

Digital Archive Strategies and Solutions for Radio Broadcasting

In this article we review coverage of digital archiving issues for radio broadcasting from the 116th AES Convention, held in May in Berlin, Germany.

A WORKSHOP ON RADIO ARCHIVING TOOLS: STRATEGIES AND SOLUTIONS

Klaus Heidrich, chair, explained that this workshop, designed to deal specifically with the needs of radio broadcasters and originally entitled "Comparison of Existing Archiving Tools," had been given the subtitle of Solutions and Strategies. He was familiar with a number of recent digital archiving projects. The successful ones always had an overall strategy designed to provide a solution to the archiving problem. This ultimately gave rise to specific tools. The field as a whole, however, did not conform to a fixed scenario but was something of a moving target.

Heidrich began the workshop by introducing the panel: Ernst Dohlus, Bavarian Radio, head of production and playout; Wingolf Grieger, Nord Deutscher Rundfunk (NDR), system coordination for digital archives; Niko Waesche, IBM Business Consulting Services; Rainer Kellerhals, Tecmath AG, executive VP of product and solutions; and Karl Pieper, general manager of VCS Media Broadcasting Solutions.

INTRODUCING THE QUESTIONS

Heidrich evaluated reasons for introducing digital archives and asked some important questions that he hoped the panel would address. Commonly used

terms in this field are audio archives, content management, and media-asset management; but the distinction between these terms is not always completely clear. The term essence is often applied to the content that is stored, whereas metadata is used to structure and describe the content. Information about the right to use the material completes the picture.

The task of migrating to a digital archive involves dealing with legacy storage media such as tapes and the handling of legacy database structures that require the addition of new metadata. Last but not least there is the question of "how to get along with rights management," which is usually the responsibility of administrative departments and which, on the basis of current experience, may not always be in the form of concise software and systems specifications that can be integrated with a digital archive.

There is a wide variety of different legacy solutions already in use for database management, including more recent relational databases. The question arises as to whether it is possible simply to add mass storage for digital content to existing database structures. In general, it turns out that this is not adequate. In fact, it is necessary to incorporate features that are specific to digital archives, such as browsing and coping with the range of different audio codecs involved. Interfaces to the digital archive solution are crucial to the success of a system. Business management systems for rights management also need to be integrated. A digital archive solution therefore requires a qualified integration concept

to ensure that everything works together.

The wide use of digital playout systems has been the driver for incorporating IT tools into the archiving process. Heidrich asked whether we are seeking a solution that improves quality and saves money, or perhaps something to preserve existing analog assets, or maybe a business requirement to serve the next generation of digital platforms for on-demand program delivery. Or is it a market-driven change led by the advance of technology?

Digital archive projects tend to be fairly expensive, but there are at least three elements involved. There is the equipment, of course, but there is also the time to transfer the legacy archive. Then there is the cost of ownership over the lifetime of the system. The question remains, can we justify such an investment? Can business processes and workflow be improved, and can new business be generated as a result?

The topic of convergence has been around for years (that is, convergence between the different broadcasting and content delivery media made possible by information technology). Is the introduction of digital archives closely related to convergence? From an operational point of view, it is interesting to consider whether the operational process is driving the solution or whether the solution is driving the process. Related to this is the question of whether an archiving solution can really be an off-the-shelf product. In particular, does an archive system have to be broadcast-specific or can it be a generic solution?

Standards, common practice, and common middleware solutions are all issues to be considered, as are standard interfaces and metadata structures. Convergence is a thorny subject and liable to fill many workshops in its own right. However, it was considered very important. "Everything should be made as simple as possible but not simpler!"

WHY SHOULD A PUBLIC RADIO STATION ADOPT A DIGITAL ARCHIVE?

Ernst Dohlus from Bavarian Radio asked why a public radio station in Germany should adopt a digital mass storage system and transfer its program archive to this format. Such a project is inevitably complicated and expensive. All of the companies offering content-management systems inevitably have to develop software before they can deliver products. They often run short of time and postpone delivery dates. There is no model solution for radio archives, so each project involves new software development. Radio broadcasters are often offered solutions that were originally developed for TV operations, and these are rarely suitable without considerable modification. Digital archiving solutions can therefore almost never be used right out of the box. Complex interfacing problems arise, making the introduction of these systems complicated and expensive. So why are we so attracted to such solutions, and why would we want to participate in such a risky venture? It is because the wave of technical developments drive us forward. Vintage equipment, for example tape machines, is becoming increasingly rare and difficult to maintain and operate.

Although some old tapes are gradually deteriorating, in fact most of the content in existing archives does not suffer from this problem, so media degradation is not the key driving factor. Actually, it is changes in broadcast production pulling us to a new solution. Journalists, editors, and producers are used to browsing for material on computers, and it is becoming common practice to find all this material online in one way or another. It is no longer considered possible to do effective production work without the introduction of a mass-storage archive for radio

content. For example, in future, the music industry will deliver audio files instead of CDs, and this content has to be kept in an archive unless it is only to be used temporarily.

NDR has calculated that its mass-storage project—involving 90,000 hours of old tape material and 2,000 hours of new production per year—requires an investment of around 7M Euro as well as 10M Euro for the outsourcing of digitization. Each hour of archived material, therefore, costs about 180 Euro. Why should we make this kind of investment? Are we attempting to perform a cultural role, digitizing the cultural radio heritage of the nation? Probably not, considering the competitive business of radio today, but certainly we should preserve important examples from radio history. Are we digitizing in order to facilitate new business solutions, such as educational programs on demand? This may be the case for commercial broadcasters, but public radio in Germany is limited in the extent to which it can undertake such commercial operations. The copyright laws also make such enterprises complicated. For example, repeat fees to artists are common, especially for older material when contracts did not explicitly state that material could be reused many times. No, the primary reason for moving to a digital archive is the way in which production techniques and delivery methods are changing, driven by new technology, making conventional recording equipment and techniques largely obsolete.

A SOLUTION OR THE START OF NEW PROBLEMS?

Wingolf Grieger asked the provocative question: "Is a solution actually a solution or is it the start of new problems?" The project at NDR, active since 2002, is called Digital Long Term Archive (or DELA, to abbreviate the German title). Historically, and still to some extent today, people have been employed simply to find material in the vaults and bring it out. Finding it usually involves a database of some sort. As more and more material with more channels is generated, the conventional archive finds it increasingly difficult to cope and will eventually be paralyzed by complexity. Unless the archive is digitized there will no longer

be a functional archive, and a radio house without an archive "is like a kitchen without a pantry."

The question of workflow was addressed by Grieger, in the form of a question as to whether the workflow of a traditional carrier-based archive is compatible with the workflow of a digital system. Traditional material has to be stored, documented, and ultimately brought out again for broadcasting. In the digital world there is no huge vault containing tapes and disks, but a mass-storage archive of some terabytes occupying about 6 square meters. In fact almost none of the documentation at NDR had to be changed with the switch to a digital archive. They are still using the same database as before. The new system actually made the process of documentation easier.

An important test of a system is in its approach to the handling of errors; this involves primarily the errors of operators as opposed to software errors. The archive number of an item, for example, acts as the unique key to the content and its location in the traditional archive. It acts in a similar way in relation to audio files in the new system. However, if an operator errs in the specification of this number, then a tape might be lost in the archive, and human ingenuity is needed to find it, either by accident or design. Such errors will always occur, and in the digital domain a tool is necessary to enable humans to intervene in similar ways so as to be able to correct errors and search for things that may have been erroneously labeled. Grieger therefore stressed the need not to lose the possibility for human control.

THE MANUFACTURERS SPEAK

Niko Waesche from IBM wanted to reinforce the point made by Dohlus that digital archives are of vital strategic importance to the future of broadcasters. Process change in broadcasting is the key issue. Archives have to be shifted from being a "luxury item" to a core technology. However, if operating costs after the implementation of a digital archive become greater than they were before, many of the advantages will be negated. Owing to the way in which technology has changed the process of archiving, it is no longer separate from the remainder of the

broadcast production process. Because of the integration between storage and production, and because of media convergence, the concept of a separate archive is no longer particularly meaningful. While the idea of creating a cultural-heritage archive may indeed be a luxury, it can be a by-product of implementing a digital archive for process change reasons. Radio broadcasters may be able to “hit two birds with one stone.” They may also find that their overall operating costs can be reduced. Waesche was not in favor of the so-called turnkey solution, partly because of the negative aspects of being locked into a single supplier with a niche solution. Technical support and skills could become problematic later on.

Rainer Kellerhals from Tecmath, a specialist supplier of radio asset-management systems, spoke about some of the issues that he had encountered in projects to date. Typically, he said, a key issue relates to the cataloging system already employed by the organization, some of which have been in place for years. If the existing system is to be kept, then the question is one of how to interface to that system and how the user interface is to be integrated. If the organization chooses to replace the existing cataloging system, then the challenge is to find a suitable supplier and develop a way of migrating the metadata from one catalog to the other. One of the biggest problems, as Ernst Dohlus pointed out, is to find a suitable product that can handle the task. The market for software solutions to these issues is relatively small, probably only consisting of hundreds of companies in the world.

Audio editing systems need to be interfaced to digital archives, but editing systems have typically been stand-alone black boxes that have not had the necessary APIs (application programming interfaces) to interface to other systems. Gradually this is changing as manufacturers are being encouraged by users to make this more possible.

It is well worth spending the additional time on a project to ensure a high level of integration and automation, as it will give benefits in the long run in terms of usability and speed of operation. The more the introduction of a digital archive goes beyond a technology deployment issue and becomes

a “change management” process the better, because changes in workflow are implied. Involving the users in the project and educating them during the process is vitally important to its ultimate success. Data entry has to be carefully considered. Repeatedly having to enter similar data is very time consuming and inefficient, so options to use master data are needed.

In the short term the benefits to the organization will tend to be relatively soft in terms of return on investment (for example, greater ease of use), but in the long term some more concrete benefits are likely to accrue. In any case, it seems there is no way of avoiding the introduction of digital archives—they are here to stay.

Karl Pieper from VCS Media Broadcasting Solutions began by speaking about the financial aspects of digital archives. Much of the work of transferring material to an archive is time-consuming and labor-intensive, and more automatic tools are required to assist in this process. In fact, the investment element of a digital archive project (that is investment in new technology) is generally a much smaller component of the total cost than that of conversion (of old tape-based material into a digital form).

The archive should really be a servant to broadcasting production rather than the other way around. The technology should not drive you into complicated changes of workflow, because this makes the acceptance of a system much easier. If you choose the right system architecture, then a ratio of about 80% off-the-shelf products to 20% customization or adaptation should be possible. However, the environment in which an archive operates will change as the years go by (for example, new metadata structures may be introduced), so the system must be adaptable to such changes, and interfaces should be flexible enough to accommodate changes in configuration. We know that mass-storage technology will change in the future. Therefore a radio archive solution really has to be independent of any particular mass-storage type or device. Standards develop slowly, and it is probably best not to wait for them but to enable the system to be adapted to new standards as they arise.

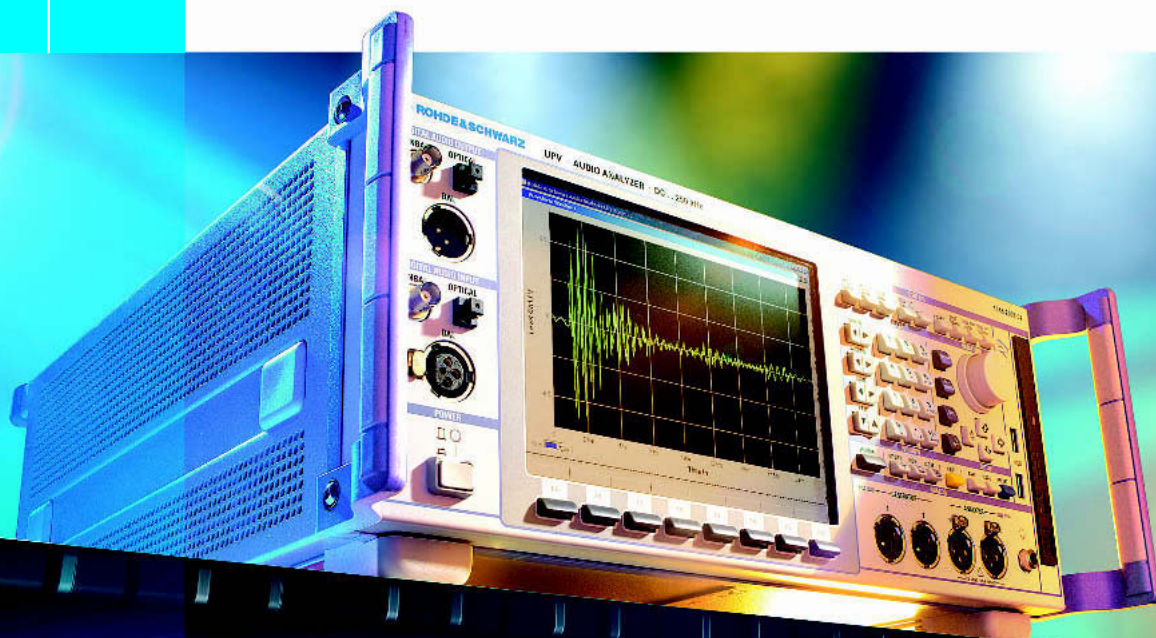
THE DELEGATES DISCUSS

During a discussion session delegates considered the topic of what to keep and what to throw away when transferring an old archive. Who will be the judge of quality? Who dares to make that decision? In fact, it might be cheaper to digitize everything rather than to employ people to decide what to keep. After having done this at NDR, some totally unexpected content has subsequently been used by producers. Some of that material would almost certainly not have been transferred if a decision had been made in advance. Once archive content becomes available on an online database it is often surprising what people decide to use.

As time goes on the functionality that can be delivered by out-of-the-box products is getting greater, thereby reducing the amount of custom work required. Digitization-on-the-fly or on-demand could be quite a popular approach with broadcasters, whereby tape-based archive material is digitized only when it is requested by a producer or editor. However, in many cases this process is farmed out to agencies that might need a few weeks notice to carry out the digitization. At NDR this process accounts for only 3 to 5% of their digitization activity. Ernst Dohlus pointed out that while they knew that perhaps half of their archive might be used at some point in the future, they did not know which half. Hence, digitization on the fly might provide some clue as to the most desirable material. Nonetheless, a producer would not necessarily know what was worth using if he could not listen to it in advance; and he would not find it if it had not been digitized, which makes a strong argument for digitizing as much as possible up front. The ability to browse the archive has often provided inspiration to journalists to make a program on something they might never have thought of before. A mixed approach to digitization is needed.

A question was asked about how the radio houses will continue to be able to play back tape-based material, now that tape machines are becoming obsolete. It was reckoned that in fact there are very large numbers of analog tape machines still in existence and that sufficient numbers of them can be

Pure audio performance



Please visit us at the AES

117 Convention in San Francisco

Booth 1235

The R&S®UPV is the most advanced audio analyzer in the world

Rohde & Schwarz presents the new R&S®UPV Audio Analyzer. Its performance makes it the new reference standard – there is simply nothing else like it.

The R&S®UPV pushes the limits of technology for broadcasting and R&D, and is an extremely fast instrument for the production line. It handles high resolution digital media just as easily as analog measurements up to 250 kHz.

With all of its performance the Windows XP-based R&S®UPV is simple to operate. It's also expandable, and remarkably affordable. Take a look at its capabilities, and then contact us to find out more.

- ◆ For all interfaces – analog, digital and combined
- ◆ Real two-channel signal processing for maximum measurement performance
- ◆ Digital AES/EBU interface with sampling rate up to 192 kHz
- ◆ 250 kHz analysis bandwidth
- ◆ Recording and replaying of audio signals
- ◆ Overlapping FFT analysis
- ◆ Expandable, with easy addition of further audio interfaces like I²S
- ◆ Built-in PC with Windows XP operating system
- ◆ Intuitive, ergonomically sophisticated user interface for efficient operation

www.upv.rohde-schwarz.com



ROHDE&SCHWARZ

pushing limits

kept going for at least the next ten years to make this possible.

Another questioner asked about quality levels when material is digitized and what is being done with the tapes after digitization? In response, one of the German broadcasters said that they are storing material in linear PCM format at 16 bits and 48 kHz and that currently they kept the tapes afterward. In fact, deleting a tape is more expensive than one might think, and one cannot be sure that the digitization will have been done correctly, it was suggested. Another broadcaster said that they have decided to keep the tape for two years after digitization, after which they will destroy it. One delegate questioned the quality of 16-bit, 48-kHz digital audio for the treasures that might be stored in some of these archives. The broadcasters expressed satisfaction with the current standard however, and said that it had been decided at a point in time when the economics of storage were a key issue. The quality of analog-taped material in the archive was not so high that it pushed the limit of the digital standard employed. Of course archives have to be reliable, so that material is not lost or damaged, and they need to use a quality of mass-storage media that is comparable to that used in banks and other similar secure environments. Methods of backup and automatic copying are also employed.

Summing up the workshop, Heidrich asked the panel for a final statement after having heard what all the others had to say. They agreed that the specification phase of the project is absolutely crucial and that broadcasters have to be clear about the business goals for the project to ensure that you get more honest answers from potential suppliers. Journalists at NDR really love the digital archive that has been introduced, so it has been a great success in the eyes of the production staff. The same type of questions discussed here arose when digital production and playout systems were introduced some ten years ago. These now seem to be very much the norm.

TAKING CARE OF TOMORROW BEFORE IT IS TOO LATE

A paper (AES number 6009) by Nicolas Hans of Dalet Digital Media and

Johan de Koster of Radio Netherlands, also presented at the AES 116th Convention in Berlin, suggested some new approaches to a pragmatic archiving strategy based on case studies of real projects. They began with some striking statistics suggesting that by 2020 it will be possible to store 1.4 million hours of audio content online for a cost of less than 100 Euros. One petabyte (one thousand terabytes) of storage will hold 165 years worth of continuous broadcast material encoded at 16 bits, 48 kHz. To store the equivalent in analog-tape form would require 2.5 million items spread over 80 km of shelves.

Echoing the comments of one of the workshop panelists noted earlier, these authors suggest that archivists need to move out of their traditional environment—the basement archive—and into the production space. Integrating the archive with the production workflow ensures that metadata are properly and quickly collected as part of the production process and that search and retrieval are made easier. In fact, it may even be possible to generate some metadata automatically in the future, thanks to the development of systems that analyze the audio material for speech and other content, generating metadata to enable the cataloging and searching of radio content. An example of this is described in a separate paper by Löffler et al., “Automatic Extraction of MPEG-7 Audio Metadata Using the Media Asset Management System iFinder,” from the AES 25th International Conference.

Although many institutions regard archiving onto CD as the most cost-effective approach, the handling of these media is an expensive operation. For example, the CD has to be burned, labeled, barcoded, and shelved. Consequently the archiving of one year’s worth of audio costs nearly 40,000 Euros, of which 80% is labor. The corresponding storage in an integrated online system using hard drives amounts to about half that amount, where the handling operations are limited to triggering file copies. The cost of the equivalent amount of analog tape storage is close to ten times the cost of CD storage, even if the space could be found. The key to the authors’ argument is that “virtual car-

riers” are the way of the future. In other words, the search for the ideal long-term physical carrier for archiving is a relatively fruitless exercise because new media continue to come along. An archive becomes a logical space that is independent of the production environment, whose physical attributes may need to change as time passes.

One of the main problems of digitization is when and where to start. New material continues to be added to the archive while we are digitizing the older material. Considering what should be digitized and when, another question raised in the workshop, Hans and De Koster referred to the experience at CBC in Canada, where it had been found that 50% of requests were for material that had been broadcast in the preceding 12 months, so the organization stopped adding tapes to its archives by first digitizing its ongoing broadcasts. The point they make is that such an approach gives rise to a fast return on investment. Digitizing current programs requires much less expenditure than transferring past recordings.

The digitization of archives can also lead to new business opportunities and increased use of the archived material, as suggested by workshop panelists. For example, at RAI in Italy the video archive experienced an 85% increase in use after digitization. Furthermore, a number of new digital television channels dedicated to historical recordings emerged.

Based on this workshop we can see that the digitization of radio archives is here to stay. It is not a question of if but when. Many radio houses have already embarked upon this operation, and more systems are available. The increased production flexibility and business opportunities that arise help to justify the high costs of such projects.

Editor’s note: The two papers mentioned in this article—and all other AES *Journal*, convention, and conference papers—are now available for online purchase. Go to <http://www.aes.org/journal/search.cfm> for *Journal* articles or <http://www.aes.org/publications/preprints/search.cfm> for convention and conference papers.