquired from them will no longer be accessible unless ripped to a non-DRM format before that date.

After bad press and negative reaction from customers, on October 9, 2008, Walmart decided not to take its DRM servers offline.

Fictionwise / Overdrive - In January 2009, OverDrive informed Fictionwise that they would no longer be providing downloads for purchasers of e-books through Fictionwise as of 31 January 2009. No reason was provided to Fictionwise as to why they were being shut down. This prevents previous purchasers from being able to renew their books on new devices. Fictionwise is working to provide replacement ebooks for its customers in alternative, non-DRM formats, but does not have the rights to provide all of the books in different formats.

Ads for Adobe PDF - Also in January 2009, Adobe Systems announced that as of March 2009 they would no longer operate the servers that served ads to their PDF reader. Depending on the restriction settings used when PDF documents were created, they may no longer be readable.

Historical note

A very early implementation of DRM was the Software Service System (SSS) devised by the Japanese engineer Ryoichi Mori in 1983 and subsequently refined under the name superdistribution. The SSS was based on encryption, with specialized hardware that controlled decryption and also enabled payments to be sent to the copyright holder. The underlying principle of the SSS and subsequently of superdistribution was that the distribution of encrypted digital products should be completely unrestricted and that users of those products would not just be permitted to redistribute them but would actually be encouraged to do so.

Chapter- 7

Fair Use

Fair use, a limitation and exception to the exclusive right granted by copyright law to the author of a creative work, is a doctrine in United States copyright law that allows limited use of copyrighted material without requiring permission from the rights holders. Examples of fair use include commentary, criticism, news reporting, research, teaching, library archiving and scholarship. It provides for the legal, non-licensed citation or incorporation of copyrighted material in another author's work under a four-factor balancing test. The term *fair use* originated in the United States. A similar principle, fair dealing, exists in some other common law jurisdictions. Civil law jurisdictions have other limitations and exceptions to copyright.

Fair use under United States law

The legal concept of "Test copyright" was first ratified by the Kingdom of Great Britain's Statute of Anne of 1709. As room was not made for the authorized reproduction of copyrighted content within this newly formulated statutory right, the courts created a doctrine of "fair abridgment" in *Gyles v Wilcox*, which eventually evolved into the modern concept of "fair use," that recognized the utility of such actions. The doctrine only existed in the U.S. as common law until it was incorporated into the Copyright Act of 1976, 17 U.S.C. § 107.

17 U.S.C. § 107

Notwithstanding the provisions of sections 17 U.S.C. § 106 and 17 U.S.C. § 106A, the fair use of a copyrighted work, including such use by reproduction in copies or phonorecords or by any other means specified by that section, for purposes such as criticism, comment, news reporting, teaching (including multiple copies for classroom use), scholarship, or research, is not an infringement of copyright. In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include:

1. the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes;

2. the nature of the copyrighted work;
3. the amount and substantiality of the portion used in relation to the copyrighted work as a whole; and
4. the effect of the use upon the potential market for or value of the copyrighted work.

The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

The four factors of analysis for fair use set forth above derive from the classic opinion of Joseph Story in *Folsom v. Marsh*, 9 F.Cas. 342 (1841), in which the defendant had copied 353 pages from the plaintiff's 12-volume biography of George Washington in order to produce a separate two-volume work of his own. The court rejected the defendant's fair use defense with the following explanation:

“ [A] reviewer may fairly cite largely from the original work, if his design be really and truly to use the passages for the purposes of fair and reasonable criticism. On the other hand, it is as clear, that if he thus cites the most important parts of the work, with a view, not to criticize, but to supersede the use of the original work, and substitute the review for it, such a use will be deemed in law a piracy...

In short, we must often... look to the nature and objects of the selections made, the quantity and value of the materials used, and the degree in which the use may prejudice the sale, or diminish the profits, or supersede the objects, of the original work. ”

Once these factors were codified as guidelines in 17 U.S.C. § 107, they were not rendered exclusive. The section was intended by Congress to restate, but not replace, the prior judge-made law. Courts are still entitled to consider other factors as well.

Fair use tempers copyright's exclusive rights to serve the purpose of copyright law, which the U.S. Constitution defines as the promotion of "the Progress of Science and useful Arts" (Art. I, § 8, cl. 8). This principle applies particularly well to the case of criticism and also sheds light on various other limitations on copyright's exclusive rights, particularly the *scenes à faire* doctrine.

Purpose and character

The first factor is regarding whether the use in question helps fulfill the intention of copyright law to stimulate creativity for the enrichment of the general public, or whether it aims to only "supersede the objects" of the original for reasons of personal profit. To justify the use as fair, one must demonstrate how it either advances knowledge or the

progress of the arts through the addition of something new. A key consideration is the extent to which the use is interpreted as *transformative*, as opposed to merely *derivative*.

When Tom Forsythe appropriated Barbie dolls for his photography project "Food Chain Barbie," Mattel lost its claims of copyright and trademark infringement against him because his work effectively parodies Barbie and the values she represents. But when Jeff Koons tried to justify his appropriation of Art Rogers' photograph "Puppies" in his sculpture "String of Puppies" with the same parody defense, he lost because his work was not presented as a parody of Rogers' photograph in particular, but of society at large, which was deemed insufficiently justificatory.

However, since this case, courts have begun to emphasize the first fair use factor—assessing whether the alleged infringement has *transformative* use as described by the Hon. Judge Pierre N. Leval. More recently, Koons was involved in a similar case with commercial photographer Andrea Blanch, regarding his use of her photograph for a painting, whereby he appropriated a central portion of an advertisement she had been commissioned to shoot for a magazine. In this case, Koons won; the case sets a favorable precedent for appropriation art where the use is deemed transformative.

The subfactor mentioned in the legislation above, "whether such use is of a commercial nature or is for nonprofit educational purposes," has recently been deemphasized in some Circuits "since many, if not most, secondary uses seek at least some measure of commercial gain from their use." More important is whether the use fulfills any of the "preamble purposes" also mentioned in the legislation above, as these have been interpreted as paradigmatically "transformative." Although Judge Pierre Leval has distinguished the first factor as "the soul of fair use," it alone is not determinative. For example, not every educational usage is fair.

Nature of the copied work

Although the Supreme Court of the United States has ruled that the availability of copyright protection should not depend on the artistic quality or merit of a work, fair use analyses consider certain aspects of the work to be relevant, such as whether it is fictional or non-fictional.

To prevent the private ownership of work that rightfully belongs in the public domain, facts and ideas are separate from copyright—only their particular expression or fixation merits such protection. On the other hand, the social usefulness of freely available information can weigh against the appropriateness of copyright for certain fixations. The Zapruder film of the assassination of President Kennedy, for example, was purchased and copyrighted by *Time* magazine. Yet their copyright was not upheld, in the name of the public interest, when they tried to enjoin the reproduction of stills from the film in a history book on the subject in *Time Inc. v. Bernard Geis Associates*.

Following the decisions of the Second Circuit in *Salinger v. Random House, Inc.* and in *New Era Publications Int'l v. Henry Holt & Co.*, the aspect of whether the copied work

has been previously published suddenly trumped all other considerations because of, in the words of one commentator, "the original author's interest in controlling the circumstances of the first public revelation of his work, and his right, if he so chooses, not to publish at all." Yet some view this importation of certain aspects of France's *droit moral d'artiste* (moral rights of the artist) into American copyright law as "bizarre and contradictory" because it sometimes grants greater protection to works that were created for private purposes that have little to do with the public goals of copyright law, than to those works that copyright was initially conceived to protect. This is not to claim that unpublished works, or, more specifically, works not intended for publication, do not deserve legal protection, but that any such protection should come from laws about privacy, rather than laws about copyright. The statutory fair use provision was amended in response to these concerns by adding a final sentence: "The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors."

Amount and substantiality

The third factor assesses the quantity or percentage of the original copyrighted work that has been imported into the new work. In general, the less that is used in relation to the whole, e.g., a few sentences of a text for a book review, the more likely that the sample will be considered fair use. Yet see *Sony Corp. v. Universal City Studios* for a case in which substantial copying—entire programs for private viewing—was upheld as fair use, at least when the copying is done for the purposes of time-shifting. Likewise, see *Kelly v. Arriba Soft Corporation*, where the Ninth Circuit held that copying an entire photo to use as a thumbnail in online search results did not weigh against fair use, "if the secondary user only copies as much as is necessary for his or her intended use." Conversely, in *Harper & Row, Publishers, Inc. v. Nation Enters.*, the use of fewer than 400 words from President Ford's memoir by a political opinion magazine was interpreted as infringement because those few words represented "the heart of the book" and were, as such, substantial.

Before 1991, sampling in certain genres of music was accepted practice and such copyright considerations as these were viewed as largely irrelevant. The strict decision against rapper Biz Markie's appropriation of a Gilbert O'Sullivan song in the case *Grand Upright Music, Ltd. v. Warner Bros. Records, Inc.* changed practices and opinions overnight. Samples now had to be licensed, as long as they rose "to a level of legally cognizable appropriation." In other words, *de minimis* sampling was still considered fair and free because, traditionally, "the law does not care about trifles." The recent Sixth Circuit Court decision in the appeal to *Bridgeport Music* has reversed this standing, eliminating the *de minimis* defense for samples of recorded music, but stating that the decision did not apply to fair use.

Effect upon work's value

The fourth factor measures the effect that the allegedly infringing use has had on the copyright owner's ability to exploit his or her original work. The court not only

investigates whether the defendant's specific use of the work has significantly harmed the copyright owner's market, but also whether such uses in general, if widespread, would harm the potential market of the original. The burden of proof here rests on the defendant for commercial uses, but on the copyright owner for noncommercial uses. See *Sony Corp. v. Universal City Studios*, where the copyright owner, Universal, failed to provide any empirical evidence that the use of Betamax had either reduced their viewership or negatively impacted their business. In the aforementioned *Nation* case regarding President Ford's memoirs, the Supreme Court labeled this factor "the single most important element of fair use" and it has indeed enjoyed some level of primacy in fair use analyses ever since. Yet the Supreme Court's more recent announcement in *Campbell v. Acuff-Rose Music, Inc.* that "all [four factors] are to be explored, and the results weighed together, in light of the purposes of copyright" has helped modulate this emphasis in interpretation.

In evaluating the fourth factor, courts often consider two kinds of harm to the potential market of the original work: First, courts consider whether the use in question acts as a direct market substitute for the original work. In the judgement of the Supreme Court in *Acuff-Rose Music* they decisively stated that, "when a commercial use amounts to mere duplication of the entirety of the original, it clearly supersedes the object of the original and serves as a market replacement for it, making it likely that cognizable market harm to the original will occur." In one instance, a court ruled that this factor weighed against a defendant who had made unauthorized movie trailers for video retailers, since his trailers acted as direct substitutes for the copyright owner's official trailers. Second, courts also consider whether potential market harm might exist beyond that of direct substitution, such as in the potential existence of a licensing market. This consideration has weighed against commercial copy shops that make copies of articles in course-pack for college students, when a market already existed for the licensing of course-pack copies.

Courts recognize that certain kinds of market harm do not oppose fair use, such as when a parody or negative review impairs the market of the original work. Copyright considerations may not shield a work against adverse criticism.

Fair use and professional communities

Courts, when deciding fair use cases, in addition to looking at context, amount and value of the use, also look to the standards and practices of the professional communities where the case comes from.

Practical effect of fair use defense

The practical effect of this law and the court decisions following it is that it is usually possible to quote from a copyrighted work in order to criticize or comment upon it, teach students about it, and possibly for other uses. Certain well-established uses cause few problems. A teacher who prints a few copies of a poem to illustrate a technique will have no problem on all four of the above factors (except possibly on amount and substantiality), but some cases are not so clear. All the factors are considered and

balanced in each case: a book reviewer who quotes a paragraph as an example of the author's style will probably fall under fair use even though he may sell his review commercially. But a non-profit educational website that reproduces whole articles from technical magazines will probably be found to infringe if the publisher can demonstrate that the website affects the market for the magazine, even though the website itself is non-commercial.

Free Republic, LLC, owner of the political website freerepublic.com, was found liable for copyright infringement in *L.A. Times v. Free Republic* for reproducing and archiving full-text versions of plaintiffs' news articles even though the judge found the website minimally commercial. She held that "while defendants' do not necessarily 'exploit' the articles for commercial gain, their posting to the Free Republic site allows defendants and other visitors to avoid paying the 'customary price' charged for the works."

The April 2000 opinion ruled concerning the four factors of fair use that 1) "defendants' use of plaintiffs' articles is minimally, if at all, transformative," 2) the factual content of the articles copied "weighs in favor of finding of fair use of the news articles by defendants in this case," though it didn't "provide strong support" 3) concerning the amount and substantiality prong, "the wholesale copying of plaintiffs' articles weighs against the finding of fair use," and 4) the plaintiffs showed that they were trying to exploit the market for viewing their articles online and defendants didn't rebut their showing by proving an absence of usurpation harm to plaintiffs. Ultimately the court found "that the defendants may not assert a fair use defense to plaintiffs' copyright infringement claim."

Fair use as a defense

The Supreme Court of the United States described fair use as an affirmative defense in *Campbell v. Acuff-Rose Music, Inc.*. This means that, in litigation on copyright infringement, the defendant bears the burden of raising and proving that his use was "fair" and not an infringement. Thus, fair use need not even be raised as a defense unless the plaintiff first shows (or the defendant concedes) a "prima facie" case of copyright infringement. If the work was not copyrightable, the term had expired, or the defendant's work borrowed only a small amount, for instance, then the plaintiff cannot make out a *prima facie* case of infringement, and the defendant need not even raise the fair use defense.

Because of the defendant's burden of proof, some copyright owners frequently make claims of infringement even in circumstances where the fair use defense would likely succeed in hopes that the user will refrain from the use rather than spending resources in his defense.

Because paying a royalty fee may be much less expensive than having a potential copyright suit threaten the publication of a completed work in which a publisher has invested significant resources, many authors may seek a license even for uses that copyright law ostensibly permits without liability.

The frequent argument over whether fair use is a "right" or a "defense" is generated by confusion over the use of the term "affirmative defense." "Affirmative defense" is simply a term of art from litigation reflecting the timing in which the defense is raised. It does not distinguish between "rights" and "defenses," and so it does not characterize the substance of the defendant's actions as "not a right but a defense."

In response to perceived over-expansion of copyrights, several electronic civil liberties and free expression organizations began in the 1990s to add fair use cases to their dockets and concerns. These include the Electronic Frontier Foundation ("EFF"), the American Civil Liberties Union, the National Coalition Against Censorship, the American Library Association, numerous clinical programs at law schools, and others. The "Chilling Effects" archive was established in 2002 as a coalition of several law school clinics and the EFF to document the use of cease and desist letters. Most recently, in 2006, Stanford University began an initiative called "The Fair Use Project" (FUP) to help artists, particularly filmmakers, fight lawsuits brought against them by large corporations.

In 2009, fair use appeared as a defense in lawsuits against filesharing. Charles Nesson argued that file-sharing qualifies as fair use in his defense of alleged filesharer Joel Tenenbaum. Kiwi Camara, defending alleged filesharer Jammie Thomas, announced a similar defense.

On September 2, 2009 Israeli District court ruled out a detailed decision not allowing disclosure of "John Doe"'s details for the request of the FA Premier League based on several reasons, but the most interesting were that "fair use" under the new Israeli law of 2007 (which is based on the US 4 factors test) is a right and not merely a defense. The court specifically states that the public may have base for a legal cause of action if its fair use right is infringed by the copyright holder. Other important decision in said judgment is the fact that the court finds streaming Internet filesharing site of live soccer games not infringing copyright as this use is fair use (mainly due to the importance of certain sport events and the public's right). The court analyzes the 4 factors and decides that due to such importance of sporting games (and other less important factors), such use is fair.

The economic benefit of fair use

A balanced copyright law provides an economic benefit to many high tech businesses such as search engines and software developers. Fair Use is also crucial to non-technology industries such as insurance, legal services, and newspaper publishers. On September 12, 2007, the Computer and Communications Industry Association (CCIA), a group representing companies including Google Inc., Microsoft Inc., Oracle Corporation, Sun Microsystems, Yahoo and other high tech companies, released a study that found that Fair Use exceptions to US copyright laws were responsible for more than \$4,500 billion dollars in annual revenue for the United States economy representing one-sixth of the total U.S. GDP. The study was conducted using a methodology developed by the World Intellectual Property Organization. The study found that fair use dependent industries are directly responsible for more than 18% of U.S. economic growth and nearly 11 million American jobs. "As the United States economy becomes increasingly

knowledge-based, the concept of fair use can no longer be discussed and legislated in the abstract. It is the very foundation of the digital age and a cornerstone of our economy,” said Ed Black, President and CEO of CCIA. “Much of the unprecedented economic growth of the past ten years can actually be credited to the doctrine of fair use, as the Internet itself depends on the ability to use content in a limited and nonlicensed manner.”

Fair use and parody

Producers or creators of parodies of a copyrighted work have been sued for infringement by the targets of their ridicule, even though such use may be protected as fair use. These fair use cases distinguish between parodies (using a work in order to poke fun at or comment on the work itself) and satires (using a work to poke fun at or comment on something else). Courts have been more willing to grant fair use protections to parodies than to satires, but the ultimate outcome in either circumstance will turn on the application of the four fair use factors.

In *Campbell v. Acuff-Rose Music, Inc.* Supreme Court recognized parody as a fair use, even when done for profit. Roy Orbison's publisher, Acuff-Rose Music Inc., had sued 2 Live Crew in 1989 for their use of Orbison's "Oh, Pretty Woman" in a mocking rap version with altered lyrics. The Supreme Court viewed 2 Live Crew's version as a ridiculing commentary on the earlier work, and ruled that when the parody was itself the product rather than used for mere advertising, commercial sale did not bar the defense. The *Campbell* court also distinguished parodies from satire, which they described as a broader social critique not intrinsically tied to ridicule of a specific work, and so not deserving of the same use exceptions as parody because the satirist's ideas are capable of expression without the use of the other particular work.

A number of appellate decisions have recognized parody as a protected fair use, including both the Second (*Leibovitz v. Paramount Pictures Corp.*) and Ninth Circuits (*Mattel v. Walking Mountain Productions*). Most recently, in *Suntrust v. Houghton Mifflin*, a suit was brought unsuccessfully against the publication of *The Wind Done Gone*, which reused many of the characters and situations from *Gone with the Wind*, but told the events from the point of view of the slaves rather than the slaveholders. The Eleventh Circuit, applying *Campbell*, recognized that *The Wind Done Gone* was a protected parody, and vacated the district court's injunction against its publication.

Fair use on the Internet

A US court case in 2003, *Kelly v. Arriba Soft Corporation*, provides and develops the relationship between thumbnails, inline linking and fair use. In the lower District Court case on a motion for summary judgment, Arriba Soft was found to have violated copyright without a fair use defense in the use of thumbnail pictures and inline linking from Kelly's website in Arriba's image search engine. That decision was appealed and contested by Internet rights activists such as the Electronic Frontier Foundation, who argued that it is clearly covered under fair use.

On appeal, the 9th Circuit Court of Appeals found in favor of the defendant. In reaching its decision, the court utilized the above-mentioned four-factor analysis. First, it found the purpose of creating the thumbnail images as previews to be sufficiently transformative, noting that they were not meant to be viewed at high resolution like the original artwork was. Second, the fact that the photographs had already been published diminished the significance of their nature as creative works. Third, although normally making a "full" replication of a copyrighted work may appear to violate copyright, here it was found to be reasonable and necessary in light of the intended use. Lastly, the court found that the market for the original photographs would not be substantially diminished by the creation of the thumbnails. To the contrary, the thumbnail searches could increase exposure of the originals. In looking at all these factors as a whole, the court found that the thumbnails were fair use and remanded the case to the lower court for trial after issuing a revised opinion on July 7, 2003. The remaining issues were resolved with a default judgment after Arriba Soft had experienced significant financial problems and failed to reach a negotiated settlement.

In August 2008 U.S. District Judge Jeremy Fogel of San Jose, California ruled that copyright holders cannot order a deletion of an online file without determining whether that posting reflected "fair use" of the copyrighted material. The case involved Stephanie Lenz, a writer and editor from Gallitzin, Pennsylvania, who made a home video of her 13-month-old son dancing to Prince's song Let's Go Crazy and posted the video on YouTube. Four months later, Universal Music, the owner of the copyright to the song, ordered YouTube to remove the video enforcing the Digital Millennium Copyright Act. Lenz notified YouTube immediately that her video was within the scope of fair use, and demanded that it be restored. YouTube complied after six weeks, not two weeks as required by the Digital Millennium Copyright Act. Lenz then sued Universal Music in California for her legal costs, claiming the music company had acted in bad faith by ordering removal of a video that represented fair-use of the song.

Common misunderstandings

Fair use is commonly misunderstood because of its deliberate ambiguity. Here are some of the more common misunderstandings with explanations of why they are wrong:

- *Any use that seems fair is fair use.* In the law, the term *fair use* has a specific meaning that only partly overlaps the plain-English meaning of the words. While judges have much leeway in deciding how to apply fair use guidelines, not every use that is commonly considered "fair" counts as fair use under the law.
- *Fair use interpretations, once made, are static forever.* Fair use is decided on a case by case basis, on the entirety of circumstances. The same act done by different means or for a different purpose can gain or lose fair use status. Even repeating an identical act at a different time can make a difference due to changing social, technological, or other surrounding circumstances.
- *If it's not fair use, it's copyright infringement.* Fair use is only one of many limitations, exceptions, and defenses to copyright infringement. For instance, the

Audio Home Recording Act establishes that it is legal in some circumstances to make copies of audio recordings for non-commercial personal use.

- *It's copyrighted, so it can't be fair use.* On the contrary, fair use applies *only* to copyrighted works, describing conditions under which copyrighted material may be used without permission. If a work is not copyrighted, fair use does not come into play, since public-domain works can be used for any purpose without violating copyright law.
 - Note: In some countries (including the United States of America), the mere creation of a work establishes copyright over it, and there is no legal requirement to register or declare copyright ownership
- *Acknowledgment of the source makes a use fair.* Giving the name of the photographer or author may help, but it is not sufficient on its own. While plagiarism and copyright violation are related matters—both can, at times, involve failure to properly credit sources—they are not identical. Plagiarism—using someone's words, ideas, images, etc. without acknowledgment—is a matter of professional ethics. Copyright is a matter of law, and protects exact expression, *not* ideas. One can plagiarize even a work that is not protected by copyright, such as trying to pass off a line from Shakespeare as one's own. On the other hand, citing sources generally prevents accusations of plagiarism, but is not a sufficient defense against copyright violations. For example, reprinting a copyrighted book without permission, while citing the original author, would be copyright infringement but not plagiarism.
- *Noncommercial use is invariably fair.* Not true, though a judge may take the profit motive or lack thereof into account. In *L.A. Times v. Free Republic*, the court found that the noncommercial use of L.A. Times content by the Free Republic Web site was in fact *not* fair use, since it allowed the public to obtain material at no cost that they would otherwise pay for.
- *Strict adherence to fair use protects you from being sued.* Fair use is an affirmative *defense* against an infringement suit; it does not restrain anyone from suing. The copyright holder may legitimately disagree that a given use is fair, and they have the right to have the matter decided by a court. Thus, fair use does not guarantee that a lawsuit will be prevented.
- *The lack of a copyright notice means the work is public domain.* Not *usually* true. United States law in effect since March 1, 1989, has made copyright the default for newly created works. For a recent work to be in the public domain the author must specifically opt-out of copyright. For works produced between January 1, 1923 and March 1, 1989, copyright notice is required; however, registration was not required and between January 1, 1978 and March 1, 1989 lack of notice is not necessarily determinative, if attempts were made immediately to correct the lack of notice. Any American works that did not have formal registration or notice fell into the Public Domain if registration was not made in a timely fashion. For international works, the situation is even more complex. International authors who failed to provide copyright notice or register with the U.S. copyright office are given additional contemporary remedies that may restore American copyright protection given certain conditions. International authors/corporations who fail to meet these remedies forfeit their copyright. An example of a company who failed

- to prove copyright was Roland Corporation and their claimed copyright on the sounds contained in their MT-32 synthesizer.
- *It's okay to quote up to 300 words.* The 300-word limit is reported to be an unofficial agreement, now long obsolete, among permissions editors in the New York publishing houses: "I'll let you copy 300 words from our books if you let us copy 300 words from yours." It runs counter to the substantiality standard. As explained above, the substantiality of the copying is more important than the actual amount. For instance, copying a complete short poem is more substantial than copying a random paragraph of a novel; copying an 8.5×11-inch photo is more substantial than copying a square foot of an 8×10-foot painting. In 1985, the U.S. Supreme Court held that a news article's quotation of approximately 300 words from former President Gerald Ford's 200,000 word memoir was sufficient to constitute an infringement of the exclusive publication right in the work.
 - *You can deny fair use by including a disclaimer.* Fair use is a right granted to the public on all copyrighted work. Fair use rights take precedence over the author's interest. Thus the copyright holder cannot use a non-binding disclaimer, or notification, to revoke the right of fair use on works. However, binding agreements such as contracts or license agreements may take precedence over fair use rights.
 - *If you're copying an entire work, it's not fair use.* While copying an entire work may make it harder to justify the amount and substantiality test, it does not make it impossible that a use is fair use. For instance, in the Betamax case, it was ruled that copying a complete television show for time-shifting purposes is fair use.
 - *If you're selling for profit, it's not fair use.* While commercial copying for profit work may make it harder to qualify as fair use, it does not make it impossible. For instance, in the 2 Live Crew—*Oh, Pretty Woman* case, it was ruled that commercial parody can be fair use.

Influence internationally

While many other countries recognize similar exceptions to copyright, only the United States and Israel fully recognize the concept of fair use.

While influential in some quarters, other countries often have drastically different fair use criteria to the US, and in some countries there is little or no fair use defense available. Even within Europe, rules vary greatly between countries. Some countries have the concept of fair dealing instead of fair use. However many countries have some reference to an exemption for educational use, although the extent of this exemption may vary widely.

Fair dealing in Canada

The *Copyright Act* establishes fair dealing in Canada, which allows specific exceptions to copyright protection. The open-ended concept of fair use is not observed in Canadian law. In 1985, the Sub-Committee on the Revision of Copyright rejected replacing fair dealing with an open-ended system, and in 1986 the Canadian government agreed that

“the present fair dealing provisions should not be replaced by the substantially wider ‘fair use’ concept.”

CCH Canadian Ltd. v. Law Society of Upper Canada [2004] 1 S.C.R. 339, 2004 SCC 13 is the landmark Supreme Court of Canada case that establishes the bounds of fair dealing in Canadian copyright law. The Law Society of Upper Canada was sued for copyright infringement for providing photocopy services to researchers. The Court unanimously held that the Law Society's practice fell within the bounds of fair dealing.

Fair use in Israel

In November 2007, Israel passed a new Copyright Law that included a US style fair use exception. The law, which took effect in May 2008, permits the fair use of copyrighted works for purposes such as private study, research, criticism, review, news reporting, quotation, or instruction or testing by an educational institution. The law sets up four factors, similar to those of section 107 under American law, to determine whether a use is fair use.

Fair use in South Korea

The Korean Copyright Act newly amended in 2009, in articles 23~38 of section 4-2 (Limitation to the author's property rights), defines the exceptional use of copyrighted material without permission from copyright holders. However, a broad concept of fair use as in the above countries still does not exist in the Korean Copyright Act.