Quality and Quantity in Multimedia Databases

Dr Péter Jacsó
Associate Professor, Graduate School of Library and Information Studies
University of Hawaii

Having the privilege to be a keynote speaker at the Australasian Information Online & On Disc Conference I could choose a subject that I felt would be appropriate for a keynote speech.

Quality of databases has been a concern for many librarians, information specialists, and end-users for a long time, and I myself have spoken and written about it often. Though the quality issues of traditional databases have not been resolved, this time I would like to focus on the problems of the latest genre of databases, the multimedia databases. In addition to the errors of commission and errors of omission in the textual databases the multimedia information sources bring up further problems in the area of non-textual media.

I will discuss three major aspects of multimedia elements in databases: functionality, quality, and quantity. I will illustrate my speech with several examples from multimedia databases, mostly encyclopedias, travel guides, atlases, and other reference databases that provide the most fertile soil for multimedia examples. Most of my examples relate to Australasia, as I was

looking for audio and video information in a number of databases in preparing for this keynote speech.

The conclusions, of course, would apply to other topical subjects as well, and these are just the multimedia-specific criteria that should be considered along with the accuracy, consistency, completeness and currency of the source texts and the indexes of the database. Obviously, these multimedia examples could not be included in the print version of the keynote speech.

What is multimedia?

We first need to define what qualifies for multimedia? The underlying problem is that there is no agreement in this essential question. Multimedia has become a very inflated term. There are many database publishers who are ready to label their products as multimedia just because there are images in the database.

This seems to be as inappropriate as calling a book that has pictures, photographs, drawings, and maps a multimedia kit. We just call these illustrated books, photo albums, picture books, or atlases.

The same should apply to databases that incorporate these pictorial elements. A few publishers follow this correct terminology, such as World Book Inc. that has recently published the *Illustrated World Book Encyclopedia* with very high-quality pictures.

In my opinion only those databases are genuinely multimedia that offer functionally relevant, substantial amounts of motion pictures (animations or video clips) or sound along with text and/or static images in an integrated way. The quality of multimedia component is crucial, of course, but it cannot be incorporated into the definition. However, there are a number of parameters that can help us to differentiate between low quality and high quality multimedia without even looking at or listening to those multimedia elements. Of course, much of the quality

aspect remains subjective. In the following I will discuss the criteria of functionality, quality and quantity.

Functionality

Multimedia for the sake of multimedia does not make much sense. I found the multimedia elements in the first edition of ZCI's Multimedia Powertalk database dysfunctional, and the entire publication of very little use as a database — independent of my distrust of gurus preaching the get-rich-in-your-lunch-break gospel. This database is the CD-ROM version of four books by infomercial superstar Anthony Robbins and noted author Paul Zane Pilzer. We hear their dialogue and narration as the text screen scrolls down, but it does not seem to enrich the text unless you crave the barking style of car commercials. The pictures of the authors, mostly of the photogenic Robbins, pop up on the screen shot from as many angles as supermodels for a cover picture of a fashion magazine. They add nothing to this business-oriented database.

Animation can add a lot of information to the text as illustrated by the continental drift animation that shows how Australia parted from Antarctica in *Encarta*. However, one could not stop wondering why did we need an animation for the boiling of the water in *Compton's Encyclopedia*, and along with some other irrelevant animation sequences it was removed from the 1995 edition. The very short, comic book style animations in *Multimedia Monarch Notes* are meant to illustrate the works of Homer, Hawthorne, Orwell and Whitman but seem to be an ill-fated 'prunk' in this literary database. They are not only dysfunctional but also distracting.

Narration of a text that appears on the screen can be very functional when it reinforces and enhances the text. This is the case with the *American Heritage Talking Dictionary* that pronounces every entry word, or the *Discis* storybooks for children,

Macmillan's Children's Dictionary that help to improve reading abilities. Broderbund's storybooks go one step further allowing the child to switch between English, Spanish and Japanese language narration. I would have not known that the emphasis is on the first syllable in Canberra had I not heard it pronounced when looking at the map of Australia, and clicking at the name of the capital in the *Encarta Encyclopedia*. It is another question that pronunciation is provided only for Brisbane and Canberra.

In the *American Vista* database it is highly informative that regional speech samples are provided to illustrate how a Hawaiian man, or an Appalachian woman talks. On the other hand, it is misleading when the same sound recording is used in the *World Vista* database for the speech illustration of Australia, New Zealand, Zimbabwe and Botswana. In *Encarta* there are speech samples to contrast French and Canadian French, and Castillian Spanish with Spanish as spoken in Mexico and Argentina, but again there are no English speech samples of Australia, Canada or New Zealand.

Music can also enrich the text and images that are displayed, or can create the atmosphere. In ZCI's *History through Art* series the renaissance music is in perfect harmony with the pictures of da Vinci, Botticelli, and Caravaggio. In *Encarta*, and *Grolier's Encyclopedia* the pictures of the musical instruments are accompanied by their sounds. In *Cinemania* '95 the sample clip from Casablanca is made perfect by playing the song 'As Time Goes By' as Humphrey Bogart confronts Sam, the pianist. On the other hand, the continuous elevator music in the *Great Cities of the World* database is dysfunctional, and annoying when the same tunes are played for Sydney, Rio, and San Francisco.

Lack of audio can render video clips dysfunctional. Had I not recognised some of the sights I would have been puzzled by the flicks without narration of Melbourne, Taipei, Hong Kong, Tokyo, and Manila in the Asia Alive database. One may argue that such flicks are still better than the videoless descriptions of Sydney, Auckland, Singapore, Osaka and Jakarta in the very same database.

Dysfunctionality may result when the audio and/or the video information contradicts the text, or is wrong. In *World Atlas* Bondi Beach of Sydney is mispronounced, and so is the name of the island of Kauai in *Encarta*. It is confusing in the *Great Cities* database when the caption identifies the building as Art Gallery of New South Wales, but on the building itself you see State Library of New South Wales.

The caption in the *Cute and Cuddlies* database reinforces the misbelief that the koala is a bear, and so does the text in the *Heinemann Children's Encyclopedia*. A dark-furred polar bear in *Grolier's Encyclopedia* is confusing for a child who reads in the caption about the importance of the white fur of the polar bear. While the *Mayo Clinic Total Heart* database has plenty of relevant and high quality illustrations, a crucial picture for the CPR operation is not only amateurish but physiologically impossible. The person who helps the victim either has no thumb on the left hand that holds the head, or managed to penetrate the skull up to the thalamus.

Omission of information can be misleading. It is disappointing in *TIME Almanac* that the map of Europe fails to identify Denmark, Luxembourg, and Monaco. Many of the maps of Australia ignore its rivers.

Quality

In addition to functionality and accuracy quality is also a prerequisite for any multimedia elements to be of value. Many of the multimedia elements have quality parameters that are easy to describe, and use for comparison in addition to the highly subjective aesthetic impressions. In the pictorial category, maps have a special status as some extra criteria may be used such as the number of cities, rivers, mountains, deserts indicated and named on the map, the identification of the neighbouring countries, or the localisation of the country or region on the globe.

Primitive drawings can be appalling, such as the map of Australia in the Quanta version of the *CIA Handbook*, especially in contrast with the many beautiful and rich in details renderings of the continent in other databases such as in *Asia Alive, Grolier, Infofinder* and *Encarta*. These databases stand out with their map quality in many respects. They are not only of high fidelity but also adequately detailed, listing many more cities, indicating and naming rivers and deserts on the map of Australia. *World Atlas* offers unique relief maps, and topographic maps of the continent.

In contrast, most of the maps in *Compton's* are less than perfect, the legends are hardly legible, and too few cities are identified. In *World Vista* the city maps for Melbourne and Sydney are of high quality but they were not scaled to the typical monitor so you really do not get an impression about the layout of the cities.

They are still far more useful than maps of Sydney in the *Great Cities* database where the colouring of the map decreases legibility, and the fonts are primitive. *Encarta* stands out among the encyclopedias by offering maps that have hot buttons to pronounce the name of some of the cities shown on the map, and unusual ones like the climate map, the exploration maps of Australia, and the globe with the country highlighted at the top of every country article.

Other pictures, such as drawings, paintings, portraits and photographs, are more difficult to describe in a tangible way, though at least two criteria can be applied for all pictures. One is the colour depth, the other is the size (resolution) of the image. Images most typically are displayed using a colour palette of 2, 16, 256, 32 000, 64 000 or 16 million colours. These are also known as 1-, 4-, 8-, 15-, 16-, and 24-bit images. The two-colour (black-and-white) images are not necessarily inferior to the colour images as is illustrated by the nice black-and-white image of President Nixon from *Bookshelf* versus his rather poor colour images in the *US Presidents* database of Quanta Press. In the colour category *Information Finder*'s portrait is top notch, equalled only by the one in the *Everywhere USA Travel Guide* database.

In addition to the colour depth the size of the image is critical. In many databases images are much smaller than the 11 x 8 inch typical screen. Image size is usually expressed by the number of horizontal and vertical picture elements (pixels), i.e. a full screen image is described as 640 x 480. In many databases the images are only of quarter size, 320 x 240 pixels (5.5 x 4 inch), or even smaller 160 x 120 (2.25 x 2 inch). Though these pictures can often be enlarged they result in heavily pixellatedpictures.

The series of pictures about the Sydney Opera House clearly shows the differences though all of them use a 256 colour palette, and were displayed and captured with the same video subset. All the pictures were enlarged from the original size only to the extent that the application software allowed it, i.e. were not stretched by Windows cursor control. While *Compton's* version of the Opera House became very pixellated on enlargement, *Grolier's* version did not deteriorate significantly, and *Encarta's* rendition of the Opera House was of the same quality after enlargement. One of the reason for this difference is that *Encarta* uses the fractal image compression technology that does not decrease the sharpness of the image as it is being enlarged. Beyond the colour depth and the resolution, the traditional photo-artistic qualities determine the quality of the picture. Though these are not quantifiable qualities they are unmistakable when you look at two pictures from the same database. While *Heinemann's Children's Encyclopedia* had a good quality picture of the Opera House, the photograph of the bay itself is rather poor from an artistic point of view. While the first map of Australia in *World Vista* is plain horrible, the other one is quite rich in details and colour coding, and the photo of Milford Sound in New Zealand is breathtakingly beautiful both as a photo and as a computer image. The cost of a database is not in direct correlation with quality. One of the best maps of Australia comes from the *Medio Magazine* database that costs $5 per issue.

The quality of video and animation is determined by the same criteria as described above for still pictures and by the frequency of frame display. This is logical as motion picture is essentially the very quick display of still pictures (frames), 24–25 per second

in Europe, and Australia, and 30 per second in the USA. It is the video component where the most compromises are made. Videos are often captured in a small frame (160 x 120 pixels) that is one-eighth of the screen, and at a rate of 10–12 frames per second. This yields usually blurry, jerky videos of little value, such as the video about Manila in the Asia Alive database.

In contrast, the same size pictures at faster frame speed in *World Atlas* are much more smooth. In *Grolier* the flick about Australia is not really a video but a narrated slide show of fairly good quality and enjoyable duration of about two minutes.

Audio quality is measured by somewhat similar criteria. Sound can be recorded at 4-, 8- and 16-bit depth (sampling size), and samples can be taken 11, 22, and 44 000 times per second (sampling rate). To relate this to a known quality CD audio recordings are created by taking samples 48 000 times a second (48 KHz sampling rate), and recording the characteristics of the sound in 16 bits. While narration is perfectly adequate at 4-bit depth, and 11 KHz frequency, and bird songs are okay at 8-bit and 22 KHz frequency, high quality recordings must be made at 16-bit depth with 44 MHz frequency. All these audio files take up significant space, and many database publishers cut corners also in this regard.

The most common one is to substitute a real musical performance by a synthesised one, known as the MIDI format. The best way to illustrate this is by playing the Brandenburg Concerto from *Compton's Encyclopedia* in both formats. The difference is convincing even for the absolutely untrained ear. As MIDI files take up hundreds of times less space than true sound recordings it is no accident that all the sources I had consulted for this speech had the Australian (and all other) national anthem in MIDI format of about 45 seconds. I would not be surprised if you had difficulty in recognising it. While most of the musical sounds are in MIDI format in all the databases, the famous speeches, the bird songs, the narrations, the pronunciations, and the folk songs are in WAVE file format which produces much higher fidelity, as illustrated by the sound of the laughing kookabura, the red bird of paradise, as well as by the Aboriginal ceremonial chant

in *Encarta*. It is essential, of course, to have hardware that can play back the multimedia elements at their best level. While a 16-bit sound card cannot improve an 8-bit recording, a 16-bit recording certainly deteriorates when played back on a cheap sound system featuring toyshop quality speakers. Similarly, a single speed CD-ROM player will skip too many frames when playing back video clips optimised for double-speed players.

Quantity

We all know that quantity does not make quality, but there must be a minimum amount of sound and/or motion picture to qualify a product as multimedia. The CD-ROM edition of *Money Magazine* cannot qualify even if it has two video clips of three to four minutes' duration that welcome the user, and give an overview of the features, respectively. They are not an integral (or even necessary) part of the product. On the other hand, a very large number of images is not necessarily an asset. *Compton's* used to boast in ads to have the most images (about 13 000). Many of those images were very small, postage stamp size pictures of low quality. For the 1995 edition *Compton's* reduced the number of images to 7000 by eliminating the worst ones, though some managed to survive, as illustrated by the photo series of the jaguar.

We would badly need some objective, tangible and quantitative way to describe the amount (and the quality) of multimedia elements of a database. The problem with quantifying the multimedia elements of a database is that different measurements are used, and many of them are so uninformative as if we were to describe a movie as 2300 feet of celluloid instead of 82 minutes of play time. *Cambrix*, in its excellent Adventure multimedia travel catalogue, indicates that it has 400 MB of video. This is an impressive number for a 660 MB medium, but it is not informative. 400 MB of video could mean 45–50 sec of video or almost an hour of video depending on the colour depth, the frame size,

the frame per second rate, and the compression technique used. It would be much better if the parameters of the video clips were specified as in Table 1.

Table 1 Video clip parameters

Number of video clips	8
Colour depth	256 colours (70 clips) 16 colours (10 clips)
Picture size	320x240 pixels
Recording speed	15 frames per second
Total duration	50 minutes

Similarly, for the audio information it would be more informative to specify the number of soundbites, the ratio between MIDI and WAVE files, and the recording depth and frequency of the latter, as in Table 2.

Table 2 Audio information

Number of MIDI soundbites	740
Total duration of soundbites	9:25
Number of WAV soundbites	20 8-bit, 11 KHz: 70% 16-bit,44 KHz: 30%
Total duration of soundbites	0:24

There will still remain a number of intangible, unquantifiable elements that can be only explored by using the database. At least the quantifiable characteristics should be made known to the potential customers. There is no doubt that many of them will not understand such issues as frame rates, colour depth, etc., but this used to be the excuse for withholding technical details about cars and nutritional information about foods, and look where we are now. Within a few years the average customers will be much more educated about multimedia, and will demand to know such information before parting with their money.