NATIONAL ACADEMY
OF SCIENCES OF UKRAINE
INSTITUTE FOR INFORMATION RECORDING

НАЦІОНАЛЬНА
АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ РЕЄСТРАЦІЇ ІНФОРМАЦІЇ

ДОВГОТРИВАЛЕ ЗБЕРІГАННЯ ЦИФРОВОЇ ІНФОРМАЦІЇ

ПРОЕКТ
«УКРАЇНСЬКА НАУКОВА КНИГА ІНОЗЕМНОЮ МОВОЮ»

КИЇВ
АКАДЕМПЕРІОДИКА
2018

LONG-TERM STORAGE OF DIGITAL INFORMATION

PROJECT
«UKRAINIAN SCIENTIFIC BOOK
IN A FOREIGN LANGUAGE»

KYIV
AKADEMPERIODYKA
2018

UDK 004.33
L87
doi: https://doi.org/10.15407/akademperiodyka.360.148

Authors: V.V. PETROV, Z. LE, A.A. KRYUCHYN, S.M. SHANOYLO, M. FU, Ie.V. BELIAK, D.Yu. MANKO, A.S. LAPCHUK, Ye.M. MOROZOV

Reviewers:

P.F. OLEKSENKO, corresponding member of NAS of Ukraine, Dr. Sci., Prof., Head of Department of the V.Ye. Lashkaryov Institute of Semiconductor Physics of the NAS of Ukraine

Yu.A. KUNITSKIY, Dr. Sci., Prof., Head of Department of General Physics of the Taras Shevchenko National University of Kyiv

Approved to press by the Editorial Publishing Board of Institute for Information Recording of National Academy of Sciences of Ukraine (August 29, 2017, Protocol No. 11)

Publication was funded in the frame of the Targeted Complex Program "Creation and Development of Scientific Publishing Complex of the National Academy of Sciences of Ukraine"

L 87 **Long-term** storage of digital information / V.V. Petrov, Z. Le, A.A. Kryuchyn, S.M. Shanoylo, M. Fu, Ie.V. Beliak, D.Yu. Manko, A.S. Lapchuk, Ye.M. Morozov ; National Academy of Sciences of Ukraine, Institute for Information Recording. — Kyiv : Akademperiodyka, 2018. — 148 p.

ISBN 978-966-360-360-5

Long-term storage of digital information is an important scientific and technical task in the conditions of rapid growth of the amount of information presented in digital form. A key poit of the problem solving is creation of special media for long-term storage of strategically important information, scientific and technical information and information representing national cultural heritage. Special type optical media development from highly stable materials for long-term storage of information is considered.

UDK 004.33

ISBN 978-966-360-360-5

INTRODUCTION

In the last decade in connection with the intensive introduction of information technologies in all areas of modern society and, consequently, the transfer of information in digital form, there is a very important problem of long-term storage of data in digital form. Before the era of Informatization, this issue was not as relevant as information recorded on paper, parchment and other media had a shelf life significantly greater than the average life expectancy of the person [1]. The growing volume of information represented in digital form, leads to the fact that the question of how storing information on digital media is becoming increasingly important. The most important part of these data, scientific discoveries, historical facts, books, movies, and more — to mankind it is important to preserve for many centuries. However, modern electronic storage media are not able to exist for a long time. The period of storing data on optical disks is on the order of 10-15 years, and other type memory also are durable. In addition, such information carriers are quite fragile and any external impact can easily collapse [2].

There are a number of engineering tasks and applications that need long-term storage of information. More and more complicated requirements for complex engineering structures, increase the demands on their reliability. Diagnosis of the condition of bridges, nuclear reactor vessels, piping, etc. is carried out using the methods of acoustic emission. Due to the fact that the lifetime of such structures is tens or even hundreds of years, it is necessary for the whole of this period to maintain a sufficiently large volume of the original spectra of acoustic emission, improve their processing, to get more information about the changes of these spectra, and hence the condition of the facilities. For example, financial documents, bank information has a shelf life of 50-100 years, medical information on human health must be maintained during its life, that is at least 80 years, and the information, which is a global scientific, historical and cultural heritage in General has to be kept forever. If earlier it was considered necessary and was able to provide long-term storage of only those documents whose importance was clear to us initially, but now there is the opportunity to go to the archives much wider range of content, the value of which is not known in advance, but which may be useful in the future. This reflects the change in the understanding of how specific businesses and society in General the importance and value of information resources, and on the other hand, the growing technical capacity to store huge amounts of information [91-95, 175-185]. Currently, the preservation of data is mainly obtained on paper, microforms or digital media with permanent restoration. However, there is now considerable concern regarding the potentially catastrophic risks of information loss. In general, it is now recognised that the different techniques of information storage are not long term secured against the risks of exceptional events.

Scientists and engineers not for the first time look to create media capable of storing information over time, comparable to the existence of mankind. The problem of long-term archival storage of data relevant to the state archives, state library funds, medical institutions, industrial enterprises. To date, the high cost of storage associated with the storage sizes, as well as the need of rewriting to new. Long-term preservation of digital information is an important issue in modern world. From the whole society to one single person, there always are some data having long-term value to be preserved [91-94, 175-185].

The world is becoming increasingly digitised, with more and more sources of data needing preservation. The sheer volume of data is a challenge that impacts most of the other challenges thus making them inherently larger and more complex. The costs of storage, downloading and ongoing maintenance, for example, are much more than they'd be with smaller volumes of data. Deciding what should be preserved and when to take preservative action becomes more complex with a larger volume of data and a wider range of storage media. This in turn increases the risk of failing to preserve items that will one day turn out to be of historical value. There is also a higher risk of data becoming indiscoverable due to the associated risk of poor metadata. Large volumes of data cannot be handled on a one-by-one basis, as can analogue materials. While the extent to which digital material can be handled at an individual level will vary according to purpose, funding and so on, in general, the preservation methodologies applied to large amounts of data will need to be scaled to fit the needs and objectives of the preserving institution. More data will inevitably mean increased reliance on automation and the development of new workflows to handle the data. Along with that comes the need for increased computer power and storage. Large volumes of raw data require tools to scale or present it in more manageable forms. This is likely to lead to expectations that the preserving institution will also provide the infrastructure to mine or manipulate the data, much as they already do with existing digital materials, such as newspapers. One of the current realities of increasing data volumes and storage expectations is that preservation software, front-end software or the storage system itself may be unable to cope with large files, multiple versions of the same file, or just with large volumes. This may require new infrastructure and software. In a world where archival data is accumulating at a rate of 250 petabytes each year, safely storing that vast amount of information has become a critically important issue, especially for fast-growing enterprises like Facebook, YouTube and other social media platforms.

CHAPTER 1

LONG-TERM DATA STORAGE PROBLEMS

1.1. Actual tasks of data storage industry

All branches of natural sciences report one common result: our world is continuously changing. We live in a dynamic world and structures that seemed to be fixed are actually evolving; however on a time scale much longer than human life or the human memory.

Natural sciences uncovered many of the rules that govern this change. Some processes are rather simple and well understood, such as the lifecycle of a star of the mass of our sun. Others are very complicated, such as the varying orientation of the magnetic poles on earth, which are not well understood.

The rules themselves do not seem to change, however. Cosmologists make observations and theories on the development on the very grandest scales of the universe, and on the longest time intervals. This is possible because the speed of light is finite, therefore, when astronomers aim their instruments at distant objects they also look into the past. What can be seen suggests that the rules have been the same since the birth of our universe, 13.7 billion years ago.

The memory of our personal lives goes back a little beyond our actual lifespan: we heard stories told by our parents and grandparents, only occasionally has someone the luck of great-grand parents. Beyond theses generations the past becomes very abstract. We oversee only a small fraction of the roughly 200,000 years Homo sapiens sapiens existed. The oldest written script seems to be Cuneiform script developed by the Sumerian culture which dates back to the 23rd century before our time. The Gilgamensh epos is a prominent example of very early texts. From this epos we get an impression about the ideas of the people of their time; hence we oversee only a little more than 4,000 years of history of human thinking — this is a mere 2% of the existence of Homo sapiens sapiens.

Notwithstanding these observations, we have a keen interest in our history. New discoveries of ancient artifacts make headlines in

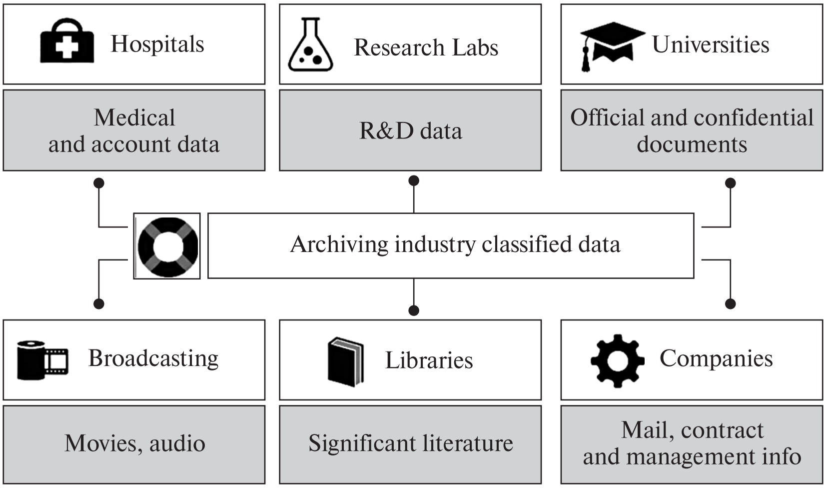


Fig. 1.1. Background of data archiving

the papers, and the discovery of a civilization under Antarctica's ice, no matter how old, would be a sensation, even more so if it were not of human origin. This fact constitutes a strong argument in support of storing data on human kind for a very long time [89-94].

Determining how long something will last has long been avery important area of study for science and technology, particularly in materials and coatings. Many advances have beenmade, and much is known today about how to reliably predict the life expectancy (LE) of a product, based on the materials used to make it and the conditions of its use. These advances are readily applied to the field of data storage.

The processes of relevance for high-density data storage over very long timescales are due to the expansion of the universe, astrophysics which describes the formation and evolution of stars and dynamic processes within our galaxy, the local cluster of galaxies and the solar system, geological processes on earth such as erosion, mountain formation, plate tectonics, volcanism and atmospheric processes, and finally biological evolution.

A target timescale for the human document project is of high importance: the timescale defines the properties of the media and the system and will tell us something about the potential readers—human or non-human, and on the possible locations where the system can be stored.

It is found that one million years is a sensible aim. However, it is problematic to store a very large amount of data on earth. Such a system might be implemented on the moon. It might make sense to fabricate many copies of small amounts of information and distribute them over the earth leading to a high probability that a few will be found in a million years time containing retrievable data.

To ensure that knowledge about human life is available for many future generations or even future lifeforms we require a form of data storage suitable for storage at extreme timescales.There are of course some requirements for such a data storage system. The system should be able to survive for at least the required time without losing its content [89-95].

As shown in Fig. 1.1, there are a variety of data that need to be archived, depending on the business or operation, for example, health records in hospitals and clinics, and design plans of buildings and machines. In North Americalaws and regulations require that companies keep evidence of their business activities for a long period of time—for example, five, ten, or 15 years.

The most fundamental component of digital preservation is managing the digital objects in archival repositories. Preservation Repositories must archive digital objects and associated metadata on an affordable and reliable type of digital storage. There are many storage options available; each institution should evaluate the available storage options in order to determine which options are best for their particular needs [3].

Cost is probably the single most important factor when considering long term storage. Cost may be a limiting factor in the number objects that are preserved. Storage costs, even if they are declining, may influence decision makers to select a low-cost storage option, at the expense of essential preservation needs. Cost often lists as the first factor in choosing a storage media option.

1.2. History of analog data storage technologies

We have as much information from the ancient world as we do because of its ability to survive neglect. Only a tiny fraction has survived. Works by Homer, Aristotle, Biblical authors, and many others are irretrievably gone, and much of what we do have is the result of lucky accidents, with documents placed where they wouldn't deteriorate. Even when monks conscientiously kept documents safe in their monasteries, all they had to do was protect them from damage, not actively update them. Today's leading forms of digital storage simply can't survive that degree of neglect [9].

Information preservation is one of the most important issues in human history, culture, and economics, as well as the development of our civilization (Fig. 1.2, 1.3).

While earliest information was recorded in carvings on stone, ceramic, bamboo, or wood, the development of civilization paved the way for new storage media and techniques for recording information, such as writing on silk or printing on paper. Eventually we were able to put photo- graphic images on film and music on records. A revolutionary change occurred in the information storage field with the invention of electronic storage media.

Keeping information for a long time has always been a challenge. Thermodynamics doesn't favor information lasting a long time and so to make that happen people have to spend effort and energy. Deciding how to create a long-term archive involves choosing the right storage system with the right technology under the proper environmental conditions. This can be combined with migration and replication

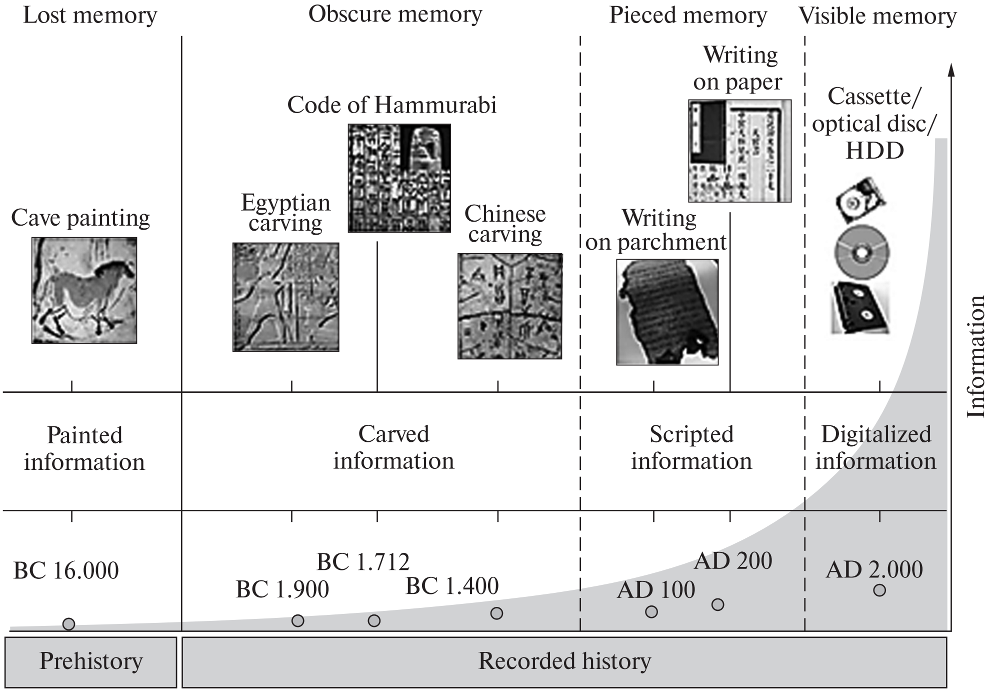


Fig. 1.2. Historical evolution of the recording of information in human society. AD, Anno Domini; BC, Before Christ; HDD, hard disk drive [59]

practices to improve the odds of keeping content useful and accessible for an extended period of time.

Italian researchers may have discovered the oldest nativity scene ever found, predating Christian nativity art by about three millennia. The rock painting depicts a newborn between parents, a star in the east, and two animals. It was discovered on the ceiling of a small cavity in the Egyptian Sahara desert [4].

Five key factors of loss of information, records and knowledge of landfills and contaminated site shave been identified. First of all there are technical or environmental factors, such as the destruction of archives, insufficient updates or lack of records such as maps or plans, or the loss of memory of ancient disposal sites due to simple site degradation. A second group of factors are related to economic conditions, particularly the lack of funds to fulfill the requirements of synthesis of information or preservation of documents and archives.

Human factors as a of loss of information, records, knowledge and memory are also very important: some factors are e.g. loss of knowledge and memory when staff leaves before knowledge is written down (e.g. retirement, change of job), ignorance or incompetence, the deliberate intention to forget about something annoying, or even voluntary criminal records and knowledge destruction. Structural factors describe structural deficiencies or factors leading to social ruptures during crisis or basic

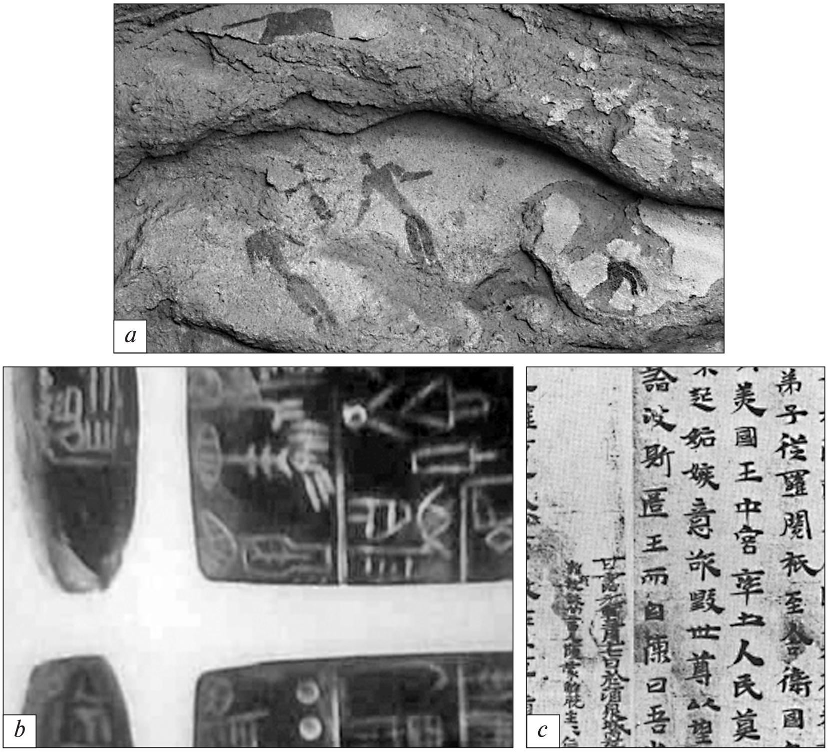


Fig. 1.3. Historical examples of information preservation: *a* — italian researchers may have discovered the oldest nativity ever found; *b* — the oldest known “written” document in the world is the “Kish Tablet”; *c* — a page from the oldest surviving paper book, Phi Y¨u Ching, written on “liu-ho” paper made in Liu-ho

changes in the organization of administrations or companies. Another category of factors that has to be taken into account is the lack of regulation or laws with regard to how to deal with information, records and knowledge.

Depending their nature and the way they are preserved, media are degrading at a different speed. Magnetic tapes are degrading quicker than film. Discs have proven by the past that they are very easily polluted and deteriorated by chemical or ambient agents.In fact, the capacity of magnetic tapes, films and discs for keeping their information decreases very slowly with the time up to the moment it decreases very quickly. A classical case is the vinegar syndrome that it is caused by the degradation of acetate films. As soon as it smells vinegar, a film is difficult to play as it starts shrinking and twisting. Its degradation accelerates and it becomes rapidly impossible to use. The same case happens with magnetic tapes. Such tapes are all more or less hydrophilic and this characteristic leads to the degradation of binders that in turn leads to particles cutaway and difficulties to read tapes. The combination of technology obsolescence and media degradation combine together to close the digitization window. Depending the case, the media or the player may disappear first. Whatever the case, the result is similar. Without a playing device, it is impossible to read a media. And a deteriorated media cannot be played, whatever the equipment you have. As these 2 phenomena occur at about the same time, it reinforce predicting 2025 for the end of the digitisation window. After that, some magnetic formats will still be available for digitisation at a higher cost while other may be just impossible to transfer. For films, the problem is different as preservation can be done better and film scanners are still manufactured today. This must be tempered by the fact that the soundly recorded on magnetic tapes.

Today we can store more information than ever before, but its durability has gone down. File formats become obsolete. So do storage devices, and they fail physically. Anything we put on a disk today will almost certainly be unusable by 2050. The year 3016 just seems unimaginably far. Yet we still have records today from 1016, 16, and even 984 B.C.E. How can our records of today last a thousand years? [9]. Digital data errors can be introduced by the communications system transporting data from one place to another, by the mechanical systems writing and reading the data onto media, by deformations in the media such as spots or micro-level warping, and a host of other causes related to the storage media. From a narrow storage perspective, a primary factor influencing the number of data errors is the storage density of the medium. Regardless of the medium, storage of digital information has always included some kind of error detection and error correction mechanism so that data can be retrieved error-free. A number of utility programs have been written for magnetic disk based systems to help users determine the location of these errors, to relocate data to other areas of the disk, and to reconstruct the data that has become partially damaged. Unfortunately, similar utilities suitable for the general user do not exist for optical media. The close mechanical tolerances in optical media and systems require very powerful error detection and error correction schemes to ensure reliability of retrieved data. Optical systems typically provide a statistical probability of error of only one byte out of every one billion bytes. The application of error detection and error correction schemes to achieve this level of reliability is automatic and transparent to users. However, the error correction schemes are limited to handling error rates below five out of every 10,000 bytes. Once this limit has been exceeded, the error correction scheme can no longer compensate for or guarantee correction of all errors, and the optical medium essentially becomes useless. One solution is to have electronics that are capable of providing access to error detection/correction data so that monitoring techniques can be used to monitor the gradual degradation of the media before the level of errors becomes catastrophic. Utility programs could be written to capture this information on a periodic basis and provide the user with a profile of the optical media [9].

Although optical disc media have a lifespan of over 50 years, the recorded data degrade over time and cannot be read. To prevent this, we optimized the power of the laser output for writing data to optical discs. However, recording marks degrade as a result of physical and chemical changes over time, and noise increases. As time passes, the marks break down and begin to blur. As blurring becomes larger, neighboring marks overlap, and they cannot be read. As a result, data is lost. Because of this, the gaps between marks must be protected as best as possible. Big marks are great for reading data immediately after it is written. However, the tradeoff is that they become hard to read because of breakdown over the years [11].

One other topic that must always be addressed in archiving of digital data is format obsolescence. There are many examples today of data-storage formats which are now obsolete and are therefore very difficult or impossible to read [10]. Today, preservation for information purposes is becoming evident. Many people understand more today the value of archives and their potential uses. But this understanding varies from one culture to another. In the past, during the analogue era, the copying of content on a new physical media was the only technological way of preservation. Material was copied onto new media, while the old was thrown in the trash. Twenty years later, the same cycle had to be repeated. This technology has now been dropped in favor of digitization and we can today clearly differentiate two types of preservation:

Preventive preservation, which consists in keeping contents in their original media, as it stands. Curative preservation, which consists in digitizing material under the best possible conditions.

Preventive preservation. This consists in keeping content in an appropriate environment, one that minimizes ageing of media. It requires a keen knowledge of the media and heavy-duty resources. It is a specialist business that most owners of archive data cannot in general handle themselves, due to the high level of expertise required.

Curative preservation. This is the only legitimate way to preserve audiovisual heritage. This technology consists in digitizing content, after the correct preparation to secure the best possible transfer quality. It is now reasonable to assume that digitized content can be preserved indefinitely, providing that certain precautions are taken. Also, digitization allows ubiquity of information, preserving it from disasters like fire.

Like software, hardware is prone to obsolescence, but also to mechanical failure. Hardware may be damaged by carelessness, neglect, overuse, or inappropriate storage. Batteries may be left in place during storage and cause unintended damage, not just to the hardware itself but also to any media which may have been left in the hardware. Digital media such as floppy disks, USBs and hard disks, are more vulnerable to deterioration and obsolescence than analogue media. Research into the average life of digital and analogue media indicates that digital media has an average lifespan of 3-50 years, while analogue has a range of 10-2200 years. This means that preservation action needs to happen much earlier in the lifecycle for digital media than it does for analogue. The care of digital media provides a further challenge. Obviously the files themselves should be copied onto sustainable media, in line with good practice; however, if the original media itself is retained (as best practice also dictates it should), it will need ongoing maintenance in terms of storage, cleaning, protection from magnetic fields, etc. This is the case whether or not the media will be in active use. An increasingly pervasive means of data storage today is the cloud. While the cloud was not designed for archiving, the reality is that it is used for this purpose by some smaller institutions. Preserving items via third-parties does pose a higher risk of loss should something go wrong. The third-party may go out of business — Nirvanix and Megacloud are two such examples — so an exit strategy needs to be set up for such scenarios. There are additional copyright, licensing and security issues with cloud storage, and privacy of personal data may also be more at risk [5].

1.3. Long-term data corruption reasons

Humanity has a data storage problem: More data were created in the past 2 years than in all of preceding history. And that torrent of information may soon outstrip the ability of hard drives to capture it [6].

The digital universe is doubling in size every two years and will grow tenfold between 2013 and 2020 (Fig. 1.4). Preserving this vast output is difficult, not just for its extent, but because much of it is ephemeral.

Besides scientific data, huge amount of commercial and personal data also have long-term value to preserve. In the year of 2006, Storage Networking Industry Association (SNIA) takes a survey on long-term digital preservation among 276 organizations [7]. Fig. 1.4 represents some results of the survey. About 90% of those organizations have digital data that must be preserved for more than 10 years. And 81% of responders even have data preservation requirement over 50 years. In addition, 18% of the respondents said that the amount of their data to be preserved is larger than 100 TB [8].

In the future volume of data is projected to grow at a 57% annual growth rate, faster than the expected growth of storage capacity. Moreover, new regulatory requirements mean that a larger fraction of this data will have to be preserved. All of this translates into a growing need for cost-effective digital archives.

The huge data set generated from large scale web applications, e.g., social networking, also need to be preserved. It is reported that most of the pictures uploaded to Facebook are never or rarely re-accessed, e.g., 97% of content only receives 29% of requests [9]. However, due to the huge potential value of those pictures, the company prefers to preserving those cold data rather than deleting them. Right now, Facebook starts build data center to maintain those cold data. The BD based archive system is one of their choices [10].

Government agencies at all levels are facing a challenge in the data-driven era. They need to decide what to do with their mission, operational and citizen data. What government agencies used tape storage for in the past has evolved with new compliance, regulatory, continuity of operations, cyber security forensics and data recovery requirements. These factors are driving a new look at a long-term archive storage tier strategy [11].

The rapid growth and development of the Internet, as well as growing use of social networks, are central characteristics of today's society. So, too, is the rapid expansion of the Internet of Things (IoT), as security cameras and many other devices and appliances connect to, and communicate over, the Internet. These developments, in turn, are generating enormous quantities of digital data. Not so long ago, businesses typically ran their own internal servers and managed all their data on-site. But the growth of data volumes and the advent of cloud computing are driving many companies to switch to cloud storage-storing their data offsite at data centers, and accessing it through the Internet. Large centers, such as those used by global shopping sites and social networks, retain large quantities of data for prolonged periods, with relatively few deletions. The recent development of new parsing technologies will open new possibilities for effectively using and reusing these types of big data [12].

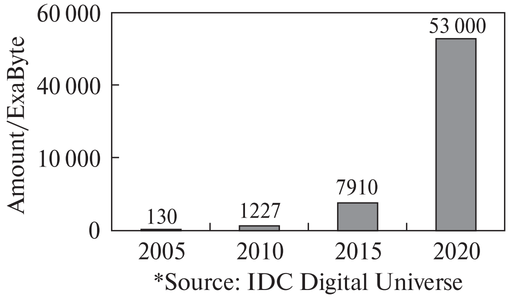


Fig. 1.4. Digital information capacity trends [6]

Since the early days of the web there have been myriad projects launched to archive and preserve the digital world that increasingly powers our global society. Perhaps the best known is the Internet Archive, which has been crawling and preserving the open web for more than two decades. As of last October the Archive had preserved more than 510 billion distinct URLs (images, videos, style sheets, scripts, PDFs, Microsoft Office files, etc.) from over 273 billion web pages gathered from 361 million websites and taking up more than 15 petabytes of storage. Much of this collection is available through the Archive's public-facing Wayback Machine that allows you to plug in any URL and see all of the Archive's snapshots capturing its evolution over the past 20 years. With such an incredible repository of global society's web evolution, we could see more applications of this unimaginable resource [13].

1.4. Storing massive quantities of long-term data

Long-Term Digital Preservation encompasses of the methods and technologies that ensure that valuable digital information can be correctly preserved in time, so it remains accessible and usable by interested parties, in spite of media failure and technological change [14]. Like many other cases in technology, digital preservation is a layered problem, with the following layers:

- Physical storage: the physical media where information is recorded.
- Data formats: the logical structure of the stored information.
- Software and applications: the executable context where information is managed.
- The Web: the most important model of distributed information.
- The Cloud: the most significant model of distributed computation.

However, while in some other technology cases the layered structure hides complexity from one layer to the next, in the case of digital preservation, there is a "multiplication" effect that makes the problems harder.

Physical storage. The first layer of any data preservation system is the physical medium where data is stored. Any digital media is susceptible of physical decay: CDs & DVDs use dyes that degrade, and magnetic media is susceptible to magnetic pre-eminence decay, for example. And mechanic elements in readers of both types of media are prone to failure.

Even so, data about the average life expectancy is hard to determine, as it depends on many factors: how many times it is accessed, the care with which it is handled, the storage conditions, and so on.

One special case that has been widely studied is one of the most used mechanisms for storing information nowadays, hard disks. Backblaze, a cloud storage company, has done extensive research on the failure rate of the hard disks they use, and it has found that the expected median life of hard disk is 6 years.

As we see, life expectancy can be worse than analog media, like paper. This explains that there has been a lot of investigation around physical storage media with much longer life expectancy, in different degrees of development:

• Sony and Panasonic have defined the Archive Disk, an evolution of the Blu-Ray disks that may store up to 1 Terabyte of information [15].

• M-Disk, a private company, has worked on a writable DVD media based on inorganic mineral layer, that enhances its durability, up to 1000 years [16].

• Hitachi has announced the capability of storing information in slivers of quartz that may last for hundreds of millions of years [17].

• The French nuclear waste agency, ANDRA, has developed a hard disk prototype [18] based on sapphire with information visually engraved in platinum that may last for a million years.

• The University of Twenties has done research in another "million year" storage disk, using tungsten and silicon nitride [19].

• Researchers at the University of Southampton (UK) propose an optical data storage system [20], where data is stored in nanogratings in quartz, created by ultra-short pulses from a femtosecond laser, with a life of 3 x 1020 years.

• But the prize for the most exotic form of long-term storage goes to DNA [21]. Researchers are encoding information using DNA, and taking advantage of its stability and self-reproducibility to store information securely for tens of thousands of years.

Another important point to have in mind is that all storage media needs hardware or software to be read. Sometimes, the trickiest part is not that the medium is in good shape to be read, but that the reading mechanism still exists. Some of the research in the previous list use ingenious methods, like codifying the information using visible QR codes, which in addition provide error-correction capabilities. Or, like the solution for ANDRA, data is stored in analogue form (readable page images), being readable just using optical magnification.

These new extended storage capabilities, and the explosive growth in media stored by Internet giants, have demanded them to automate their capabilities in what is called "cold storage". For example, Facebook uses self-built massive robot systems, using cheap Blu-ray disks. And Amazon, with its Glacier storage solution, offers similar capabilities. These kinds of automated setups could be used for long-term archiving purposes.

Also note that not only the physical storage medium needs to last, but also all the surrounding infrastructure, like datacenters where data is stored, needs to be designed taking a long term view on its survivability. There is no gain in storing costly sapphire disks in a building that may be crippled by earthquakes or floods [91-94, 175-185].

The best place to keep your data is on a running system and as that system ages, move it to a newer running system. The standards and media involved will change and as we get further and further along, even reliable media becomes difficult to recover data from because the technology that reads that media is scrapped in favor of newer technology.

Long-Term Digital Preservation can be considered the most complete form of all the approaches to preserve data in time, as it is shown in the following table (based in the definitions on the SNIA dictionary).

File formats have long been considered one of the biggest risks in digital preservation. However, this has not proven to be the overwhelming danger that it was initially perceived to be. In large part, this is due to the availability of open file formats, resulting in the formats being supported by more software applications. Proprietary file formats continue to pose a challenge, as their specifications are less likely to be openly available. Making software compatible with such formats, or converting the files themselves into a more open format can only be done with permission from the patent holder. This complicates long term preservation of such files, as the files may not be able to be migrated or normalized to a more accessible format. To keep such files accessible, the software that renders the file may also need to be preserved; which in turn brings its own set of issues. Many files deviate in some way from their official specification, so even if the specification is available, it may not necessarily be possible to convert the file to an open format. Additionally, not all file formats are suitable for long term preservation, even if they have an open specification. Some lossy and compressed file formats pose a higher risk of total loss if even a single bit is lost. Some types of digital media have a generally agreed archival format; TIFF is the accepted format for images, for example. However, not all media types have an archival format, including videos. While this issue will likely resolve over time, preserving institutions must in the meantime use their best judgement about what preservation file formats to use in such cases.

Metadata is probably the most important aspect of digital preservation. Materials with poor metadata may be undiscoverable, their authenticity unverifiable and their context unclear. Thus, they may not be as usable as they otherwise would. Inadequate or missing structural metadata will also impact on rendering. Preserving materials without good metadata is pretty much the same as throwing them away; along with all the resources expended in "preserving" them. It is important to gather metadata at time of creation if at all possible — some context may be lost over time. The importance of this has been recognized within the research field, with more effort now being put into encouraging researchers to create data management plans at the start of projects.

Some file formats support inbuilt metadata, e.g. TIFF and PDF. However, often only a very limited set of descriptive fields are supported, and sometimes the metadata is incorrect or absent altogether. PDFs, in particular, are at risk of containing poor or inaccurate inbuilt metadata. A very common issue, in my experience, is the 'author' name being that of the person who converted the file to PDF rather than the actual author of the content. Bad metadata like this may have a detrimental impact on the long-term fundability of the file. Having said that, my experience has been that Google is still able to retrieve such files, as long as the PDF itself is searchable. The act of preserving a file may alter it in some way that ultimately impacts its rendering or authenticity. It is important that any changes of any kind made to a file or its derivatives are well documented in the metadata [22].

Different technologies for different needs: We have seen that each of the storage technologies has different characteristics, making them suitable for different needs. Data owners need to thoroughly consider their requirements in terms of e.g. acceptable data retrieval time, security level and budget limitations during the archival period. The following table provides an overview of some of the attributes of each of the storage technologies. It is important to keep in mind that this is a general overview of technologies constantly in development. Within each category, there are differences between vendors and products.

Data must be stored and managed with appropriate systems and on an appropriate carrier. There are digital asset management systems or digital object storage systems available that meet the requirements of digital preservation and sustainability programmers, some approaches to which are discussed below. Once requirements have been determined, they should be thoroughly discussed with potential suppliers. Different systems and carriers are suited to different needs and those chosen for preservation programmers must be fit for their purpose. The overall system must have adequate capabilities including: Sufficient storage capacity: Storage capacity can be build up over time, but the system must be able to manage the amount of data expected to be stored within its life cycle.

As a fundamental capability, the system must be able to duplicate data as required without loss, and transfer data to new or "refreshed" carriers without loss.

Demonstrated reliability and technical support to deal with problems promptly. The ability to map file names into a file-naming scheme suitable for its storage architecture. Storage systems are based around named objects. Different systems use different architectures to organize objects. This may impose constraints on how objects are named within storage; for example, disk systems may impose a hierarchical directory structure on existing file names, different from those that would be used on a tape system. The system must allow, or preferably carry out, a mapping of system-imposed file names and existing identifiers. The ability to manage redundant storage. As digital media have a small, but significant failure rate, redundant copies of files at every stage are a necessity, especially the final storage phase.

A level of automated error checking is normal in most computer storage. Because archival digital materials must be kept for long periods, often with very low human usage, the system must be able to detect changes or loss of data and take appropriate action. At the very least the strategies in place must alert collection managers to potential problems, with sufficient time to allow appropriate action [164].

While there are many strategies for preserving the availability of digital records, there is no single solution, no best practice, and no established policies or procedures that meet widespread needs. The issues fall into four categories: storage media, hardware, software, and governance. Many variables affect the lifespan of storage media, and no comprehensive, scientific evaluations advise the consumer. However, it is clear that each medium has an Achilles' heel. For example, prudent care reduces, but does not prevent, deterioration of digital linear tape, such as:

1. increasingly brittle tape;
2. failure of the adhesive that attaches the magnetic particles to the tape;
3. exposure to magnetic fields.

Optical disks use organic dyes that biodegrade over time, and few blank disks come with date of manufacture. The conditions under which disks are shipped and stored are generally uncontrolled. The door will never close on improvements in records preservation. For example, the current ultimate in size reduction and longevity is Rosetta HD from Norsam Technologies, which writes analog or digital records with an ion beam onto stable media, such as nickel. The engraving is 10 microns wide, which reduces character size more than 20 times from standard microfilm. The nickel is stable, unaffected by temperature, humidity, magnetic fields, and more [23].

Very important criterion regarding digital preservation is the average lifespan of digital media. Selecting long-lived media for archiving digital content affects not only the end costs, but the long-term safety of the objects as well. Typical digital storage media have an expected lifespan of 3-10 years, though failure could occur at any time. Short-lived media, when combined with ineffective backup procedures, can result in the permanent loss of digital content [23].

1.5. Digital preservation cloud services

The Cloud is a term used for storage services where the data is stored in virtualized pools operated by third parties. These hosting companies operate large data centres, filled with servers, which could be located anywhere in the world. They are normally subject to strict security measures such as restricted access, environmental control and emergency backup power supply. For the end user, the Cloud offers clear advantages such as instant access, ease of use with no need for maintenance, and it is inexpensive to buy or lease storage space. This makes it well suited for basic storage needs. However, for the valuable data that requires secure, long-term preservation the Cloud involves some major concerns related to security, privacy and ownership. Ownership and privacy concerns: The Cloud involves entrusting Data Centre with servers a third party to store the data. This means the service provider has access to

Fig. 1.5. CERN server farm [29]

the data and may accidentally or deliberately disclose or alter the information. The service provider may also go out of business, posing questions as to who owns the servers and the data on them. The data owner will in this situation have limited control, and the data is inaccessible from the moment the network connection or power supply is switched off. Amazon Glacier is an online file storage service developed for long-term storage of data that does not require instant access. Like other Cloud services, Glacier is based on hard drives, but the data retrieval time is 3 to 5 hours according to Amazon. Storage is cheap, but users are charged for retrieving the data hence the risk of vendor lock-in should be considered. The privacy, security and ownership concerns related to cloud storage also apply to Amazon Glacier. Security issues: Data accessible through the web is vulnerable to hacking. Although the service providers strive to increase the security level, there is no guarantee that the data is out of reach of individual hackers, companies or national security agencies that might have an interest in accessing the information. Some data centres offer offline back-up, normally in the form of magnetic tape, in addition to cloud storage. This makes the data more secure, but a migration-based archiving strategy is then needed when the time perspective is long [24].

Cloud storage for the various operations in the media and entertainment industry is projected to grow from $2.5 B in 2016 to over $20 B by 2021 according to the 2016 Digital Storage for Media and Entertainment report from Coughlin Associates [25].

The Internet is playing an increasing role in the media and entertainment industry. It provides rapid transport of information between connected devices and data centers, it is the basis for today's growing over the top (OTT) content delivery, and it is the basis of collaborative workflows that span space and time and a move to operating rather than capital storage investments [25].

Even as cloud delivery models represent a lot of advantages for business and end-users, the considerations about digital preservation tend to be secondary, if at all present, in the design of cloud services. Similar problems to those mentioned previously in the case of web content extend (and get even worse) in the case of cloud solutions. We can mention:

• In some cases, the user is unable to have copies of the data stored in the cloud service.

• Control over data content, format, and associated metadata is limited.

• The execution environment can be very complex (or just impossible) to reproduce in a controlled manner.

• Execution environment can be distributed, as a combination of different services that may evolve in different ways.

• Cloud services may disappear, without previous warning. And, in some cases, they can take user's data with them.

Obviously, some of these problems may have an impact that goes beyond the needs of digital preservation. For example, some governments are trying to avoid "cloud lock-in" using SaaS solutions built over open source solutions, known as OpenSaaS. And some companies are looking to extend the concept of "software escrow services" to cloud applications (Fig. 1.5).

But there is a brighter side to cloud regarding digital preservation: to provide Long-Term Digital Preservation services in a cloud-based model, or, following the typical cloud naming scheme, LDPaaS [26]. It may sound a bit contradictory, but the model has some important advantages:

• It brings the advantage of cloud solutions (elasticity, virtualization, pay-as-you-go) to a complex field, as we have seen.

• It will bring Long Term digital preservation capabilities to smaller archives, that couldn't deploy them internally.

• In fact, big initiatives (at the national or transnational level) could be the providers of LDPaaS solutions for these smaller entities, and that would help to achieve a certain level of standardization.

Online archiving is certainly an option, but even in the age of ubiquitous broadband, online storage is relatively slow, even slower than optical in many cases and relatively expensive and unavailable when communications systems are down [27-30]. However, there are drawbacks. First off, though the means of delivery may seem magical and your data is often referred to as being safely stored "in the cloud", in reality, it's stored on someone else's hard drives or other media. It's as safe as a given service has made it. Then there's the ongoing cost in the form of monthly fees, and in some cases transfer charges. Also, speed and availability are limited by your online connection (DSL often has very slow upload speeds) and when your service is down, your archive is unavailable.

1.6. Conclusions

• Storage and database management is a vast field with many decades of results from scientists and researchers. The growing volume of information represented in digital form, leads to the fact that the question of how storing information on digital media is becoming increasingly important. Modern digital storage media are not able to exist for a long time.

• Scientists and engineers look to create media capable of storing information over time, comparable to the existence of mankind. To date, the high cost of storage associated with the storage sizes, as well as the need of rewriting to new. The world is becoming increasingly digitized, with more and more sources of data needing preservation.

• There is a higher risk of data becoming indiscoverable due to the associated risk of poor metadata. Large volumes of data cannot be handled on a one-by-one basis, as can analogue materials. The preservation methodologies applied to large amounts of data will need to be scaled to fit the needs and objectives of the preserving institution.

• Big data volume stresses the storage, memory, and computational capacity of a computing system and often requires access to a computing cloud.

CHAPTER 2

LONG-TERM DATA REPRESENTATION METHODS

2.1. Advantages and disadvantages of digital data recording

The real sense of digital technology is how simply large amounts of data can be stored and transferred, and how small that storage can be. This easiness of use and high density has convinced many to use it as the primary data archival technique within their company for technical and design data files. Minimal storage space, fast retrieval and cost are all compelling benefits. And we need to remember that sometimes all that data can be lost with a failed hard drive, a few keystrokes or system malfunction. The underlying difficulties with digitally stored data are becoming widely apparent to many because of well-established digital storage pattern. The digital preservation field is evolving quickly. All focal areas are changing and best practices now are still under debate. Preservation efforts must address preservation of the technologies of the past and also of the future. The quickly increasing volume of data requiring preservation makes other digital preservation challenges naturally more complex. Also the shorter lifespan of digital materials makes the need for long effective preservation action more urgent. When a file is digitally stored, the data is requirements involved in long-term data preservation.

National Archives of the Netherlands presented a report on preferred file formats [31]. There are 9 categories of their preferred and acceptable formats for. The blog post take up the question about the "spreadsheet" category for which it lists the following preferential and acceptable formats: Preferred: ODS, CSV, PDF/A. Acceptable: XLS, XLSX. Justification/explanation for using PDF: PDF/A — PDF/A is a widely used open standard and a NEN/ISO standard (ISO:19005). PDF/A-1 and PDF/A-2 are part of the "act or explain" list. Note: some (interactive) functionality will not be approachable after conversion to PDF/A. If this functionality is deemed essential, this will be a reason for not choosing PDF/A. There may be situations where PDF/A is the best, but choosing a preferable format should "take into account the purpose for which a spreadsheet was produced, its content, its intended use and the intended (future) user(s)". While analog media have better proven long-term security, they may not provide adequate representations of the original object. This solution may also lead to severe loss in functionality and introduction of the original digital object. For example, it is not possible to print out an interactive video, to preserve a multimedia document as a flat file or to microfilm the equations embedded in a spreadsheet. Digital data written as 0 and 1 obviously need decoding to get back the original file. To make data retrieval future-proof, declarations of how to decode the information is written as human readable text on the storage medium [116]. Electronic documents do not have the same longevity as physical materials. Tis data, stored in libraries, archives and museums have been preserved for centuries. A lot of text and images were transmitted traditionally on paper. Books, newspapers, prints, maps, music scores, manuscripts and other paper objects are housed in closed storage facilities to avoid damages through excessive handling. However they may be regarded as relatively safe. The electronic items, digitized or born digital, are more susceptible and fragile to different factors that may jeopardize their integrity and longevity. However digital objects can be copied without severe quality degradation and transmitted distantly. They do still face quality loss challenges, such as damages to entire disks or random bit errors. Therefore the lifetime of digital media is much less predictable than the one of time proved media.

The problems with digital storage right from the start materialized as problems with the physical media itself. The next problem underlying storage media is also one of physical formatting and operating system compatibility. Very few computer systems find out operational life spans longer than 5 years with the original users. Now it is often only 3 years, yet data usually has enduring value that spans some equipment generations and operating system changes. This fundamental time problem can be really serious, as it has no universal solution in modern digital landscape. To avoid software failure digital data owners often use standards-based protocols for access to data storage. Here different storage sites are running different realizations of the storage software. Thanks to such solutions the reliability and integrity of data does not depend on the same considerations of any single implementation. All digital methods and media have an inescapable error rate.

Currently in hard drives it's a pretty low read or write error rate, and it is never zero. As the data is copied and transferred, this collects, and never decreases. In a large video file, it's maybe a bad pixel and literally un-detectable, but in executable code, data corruption may be fatal. As media and hardware ages, the rate increases. Multiple stored copies with hash totals can help deal with an error, but can also lead to three copies, all lightly different for unknown reasons, which is not an improvement. Just be aware this is a low-level background irritation to deal with.

In systems designed to provide long-term access to archived electronic documents a basic reference is avoiding failure which accounts for information loss. A common technique is periodic media refresh comprising reading in the digital data, checking for errors. For this traditionally used error correction techniques and rewriting on new media. It's important to stress that the stored data can be read in the future. Some repositories store multiple representations of the same document or use archival quality file formats having specifications precise enough to build a credible interpreter or make conversion. To avoid software failure digital data owners sometimes use standards-based protocols, which can access to data storage. There different storage sites are running different implementations of the storage software. Thanks to such capabilities the integrity and reliability of data does not depend on the integrity and reliability of any single implementation.

However, carrier instability is only part of the problem. Such machine-readable documents, like audio and video recordings depend on the availability of format-specific replay equipment, some of considerable delicacy. Now we have experienced ever shorter commercial life cycles of dedicated audio and video formats. Thanks to the technical development over the past 20 years we can use such possibility.

Whenever a format had been superseded by the next, industry fleetly stopped production of new equipment, duplicates and professional service support.

This problem is not limited to text files, but applies equally some kinds of electronic documents like photos, video, audio.

It is already the case that disk drives capable of reading 5 1/4 inch floppy disks are not readily available. So, one concern leading to the use of the term is that documents are stored on physical media which require special hardware in order to be read and that this hardware will not be available in a few decades from the time the document was created.

The digital medieval period or real Dark Ages is the perception of a possible future situation where it will be impossible or difficult to read multimedia or historical electronic documents, because they have been recorded in an obsolete and obscure file format. The name derives from the term Dark Ages means that there would be a relative lack of written record, as documents are transferred to digital formats and original copies lost. An early reference of the term was at a conference of the International Federation of Library Associations and Institutions (IFLA) in 1997 [32]. The Digital Dark Age also applies to the problems which appear due to obsolete file formats. In such instance it is the lack of the necessary software which causes problems when retrieving stored documents. This is primarily problematic when proprietary formats are used, so, it might be impossible to make appropriate software to read the file. A popular example is with NASA, whose early space records have suffered from a Dark Age issue more than once. For a long time magnetic tapes from the 1976 Viking Mars landing were unprocessed. When later analyzed, the data was unreadable as it was in an unknown format and the original programmers had either died or left NASA. The images were finally extracted following many months of puzzling through the data and researching how those machines functioned [33].

Also we can remember the BBC Domesday Project in which a survey of the nation was compiled 900 years after the Domesday Book was published. While the information in the Domesday Book is still actual and understandable today, there were great fears that the discs of the Domesday Project would become unreadable. It can happen because computers capable of reading the format had become rare and drives capable of accessing the discs even rarer. In 2002 the system was emulated using a system called DomesEm by the CAMiLEON project. This supposes the information on the discs to be accessed on modern computers [34]. The Apollo 11 missing tapes were recorded from Apollo 11's slow-scan television (SSTV) telecast in its raw format on telemetry data tape during the time of the first Moon landing in 1969 and later lost. A team of retired NASA employees and contractors unsuccessfully tried to locate the tapes in the early 2000s. The data tapes were recorded as a backup in case the live television broadcasts failed for any reason. NASA ground receiving stations performed real-time scan conversion to the NTSC television format in order to broadcast the SSTV transmission on standard television [35].

The means of safeguarding digital documents are significantly different from those employed for "traditional" documents. At the same time, few countries have as yet adopted a national policy regarding digital information, and remain unaware of the risk of disappearance of commonly used means of transmitting and storing digital information, such as websites, email or databases.

2.2. Digitization of analog data

The question of how long knowledge can be expected to last and for the immediate future the human factor seems to be governing factor. But consideration the challenges and chances of maintaining knowledge across emphatic temporal distance, and in respect to the options of reconstructing damaged or lost knowledge, an "archaeological" perspective is useful which focuses on the non-human agencies of knowledge convention. Media theory here is helpful since it addresses both the engineering (techno-mathematical) and the philosophical (epistemological) questions involved. For the analysis of the technologics of knowledge tradition, a focus on both the material (technical) forms which are subject to physical entropy and on the immaterial (logical), almost time-invariant codes of transmission is required: material embodiment ("markers") versus logical implementations (archives), and the physical versus the symbolical mode. "Symbolic" does not refer to symbolism in its traditional sense of metaphorical meaning, but to discrete characters in coding information (ranging hitherto from Arabic numbers and alphabetic letters up to the binary code of Zeros and Ones embodied as Low and High voltage levels in electronic computing). The current shift from material memory as cultural premise to techno-mathematics as the dominant form of cultural communication corresponds with a different kind of temporality. Cultural memory once intended for eternity transforms in to an ongoing practice, economy and aesthetics of short-term intermediary storage, such as repeated data migration, "the enduring ephemeral" [36].

The idea of storage information using VM appeared. First it was introduced in 1994, and the QR Code has gained wide acceptance in such actual industries as warehousing, manufacturing and logistics, healthcare, retailing, life sciences, office automation and transportation, etc.

The QR (Quick Response) Code is a two — dimensional (2 — D) matrix code that belongs to a larger kind of machine — readable codes, all of which are often referred to as barcodes, irrespective of whether they are made up of bars, squares or other — shaped elements. Compared with 1 — D codes, 2 — D codes can hold a larger amount of data in a dense space, and compared with other 2 — D codes, the QR Code can hold much more data still. An attached error — correction method and other unique characteristics allow the QR Code to be read more reliably and at higher speeds than other codes. Like written language, barcodes are visual reasoning of information. Unlike language, which humans can read, barcodes are designed to be read, decoded and understood by computers, using machine — vision systems consisting of optical laser scanners or cameras and software with barcode — interpreting function. The rules of barcode constructing (its grammar) and the character set it uses (its alphabet) are called its symbology.

Fast omnidirectional scanning : Position — detection patterns in three corners of a symbol allow the QR Code to be read from any angle within 360 degrees, disposal the need to align the scanner with the code symbol. The position — detection patterns also eliminate any background interference, ensuring stable high — speed reading.

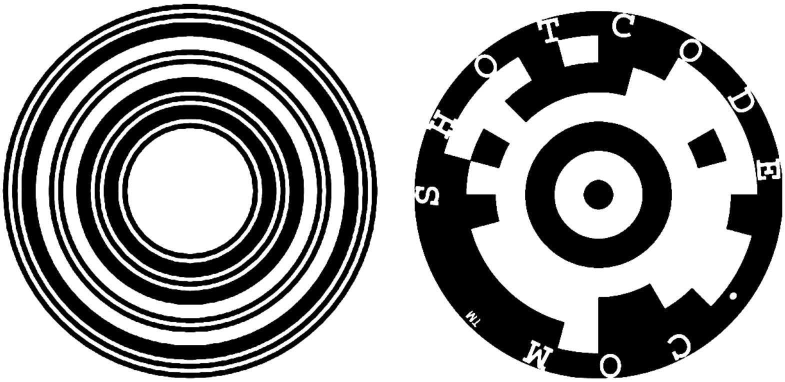
We can view many different quick codes, developed for different purposes. The main among them and their characteristics are presented in Table 2.1.

The actualization was the quick response (QR) code [37, 38], and it can be easily decoded by modern smartphones. The level of QR code containing the largest

Table 2.1. **ISO/IEC Standardized symbols**

Standart	PDF417	Data matrix	Maxi code	QR code	Aztec code
Developer (Country)	Symbol (USA)	CI Matrix (USA)	UPS (USA)	DENSO (Japan)	Hand Held Products (USA)
Code type	Multi-low	Matrix	Matrix	Matrix	Matrix
Data size, Alphanumeric	1850	2355	93	4296	3067
Characteristics	High capacity	High capacity, small space	Fast reading	High capacity, small space, fast reading	High capacity
Main market	Identity documents, Tickets	Food administration, Medical	Logistics	All industries	Airline, railroad
Standard	AIMI ISO	AIMI ISO	AIMI ISO	AIMI ISO JIS	AIMI

Fig. 2.1. Optical microscope images of the same QRcode



a *b*

Fig. 2.2. Circular barcode structure: *a* — Circular 1D barcode, *b* — ShotCode

amount of information can lose up to 7 % of the data before the code becomes almost unreadable. For the encoding of the final disk, it is likely that a coding scheme would be required which focuses on easy decodability.

By keeping the size of the QR code low, it is possible to read out the disk by an optical microscope (Fig. 2.1). The entire disk was covered with a centimeter sized QR code, and each code pixel consists of a set of much smaller QR codes with pixels of a few micrometers in size. The initial effort to create a medium containing embedded data which is able to survive for 1 million years is promising. The optical readable data in the form of QR codes was able to outlast the temperature up to 713 K.

Readout of bar-code or QR code from optical disk surface argues that there are applications where circular barcode constructions are more preferential. Circular 1D barcodes were broadly used for tagging CD/DVD items [91, 177]. Structure of 1D circular barcode (Fig. 2.2) consists from series of concentric circles and typically based on a standard barcode symbology. Such kind of barcode is readable by all the devices used to decode traditional barcodes.

Photoluminescent circular 1D barcode was one of the elements of multilayer optical disk's photoluminescent protective layer, together with photoluminescent im-

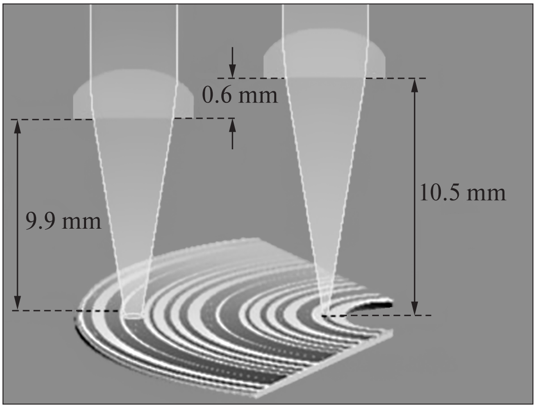


Fig. 2.3. Photoluminescent circular 1D barcode recording process [39]

age, microtext, additional and corrupted sectors area [91, 183-185]. Multilayer disk driver's objective lens is supposed to have a 2 mm vertical shift range which allows reading and recording circular 1D barcode at the same mode as disk data recording and readout (Fig. 2.3).

ShotCode was a first circular 2D barcodes [39], developed to be read by low resolution cameras, webcams or mobile phones. ShotCode had limited payload (5 to 6 bytes), so it has been widely replaced by QR codes nowadays. But researchers also have an interest in circular barcodes, developing new formats for codes with a data density comparable to QR codes. One of the brightest projects was shown at [39]. Researchers resolved to start from the requirements of minimum payload value of 25 alphanumeric characters (200 bits according to ASCII characters code). It was assumed that the barcode design has to support different sizes (1mm 10 cm) and be stable to image corruption or noise. The three design types of the circular 2D barcodes differ only in the number of the start marker repetitions and radial zebra patterns (Fig. 2.4).

The barcode includes:

- solid black ring: for locating the barcode in an image;
- white circle which surrounds black ring;
- square markers for a rough perspective correction;
- start marker which defines the beginning and determine the outer radius of the barcode;
- angular zebra pattern to determine the angles and achieve accurate perspective correction;

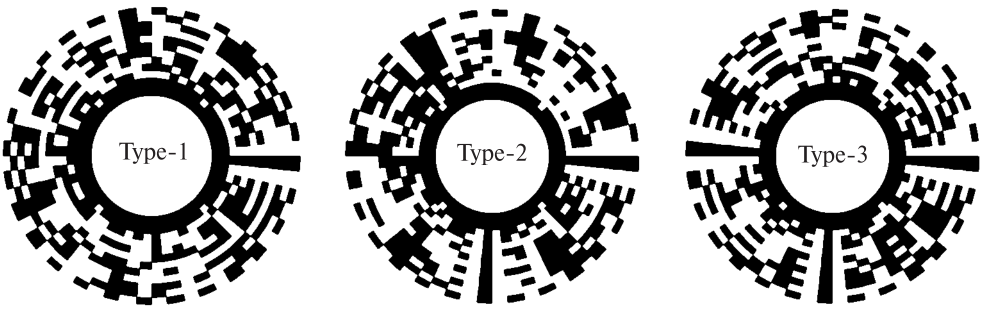


Fig. 2.4. Three types of generic circular2D barcodes [39]

- radial zebra pattern which determines the radial timing of the rings;
- data zone which store useful data and error correction codes.

Decoder is able to decode 5 mm 300 dpi circular 2D barcodes with a small average bit error rate, and it can be corrected with an error correction code.

Problems going together with long term data storage are analyzed and educed. In this paper we proved the possibility of long-term visual data storage. Also we showed the possibility of usage the hybridized digital-visual method for dependable long-term data storage.

2.3. Digital data archiving methods

It's important to save data on media that will be readable in the future, but also to use file formats that'll be approachable later. This is accurately the problem. If you need to preserve documents long-term, it's best to break them out of proprietary formats like Photoshop and Microsoft Office's *.doc and *.xls to Open Document Format (ODF), Open XML or PDF. If it is available, photographers should consider saving images in original raw format, uncompressed TIFF format, or another formats like PNG or JPEG. For audio, Also, if it is possible, you can save the highest uncompressed WAV or AIFF. Video is harder, but lossless MPEG-2 seems to be a good choice. If you want to geek out, the Library of Congress maintains a numerous mention on the pros and cons of archival file formats [40].

Of particular value to us here is the issue of CAD/EDA software; no other single user application group has had as much deliberate format changing and forced incongruity. This has proved to be radically destructive for users. Virtually all the CAD/EDA low-cost software makers were bought up by more costly vendors, and then the inexpensive versions were systematically removed.

Each version of new software was subsequently more expensive, these versions were deliberately limited in terms of backward compatibility. This single set of design data comprising the bulk of critical intellectual property owned by a company, is the most at risk for obsolence and inconsistency going forward. The new corporate CAD entity cannot even offer any software, compatible or operationally similar with your existing files, yet appeals to your "product loyalty" to deal with them. A similar trend has also occurred with common word processing files. Endless application and format churning has continued, to virtually no useful purpose except to frustrate office workers worldwide, and create the impossibility of useful training. Eventually, at this stage many users are simply turning to the free and more stable Open Office/Libre Office suites offering full file compatibility, and exiting the costly office software rat race. Sometimes the 3+ year old existing software tool is perfectly fine, and so forcing new buys in applications has become a real performance in the digital marketplace. The classic example is the Windows XP operating system, which is one of the largest user bases in the world. A new release is promoted with some new features and fixes, which are cleverly bundled with some critical incompatibilities. If some new copies are bought, new data or imported old data can no longer be worked on with existing older software, forcing still more purchases. It is a good plan for the vendors but not so good for users. And it is ultimately highly destructive for all, as files and program operation completely fail to achieve any long-lasting stability, destroying high productivity and effective training.

2.4. Data preservation strategies

Part of the archiving compensation is identifying and moving inactive data out of current online production systems into a long-term holding area. In the business it was named "cold storage", and it is implies that an explicit tradeoff has been made to input a slower response time to retrieve that older data than the quick access needed for more timely information saved online. Cold storage can include social media data, photo files of users from previous years. All this material must be kept for posterity, but is very rarely accessed. It is an economic decision made by vendors who choose to free up their more expensive online storage by relocating content that is rarely used. Cold storage usually comes down to a choice among hard disks, tape-based media, solid-state drives, optical media and cloud storage (Fig. 2.5). The decision about what medium to use is complex and varies depending on use, ease of operation, cost, durability and the user's previous experience [41].

As data volumes rise, so do storage costs-making it basic to implement storage systems that distinguish between frequently accessed, occasionally accessed and infrequently accessed data (or hot, warm and cold data) selecting the best storage media for each.

Hot storage demands high performance and is best implemented with high-cost (flash) memory. With warm and cold data performance takes a back seat to requesting of long-term reliability, the ability to maintain large size of data at relatively low costs, and the ability to protect data integrity in "green environments" with limited environmental controls. Sony is proved that optical disc storage fills all of these requirements, and consequently ideal for warm and cold storage [41].

Google announced Nearline, a product offering long-term storage of data that doesn't change. There is a growing problem of storage requirements for companies incorporating technology into their daily routines. There are only few companies today that don't have a need for digital storage. Nearline offers a simple pricing model; you are currently charged 2.6 cents per month per gigabyte of storing data. For the case study I was doing, the company had 1.5 terabytes which translates to $39 per month for storage. Companies are no longer worrying about where to store physical files. But they have a problem how to back up and store data online that is no longer necessary to access on a daily basis but needs to be there just in case.

The Digital Preservation Network (DPN) is the only large-scale digital service of preservation that is built to last beyond the life spans of individuals, organizations and technological systems [43]. It shared collectively across the academy that protects local and consortia preservation efforts against all types of catastrophic misfortune. The DPN is a planned "dark archive" for preserving not actively used or accessed materials, but they can be made available for use at any time from multiple storage facili-

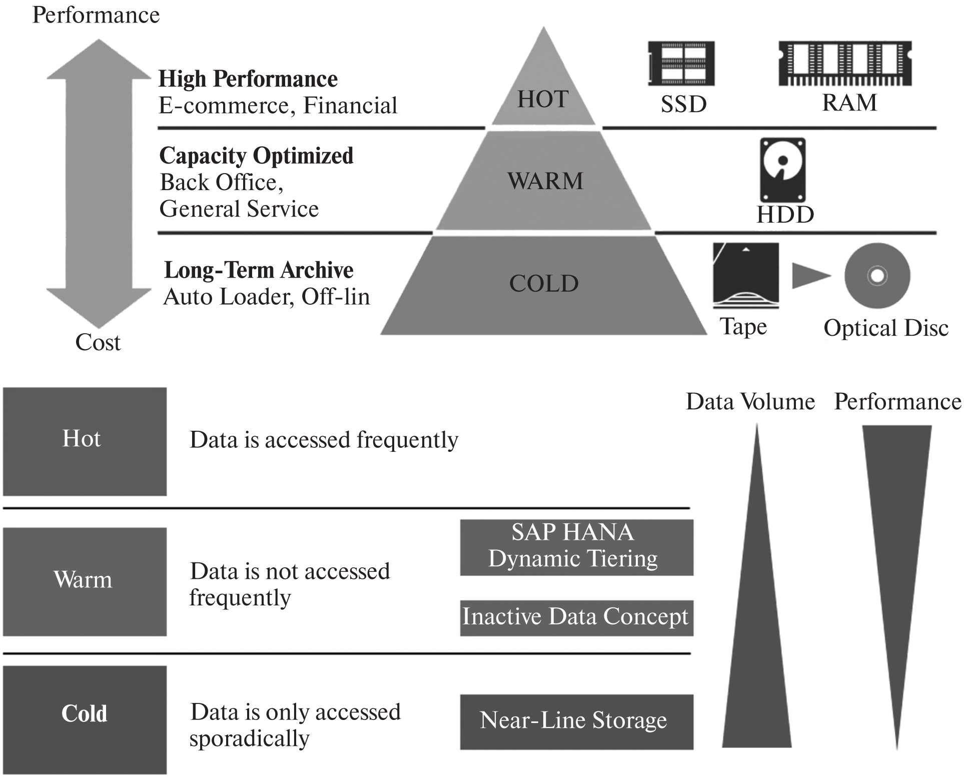


Fig. 2.5. Storage system hierarchy [41]

ties. The DPN guarantees academic institutions that their scholarly resources will be available in the event of any type of change in administration or physical institutional environments. The DPN ensures the survival, ownership, and management of preserved digital content in the future by establishing a redundant and varied technical and legal infrastructure at multiple administrative levels [43].

The DPN ensures the secure preservation of stored content by leveraging a heterogeneous network that spans diverse technical, geographic and institutional environments. Local repositories become contributing nodes, which put in new collections. The DPN creates several federated, replicating nodes, which are digital repositories for the contributing nodes with a long-term preservation. The replicating nodes contain redundant dark copies of all deposits that can be made available in cases of critical loss. The diversity of the DPN nodes decreases the risk of a single point of failure. Objects are replicated across nodes that embody physical, technical, organizational and also political diversity. A single point of failure cannot jeopardize centuries of scholarship.

DPN's preservation process can be presented in five steps:

1. Content is deposited into the system via an Ingest Node;

2. Content is replicated to at least two other Replicating Nodes and stored in varied repository infrastructures;

3. Content is checked via bit auditing and repair services to ensure the content remains the same over time;

4. Corrupted or destroyed content is restored by DPN;

5. As Nodes enter and leave DPN, preserved content is redistributed to maintain the continuity of preservation services into the far-future.

The DPN ingest and replication nodes are in themselves preservation archives that utilize best practices to secure the holding data. DPN members can be convinced that the content deposited into the system is secured for the next generation of scholars and beyond [43].

2.5. Data encryption algorithms

The most signifying component of digital preservation is managing the digital objects in archives. These Preservation Repositories must archive data and associated metadata on an affordable and reliable type of storage. There are many digital storage options available, and each institution should evaluate the available storage options in order to determine which options are best for their own needs [43].

The Migration Time Frame is the important criterion for the best storage option that involves the lifespan of the media and the resulting costs of digital preservation. Every digital medium and system has a limited lifespan. The system will become obsolete or the media will eventually fail. Most digital media need to be refreshed or migrated regularly in order to preserve digital objects beyond the expected lifespan of the media [43].

A basic issue in digital data preservation is that there still has to be something physical to hold the data. Digitization breaks the link between content and specific carriers, but there still has to be the carrier. Modern storage has a short lifetime. Future technology is not likely to last longer than data tape lifetimes or current hard drive. There is general agreement that format deteriorating and the lack of replay equipment is a greater adversity than carrier decay. The time window left to transfer contents from analogue and single digital carriers to digital repositories successfully is appreciated to be not more than 20 years. Digital preservation is identically demanding, as it requires an ongoing investment to keeping digital information energetically alive. Appropriate professional management software and modern storage technology is expensive, all of them needs subsequent renewal at least in the pace of migration intervals, which are generally in the order of five years.

Digital Preservation is much more than only backing up digital content [43]. We can see three main technical approaches to digital preservation: technology preservation, emulation and data migration. The first two examples used to access the object, either maintaining the original hardware and software or using current technology to replicate the original environment. The work on “persistent archives” established on the articulation of the essential characteristics of the objects to be preserved may also

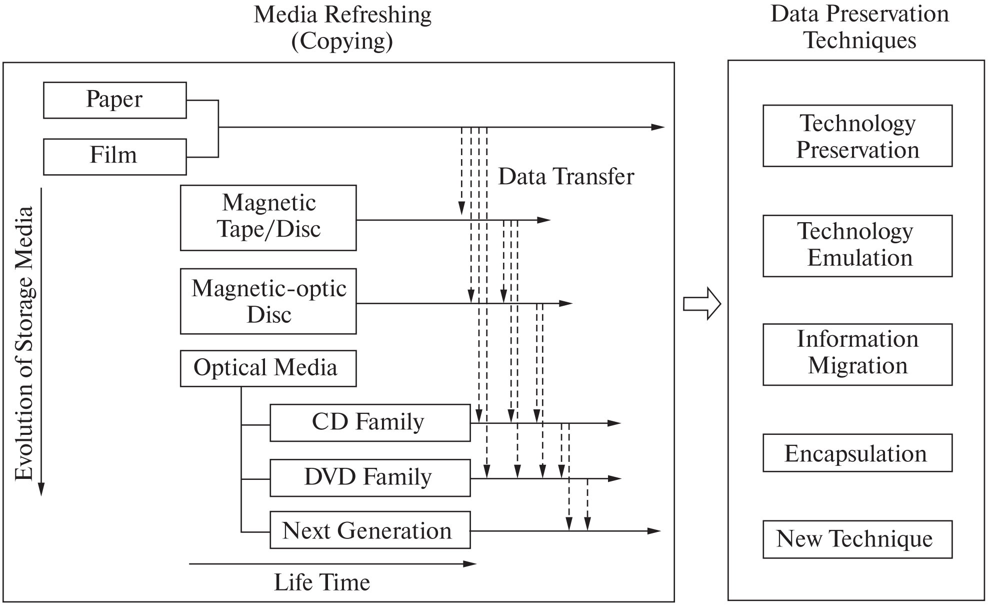


Fig. 2.6. A progression of information preservation strategies [48]

be of interest [44]. The data storage media evolution and the preservation technology development can be described as shown in Fig. 2.6. This diagram lists the various media used in digital and analog data storage and the techniques needed to ensure that the data on them is preserved. It also highlights the trend from analog to digital/optical storage media and demonstrates the transfer of data from one media generation to the next.

While it is easier to create, amend, and distribute digital data, the media storing this data such as optical discs are not as intense as traditional analog media such as film or paper. Modern information preservation requirements show, that this paper will focus on the technical strategies aspects used in digital information preservation [44].

Existing preservation strategies can be broadly classified into two main approaches (Fig. 2.7).

The first is the more conservative access where the original technological environment is fully preserved for decoding the digital information in the future. It can be further divided into two preservation techniques. The first is to preserve the working replicas of all computer hardware and software platforms for using in the future. This is referred to as the technology preservation strategy [42, 43]. The other is to program the newer digital systems to emulate on demand the older obsolete platforms and operating systems. This is technology emulation strategy.

The key problem influencing the long-term usability of digital records to develop a migration strategy for moving records to new technologies and media as older ones are displaced. All this digital records are technology dependent and therefore technology obsolescent is likely to be the most serious interference to the long-term

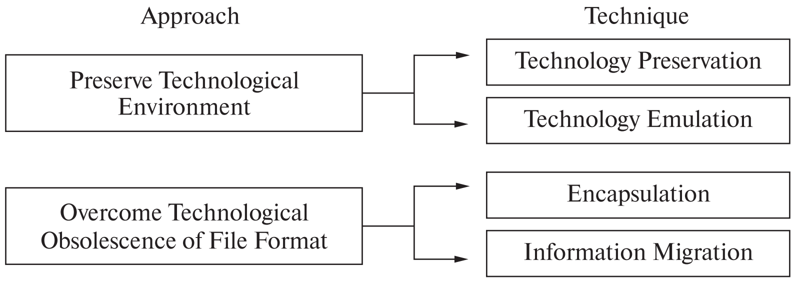


Fig. 2.7. Existing preservation approaches [48]

usability of digital records. But the development and implementation of a migration strategy to ensure that digital records created today can be processed by computers and intelligible to humans in the 21st century is absolutely essential.

Migration strategies focus on the sustenance the digital objects in a form that is accessible using current technology. In this variant, objects are periodically transferred between some technical environments, newer one, and while as far as possible maintaining the content, usability context and functionality of the original. These migrations may require the object copying from one medium (or device) to a new one and/or the transformation of the object from one format to another. Some kind of migrations may require only a relatively simple format transformation. A migration to a very different environment may involve a complex process with considerable design effort [46]. The key problem affecting the long-term usability of digital records is to develop a migration strategy for records to new media and technologies as older ones are displaced.

2.6. Multimedia data storage systems

The impending fact is that digital records are technology dependent and therefore technology obsolescent is likely to be the most serious disturbance to the long-term usability of digital records. Due to this fact the development and implementation of a migration strategy to ensure that digital records created today can be processed by computers and intelligible to humans in the 21st century is exactly essential [46].

Fig. 2.8 shows a schematic diagram to select the suitable preservation techniques for digital resources according to complexity of digital and the type, also information and the availability of the data format with its usage. The capture of metadata becomes a critical part of a migration-based preservation strategy. Metadata needs the supporting of the object management and migration process. But migration inevitably leads in the longer term, to some changes in, or losses of, original functionality. This is relevant to the interpretation of the object, users will rely on metadata about the migration process- and about the original object and its transformations — to realize some understanding of the functionality realized in the original technological environment

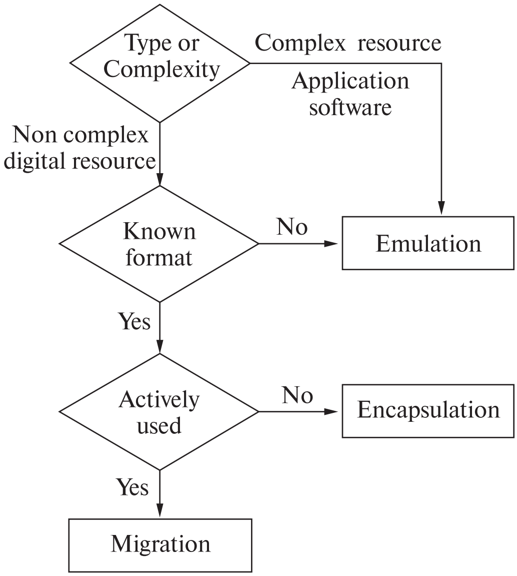


Fig. 2.8. Diagram for selection of digital preservation techniques [46]

[46]. All the potential problems with data migration motivate various organizations to delay the deployment and purchasing of new technology. Such delays can be detrimental in and of themselves, because older hardware may require more hands-on maintenance, customary has lower performance and is more prone to failure. Most organizations search to deploy new technology to except such issues. But delays in implementing new technology have a business risk. An additional point is that delaying deployment of a new storage device that has already been purchased or leased raises its effective cost, as the company is amortizing the cost of both the old and new devices or is paying lease fees for all kinds of devices [46]. Long-Term Data Preservation could be the most advanced form of data management, so governance models & architecture are essential for it to work. The starting point is the OAIS ISO 14721: 201243, that is the Reference Model for Open Archival Information Systems (OAIS). The reference model of OAIS provides a conceptual framework for the understanding and increased awareness of archival concepts needed for the long term digital information preservation. A conceptual reference like it doesn't imply any specific methodologies or technologies, but more the definition of the functional entities that take part in these types of systems (Fig. 2.9).

More concrete architectures are derived from this conceptual overview. The good example of a derived architecture from OAIS is the one defined in the Project CASPAR (Cultural, Artistic and Scientific knowledge for Preservation, Access and Retrieval) [46]. Data migration shows the process of making an exact copy of an organization's current data from one device to another, mainly without disabling or disrupting active applications. After that this process redirecting all input/output (I/O) activity to the new device. We can see a variety of circumstances that might cause an organization to undertake a data migration, including:

- Replacement or upgrade of all server or storage technology;
- Consolidation of server or storage;
- The data center relocation;
- Server or storage equipment maintenance, including workload balancing or other performance-related maintenance.

The cost of migrating from one media generation to another, or from one media type to another can be significant. Migration would include the costs of the new

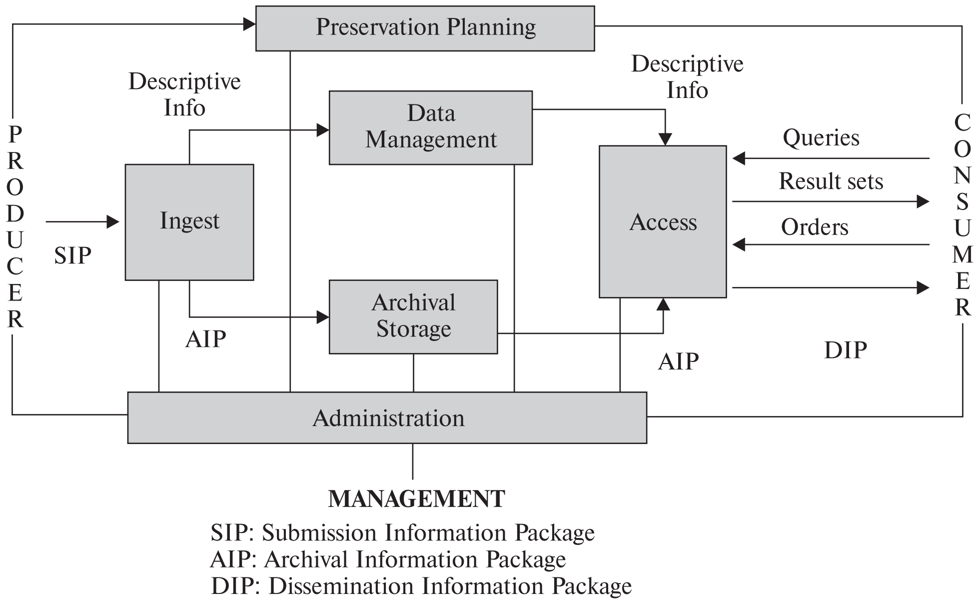


Fig. 2.9. Open Archival Information Systems Functional Entities [46]

media and systems. Therewith, it must include the personnel costs to manage the replacement and verification processes so that there is no data loss or degradation. Modern hard drives have a limited lifespan, so, they must be replaced regularly [92]. Data migration is an important event consuming significant budget and labor, and occurs very regularly. The conjunction of the frequency of and resources consumed in a data migration results in data migration taking an important amount of the IT budget. As storage infrastructures become more complex and large, data migrations are also becoming more complex, risky and labor intensive. Few organizations must begin managing this growing portion of their IT budgets more effectively.

Key findings illustrated the migration expense and best practices to reduce all the risks include:

- Enterprise storage migration costs can exceed US $15,000 per terabyte migrated;
- Storage migration projects required 4 to 6 hours per host, from internal organization resources. Of these hours, 4 to 5 hours were used to plan the migration and 1 to 2 hours (~30%) were used to execute the migration;
- Migration project expenditures are on average greater than 200% of the acquisition cost of enterprise storage. With an average of 4 years useful life, the annual operating expenses associated to migration represent >50% of attainment cost;
- Duration of the migration is mainly due to limited maintenance windows. Common migration techniques need application outages due to either SAN rezoning and/or host reboot activities [47-49].

The migration/remastering costs for 5 PB of content over 75 years is much less for an optical system with the media replaced every 50 years rather than more frequent tape and HDD commutation [47-49]. As organizations are continually migrating data, they are not considering data migrations a core competency. It is usually the result of another event in the data center, such as technology refresh, an application upgrade and data center consolidation project. Data migration projects can be rather complex, with large-scale projects requiring many in-house and contractor personnel.

Consequently the labor, consulting, software and hardware for data migration have become a large market. The overall market for data migrations can be calculated by identifying the amount of data migration activity that results in grand data migrations. Some data migrations become a result of refreshes of technology. So, an average enterprise company has an average of 1,400TB of network attached storage (NAS) and nearly 8PB of storage across their storage systems. We remember that the useful life of most storage systems is 3 to 5 years, and this size of organization will often have multiple storage systems at the end of their useful life spans and requiring a refresh. Nearly 40% of the organizations surveyed plan a technology refresh in the upcoming year, up from 26% the year before. Eventually these large enterprises would be always conducting a data migration of multiple storage devices at any given time.

Cost is one of the most important factor when consider long term storage. Archive storage costs must be conceivable, else content owners will not submit materials for preservation. Storage costs, even if they are declining, may influence decision makers to select a low-cost storage option at the expense of basic preservation affairs [47-49].

2.7. Conclusions

• Every digital method of information storage has an inescapable error rate. Read and write error rate could be low but it is never zero. As the data is copied and transferred, this accumulates. In the case of executable code data corruption can be fatal. As hardware and media ages, this rate increases. Multiple stored copies with hash totals can help deal with this error, but can also lead to several copies, all slightly different for unknown reasons, which will not be an improvement.

• Media instability is only part of the problem. As machine-readable documents recording depends on the availability of format-specific replay equipment, some of considerable sophistication. Photoluminescent circular 1D barcode was one of the elements of multilayer optical disk's photoluminescent protective layer, together with photoluminescent image, microtext, additional data sectors and corrupted sectors area.

• Part of the archiving equation is identifying and moving inactive data out of current online production systems into a long-term holding area, or "cold storageCold storage is an economic decision made by vendors.

• The most fundamental component of digital preservation is managing the digital objects in archival repositories. Preservation Repositories must archive digital objects and associated metadata on an affordable and reliable type of digital storage.

CHAPTER 3

PHYSICAL LIMITS OF DATA STORAGE LIFETIME

It is noteworthy that high quality paper copies continue to be one of the most durable techniques for data storage. Recordable digital media of all types is just not very expressive by contrast, even microfilm is better. Some media types can be erased at any time, which poses an ongoing serious threat to secure data storage [51]. The most traditional failure mechanisms for materials (excepting mechanical wear) include oxidation, corrosion, and breaking of chemical bonds. Every failure mechanism is exacerbated by elevated temperature, humidity, and exposure to light, so, any controlled environment that is intended for archival storage always includes controlled temperature, humidity and light. We know three basic technologies available for storing digital data: magnetic (including magnetic tape and hard-disk drives), optical (including CDs, DVDs and BDs) and solid-state (consisting primarily of flash memory). These technologies use well known materials and processes to manufacture the storage media, and have known failure mechanisms [50]. The U.S Military is still using outdated floppy disks, an old fashioned form of memory storage, to coordinate its nuclear operations, a new report has uncovered. Some of these disks are over 30 years old, but they are predominate in operational activities such as the potential use of intercontinental ballistic missiles, nuclear bombs and tanker support aircraft. The report, entitled "Federal Agencies Need to Address Aging Legacy Systems", urges agencies to replace their current systems with new technology. "Federal legacy IT systems are becoming increasingly obsolete: Many use outdated software languages and hardware parts that are unsupported", the report said [51].

3.1. Lifetime of data storage based on hard drive disks

For years, magnetic media based solutions, such as tape and hard disk drive (HDD) based archive systems; monopolize the data archiving market due to their low cost and high capacity. However, in the era of big data, rapidly increasing volume, velocity, and diversity of data set

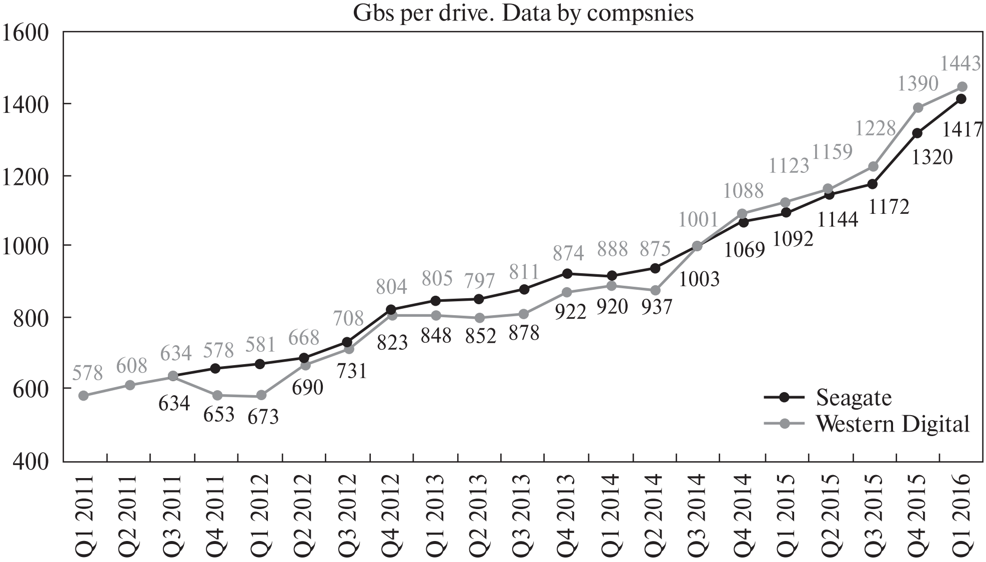


Fig. 3.1. Increase in areal density on hard magnetic disks [52]

bring numerous challenges to the archive systems in various aspects, such as capacity, performance, cost, reliability, power consumption, etc. [47-49].

Some active archives often use hard disk drives because hard disk drive arrays can be continually connected to the storage network, allowing relatively rapid access to content. Active archives with hard disk drive can also be combined with flash memory to supply with better overall system performance. But even the best hard disk drives do not last forever. They can wear out with continued use and even if the power is turned off the data in the hard disk drive will eventually decay due to thermal erasure. In this case we face with thermodynamics, the real enemy of data retention. In practice hard disk drive arrays have built in exorbitance and data scrubbing to help retain data for a long period. It is probably good suggestion to assume that HDDs in an active archive will last only 3-5 years and will need to be replaced over time. Using commodity HDDs, open source software and commodity computer components companies such as BackBlaze inform that they build a 180 TB HDD system for $9,300. That comes to $51.67 per TB or about $0.05 per GB [48]. The data is recorded by magnetizing a thin film of ferromagnetic material. A hard disk is a rapidly rotating disk used for storing digital information. A read/write head on an arm accesses the data when the disk is spinning. Hard disks provide instant access to the data. The reading and writing device is an inner part of the storage medium. High failure rates make hard disks Hard Disk Drive inexpedient for long-term preservation. A recent study projects an average lifespan of six years [49].

Modern technology used to make advanced types of hard drives. They create it possible to substantially increase the recording density on magnetic disks. The use advanced techniques of recording have allowed to process storage devices that have become now the main type of memory for storing big data. Fig. 3.1 presents the technology that was used to work up new types of hard drives over the past thirty-five years [49]. Promising constructions to achieve high density recording are pattern magnetic media [49]. Achieved areal densities of 10 Terabit–per–Square Inch pattern structures are obtained by the method of Bloc Copolymers [52].

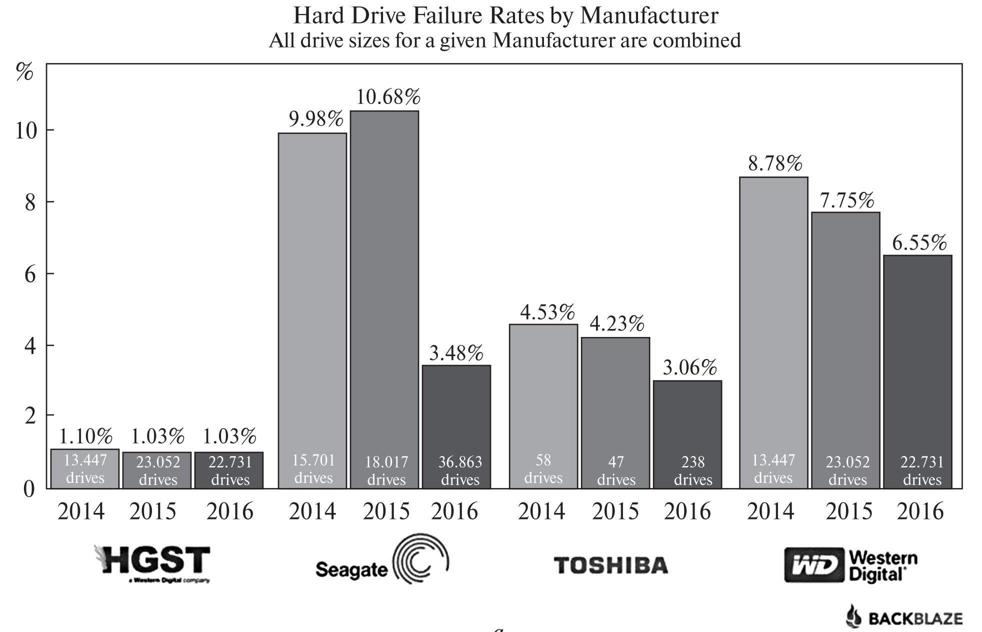
The long-term future of HDDs likely rests with high capacity HDDs, especially in data centers serving cloud storage applications. Further still looking at the declining shipments in high performance HDDs and computer HDDs the best path for development would appear to be a high a capacity drive as possible. In its earnings call Seagate said that it will ship 14 and 16 TB HDDs over the next 18 months. The company has indicated it would like to ship 20 TB HDDs by 2020. The ATSC areal density roadmap indicated that 100 TB 3.5-inch HDDs could be real by about 2025 [52].

In recent years compiled a prominent amount of statistical data on the failure in data centers, hard disk drives [49, 52]. It seems that hard drives have three distinct failure “phases”. In the first phase lasting 1.5 years, hard drives have an annual failure rate of 5.1%. For the next 1.5 years, the annual failure rate drops to 1.4%. After three years, the failure rate explodes to 11.8% per year. This means that around 92% of drives outlast the first 18 months, and almost all of those (90%) then go on to reach three years [47-49, 52]. There are a lot of various assumptions of the life time of the HDD. Hard drives have typical lifespans of two to eight years, depending on their environment and how they’re treated (Fig. 3.2).

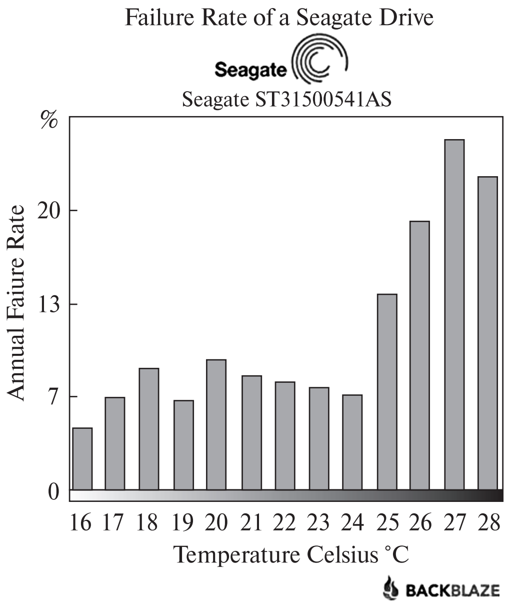
Hard drives of computers are the most fragile parts of computers. All of them break down more often than other connected to computing device. We can see so many moving parts inside of hard drives that it is no wonder they destroy so easily [53].

Complex electromechanical system of HDD has the distances between the moving parts is tens of nanometers spinning at high speeds and rely on complex components and circuitry [52, 53]. The main cause of HDD failure lies in the fact of electromechanical systems. We can also accentuate a strong influence on thin-film recording medium of the electromagnetic fields. The common failure mechanisms in the hard drives current generation are related to the head-disc interface. A head crash are usually exposed by severe data loss, and data recovery attempts may cause further damage if not done by a professional with proper equipment.

Drive platters are laid over with an extremely thin layer of non-electrostatic lubricant, and the read-and-write head will simply glance off the surface of the platter should a collision occur. This head hovers mere nanometers from the platter’s surface which makes a collision an acknowledged risk. There are a lot of different sources of these problems, between them are temporary interface disruption, handling damage, media damage, and thermo-mechanical stability of the read and write structures. Sometimes the root-cause is a customer issue like extrinsic contamination, mechanical shocks, and condensation. And sometimes the root-cause is a manufacturer issue like intrinsic contamination, servo errors, particulates, increases in lube thickness, localized weak or thin regions in the carbon overcoat, head contacts with the disc, etc.



a



b

Fig. 3.2. Hard drive failure rate as a function of time and temperature [53]

Faulty air filter is another cause of failure, which equalizes the atmospheric pressure and moisture between the drive enclosure and its outside environment. If the filter fails to capture a dust particle, the particle can land on the platter. It causes a head crash if the head happens to sweep over it. After that particles from the damaged platter and head media can cause one or more bad sectors. These, in addition to platter damage, will quickly render a drive useless. A drive also includes controller electronics, which occasionally fail. It may be possible to recover all data [47-49, 52, 53]. It should be noted that Hard drives interfaces become obsolete quickly.

Research and development to improve the reliability of HDD will increase the time of the occurrence of faults Specified failure rates for today's hard drives are improved from earliest generation products. In 1993 year the most manufacturers specified mean times between failure (MTBF) between 100,000 and 250,000 hours. Seagate's recent desktop product specifications indicate a MTBF of 600,000 hours, and 1.2 million hours for their enterprise products [54].

On the durability of the HDD is greatly influenced by operating circumstances. Research conducted by the company Backblaze showed that the greatest influence on the durability of HDD temperature at which the drive is operated [54]. Operating a disk drive in a more strained environment, such as an industrial location with higher temperatures and humidity, vibration and corrosive gases, could result in an increase in the annual failure rate and a decrease in the actual lifetime. The useful effect will be based on the mechanisms that initiate failure and the specific environmental parameters.

Future trends in hard drive reliability may be towards decreasing failure rates and increasing design lifetimes. It is actual for potential markets outside the home/office (automotive, industrial, avionics, etc.) where long-term (>10 years) reliability can be as a significant driver as cost or performance. Intromission in these markets is currently minimal, but with increasing memory requirements. Especially it is connected with the development of drive/fly-by-wire technology and the increasing use of sensors, significant growth is expected. The area of prognostics is one of the area where the disk drive industry is far ahead of the electronics industry. Self-Monitoring, Analysis, and Reporting Technology, or SMART, is a capability that allows specific disk drive characteristics are monitored (by the drive, without host software overhead) and compared to "alert levels". Should an alert level be exceeded, the frequency of monitoring usually grows and a status indicator is written to a register location within the drive. This capability is proposed through all the major disk drive manufacturers. SMART can be helpful in predicting and scheduling maintenance to maintain high reliability. But it is unable to detect and respond to assembly-level wear out mechanisms, such as solder joint fatigue or dendritic growth [55].

Most hard circle drives (HDD) last somewhere between three and five years before some of their segment is fizzled. That doesn't generally mean the drive is hopelessly busted, yet three to five years is still about to what extent they last, whether you're discussing an interior drive for a server or desktop, or an outer hard circle drive. With the greater part of the moving parts inside, something will in the end quit working. As with any media putting away imperative information, it's critical to utilize quality equipment [55].

High capacity 8 or 10 TB disks offer superior capacity and cost but two fundamental issues have to be addressed to create a secure and cost effective Media Archive Power Costs & Security. Now and five years ago popular opinion was that hard disks would never be able to be emulous of tape. It was connected with the cost of their power and the cost of the air conditioning to get rid of the heat they dissipated. Because of this seems to be an insuperable problem, compared with a tape library, where the recording media is passive [56].

To increase the HDD lifetime proposed to use a mode of operation in which the rotation of the disk media occurs only when accessing the drive. There is the technology that spins up hard disks in the storage equipment when access is required. Also it is spins down the disk when there is no access. This method cuts power consumption by spinning down the disks, and seems to be highly energy-efficient if applied in the storage equipment that has hard disks that are infrequently accessed. Also it can be used in the disks with no access for long periods of time. It is specifically used for backups of data at regular intervals. The idea behind MAID is that the disks containing older, seldom accessed data remain idle most of the time, in such a manner saving power and wear and tear on the disks themselves. If a piece of data needs to be derived, and the disks can be quickly spun up, the data is retrieved, and the disks go back to an idle state. A MAID, which can have something near hundreds, or thousands of individual drives, offers mass storage at a cost per terabyte roughly equivalent to that of tape. Modern MAID technology is provided as an option to high-volume tape libraries.

The example of such system is the ALTO system [57], which can effectively be used to store cold data. It is high-density, enterprise-class, offline data archive solution — to allow a convenient and cost effective alternative to cumbersome robotic tape libraries and expensive clustered RAID storage setups. This system provides easy scalable capacity and security thanks to user defined replication level. Due to non-spinning disks, ALTO has triple digit extended disk life and ultra-low power consumption. Also it combines on-line and off-line data storage with use of individually managed commodity disks that can be externalized for vault storage and bank-like security. It is supported by over 30 industry leading asset management providers, and integrates with other systems through standardized API or native low level control of the system.

This system also offers the flexibility and scalability to deliver a secure, reliable long-term archive. ALTO easily expands capacity of required business without complexity in small increments for low costs. ALTO provides secure, easy and affordable technology for archiving and cold data storage [57]. Archive, based on ALTO can start as a single device at a single location. It can grow up in scale, security or geography without any limits. For Archive with high speed and Restore access ALTO is deployed at low cost on-site, connected to a local LAN.

For increased resilience, ALTO supports off-premises deployment at geographically separated location, independently of it being another building or continent [57]. ALTO takes advantage the most recent strategies for data protection, with the unique ability to add disks of any capacity from any vendor on an as-needed basis. Also it scales to 1EB (1,000PB), so we escape costly and time-consuming migration. The disk management software Prometheus allows idle disks to be stopped, saving power and allowing the data to last 100 years. Individual ALTO disks can be externalized for ultimate security. The most advanced ALTO file system ext4 allows the disks to be opened on any PC, Mac or Linux Computer. It allows easy access to any of the files and non-proprietary forensic data recovery. The actual data may be per user choice occasionally secured by transparent encryption [57]. This technique also allows the disks to be fully spun down and stopped. We see the individual disks spinning-up to archive or restore files on demand with a latency of seconds, rather than minutes, even for the largest Petabyte-class configurations. Spinning down the disks overturns the historical power consumption objection. It can greatly extends the life of the disk drives which are more likely to clock up 10 — 100 hours per year rather than 8750 hours per year for traditional disk storage. ALTO from Disk Archive Corporation is a simple actualization of a tried and tested storage concept using non-linear disks to replace the linear tape media. It can bring important performance & business benefits for Broadcasters and Media companies [57].

3.2. Lifetime of data storage based on magnetic tapes

Another widely used medium for cold storage is magnetic tape. It consists of a tape substrate coated with magnetic particles that retain data in the form of magnetic charges. The magnetic coating helps to retain electronically encrypted data in digital format [58]. Magnetic heads come into contact with the tape to write by altering the charges and read by detecting the charges. Magnetic tape offers rather high recording densities, and rather low cost per unit of capacity in comparison with other. As improvements accelerate capacity typically lead to a new generation approximately every five years—accompanied by changes in drive systems and software. As writing can go back one generation and reading can go back two, relatively frequent data migrations are required for really longer-term storage. Random access is not suitable because heads must wait for the winding mechanism to move tape to the appropriate position. Tape is also susceptible to damage from water contact, high humidity or electromagnetic waves [59]. The tape is traditionally packaged in a plastic cartridge, and tape drives read and write the data. Magnetic tape allows large amounts of digital data to be stored at a relatively low cost. It makes well suited for a mass storage purpose [58, 59].

Due to the potential improvement space of storage density and relatively low cost the tape is most likely to be the best candidate of long-term digital preservation in this decodes among magnetic storage media [47-49].

There are considerable advantages to using tape for long-term archiving, backup and disaster recovery:

- security: once the data is recorded and the cartridge removed from the drive, the data is inaccessible until the cartridge is reinstalled. This means that the data cannot be corrupted by a virus while it is offline. Also security is further enhanced by drive-level encryption;
- energy savings: once data is recorded, the medium is passive; no power is needed when it sits in a rack;
- reliability: tape media is removable and interchangeable, meaning that unlike HDDs, mechanical failure of a drive does not lead to data loss, be-cause a cartridge can simply be mounted in another drive;
- lifetime: because the medium is passive, it is extremely reliable with a long lifetime. Some tapes have been in use for forty years;

• cost: savings estimates of the total operating cost of tape backup relative to HDDs range from factors of three to twenty-three. It is actually used even if the latest developments, such as data deduplication, are taken into account. In archival applications, where deduplication cannot be used effectively, cost savings can be a little bit higher [59].

Attractive tape can either lose information by losing its attractive charge, or when the layers of the tape begin to particular. Any attractively charged capacity medium will in the long run lose its attractive charge and in this way its information. As indicated by a modest bunch of sources like eHow, Wikipedia, and Searchdatabackup.com, makers can guarantee that tape can last up to thirty years. It can make it a valuable medium for documenting. The issue with that number is that attractive tapes will just keep going that long under completely ideal ecological conditions. They should be kept in a spot where both mugginess and temperatures are steady. There is one more reasonable lifespan for attractive tape is around ten to a quarter century. They are more helpless to wear and tear if utilized much of the time. Since tape tapes and information tapes are fundamentally the same as, the lifespan of tape tapes is congruent with one of attractive tapes. Lifespan truly relies on upon the assortment of elements we've said. Some have been known not out rapidly because of over the top use, while others last upwards of thirty years. A tape endures somewhere around ten and a quarter century [59]. Less active archives with data stored for longer periods of time will be interested in storage media that can retain the information stored on them for an extended time period. The magnetic tape is digital storage media that are used for long term cold storage applications. We see many benefits in the use of magnetic tapes compared to other storage media. One of their advantages is persistence. Unlike other data storing media, as a rule tapes have a much longer useful life and are less prone to the modern drives risks. Magnetic tapes can be read securely, even after 30 years. And the average hard drive scarcely lasts only five years. Such persistence is offered neither by SSDs, enterprise HDDs or cloud computing. Physical problems like broken or damaged hard disks or logical issues like software failures or interrupted software updates can even lead to loss of data. Saved data on magnetic tapes can be read even after decades, supposing they have been stored according to the manufacturer's instructions [58].

Archiving magnetic tapes come in half-inch tape cartridges. The popular modern formats are the LTO format supported by the Ultrium LTO Program, the T10000 series tapes from Oracle/StorageTek and the TS series enterprise tapes from IBM. New magnetic tapes have a storage life under low temperature/humidity storage conditions and low usage of several decades. Also they have currently native storage capacities per cartridges as high as 8.5 TB. iNSIC projections from 2012 indicate that we could see 100 TB^+ magnetic tapes sometime after 2020 [60].

The LTO Program companies say, that LTO tapes are cheaper, more accessible and save power and cooling costs compared to all other cold storage media. The LTO tape cost is two cents per gigabyte of compressed data, which is 26 times cheaper than optical discs. Tape also has a better bit error rate than discs [61]. The addition of the

Linear Tape File System (LTFS) seems to be the biggest advance in tape storage technology. It's essentially a self-describing tape format developed by IBM to address tape archive requirements. LTFS is additional data that provides a file system interface on the tape. LTFS format of data tapes can be used independently of any external database or storage system, allowing direct access to file content data and metadata. It works like any optical disc and any computer operating system can read it. LTO data tapes written with this format can be exchanged between LTFS format systems. Any software systems that understand the format can provide users with a file system view of the media [61].

Magnetic tape as the carrier continues to improve the fact of increasing the storage density and writing speed. For ten years there is an open recording technology LTO (Linear Tape Open), which is a unified specification of the records to magnetic tapes and drives. It was joined by large manufacturers such as HP and IBM. A new generation must be compatible with the previous devices. The LTO consortium is focused on the project plan, establishing a new generation of solutions every two years [61]. Today LTO is widely used for mass storage, and the latest generations LTO 6 can store up to 2.5 terabyte. It is important to be aware that data capacity increases for each generation for long-term preservation purposes. But also this has a negative impact on the expected lifetime. LTO requires a migration-based archiving strategy in order to ensure data safety and accessibility as for other magnetic tapes. Therefore, the usage should be kept to a minimum, and it is considered best practice to migrate the data at least every 5-10 years to better tape formats. Migrations and vendor lock-in: A drawback with magnetic tape is that data retrieval is dependent on special reading devices, i.e. tape drives, which need to be maintained. New tape drives are generally not able to bring back data from older generations of tapes. And regular and endless investments in migration are needed when we use magnetic tape for long-term preservation [61].

Current tape technology attains a storage density of about 1 Gb/in2 and a cartridge capacity on the order of a terabyte (Fig. 3.3). Analyzing the limits of current tape technology suggests that tape areal density can be further pushed by two orders of magnitude, leading to cartridge capacities in abundance of 100 terabytes. So, tape becomes a very attractive technology for data archiving with a sustainable roadmap for the next ten to twenty years.

Last years great progress has been made in increasing the areal density of magnetic tapes. Sony has developed a new technology that pushes tape drives far beyond where they once were, leading to individual tapes with 185 terabytes of storage capacity. Sony employed the use of sputter deposition in order to create the new tape. It creates layers of magnetic crystals by firing argon ions at a polymer film substrate. The magnetic particles measured in at just 7.7 nanometers on average, combined with a soft magnetic under-layer. They are able to be closely packed together. Perhaps surprisingly, storage tape shipments grew 13% two years ago, and were headed for a 26% growth just last year. Sony also announced that it would like to commercialize the new material — as well as continue developing its sputter deposition methods. Then,

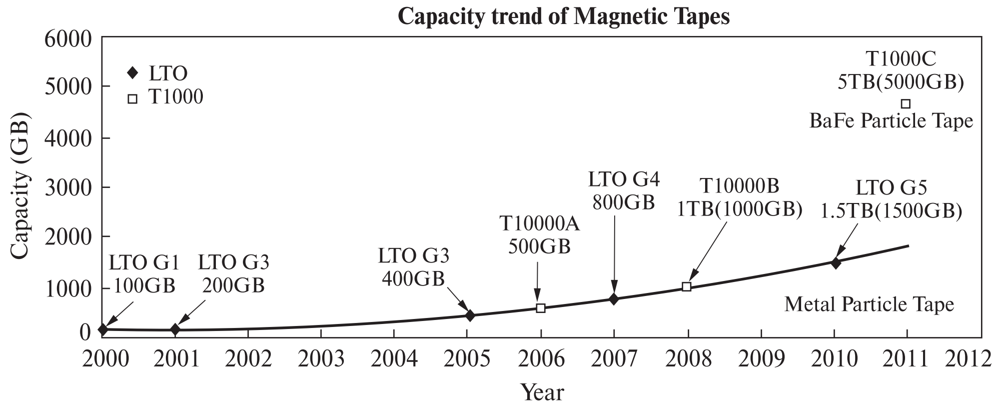


Fig. 3.3. Increasing the capacity of magnetic tapes [61]

Fujifilm has specifically announced that its 154 TB tape will have data density of 85.9 billion bits per square inch on cost-efficient linear magnetic particulate tape, the standard LTO cartridge will be capable enough to store up to 154 TB of uncompressed data with this breakthrough in data density. It is 62 times greater than today's current LTO6 cartridge capacity [62].

Tapes used in open systems typically did not hold file metadata in a form that was easy to access or modify independent of the file content data. It is actual for LTFS format, introduced in 2010. Instead, the tapes often contained files stored in a random sequence, without filenames or directories. Various external databases were used to maintain file metadata, like file names, timestamps and directory hierarchy. Along with reliability media archival life and ease of use now essentially the same with all cold storage media tape wins in the end due to its lower cost. Cloud services still use either tape, optical discs or hard drives for storage. It's just that users don't have to know the technology and pay only for the easement of having a third-party worry about operating and maintaining the system. Tape-based mediums are always going to be the no-brainer because cost is what drives the world, but only if you have accessibility with LTFS and can keep your costs down [61, 62].

When using tape for archiving purposes one has to bear in mind that tape is a sensitive material that needs proper handling and environmental conditions. Magnetic fields are obviously a threat to any magnetic storage, and even smoke and other small particles can cause loss of data and damage. It is therefore imperative to have clean operating conditions, with the humidity levels and right temperature. Magnetic tapes have a lifetime expectancy ranging between 10 to 30 years, but only with minimal usage, handled properly and stored in optimal conditions. For reliable safety information must fulfill a number of conditions: regularly re-writable tape must be replaced after a few years, long-term storage must occur in strict compliance with the terms and conditions. The ideal parameters are humidity 30-40% and temperature 15-25 °C) [80]. Every tape type showed a loss in magnetization when they were under induces stress conditions of relative humidity and higher temperature. Depending on the type of tape used, LEs ranging from 10 – 200 years if storedat 30 °C. This range decreased dramatically to 0.7 – 7 years if stored at 60°C. Testing these same magnetic tapes showed that at 50°C, with various levels of relative humidity (RH) at RH = 20%, the range was from about 0.6 – 2.8 years; at RH = 80%, this decreased artificially to a range of about 0.3 – 0.9years [61-63].

Magnetic tape should be stored as specified in ISO 18923:2000 Imaging materials — Polyesterbase magnetic tape — Storage practices. If magnetic media needs to be used on a regular basis, this should be done in accordance with ISO 18933:2012 Imaging materials — Magnetic tape — Care and treatment practices for extensional usage. Offline media should be stored in a temperature range of less than 20°C. Relative humidity is recommended for long-term storage, less than 18 °C and 40%. And the temperature and humidity should be lower for very long-term storage (decades). ISO 18923 advises that magnetic media should not be stored below 8 °C, these environmental levels must be stable. Mould will start to grow at around 60% relative humidity and if the humidity fluctuates more than 10% in 24 hours. Exposure to ultraviolet (UV) light will also hasten media degradation. Fluorescent tubes with UV filters should be used wherever possible in storage areas. They must be turned off when not in use. UV light can be measured with a light meter, if levels should not exceed 75 W/lumen [64]. Linear Tape Open (LTO) and a more modern barium ferrite data cartridge tape, like the long-term stability of magnetic particulate tapes was investigated using samples that had been stored under an accelerated condition for one year. The degradation of the barium ferrite tape was equal to or less than that of the LTO tape. But the one year acceleration test showed that the barium ferrite tape and the stability of LTO could be estimated to be more than ten years [61, 64].

We know several technological problems that limit the shelf life of magnetic tapes data. Current A-inch tape products are all magnetic, and suffer two main degradation mechanisms. First of all it is the slow relaxation of the magnetic domains. These domains are how the data is stored in any magnetic storage product. These domains are randomly oriented when not under the influence of any external magnetic field. Magnetic domains change to produce a net magnetization in one of two directions after being subjected to an ambient magnetic field. So, difference in the direction of the remaining magnetic field is the difference between 1s and 0s in digital data. With time and with fluctuations of temperature magnetic domains begin to relax, slowly reverting to their original random orientation, and slowly degrading the difference between the encoded 1s and 0s. Ultimately, so many of these bits will have degraded that reading a file back will have become impossible. There is also another main degradation mechanism of magnetic tape; it is the delamination of the recording layer from the plastic substrate of the tape. Modern A-inch magnetic tape has the recording layer, which is applied to the polyester substrate in a printing process. The binder in the recording material is what keeps the recording material bound to the substrate, but such materials are generally organic. All of them degrade with time, which leads to small pieces of the recording layer peeling off from the polyester substrate. We can see the delaminating. Wherever these pieces peel off, the data is irretrievably lost.

Fig. 3.4. The StorageTek SL8500 modular library [65]

Tape systems have evolved significantly, improving in good availability, compression, management and tape cartridge capacity. Modern tape archive systems are designed for large-scale data protection (Fig. 3.4).

They can archive environments and scale up to 2.1 exabytes of data-protection capacity [64]. The T10000 T2 Data cartridge stores 5TB of data. The T10000 T2 Sport cartridge stores 1TB of data. It is a little bit shorter and designed for faster data access. TheT10000 T2 VolSafe cartridge is a WORM tape. The T10000 T2 cartridges use barium ferrite compound on aramid. It is a heat-resistant synthetic fiber used in aerospace and military applications, e.g., for body armor—a well-known example of an aramid is kevlar. The tape is 5.2 μm thick and 1147m long. The archival life of it is rated to be 30 years and 25,000 loads/unloads. The cartridge weighs 270 g. It is certified to operate within the range of between 10 °C and 32 °C. Te best archival temperature lies between 15 °C and 26 °C. It can outstand temperatures between −23 °C and 49 °C on shipping Oracle DataSheet [65].

The StorageTek SL8500 modular library can grow from 1,450 to100,880 cartridges. Depending on the library size from 64 to 640 T10000C drives can be mounted in it. The library stores up to 542 PB of data without compression, or more than 1 EB of data with 2:1compression, in its maximum configuration. Although disk has been catching up with tape lately, tape is still200 times more energy efficient and 15 times less expensive than low-end disk. Up to 32 SL8500 libraries can be combined into a single complex managed by a single control interface, providing up to 33.8EB of storage OracleData Sheet. Tape can stay in its slot on the shelf, unbothered, without drawing any power at all for 30 years [64, 65]. All the data on hard disc must also be duplicated on data tape. The effect of the failure could be very damaging, as there are significant amounts of data on each tape, so multiple copies guard against this possibility.

A lot of information centers are widely used for long-term storage data tape storage drives. Google prefer to use tape storage drives and devices to archive and back up every email it stores for long-term data archive solutions.

That tape is less expensive, and Google has greater longevity and reliability and is more portable and compatible with a variety of data formats than hard disk drives (HDDs). The need for long-term data archive solutions that will survive well into the future is only increasing. Recent advances in the LTFS, or Linear Tape File System and tape libraries from IBM, Oracle, Quantum, Spectra Logic and others are making data access times much faster. Furthermore, Linear Tape-Open (LTO) standardization, now on its sixth iteration (LTO-6), guarantees data access across devices well into the future [61-64].

The data center for US Intelligence Advanced Research Projects Activity (IARPA) in Washington DC hold an Exabyte, or one billion gigabytes, on tape drive. This center would require US$1 billion over 10 years to build and maintain, as well as hundreds of power megawatts [125].

Long-term tape storage like Advanced Intelligent Tape (AIT) or Linear Tape-Open (LTO), usually can only write to the current and one previous tape generation, and read only two prior generations of tape. So, the tape drives are often upgraded every 10 years, making the tapes obsolete sooner than the anticipated life expectancy. And in three generations of LTO tapes, which currently is approximately ten years, the tapes and drives could become obsolete. The Library of Congress has AIT2 tapes and the tape drives. But all this drives are rather difficult to connect and to use [82].

3.3. Lifetime of solid-state storage

Streak accumulation comes in three diverse basic stockpiling media: Flash drives, SD cards, and strong state drives (SSDs). Streak drives can last up to ten years, at the same time a streak memory doesn't for the most part debase on account of its age. And in light of the quantity of compose cycles, which implies the more you erase and compose new data, the all the more rapidly the memory in the gadget will begin to damage. Since every one of these gadgets are comparable in that they all utilization streak memory, they'll all damage in a comparative design. Whether it does or not, one thing is sure: better equipment will pay off. Given the random of makers, lifespan may contrast a considerable amount beginning with one gadget then onto the next, yet streak memory gadgets estimated for more compose cycles will normally last more. With regards to equipment, holding back to spare cash won't pay off over the long haul, particularly on the off chance that you lose valuable information, which can cost you significantly more than you would've spared. And with regards to blaze SD cards and drives, you'll likely ruin them or lose in the clothes washer before whatever else happens. You can select the right equipment, and ensure any information is moved down elsewhere to make sure it keeps going. But you never know when any kind of media may fall flat [130].

Have flash memory there is one very nasty drawback preventing this type of media replaced all existing optical and magnetic storage. It is connected with the durability and reliability. It is the fact, that flash memory has a finite number of cycles of erasing and writing. The appraisal of the manufacturers show, that modern flash memory can withstand an average of 100 thousand cycles erase / write [96].

At the 2013 appeared the idea of using really bad low endurance flash memory for a cold storage archive. Flash writing is best done at elevated temperatures while data disturb and data retention favor storage at lower temperatures. The JEDEC JESD218A durability specification states that if flash power off temperature is at 25 degrees C then retention is 101 weeks — that isn't quite 2 years. So it appears conventional flash memory may not have good media archive life. It should only be used for storing transitory data. Also Cold storage should be inexpensive, but the demand for flash memory to fill computer application like consumer combined with the cost of building flash memory factories will make storing long term data on flash relatively expensive. However Bruce Moxon from Samsung reported that flash memory may play a significant role in an overall active archive where data is available for access since flash caching can help meet access requirements. Also flash memory can also be used for object caching of keys and metadata. But flash memory may not be suitable for the main storage in an archive system or for an active archive. Also flash memory can provide better overall system access to content kept in a lower cost media storage library [66].

NAND flash memory widely used as a storage medium for mobile devices, first of all because of its lightweight, silent, shock-resistant, and energy-efficient characteristics, lacks an over-write operation. NAND-based storage devices deploy a FTL (flash translation layer) to emulate the over-write operation with out-of-place updates. Clean spaces eventually become scarce as out-of-place updates write data to new unwritten spaces. A garbage collection process accompanies several NAND writes and erases. Also it is important to reduce the latency and frequency of garbage collection. The central idea is to cluster hot and cold data separately to reduce the number of valid pages in a victim block at the time of garbage collection. Reducing the number of valid pages of the victim block contributes to the reduction of garbage collection process with both the latency and frequency [67].

NAND flash EEPROM memory is being widely used in very high densities as portable storage (thumb drives), and even as replacements for mechanical disk drives in some portable computing equipment. We see the problem in the limited write endurance of this technology, its medium storage life and its various parasitic fault modes. It makes some travail for real archive storage or any long-term application. It is not optimal for more serious longer term uses, but is a good match for low-cost consumer applications [67].

The significant gap in update frequency between hot and cold data motivates to separate hot and cold data on different flash blocks to avoid unnecessary program/erase cycles. Some drawbacks of NAND flash memory such as no support for in-place updates and limited program/erase cycles, which trigger the evolution of so-

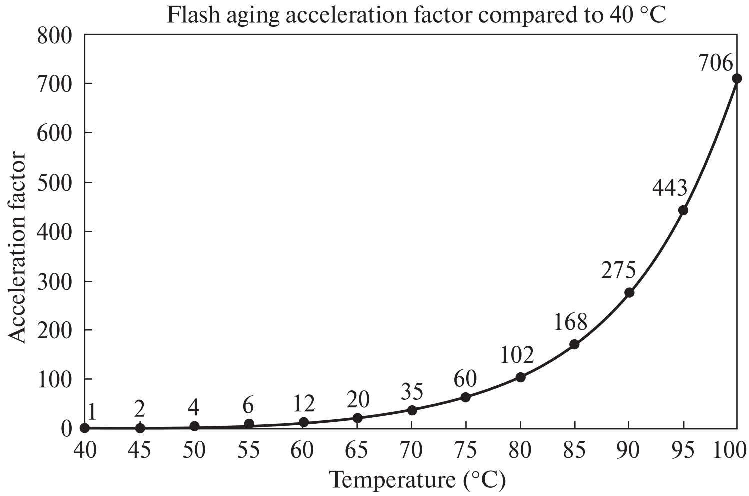


Fig. 3.5. Flash memory storage aging [70]

phisticated buffer management algorithms in order to reduce write and/or erase operations to flash memory. Some buffer management algorithms determine a request to be hot or cold based on its requested data size. But the data size could become a wrong indicator of update frequency in many applications. A new hot/cold identification scheme offered in order to increase identification accuracy. It can enhance storage performance and durability by reducing program/erase cycles. This technique uses the process identification actual for many operating systems as a hot/cold indicator [68]. External SSDs are rugged and virtually shock-proof, but the NAND they employ won't hold data forever. Electron traps like the cells leak over time. The technology is also relatively new, and no one is quite sure how long an SSD will retain data when stored unpowered. But now you won't find companies touting them for long-term backup [69].

The skill of NAND flash to store and retain data depends on the temperature which the NAND flash is subjected to during writing, and between the time of data writing to the time of data reading (Fig. 3.5). With the growth of the temperature that the NAND flash experiences, we can see the growth of charge de-trapping mechanisms acceleration that could lead to random data bit failures. NAND endurance is also impacted since endurance has an inverse relationship to data retention. At the time of programming and erasing NAND the rate of wear-out of NAND cells is affected by temperature. The temperature and duration to which the NAND flash is subjected to after programming is the most critical part in determining the acceleration factor, which relative to a temperature of 55 °C is shown in the table below.

For NAND devices specified with 1 year of data retention, storing at 85 °C will accelerate the charge de-trapping mechanism by 26 times when compared to storing at 55 °C. When the device has been cycled through the maximum number of program-erase (endurance) cycles as specified by the manufacturer, rather at the end of life of a NAND cell, data loss can occur if the NAND is stored or read over extended periods of time at high temperature. And contrary, when the NAND is stored or read at a lower temperature than 55 °C, the acceleration factor becomes less than 1. So, the NAND data retention is extended relative to the specification. At the end of NAND's rated endurance, the NAND device is usually not in jeopardy of immediate failure. It should be note, that NAND manufacturer's endurance ratings are typically specified to ensure that the number of bad blocks that occur over time will be within a predictable percentage limit. For another thing, the NAND will be able to retain data for 1 year at 55 °C in accordance with JESD47H.01. Blocks may become bad at a faster rate and the data retention capabilities of the drive become diminished beyond the endurance limit. The impulse to reliability of the drive is then dependent upon the media management capabilities of the drive controller [67, 69].

SanDisk gives an argument of 100 years on drives with corrosion protection. The theoretically estimated retention period is as much as 1300 years. SanDisk offers users the Memory Vault storage device based on modern flash memory for long-term file storage. Built-in controller permits you to erase and write any data, but after a certain number of cycles, the carrier switches to read-only mode. Such protection mechanism prevents accelerated wear of the memory cells, and they are still subject to "natural ageing". The speed of this aging is dependent on the temperature at which is accumulate the drive. Industrial standard JESD22-A117B allows you to calculate when the flash drive fails. The technique is based on the phenomenon that the life of one cell depending on the external temperature decreases exponentially [70].

Devices like the WORM can provide the long-term data storage on flash drives. Model SDCard Write Once from Toshiba has built-in protection from being overwritten. The recorded files cannot be accidentally erased. From the above-described Memory Vault this card is distinguished only by the lack of corrosion protection. SD Card Write Once can be used in specific areas, such as in the legal profession, to preserve evidence or proof, for storage of medical information [70].

Most flash storage technology is aimed at making access to media faster and boosting its longevity. Herewith, cold flash takes the opposite approach, opting for low endurance and low performance, due to the fact that photos and videos on Facebook are WORM (write once, read many) media, and it will not require several overwrites.

Cold Flash can be cheap, dense and low grade because it would be used as a WORM medium. It can take 10 times longer to write data than regular Flash memory, so, it is a little slower. After the data is written to the Flash media, it would only be "read", but not multiple rewritten. Such low-cost memory is not here yet, though Archival Flash devices are now coming to the market [71].

One way in which the data in a permanent solid-state memory device could be stored would be to represent a 1 with an intact fuse, and a 0 as a blown fuse. But the fuse material must be as long-lasting as the IC itself. We need a materials approach to solve this other half of the problem of permanence for solid-state digital data storage. We must find a material which is extremely stable, somewhat resistive but not an insulator. Also it does not grow dendrites when the fuse is blown. This would produce extremely long-lasting fuses either intact or as programmed. These materials exist, and we have been successful in producing and programming such fuses.

We can see a lot of development work to be done before ICs using these design concepts can be commercially available, but the fundamental concepts have been proven in the lab and in lifetime testing. The flash memory can be used for long-term data storage, but can not be considered as optimal choise because of high power consumption.

3.4. Lifetime of data storage based on microfilms

Most systems require a stable back-up copy to be maintained for all data (including migrated versions) in addition to online versions for longstanding use. Unfortunately, most digital storage techniques do not offer themselves for these purposes. Hard disc RAID arrays require regular operation and need to be replaced every few years. Tape drives as long-term storage of massive amounts of data require regular re-winding of tapes to maintain them readable. With current development cycles the persistence of the tapes has surpassed the support life-time for tape readers, rendering the respective tapes unreadable unless migrated to new types of tapes. This has led to the revival of a rather unexpected storage technique for digital data, like microfilming. Microfilm has proven a very durable medium, requiring no maintenance apart from appropriate storage conditions. Especially it is subject to black/white film.

Such storage medium like Microfilm has a life span of more then 100+ years. It has an advantage that a media migration has to be done less frequently. It is already used for long-term storage of scanned paper documents, texts and images. It is also in widespread use in archival institution, ensuring that expertise in its handling as well as appropriate technology is in place. In terms of cost, technology independence and stability, microfilm storage offers a promising solution for off-line storage [72].

The durability and secure read-back of analogue data create microfilm a migration-free storage alternative. Potentially it can create large cost savings and eliminate the risk of migration-related data loss. As a true WORM (Write Once, Read Many) it is literally impossible to delete or edit the information that is written to the film. So, by being an analogue medium stored offline, microfilm is also a secure solution. Hence microfilm meet many of the criteria for long-term preservation of valuable data [74].

Microfilm is 16mm or 35mm film stored in cassettes or on the open reels. It contains small images, typically in black and white, and some microfilm formats also support digital data. The two main types of microfilms are:

- silver halide film. It is the recommended alternative for long-term preservation. The image is transferred to the film by using silver emulsion bringing on a polyester strip.
- vesicular film. It creates the image on the polyester strip by using microscopic bubbles instead of silver emulsion. Such method makes it a less expensive but also a less durable solution. A microfiche has many of the same attributes as microfilm.

But it is formatted as a card rather than as a film reel. They are normally stored in open top envelopes in boxes or drawers. An ultrafiche like a compact version of a microfiche can store images at much higher densities and often made directly from computers [74].

Banks and hospitals used to convert their files to microfiche or microfilm in order to save on back up space and to make an additional copy of all their information. The data that you store as a business needs to be protected whether it is several years old or brand new. We need to have a good backup plan in place so that our film stays safe. Businesses traditionally use microfilm and microfiche as an archiving tool because it's reliable proven method to back up data over time. It's available and accessible if we lose the ability to access the technological solutions like cloud storage that is becoming more and more popular. Some of these techniques can also be applied to old family films and photograph negatives that you might want to store for your reasons. Microfilm and microfiche are two storage methods that many businesses rely on to store their data. This type of information is quite safe, and it has an expected shelf life of about 500 years if you choose the relevant materials for the film.

While films can be almost ideal way to archive information, it's always good to have a backup variant. Every time you handle the microfiche or microfilm, you expose it to conditions that may damage it. It's easier to access the information if it is stored on a database. You also can keep your films safer, by making a digital copy for your employees to reference for research or other reasons. While digital storage may not take the place of film storage, it is an additional method to your storage solutions [73].

Microfilm was designed for storing analogue images. Traditional microfilm is simply not an efficient alternative in a world overflowing with digital data [74]. It is possible to store digital data, files or documents on microfilm and to convert them back to a digital format. This offers a healthy solution for the permanent storage of the original data streams as defined in most archival regimes in long-term archives and e-Government settings. Advantages of long-term archiving on microfilm include specifically the lack of maintenance needed for maintaining the data, as opposed to most other digital storage alternatives which need regular media replacement or handling. A further advantage is the simplicity in terms of technological requirements for decoding and reading the data, as normal scanning/imaging devices coupled with OCR software are suitable to decode the information, as opposed to more complex systems required for specific tape drives or spinning disc type storage media. Due to the analogue storage medium hybrid storage of dual analogue and digital views of the data side-by-side are real, allowing easy inspection of data prior to reading and decoding into its first original form [74, 75].

Computer Output Microfilming (COM) technology provides pre-scans the original with the follow derivation of the resulting graphics file at microforms like frame microfiche or film roll fram. According to the "classical" method subsequent chemical-photographic processing microforms, depending on their type can be performed automatically and separately within the COM system.

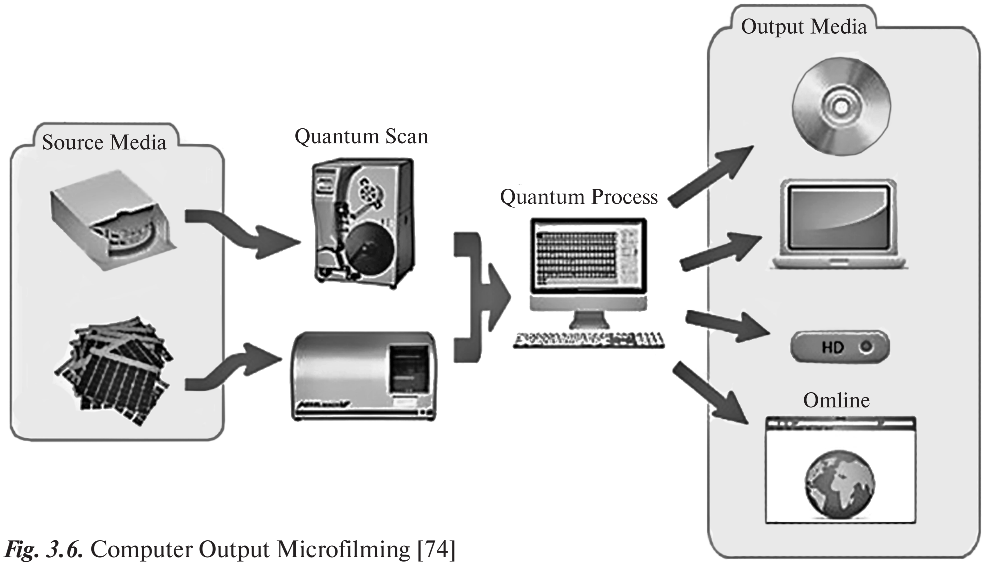


Fig. 3.6. Computer Output Microfilming [74]

You may scan the microfiche and microfilm and turn your film reels or negatives into digital copies (Fig. 3.6). This process takes some, and so you may only choose to do this to the most actual and important documents.

There is another solution, when you convert your microfilm to digital images for backup [165]. To make a graphical barcode using a special program that allows you to convert your digital document in black-and-white raster two-dimensional bar code image suitable for recording on microfilm through the COM system. Black-and-white bar codes allow achieving a relatively high density information recording. The bar-code bitmap is recorded on the microfilm stored there. Then if it is necessary it can read using conventional scanner microfilm. This method in combination with an effective error correction decoding on the output allows to accurately reproducing the information [81].

Microfilm as a preservation tool for recorded history was the real reason of its importance. A microfilm image of an historic map or a newspaper preserves that image for estimates of over 500 years, and is therefore quite enduring and stable. It is a usable tool for future generations that can be employed in tandem with other media. It is truly the best format to protect our history, because digitized images can be microfilmed, and microfilm can be digitized for ease of access.

At the new level of development of micrographic equipment we see the possibility of storing information on microfilm in a hybrid way, where the microfilm is written as digital code and analogue image. It can be used as a hybrid carrier, combining analog and digital information, because the microfilm allows you to read information and the person and the electronic device. To record a color microfilm can be used color laser COM system of a new generation Archive Laser Recorder, has achieved enough high recording density.

The new technology of digital data preservation on microfilm in the long term allow to preserve on microfilm all digital documentation and else information. In addition to the ongoing conservation of digital color and black-and-white graphic, photographic and textual documentation offers the possibility of preserving on microfilm any type of digital data like the digital audiovisual documentation, software products, documentation of three-dimensional CAD applications, etc. [75].

Piql Preservation Services is a solution that has been designed with the necessity of long-term digital preservation in mind. In result we have a system that allows digital data to be preserved efficiently and safely. It is easily retrievable independent of future access to specific technologies or vendors. Piql Preservation Services differs from traditional microfilm in enabling digital data to be stored alongside visual images. It also makes this method suitable for the digital era. The flexibility of analogue information and combining digital opens up for new prospects and additional security [75]. Piql offers Preservation Services based on a turnkey system designed to comply with the necessity of long-term digital preservation. As a result we see a technology allowing digital data to be preserved efficiently and safely, and easily retrievable independent of future access to specific technologies or vendors.

Piql uses high-resolution photosensitive film as a digital storage medium. Information is written to film as large QR-codes, each containing 8.8 million pixels. This allows any type of data to be preserved offline, on a storage medium with a documented lifetime of as much as 500 years. Piql's solution offers a lot of the same preservation qualities as microfilm, but applied to digital data. As a true WORM stored offline, it is impossible to manipulate or delete data after it is written. Digital data written as 0 and 1 obviously need decoding to get back the original file. To make data retrieval future-proof, explanations of how to decode and retrieve the data is written as human readable text on the storage medium. Valuable data can be preserved for the unforeseeable future by being self-contained with a documented lifetime of 500 years. The technology is also migration-free in terms of storage medium, reading device and file formats, provided that archival formats are used. The latter is due to the fact that file format specifications can be preserved in readable text on the storage medium.

With Piql Preservation Services, data is written on to photosensitive film, a proven longevity of which is counted more than several hundred years. The solution includes all processes and equipment which is necessary to write and read-back data. The data can be recovered with simple means such as a light source and some sort of digital camera and computer, and also a data scanner is developed for high-quality data retrieval. Piql Preservation Services allows for metadata Digital and analogue data searches and accessibility on a different level than traditional microfilm. As a seamless element within a standard IT infrastructure, users can search for the requested file and get it back in original format within minutes. The real retrieval time will vary according to each user s set-up and the size of the film. Migration-free long-term preservation: Piql Preservation Services is not suggested for data that needs to be immediately accessible to ensure business continuity. It is rather a safe, future-proof and efficient option for valuable data that requires to be preserved for future use. Piql's technology allows for metadata searches and accessibility on a various level than traditional microfilm. Users can search for the requested file and get it back in original format within minutes the storage medium as a seam less element within a standard IT infrastructure. Valuable data can be preserved for the unforeseeable future by being self-contained with a documented lifetime of 500 years. The technology is also migration-free, both in terms of storage medium, reading device and file formats (provided that archival formats are used). The latter is due to the fact that file format specifications can be written in readable text on the storage medium. Piql has developed all hardware and software required to write and read back data. Yet there is no vendor lock-in as the decoding software is open source and the data can be retrieved by using any computer and digital camera available in the future. Piql Preservation Services offers many of piqlbox: film final package. New polymer materials for a microfilm-based "Digital Optical Tape" (DOT) medium are developed to increase the reliability of microfilms [75].

There are some main reasons to microfilm records and documents. For a number of reasons microfilm is widely used for long-term preservation of data.

Long durability and no need for decoding: Microfilm has a lifetime expectancy of up to several hundred years when stored correctly at the recommended temperature and humidity levels. As the data is Microfilm reader normally analogue, data stored on microfilm can be read back by simply using a magnifying glass. Desktop readers with large screens and zooming lenses are normally used for data retrieval. Hence data retrieval is independent of specific reading devices, and the images on the film require no software decoding. The downside is that it is a manual and cumbersome process to access and reproduce data stored on time proved microfilm [72-75].

Information Security — information stored on permanent microfilm is accessible for up to 500 years if stored securely and properly; not vulnerable to internal or external hackers as with digital data or any else information stored in the cloud. Microfilm is difficult to change without detection, is highly resistant to detriment, and is hardware-independent making document recovery possible from one generation to another. Microfilm is becoming the actual archive media of choice for government or historical documents as well as genealogical histories and legal documents. First of all it is connected with the reason of it standing up to the test of time.

Document Integrity — microfilm does not aggravate with age like paper nor stand the chance of digital data loss through migration or a failure to migrate; microfilm is also not sensitive to outdated software, technological obsoleting or hacking like digital data. A microfilmed image guarantees the integrity of documents and their content from cyber theft, as well as allowing an option for information reconstruction in the event of a misfortune. Longest Lasting Media — ensures safety longer than any other media. The microfilm provides storage for information that is fully consistent with the original during the warranty period near 100 years.

Microforms, including microfiche, are not adopted for audio and video, and color images raise the cost of it, though it is possible. The same environmental factors as paper can affecting on it, but they are less resistant to high heat. Retrieval of individual records rather slow, although proper indexing and computer-aided finders improve retrieval times [74].

The principal drawback with microfilm is that the workflow for reproducing data is time-consuming and largely manual. The storage conditions are the most significant aspect of film preserving. The two key environmental factors for microfilm collections are temperature and relative humidity. The recommended humidity rate is below fifty percent, or even lower for certain types of films. The maximum storage temperature is 70 degrees, though most users recommend storing the films at lower temperature. Often companies will contract out the storage of this type of material to businesses that work to preserve this environment. This is because it can be difficult to preserve all of the necessary conditions.

For the dependable storage of files are recorded on microfilm and are constructed using specific warehouse with constant humidity and temperature. Here are examples of the most famous stores.

Since 1938, the genealogical society of Utah has been collecting historical and genealogical data on rolls of microfilm. The Granite Mountain Record Vault is the permanent repository for these numerous microfilms. It is located about one mile from the mouth of Little Cottonwood Canyon in Utah's Wasatch Range, and twenty miles southeast of downtown Salt Lake City.

As it is commonly known, The Vault is a massive excavation reaching 600 feet into the north side of the canyon. Architected between 1958 and 1963 at a cost of $2 million, it consists of only two main areas. The office and laboratory section sits beneath an overhang of about 300 feet of granite, and also houses shipping and receiving docks, microfilm processing and evaluation stations, and administrative offices. Under 700 feet of stone, the Vault proper is situated farther back in the mountain behind the laboratory section and consists of six chambers, which are accessed by one main entrance and two smaller passageways. Each of this chamber is 190 feet long, 25 feet wide, and 25 feet high. Particularly constructed Mosler doors weighing fourteen tons at the main entrance and nine tons guarding the two smaller entrances are designed to survive a nuclear blast. In the six chambers, nature maintains constant temperature and humidity readings optimum for microfilm storage. Each of these chambers contains banks of steel cabinets ten feet high. Near 2.4 million rolls of microfilm, in 16 mm and 35mm formats, and approximately 1 million on microfiche, were housed in two of the six chambers. This collection increases by 40,000 rolls every year. Optical disks with greater capacity for storage than microfilm and other alternate media are being considered for use and may make further expansion of the Vault unnecessary. The genealogical information contained on these microfilms is collected from governmental agencies, libraries, churches and consists primarily of birth, marriage, and death registers. Also researcher can find their the wills and probates, census reports and other documents that can be used to establish individual identities. Such ordinances are considered essential for the salvation of the dead-that is, those who died without hearing the full message of the gospel of Jesus Christ. Latter-day Saints use this information to assemble family group charts and pedigrees for the purpose of binding together ancestral lines of kinship through sealing ordinances performed by proxy in temples [75].

The Barbarastollen underground archive intended to preserve Germany's cultural heritage from man-made or natural disaster. It is located in a disused mine near Freiburg im Breisgau, Baden-Württemberg, Germany. It holds microfilms with about 900 million images from German museums and archives. The Barbarastollen tunnel is guarded by the Hague Convention for the Protection of Cultural Property in the Event of Armed Conflict of UNESCO. It is a "refuge intended to shelter, in the event of armed conflict, the movable cultural property" according to article 1 subparagraph b. And it is the only cultural property in Germany under special protection according to Chapter II of the convention, and one of only five such sites worldwide, the others being the Vatican City and three refuges in the Netherlands. Saint Barbara was the source of its name. The entire complex is buried under 400 meters of the stone. It is thought to survive a nuclear war, and its contents should survive for at least 500 years without any real damage. About 1,500 barrels of microfilm with a total length 31350 miles accumulated at the Federal archives of Germany, receiving documents for long-term storage since 1975. Humidity inside is 75 percent, and the average temperature is 10 degrees. The Basic law of Germany, like all the documents that accompanied its adoption stored in this great archive on the microfilm. Send a copy of the first edition in archive 3 October 2016, on the occasion of the Day of German unity [73].

The Pennsylvania Historical and Museum Commission (PHMC) manages the Local Government Security Microfilm Storage Program. The program began in 1987 and stores approximately 200,000 rolls of microfilm for state, county and local governments and even school districts. It works through the Records Services Division of the Pennsylvania State Archives (PSA) to realize this service. The private records center is leased from Iron Mountain at its underground facility in Boyers, Butler County [76]. The abandoned mines gave the opportunities for underground storage facilities since 1954. The Iron Mountain Corporation merged with the longtime owner of the facility in 1998 and realized a large storage facility located in a former limestone mine near with Boyers Corbis. The United States Office of Personnel Management (OPM) maintains a highly secure facility in an abandoned mine in Boyers, PA, which contains documents from security clearance proceedings. The United States Patent and Trademark Office stores original records in an underground storage facility in Boyers. Iron Mountain facility in Pennsylvania also houses the records of countless corporations and highly sensitive government agencies in its array of tunnels. Also it may best be known as the home to the photographic collection of Bill Gates' Corbis Corp. venture [76].

The Library of Congress movie and film archive is located in an underground bunker in Culpepper, Virginia. It was originally a gold storage unit and later a fallout shelter during the Cold War, and now the Library of Congress stores film there. It is used to ensure the survival of the nation's films through preservation and restoration. Also the center specializes in repairing and processing films of many different types and sizes. It includes nitrate films in 124 nitrate film vaults. Some of the films are over 100 years old. They preserve old and new films not necessarily the monetary value, but more for the historical value. The purpose of it is to remember what the early times were like, what we had and what we did.

Special cameras capable of photographing at reduced size are used in order to transfer an image to microfilm. Then the image is printed on the film and chemically processed in laboratory-like environments. The film processing makes the recording process more complex compared with alternative technologies. There was developed special processing equipment for the automated manufacturing of microfilms. Information about this equipment is given below.

Fujifilm's AR-1000 Document Archive System is network enabled and offers high-speed recording of 400 pages per minute. A dual film magazine minimizes manual intervention and increases workflow operations without compromising speed. Its 5.9 target on film is equal to 300 dpi output and autorotation automatically rotates and orients all images with typewritten text. The user-friendly Virtual AR™ job setup software allows for the preview of data transferred to film and proposes a document management solution from creation through storage and destruction.

Each of Fujifilm's PET-based silver halide archive media proposes optimal sensitivity for red LED, panchromatic sensitivity for the vivid duplication of images, and a resolving power of 850 lines/mm. Also archive Media AM-66 has a 2.5 mil thickness and is 16 mm × 66 m/215 ft whereas the new AM-33 has a 5 mil thickness and is 16 mm × 30.5 m/100 ft, catering to low volume applications. AM-33 will run from a separate magazine than AM-66, further buildup the proficiency of operations [162].

A very cost effective process will save essential data onto the proven media microfilm. The microfilm created will suffice all the requirements of long term archive. The SMA 16+ produces an analog digital files backup and includes hardware, software, and migration independent. Properly and stored produced microfilm under the right conditions has an ANSI certified life expectancy rating of 500 years. Microfilm cannot be manipulated through coding, unlike digital files it cannot get a virus or spyware, also it cannot be hacked by malicious parties. The SMA 16+ can handle image files such as TIFF, JPEG and PDF etc., it will accept and write bitonal, grey scale, or color images without file size restrictions. The converting process is reliable and rather quick providing excellent image quality. The original document size can vary from as small as a postcard up to engineering drawings and large maps. The SMA 16+ requires minimal user intervention. And once the converting process has begun, the system will work almost unattended. This system includes a PC workstation and intuitive operating software, and it can manage digital files. In duplex mode, the SMA 16+ can write up to 8200 images per hour to 16mm microfilm. For smaller originals like library cards or checks an optional available nesting tool can pre-arrange the images which results in a much higher frame per hour conversion rate [163].

Archive Laser system can store digital data as a bit stream on long-term stable optical film (microfilm). Also it is able to retrieve it afterwards with simple optical scanners. One of the core achievements of this research is the density of more than

1 Terabyte per reel of film. This data carrier enjoys all the advantages and characteristics of an optical, analog data carrier. Among them are: long-term durability of up to 500 years, no manipulation or virus attacks and independence from rapidly changing soft- and hardware generations [77].

Technology of the digital data long term preservation on microfilm is allowed to store digital documentation and any information. The electronic document type does not matter, since all digital files can be represented as a graphical two-dimensional barcodes and thus are composed of a set of binary data. This kind of saving information allows you much closer to solving the problem long-term insurance preservation of electronic documents [77]. But the technology long-term storage of documents on microfilm has some disadvantages. An analog technology of microfilming is based on making reduced to the sizeframe analog microfilm copies of the original. The high cost of equipment and supplies materials, the complexity of the technological process of microfilming make the use of microfilm economically inefficient way to long-term storage of data.

3.5. Lifetime of DNA-type memory

We see the necessity of establishing long-term storage technology. With the rapid transition of records from paper to digital media, with respect to cultural heritages or public documents, semi-perpetual storage technology permitting access to recorded data throughout generations without any degradation resulting from heat or humidity, is really necessary today. To meet this necessity, research is being pursued worldwide on different long-term digital data storage technologies (Fig. 3.7) including the use of semiconductors and biological DNA [78-84].

Using DNA to archive information is a really tempting possibility because it is special dense, with a raw limit of 1 exabyte/mm^3 (10^9 GB/mm^3), and long-lasting, with observed half-life of over 500 years [79, 83]. About the high potential of DNA memory of storing large amounts of data illustrated in Fig. 3.8 comparative assessment of the various types storage media capacity [84].

DNA-based storage also has the advantage of eternal relevance: as long as there is DNA-based life, there will be well-defined reasons to read and manipulate DNA. There are two standard practices in biotechnology, from research to diagnostics and therapies. The write process for DNA storage maps digital data into DNA nucleotide regulating (a nucleotide is the basic building block of DNA), *synthesizes* (manufactures) the corresponding DNA molecules, and stores them away. Reading the data involves *sequencing* the DNA molecules and decoding the information back to the original digital data. Progress in DNA storage has been rapid: in 1999, the state-of-the-art in DNA-based storage was encoding and recovering a 23 character message [52]; and then, in 2013, researchers prosperously recovered a 739 kB message [78-84]. The basic unit of DNA storage is a DNA strand that is roughly 100-200 nucleotides long, capable of storing 50-100 bits total. A typical data object maps to a great number of DNA strands. The DNA strands will be stored in "pools" that have

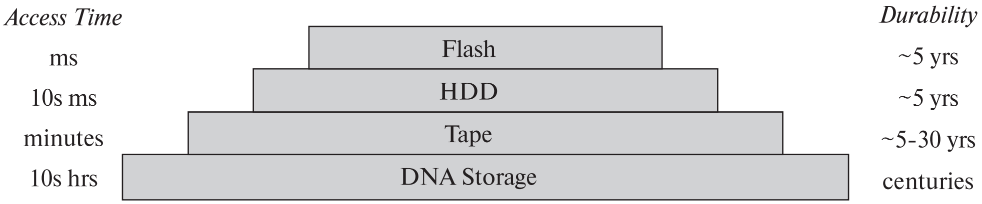


Fig. 3.7. DNA storage as the bottom level of the storage hierarchy [79]

stochastic spatial organization (Fig. 3.9). They do not permit structured addressing, unlike electronic storage media. In view of this fact it is necessary to embed the address itself into the information stored in a strand. After sequencing this way one can reassemble the original data value [81].

Recent studies have made large strides in developing DNA storage schemes by exploiting the advent of massive parallel synthesis of DNA oligos and the high throughput of sequencing platforms. It is necessary to create a strategy to store and retrieve DNA information that is robust and approaches the theoretical maximum of information that can be stored per nucleotide. The matter is that most of these experiments reported small errors and gaps in the retrieved data. The success of such strategy lies in careful adaption of recent developments in coding theory to the domain specific constrains of DNA storage. New strategy can retrieve the information without a single error. To further stress strategy was created a deep copy of the data by PCR amplifying the oligo pool in a total of nine successive reactions. It reflects one complete path of an exponential process to copy the file $218*10^{12}$ times. The original data optimally retrieved with only five million reads. Taken together, this access opens the possibility of highly reliable DNA-based storage that approaches the information capacity of DNA molecules and enables virtually unlimited data retrieval [78]. So, DNA memory has great potential for long-term storage of large amounts of any kind of information (Fig. 3.10) [78].

The attainment of that potential won't be easy. Before DNA becomes a viable competitor to conventional storage technologies, researchers will have to surmount a host of challenges, from reliably

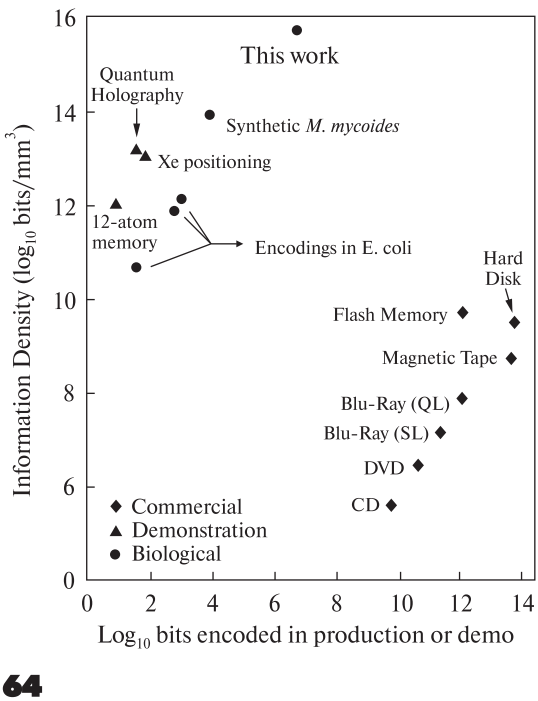


Fig. 3.8. Density of recording on different media types [54]

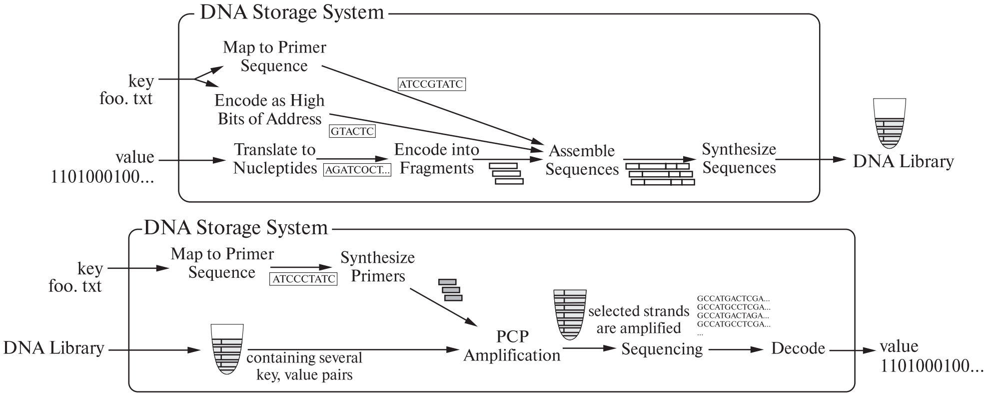


Fig. 3.9. Overview of a DNA storage system [81]

	 HARD DISK	 FLASH MEMORY	 BACTERIAL DNA
Real write speed (ms per bit)	~3,000-5,000	~100	<100
Data retertion (years)	>10	>10	>100
Power usage (watts per gigabyte)	~0.04	~0.01-0.04	$<10^{10}$
Data density (bits per cm3)	$\sim10^{13}$	$\sim10^{16}$	$\sim10^{19}$

Fig. 3.10. Storage limits of different storage systems [78]

encoding information in DNA and retrieving only the necessary information, to making nucleotide strings cheaply and quickly enough [125]. DNA exceeds digital media for storage and copying and matches something like the Rosetta Disk for longevity, the problem is the technology to make, store, and read such DNA texts is relatively high tech. Therefore it is vulnerable or unworkable in many catastrophic scenarios. It's also much less readily searchable than digital media or common indexed paper texts.

DNA has some advantages for storing digital data because of it's ultracompact form, also it can last hundreds of thousands of years if kept in a dry, cool place. And as long as human societies are writing and reading DNA, they will be able to decode it. Capable of storing 215 petabytes (215 million gigabytes) in a single gram of DNA, the system could store every bit of datum ever recorded by humans in a container about the size and weight of a couple of pickup trucks. But whether the technology takes off may depend on its cost. And compared with other forms of data storage, writing and reading to DNA is relatively slow. Unfortunately the new approach isn't likely to fly if data are needed instantly, but it also would be better suited for archival applications.

3.6. Lifetime of data storage based on optical discs

The development of optical discs with ultra-long shelf life has been a subject of violent research. Some engineering solutions have been developed and proposed for increasing the data storage lifetime on optical mediums with plastic substrates. There were developed many technologies of information storage shelf life increasing on optical mediums of different types. Development of special recording layers is the main direction of data storage reliability enhancement on write-once optical mediums. Voids in polymers and glass materials which are permanent laser-induced physical changes can approach to long-lifetime storage without data degradation. Fundamentally, laser inscriptions in glass materials can withstand temperatures of up to 1000 °C. Also they can maintain data stability and readability for up to thousands of years. Each bit cannot be smaller than $\approx \lambda/2$, where λ is a light wavelength due to the optical diffraction barrier; this limits a capacity to 50 GB per disc. The top-down nanocomposite approach proposes an alternative method. It may allow the development of SPIN methods to break the diffraction limit for ultrahigh capacity optical data storage with thousands of years of lifetime [85-90].

There are optical discs that are unquestionably the hardiest archival mediums available to consumers. For instance, the accelerated aging tests of 4.7 GB DVD+R Data Tresor disk show archiving life expectancy of 160 years [88].

Write-once BD-R HTL (High To Low) can last for 100 to 150 years given a relatively mild environment — i.e., not on your dashboard in Phoenix. Milleniatta's M-Disc BD-R and DVD+R write-once discs use an even more stable data layer that is rated for 10,000 years. Only its polycarbonate outer layers reduce that to a mere 1,000 years. Of course, it is all theoretical, but the testing MOs were censorious and performed by the government of France (BD-R), and the Navy for the Department of Defense (M-Disc DVD) [89].

It is possible to perform photolithography at the nanoscale using visible light because of the recent advances in materials science have made. One approach to visible-light nanolithography, or resolution augmentation through photo-induced deactivation, uses a negative-tone photoresist incorporating a radical photoinitiator. It can be excited by two-photon absorption. With follow absorption of light, the photoinitiator can also be deactivated before polymerization occurs, and the deactivation step can be used for spatial limitation of photopatterning. Two-beam optical lithography which utilizes a doughnut-shaped inhibition beam to inhibit the photopolymerization triggered by the writing beam at the doughnut ring, leading to reduced feature size and enhanced resolution. The fabricated feature size and resolution by two-beam OBL can break the limit defined by the diffraction spot size of the two focused beams, although both focused writing and inhibition beams result in the spot size limited by diffraction. It is the fact, that the smallest feature size and the highest resolution are confined by the mechanical strength of the solidified material. It can be far beyond the diffraction limit provided that an appropriate photoresist with high mechanical strength can be developed. Physical and chemical processes driven by multiphoton absorption make possible the fabrication of complex, 3D structures with feature sizes near 100 nm. The field of multiphoton fabrication has progressed rapidly since its inception less than a decade ago. Multiphoton techniques are now being used to create functional microdevices [78, 79]. Inspired by stimulated emission depletion (STED) microscopy superresolution lithography has become popular in the biosciences due to the dramatic resolution gains it makes possible. A alteration on confocal microscopy, STED works by the nonlinear, spatially modulated de-excitation of fluorophores to restrict fluorescent emitters to volumes or areas smaller than the diffraction limit. A straightforward extension utilizes an excitation beam, which causes a photoreaction initiating polymerization. Then the ring-shaped depletion-beam induces a de-excitation reaction, inhibiting the reaction of polymerization. It is possible to confine polymerization to a region that is smaller than the far-field diffraction limit by controlling the region of inhibition to surround the polymerization region sufficiently. In an exemplary implementation of STED microscopy [53-56], a collimated excitation beam is focused to a point by an objective lens to attain an approximately Gaussian, diffraction-limited point-spread-function (PSF) in the sample plane. Another beam, at a longer wavelength causing fluorophore de-excitation, is modified with a phase plate or grating. Then it is combined with the first beam via a dichroic mirror so that the two beams are colinear before passing into a microscope objective lens. The modified phase of the deexcitation beam causes it to have an intensity profile like a higher order Gauss-Laguerre or Hermite-Gaussian function at focus. With the main feature it is being a central null in the intensity profile. As the intensity of the deexcitation beam is enhanced relative to the excitation beam, the excitation spot at focus is more and more squeezed down in area or volume enabling diffraction-unlimited superresolution. Several lithography techniques have appeared recently that bring in the basic STED concept in reverse, to attain superresolved features in a photopolymerizable resin [85].

3.7. Conclusions

• Magnetic media based solutions (from tape to hard disk drive) based archive systems, monopolize the data archiving market due to their high capacity and low cost. However, in the era of big data, rapidly increasing volume, velocity, and variety of data set bring numerous challenges to the archive systems in various aspects, such as capacity, cost, performance, reliability and power consumption.

• Magnetic tape offers relatively high recording densities and relatively low cost per unit of capacity. As improvements in speed and capacity typically lead to a new generation accompanied by changes in drive systems and software. While writing can go back one generation and reading can go back two, relatively frequent data migrations are required for long-term storage.

• Flash memory has a serious drawback that prevents this type of media replaced all existing optical and magnetic storage, and is connected with the reliability and durability: it has a finite number of cycles of erasing and writing. Estimates of the manufacturers, modern flash memory can withstand an average of 100 thousand cycles erase /write.

• Microfilm has proven to be a very durable medium, requiring no maintenance apart from appropriate storage conditions. Microfilm as storage medium has a life span of more then 100+ years and has the advantage that a media migration has to be done less frequently. In terms of cost, stability and technology independence, microfilm storage offers a promising solution for off-line storage.

• Using DNA to archive data is an attractive possibility because it is extremely dense, with a raw limit of 1 exabyte/mm^3 (10^9 GB/mm^3), and long-lasting, with observed half-life of over 500 years. About the high potential of DNA-the memory of storing large amounts of information comparative assessment of the capacity of the storage media of various types.

• The development of optical discs with ultra-long shelf life has been a subject of intensive research. Many engineering solutions have been proposed and developed for increasing the data storage lifetime on optical mediums with high stable substrates.

CHAPTER 4

LONG-TERM USABILITY OF OPTICAL MEDIA

Optical discs such as CDs and DVDs are easy to exploitation, and both the hardware and the media are relatively inexpensive to procurement. Archivists widely use optical discs. High failure rates: Originally developed primarily as a mass consumer product, elaborate measures are needed when using optical discs for long-term preservation. But optical discs have a relatively serious risk [90]. CD and DVD recordable data is proving primarily problematic, as cheaper suppliers are continually squeezing higher quality media vendors out of the marketplace, as they did in the hard drive sphere. Hereafter, despite lofty claims of up to 100 year lifespan for the technology, real life testing has shown far shorter lifespans for this media than depended [92]. Traditionally, WORM optical media manufacturers claim five years of shelf life for blank disks and twenty to thirty years of existence after recording. These life expectancy claims are based upon test procedures that vary between manufacturers [92].

4.1. Peculiarities of optical information recording

The development of new types of different communication techniques and media has led to the changing role of optical media in the recording systems, also storing and transmitting data. Optical media may see less use for live recording, but it is still popular for archive aims. If you think of optical (CD/DVD/Blu-ray) solely as a means of movie or software delivery, it probably seems obsoleted. You might also dismiss garden-variety CD, DVD, and BD-R (LTH) [90-92, 185].

The successful use of optical disks for long term data storage is built on the characteristics and capabilities of modern optical recording. Optical media for long term data retention have significant advantages compared to other technologies of data storage, namely:

1. Recording information on optical discs is non-contact, which prevents surface of the carrier. Optical data recording has an impor-

tant advantage when it comes to data longevity. As a result the recording and playback process does not involve any contact between the media and the recording or playback mechanism. This separation between the recording or playback device and the media allows an infinite number of playback iterations, which is unique when compared to tape. The relative simplicity of the playback mechanism means that, if the data persists on the media, future optical playback systems will easy be capable of being adapted as necessary to read data stored constantly on optical discs.

2. Recorded images can be analyzed visually using optical instruments. The presentation of data on the surface of optical storage media in the form of microrelief structures allows reading the recorded information by numerous methods.

3. Transformations which occur at temperatures above 20000 C for the manufacture of optical media can be used high-stable materials.

4. Record data on optical discs is carried out in formats that ensure compatibility of recorded discs of various types.

5. For reproducing data from optical disks does not require particular software.

6. The playback device data from optical discs created with the intention of backward compatibility, implemented from CDs in CD to BD.

7. The characteristics of the optical discs create opportunities for high-reliability long-term storage of data changes in a wide range of humidity and temperatures. CDs retain the recorded information is usually longer than other digital storage media. Long experience of CDs using proved that their use is a smart choice for archiving data, images or text in digital form. Studies of long-term data storage on CD-ROMs showed that the nickel stampers, used for copying CDs, can provide long-term data storage [92, 185].

You should handle optical media with the same care as a photographic negative. Although discs are read from the bottom, the layer containing the track is actually much closer to the top of the disc because the protective lacquer overcoat is only 6–7 microns thick. The disc is an optical device and degrades as its optical surface becomes dirty or scratched. Writing on the top surface of a disc with a ballpoint pen easily damages the recording underneath. Be careful even when you use a marker to write on the disc. The inks and solvents can damage the print and lacquer overcoat on the top of the disc, and afterwards the information layer right below. Use only special markers designed for or tested as being compatible with optical media. The significant thing is to treat both sides of the disc carefully, especially the top (label) side (Fig. 4.1).

The critical portion of an optical disc is the data layer, and in fact it can be damaged relatively easily. Although in theory data layer is well protected, in reality it is not so. Because optical discs are a dense form of information storage, even small degradation can cause significant information loss skipping or repetition of tracks. Some deposits, such as fingerprints, may cause etching of the plastic surface leading to irreversible damage. The polycarbonate plastic layer can “flow” over time. There are many sources of potential damage to discs: Solvents can affect the lacquer layer and subsequently the metal layer on a CD. A common cause of optical disc failure

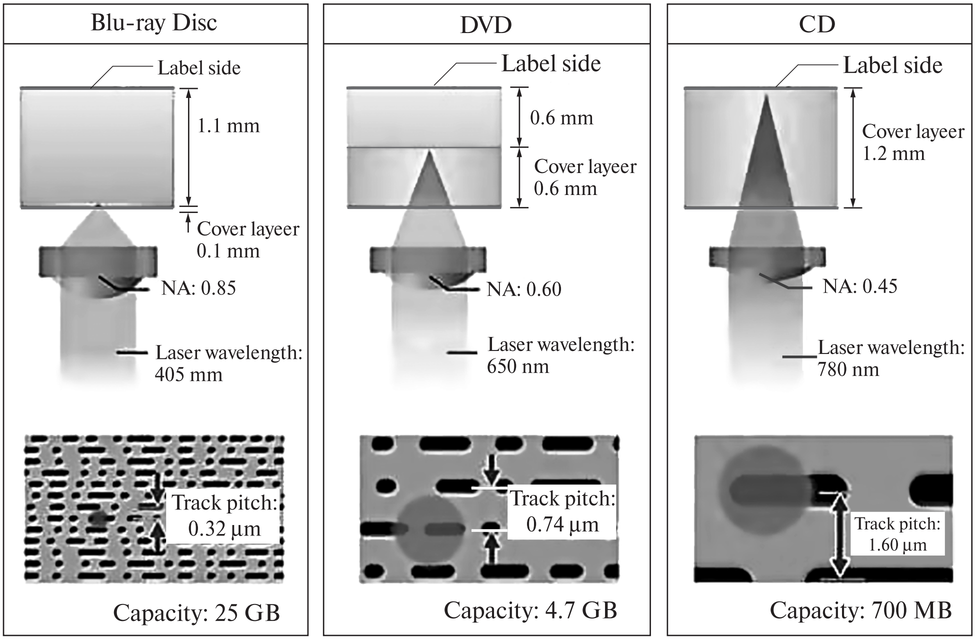


Fig. 4.1. Improving the systems of optical information recording [89]

is damage to the polycarbonate plastic layer. They are read optically, so, any marking that interferes with the light path (e.g. scratches or surface deposits) can cause problems. Inks used to print information on the label surface may corrode the plastic or lacquer layer and subsequently the metal layer. The plastic layers may slowly lose their shape, making them difficult to read. Sharp points can easily damage the lacquer and metal layers, making the disc unreadable. CDs and DVDs are particularly prone to damage to the "label" side from writing implements. As with all record media, particularly dramatic and sudden changes in the temperature and the humidity can cause degradation.

Disc becomes unreadable in a result of corrosion of the metal layer. Gold and some certain metals are more resistant to corrosion than others. Discs are susceptible to mould which can digest and damage the information layers. Avoid discs sticky tape and labels, and the adhesives can eat into the writing surface, damaging data [92, 176].

If left in an environment that allows direct extreme heat and/or sunlight the organic dye or phase changing film that holds the data will degrade quickly. As a result the disc becomes unreadable in a matter of days. Paper and cardboard envelopes tend to generate dust, so they are unsuitable for the long-term storage of optical media. Optical discs should be stored in rigid plastic jewel cases, which are reasonably dustproof and suitable for long-term storage. They are usually constructed of an inert plastic. For long-term disc storage, it is prudent to remove the booklet or paper label insert from inside the case and attach it to the outside, perhaps in a sleeve. The paper can attract moisture and produce higher moisture content in the case. Also it can spread moisture by contact with the disc. The pressure may lead to warping or deformation. Discs should not be stacked or packaged in groups so that they lean against each other. Discs stored horizontally for a years can warp. Jewel cases seem to be the ideal enclosure because they support each disc at the hub and deflect any impact from other items. Discs should be stored upright. If you have to label the disk, use a water-based felt-tip permanent marker to mark the label side. Do not use ballpoint pens or pencils. Graphite dust from pencils can interfere with the reading of the disc and damage it. Do not use adhesive labels as the adhesive can damage the disc. If an optical disc becomes dirty, dusty or has fingerprints, it is still possible to clean it before permanent damage occurs. Take great care and first blow off debris with a compressed air duster. You can then wipe them clean, using a very soft brush, a lint-free cloth or a non-abrasive photographic lens cleaning cloth. Any finger marks and oily dirt deposits on the discs data side can be removed by alcohol-based CD/DVD-cleaning fluid. You may apply the solution sparingly to the disc surface and wiped off with a lint-free cloth. Remember, that cleaning motion should never be circular (along the tracks), always brush from the center of the disc outwards. If a scratch is created while cleaning, it will do less damage cutting across the tracks than along them [100].

Due to capacity, reliability and power consumption, the optical media based archive system becomes an attractive option for long-term digital preservation. HDD (Hard disk drive) techniques have a limited lifetime of 2-5 years. But frequent data migration is needed to avoid potential data loss. However, optical disc has a longer lifetime than HDDs, which dramatically reduces the required frequency of data migration. Optical disc only consumes energy when the data are written or read out, but it does not consume any energy when the optical disc is in idle state. Optical technology greatly reduces the waste generated by frequent data migration, reduces the cost for the replacement of new units associated with short lifetimes and reduces the energy consumption in idle status. Moreover, increasing the lifetime of optical disc to greater than 50 years can increase the savings in overall expenditures, including costs for storage devices and even electricity.

Despite its slow speed, optical disc is really perfect for archiving most significant data and optical disc retains a very strong presence in the archival community [92, 185]. It must be clearly understandable that the use of recordable optical disks as sole digital target media constitutes a real risk, though it is unfortunately widespread, for instance amongst small and less wealthy institutions.

4.2. Long-term optical data storage types

Not only scratching, but several environmental factors such as dust, heat or UV light can cause severe damage to the disks. Optical discs also offer low data capacity compared to alternative technologies. The life span also depends on manufactured quality, how well it is recorded and its physical handling and storage. To lower all the risks of failure, best practice is to frequently migrate onto newer formats.

This is in sharp contrast to manufacturers who tend to claim an expected lifetime of up to several hundred years. Blu-Ray is an optical disc designed to supersede the DVD format with a higher capacity, but today it is more popular among households than professional archivists. A Blu-Ray reader is needed for data retrieval. The future accessibility of these is a challenge the manufacturer does not give an answer to [89].

New developments in optical disks for long-term storage show new possibilities to better meet the changing demands of archiving data in terms of media data rate and storage capacity [89]. Analyze the optical disks characteristics that are available for archival storage.

Specification for drives Holographic Versatile Disc (HVD) was ready in late 2004 (Table 4.1). The volume of these vehicles was expected to reach 3.9 TB, and data transfer rate — 1 Gbit/s. General Electric Company continues to lead the development of this device. In early 2009, the experts from GE first introduced its holographic storage medium is a transparent disk the size of a regular DVD which can record up to 500GB of data. The information was recorded directly in the disc material and does not require any additional layers to the dye, from holographic technology. Due to the fact it was expected that it will provide a much larger storage terms. InPhase Technologies was counting on the period of preservation of data for at least 50 years. In the disk GE-only hologram with a size of 0.3×5 micrometer represents one bit of information. All the hologram located as well as modern Blu-ray discs or DVD and the number of layers on the new format storage can be from 50 to 100. The amount of information placed on this disc is much more. Such placement of the recorded data is intended to make existing playback devices compliant with the drive types that are available today in the market. Thus, GE aims at capturing both private and corporate clients [80].

In recent years, Blu-ray Discs (BDs) and holographic discs and other high capacity optical media have emerged with the revival of optical storage, e.g., the capacity of a BD ranges from 100 to 128 GB in current generation (BDXL, high capacity recordable and rewritable discs), and that of the next generation is up to 500 GB [15]. New BDXL format provided disk capacity 128 GB (write once) and 100GB (rewritable discs). But this new format has one noticeable disadvantage: the BDXL discs will not be read by existing players.

The accomplishments achieved in the technical development of the Blu-ray Disc format were utilized to create a larger-capacity optical disc standard for preserving valuable data in the enterprise market [89]. Now the

Table 4.1. **Archival Disc Specifications**

Parameter	Value
Disc Capacity (Type)	300 GB (Appendable)
Optical Parameters	Wavelength: 405 nm, NA (numerical aperture) = 0.85
Disc Structure	Double-sided (3 layers/side), land/groove recording
Track Pitch	0.225 pm
Bit Length	79.5 nm
Error Correction	Reed-Solomon
Recording Mode	Write Once (Appendable)

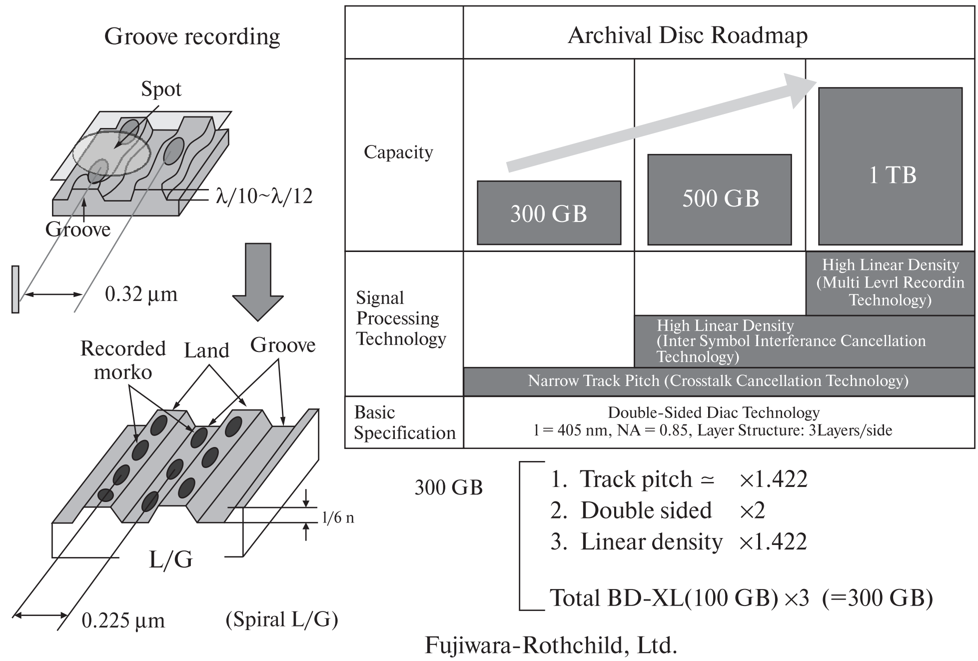


Fig. 4.2. Common information about Arhival Disc [89]

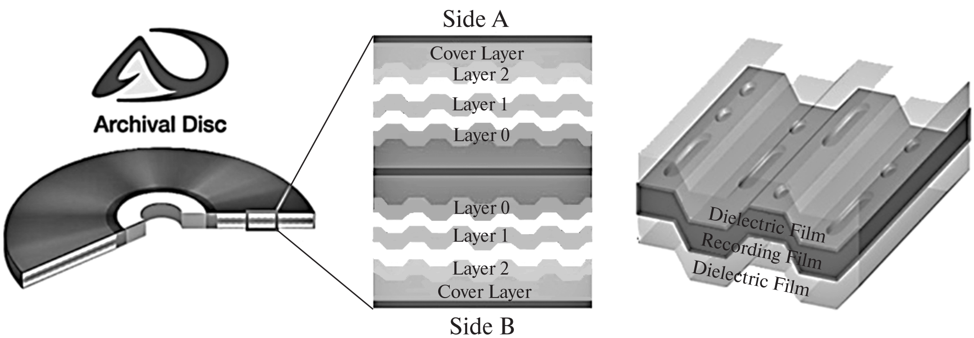


Fig. 4.3. Structure of the Archival Disc [137]

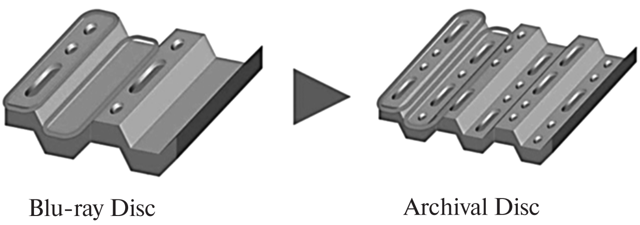


Fig. 4.4. High-Capacity Land/Groove Recording [183]

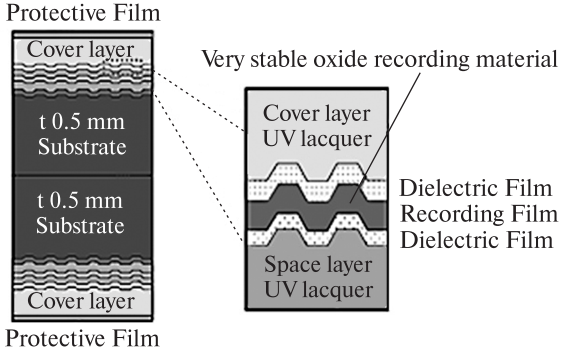


Fig. 4.4. . Cross-sectional diagram that illustrates a configuration example of an optical information recording medium [183]

main area of application optical disk is archiving data. Panasonic and Sony jointly established the "Archival Disc" standard for professional-use, next-generation, optical discs (Fig. 4.2).

A significant increase in the capacity of Arhival Disc compared to the Blu-ray Disc was achieved by applying multiple technical solutions. By implementing six discrete recording layers (three on each side), the new standard achieves a revolutionary jump in storage capacity, to 300 GB per disc.

The real values achieved in developed Blu-ray Disc format were utilized to create a larger-capacity optical disc standard for preserving essential data in the enterprise market [91, 175-185]. Today the main area of application optical disk is archiving data. Panasonic and Sony jointly established the "Archival Disc" standard for professional-use optical discs of next-generation.

A relevant increase in the capacity of Arhival Disc (Fig. 4.3) compared to the Blu-ray Disc (Fig. 4.4) was achieved by applying multiple technical capabilities. By introduction six discrete recording layers (three on each side), the new standard attains a revolutionary jump in storage capacity, to 300 GB per disc.

There are inorganic and organic pigment materials as the recording materials that are used in a direct read after write type optical information recording medium. Inorganic materials are being widely considered for the recording material. But organic pigment materials have mainly been considered for the recording material of a direct read after write type optical information recording media of the related art, with the mass storage optical information recording media of recent years.

As one of the widely considered inorganic materials, there is an inorganic material including a Pd oxide. An inorganic recording layer that includes In oxide and Pd oxide where the Pd oxide includes Pd monoxide and Pd dioxide and the ratio of Pd atoms with respect to the total of In atoms and Pd atoms is 6 to 60 atomic % is proposed in [91, 185] A recording layer with an oxide of a metal of which an absolute value of the standard free energy of oxide formation per 1 mol of oxygen is larger than that of Pd (hereinafter referred to metal X) and a Pd oxide, wherein the Pd oxide includes a Pd monoxide and a Pd dioxide, and wherein a ratio of the Pd atom to a total of the metal X atom and the Pd atom which are contained in the recording layer is 4 to 85 atomic % (Fig. 4.5). The transmittance changes of an information signal layer with an inorganic recording layer that includes a Pd oxide extremely before and after the recording of information signals. Further, an inorganic recording layer that includes one or both of In and Sn, Pd, and O is proposed in [91].

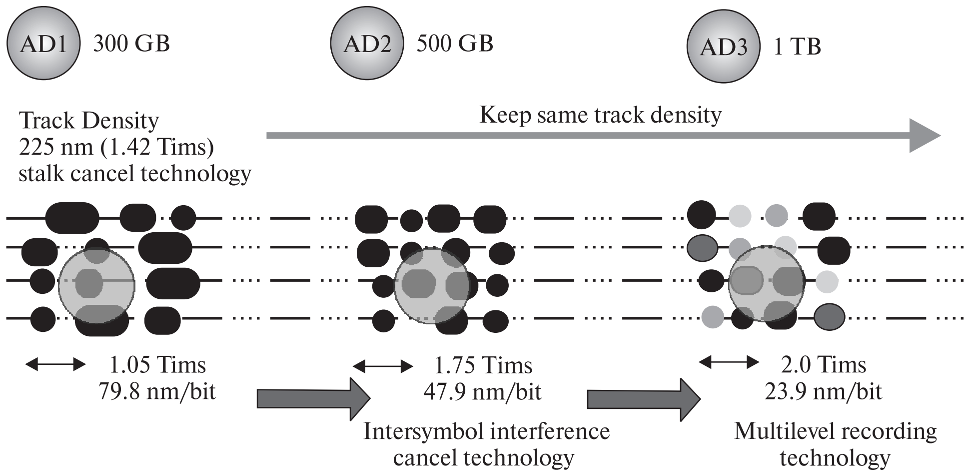


Fig. 4.6. The Archival Disc road map [88]

Fig. 4.7. Everspan system full configuration (*a*), everspan-loading robot (*b*) [93]

Passingly, in recent years, in order to further increase the recording capacity in recordable high-density optical information recording media such as DVDs and BDs, techniques for increasing the recording layers number have been broadly adopted. With a multi-layer optical information recording medium, recording and reproduction of signals with respect to the recording layer positioned at the deepest portion from the information reading face side is performed using laser light that has been transmitted through the recording layer instantly before the recording layer position at the deepest portion.

Therefore, with a recording layer other than the recording layer positioned at the deepest portion from the data reading face side, if the transmittances of the information signal layers change instantly before and after the recording of signals, since the effective laser strength of the recording layer positioned at the deepest portion changes, recording of signals at the deepest portion of the layer is not performed correctly. In particular, with an optical information recording medium with three or more layers, since the number of signal layers that are transmitted increases and the influence of changes in the transmittance which is multiplied with the number of layers with respect to the layer at the deepest portion becomes large, the more layers a medium has, the smaller the changes in the transmittance ought to be both before and after recording [87].

Nowadays it is planned to enhance the capacity (Fig. 4.6) of the Archival Disc to 1 TB [88].

A new generation of high-capacity optical discs, created jointly by Panasonic and Sony, serves as the storage media for Generation 2 of the Optical Disc Archive. Sony offers the new optical drive system in three components: a stand-alone USB-based unit, the cartridges that actually contain the optical media and an 8 GB fiber channel "library" unit. The overall system doubles the speeds of reading and writing compared to the previous generation, and is supposedly great for real-time 4K video recording. In addition to Optical Disc Archive Generation 2 development, Sony is also employing such technologies in new systems, which were designed specifically for data centers called Everspan, where explosive growth is expected in coming years (Fig. 4.7). On March 2016, Sony Optical Archive Inc. unveiled the Everspan Library System (Everspan). It was a scalable optical library system solution delivering archiving capabilities far exceeding the capacity of what is available in the marketplace now, at a fraction of the cost. Everspan utilizes 300 GB Archival Disc supporting data storage for over 100 years in data center environments. Everspan is includes three units: the Base Unit, the Robotic Unit, and up to 14 Expansion Units. Capacity can be increased by adding Expansion Units, which require a nominal increase in power expenditure, and cooling demands. Up to 64 Sony optical array drives with average transfer rate of 280 MB/s can be incorporated into the system. Everspan can store an enormous 181 PB (Petabytes) of archival data. Up to four systems can be incorporated in a single system, giving access to an astounding 724PB of total addressable storage.

Energy skill is an issue facing many enterprises now, and Everspan tackles this with one of the most efficient industry power/storage ratios. The Base Unit and the Robotic Unit draw power but all of the additional Expansion Units draw miniscule power to keep the sensors and air filters running. This design allows customers to scale up data storage, only slightly impacting the total system power. In data center environments, at approximately 9kW for a typical 181 PB system, (when the system is idle power usage drops to less than 2kW for a complete 181 PB library). The Everspan Library System is currently being evaluated by several institutions and companies [93].

4.3. Problems associated with optical long term storage systems

Major Optical Disc Failure Mechanisms: Data Layer Failure, Delamination, Mechanical distortion Disc, Warping Photochemical effects (environmental UV exposure primarily) and Tilt [93].

The available optical discs cannot realize the required level of reliability and data retention time due to the low stability of the polycarbonate substrate [32]. Metal reflective coating of the CD is described by a weak adhesion to the polycarbonate substrate, which leads to a rapid stratification of the carrier during use, particularly when sudden changes of temperature. Rewritable and recordable discs are unsuitable archival media they use basically unstable, organic dye-based data layers [126]. The three common organic dyes used in recordable discs are azo-dyes, cyanine, and phthalocyanine. In a recordable CD each dye gives the media its distinctive look depending on which metal is used for the reflective layer; azo (deep blue) has developed into different shades of blue, the original being a deep blue, and the more recent Super Azo a brighter shade of blue, cyanine (blue) dye becomes blue on silver media and green on gold media; phthalocyanine (clear light green) dye appears clear transparent on gold media, but light green on silver media [94, 176].

Rewritable DVDs and CDs operate on different principle. Rewritable discs are erasable, they can be rewritten, for all that a finite number of times. The recordable layer is made of antimony, tellurium and germanium. A laser heats the surface to two set temperatures. The higher temperature is the melting point (approximately 600 degrees centigrade), while the lower level temperature (approximately 350 degrees centigrade) becomes the crystallization temperature. Controlling the cooling rate and heating the disc produces a track of amorphous or crystalline areas. These areas will be interpreted by the reading laser like the pit/land structure of a CD-ROM due to their different reflectivities [95].

With the aim of identifying optical disks types and the conditions under which ensures long storage time of recorded data in the Library of Congress was the analysis of the optical discs characteristics after long-term use and storage [96]. The results show that the CD-R in which I use-ll organic dyes that have a shorter data retention period, compared with strain-povanim CD-ROM. Due to the higher density of the data records DVD media have a greater likelihood of loss of recorded information than CD format CD. A significant impact on the retention periods has conditions of use and storage, even the inscriptions on the labeled side of the CD-ROM lead to accelerated media degradation. It is interesting, that some discs that were made in one period and stored in the same conditions, was absolutely destroyed, while other carriers did not have any damage [97]. As a universal rule it has often been suggested that only reliable brand recordable CD and DVD are purchased, and testing has revealed a range of compliance with agreed standards even amongst them. Therefore, it is recommended that the responsible individual or institution insist on dealing with a supplier that is open about the importer or manufacturer they deal with, and who is capable to provide contact with the relevant technical personnel in the manufacturing company [164]. Yearly check of collections burned to disc since 1995 showed that between 2% and 5% of the discs failed annually [97].

In the process of WORM optical disc carrier's creation, the main attention was paid to the materials for recording media capable to achieve long-term storage of recorded data.

Tens of vitreous tellurium alloy compositions were used as recording media for WORM optical disc carriers. According to performed calculations four-component

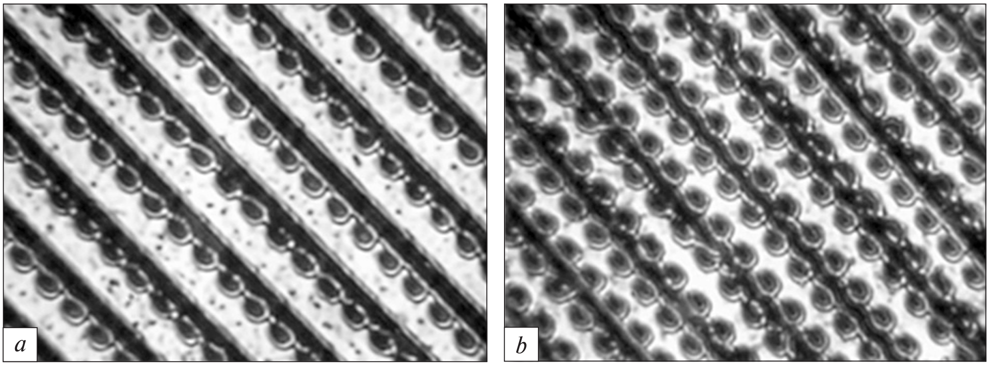


Fig. 4.8. External appearance of the recording zone on the optical carrier: data recording is carried out from one side of a guide track (*a*) and from both sides (*b*)

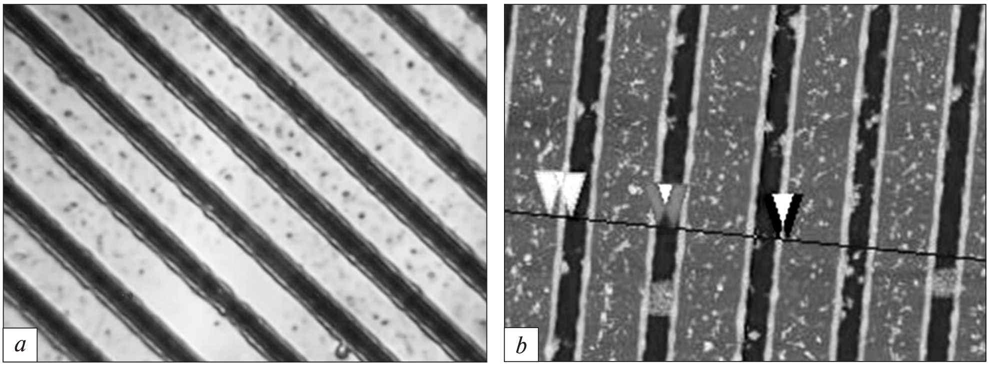


Fig. 4.9. Guide tracks on the optical disc carrier with $Te_{14}Sb_{10}Se_{61}Ge_{15}$ recording medium (track pitch is 1.6 μm). The images are obtained using: *a* — optical microscope; *b* —scanning tunnel microscope [183]

eutectic alloy $Te_{14}Sb_{10}Se_{61}Ge_{15}$ was chosen to make recording media providing long-term storage of recorded data (Fig. 4.8, 4.9). The mark of using this alloy was in necessity to apply radiation at wavelengths shorter than 550 nm for data recording.

Optical discs with recording media containing more than 60% of tellurium were made with the aim of presentation the comparative investigations of optical carriers. Preference of these carriers was that information recording on them could be carried out with infrared lasers. One of the main reasons to use as recording media tellurium alloys instead of pure tellurium films was the necessity of increasing the corrosion resistance of recording media. The dense oxide films that reduce the rate of tellurium oxidation arise on the surface of tellurium alloys with Ge, Sb, Se.

After long-term 30 years storage the optical disc carriers restrained mirror reflection, uniformity of coating within the area of data recording. The shape and sizes of pits were susceptible to slight changes: the sizes of pits are larger a little. The performed layer-by-layer Auger analysis of $Te_{14}Sb_{10}Se_{61}Ge_{15}$ recording medium manufactured 30 years ago presented that the oxygen content did not exceed 13 at. %. Oxygen was uniformly distributed along the thickness.

In the initial state (Fig. 4.9, *a*), at the oxygen background 475, 490, 511 eV (the latter is the main peak) lines of Sb and Te are not recorded. After etching the oxide film (Fig. 4.9, *b*), Sb and Te peaks are developed (457 and 498 eV). Then the oxygen content is reduced down to approximately 13 at. %. Then, superposition of the main tellurium Auger peak with the oxygen one takes place. There is observed the shift of the peak from the scale position 491 up to 498 eV.

We see, that the oxygen-passivated film of multicomponent chalcogenides preserved its chemical composition, while the oxide film was easily etched by argon ion beam. It is remarkable that oxygen distribution has a similar character both in recording media covered with a protective polymer layer and in unsealed recording media. Practicality of considerable oxygen amount in this recording medium can be related to engineering features of creating a thin recording layer. In the process of vacuum deposition a considerable amount of broken bonds arises in the structure of chalcogenide glass. As a result we can see the stage-by-stage accumulation of oxygen in this thin film. In the process of manufacturing an optical carrier, concentric guide tracks were recorded on the surface of recording medium of this carrier by focused laser radiation. Information recording was carried out on one or both guide tracks sides. It is rather technological method of forming guide tracks, and it does not require any special preparation of the information carrier substrate, the tracks can be recorded using laser-beam recorder [91].

Modern recordable optical discs do not meet true archival-quality needs, customary meaning a lifetime of ≥100 years. Failure mechanisms limited their lifetime include dye degradation, oxidation, delamination and corrosion. All these failure mechanisms we can find in the literature [91-96]. All four optical disc failure mechanisms must be addressed and resolved with purpose of producing a recordable optical disc with a minimum lifetime >100 years. All four of these degradation mechanisms are exacerbated by exposure to elevated humidity, temperature and light [95].

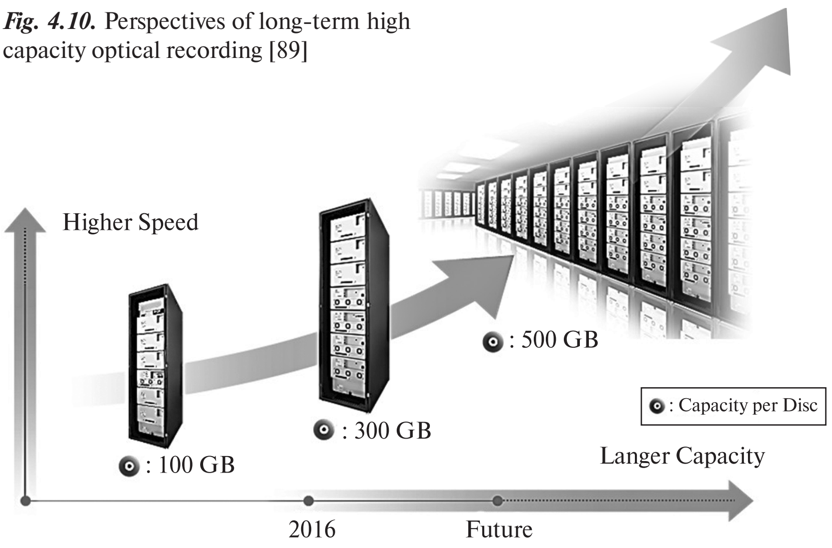
The use of recordable CD and DVD discs in libraries and archives has soon raised concerns as to their reliability as storage media. Most of these concerns have focused on the issue of life expectancy of the media themselves, concluding that different types of dyes, recording substrates, and reflective layers behave differently. Recordable CDs and DVDs are often chosen as archival carriers, but the risk of a storage system failure based on this type technology is high when compared to their approaches. The Memory of the World Sub-Committee on Technology joined forces with the IASA Technical Committee for the first time in June 2002 by inviting manufacturers of recordable CDs to the UNESCO headquarters in Paris to argue the problems actual for archival users because of their products. While several products have been marketed since which attempt to optimize their longevity and while the representatives of manufacturers expressed sympathy with the concerns of the archivists, the primary problem remained unsolved.

4.4. Perspectives of optical recording

These new extended storage capabilities, and the explosive growth in media stored by Internet giants, have demanded them to automate their capabilities in what is called "cold storage". For example, Facebook uses self-built massive robot systems with cheap Blu-ray disks. Also Amazon, with its Glacier storage solution, offers similar capabilities. Such kinds of automated setups could be applied for long-term archiving purposes [87]. The struggling electronics maker wants to persuade enterprise users to move from tape storage and hard drives to optical disc libraries by emphasizing that discs feature longer lifetimes, lower cost and higher data throughput rates [88]. Facebook had built a cold storage system from 10,000 Blu-ray discs. It holds a petabyte of digital information and cuts costs by 50% and energy use by 80% compared to a Facebook cold storage system that uses hard disk drives [91].

Panasonic Corporation recently announced the developing of freeze-ray, an Optical Disc-Based Data Archive System in collaboration with Facebook. Panasonic was able to design freeze-ray by collaborating with Facebook. They meet the growing demand for more efficient and sustainable ways to store and access cold data — infrequently or never accessed data stored for the long term — in the world's data centers [89]. An optical data archive system like "Freeze-ray™" is designed and validated to meet the needs of large scale enterprise markets and information centers. Optical media enables safe storage of critical & valuable data, with data integrity, for a long time line. It is a permanent and low cost solution to meet the cold data archive requires. The "freeze-ray" data archiving system was bornby integrating Panasonic's high-density optical disc and device technology, and library software development technology to simplify system control. The newly launched 2nd generation freeze-ray system features a storage capacity of large data near 300 GB per disc. This enables

Fig. 4.10. Perspectives of long-term high capacity optical recording [89]



secure, long-term storage of valuable data assets for 100 years or more. It also offers a controlled expandability to enable an optimal system configuration for individual requires. There is no any need for constant electric power or air conditioning costs, making it environmentally sensitive, low-cost, safe data storage solution. It is ideally suited to cold data archiving use [147]. Sony promotes the wide use of optical disc systems at large data centers, used by various IT, financial and energy (oil and gas) industries. Hereafter, Sony will continue incorporating new technical advances into archive systems. It ensures even higher reliability in the years ahead.

Optical Archive in California tries to develop new optical disc library systems for corporate clients' "cold storage", which hold data that aren't accessed often but are preserved for a long time. Between examples are regulatory or legal documents and photos on social media sites [89].

Thus modern optical carriers are at a disadvantage in relation to magnetic ones on recording density. But they have great potentialities for providing high-reliability long-term data storage (Fig. 4.10). High reliability of information storage on modern optical carriers is connected with processes of data reproduction and recording are noncontact. The optical method of reading allows executing substantially an endless number of information reproduction cycles without its damage or even distortion. The use of sustainable caddy-type cassettes prevents the touch of user hands during the loading the player and contact the surface and the layer through which a laser radiation is being focused in the process of data recording and reproduction. Such data presentation method allows reproducing them by different methods, namely, by electron-beam microscopy and tunnel.

4.5. Conclusions

- The successful use of optical disks for long term data storage is based on the capabilities and characteristics of optical recording. Optical media for long term data retention have significant advantages compared to other technologies of data storage.
- Scratching and environmental factors such as dust, heat or UV light can cause severe damage to the disks. Optical discs also offer low data capacity compared to alternative technologies. The life span depends on factors such as manufactured quality, how well it is recorded and its physical handling and storage. To lower the risk of failure, best practice is to frequently migrate onto newer formats.
- Metal reflective coating of the optical discs is characterized by a weak adhesion to the polycarbonate substrate, which leads to a rapid stratification of the carrier during use, especially when sudden changes of temperature.
- The optical method of reading allows to execute endless number of cycles of information reproduction without its distortion and damage. The use of special cassettes prevents the touch of hands of user in the process of loading the player and contact the information surface and the surface through which a laser radiation is being focused in the process of information reproduction and recording.

CHAPTER 5

TECHNOLOGIES OF OPTICAL LONG-TERM DATA STORAGE DEVELOPMENT

5.1. Long-term reliability of modern optical disks

The development of optical discs with ultra-long shelf life has been a subject of intensive research. Many engineering solutions have been proposed and developed for increasing the data storage lifetime on optical mediums with plastic substrates. Technologies of information storage shelf life increasing on optical mediums of different types were developed. Development of special recording layers is the main direction of data storage reliability improving on write-once optical mediums. Voids in polymers and glass materials which are permanent laser-induced physical changes can provide an approach to long-lifetime storage without data degradation.

UDO mediums (Fig. 5.1) uses a patented and field-proven Phase Change recording technology that produces a very stable recording surface with a mediums lifetime in excess of 50 years [89].

UDO mediums are available in Write Once (True Write Once and Compliant Write Once) and Rewritable formats (Fig. 5.2). Rewritable UDO medium is typically used in archive applications where the stability and longevity of optical media are important, but records change frequently and there is a need to delete and rewrite information. UDO Rewritable data medium is ideal solution for small office and departmental applications or in larger unstructured archives that are not subject to specific regulatory or corporate standards.

True Write Once media offers the highest level of physical record authenticity and is ideal for document classes with very long or indefinite record retention periods. Common uses of True Write Once medium include medical, financial, industrial and cultural applications that require documents be retained for years or decades and must be held to a very high standard of legal admissibility.

Compliant Write Once medium combines longevity and secure authenticity of True Write Once UDO medium with the ability to shred expired records for document classes that demand physical dis-

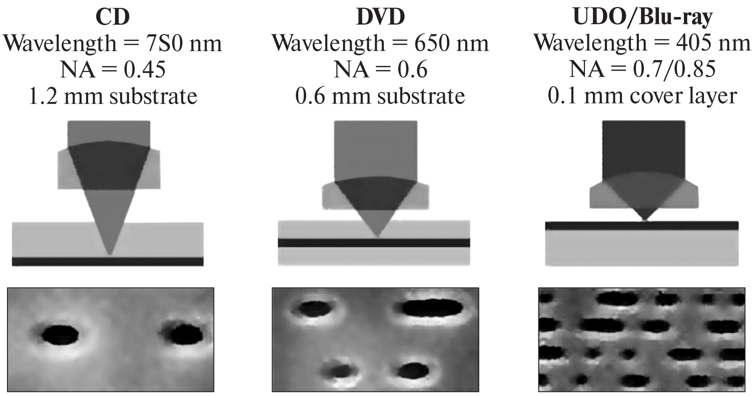


Fig. 5.1. Comparing CD, DVD, and UDO/Blu-ray [89]

Fig. 5.2. Plasmon Ultra Density optical double sided WORM disk and Plasmon's UDO drive. WORM disk can store up to 60 GB of data (UDO2) [98]

Table 5.1. **Medium Specifications**

Parameter	Value
Disk Diameter	130 mm
Disk Thickness	2.4 mm
Cartridge Size	5.25 inch - ISO Standard 135 × 153 × 11mm
Capacity	30 GB
Sector size	8 KB
Number of user sectors/side	1,838,652
Data area	27.0-62.5 mm
Recording layer	Phase change
Recording format	Land & Groove
Recording side	Both sides
Recording density	7.4 Gb/in
Medium layers	1
Data encoding	RLL (1,7)
Rewrite cycles (Rewritable Medium)	10,000
Medium life	50 + years
Archival Temperature	5-55 °C

posal. Designed specifically to assist organizations in meeting regulations on data retention and disposition while managing corporate risk, Compliant Write Once medium can be used in a wide range of financial, administration and legal applications.

Magnified UDO disk from Plasmon shows non-reflective amorphous (grey) and reflective crystalline (light/dark) bits (Fig. 5.3). The crystalline area is not a single crystal, but a multi-crystalline one. Due to different crystal alignments and interaction of the beam used to image the material, the reflective bits appear variously light and dark (Table 5.1).

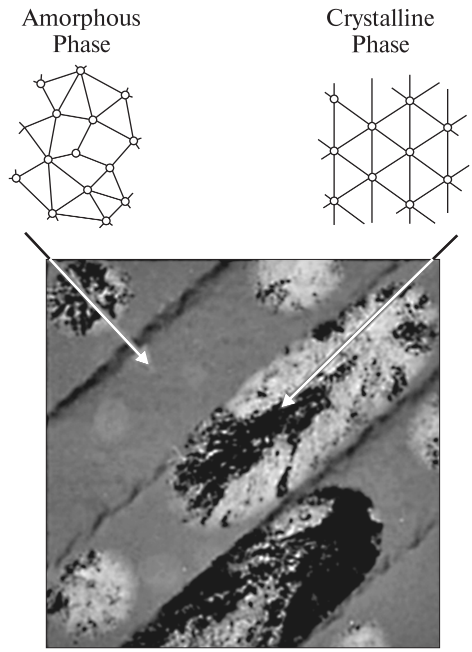


Fig. 5.3. Bits of Phase Change UDO disk [98]

The different sizes of bits are due to UDO's encoding method. Guaranteed lifetime of such mediums is 50 years [98]. In order to protect UDO medium from physical damage and contamination (dust, fingerprints, etc.), the disk is enclosed in a rugged ISO standard 5.25 inch cartridge. A unique double-shutter design prevents dust accumulation on the top surface of the media during recording (Table 5.2). Lower dust contamination means more reliable read and write operations and extended life.

Table 5.2. **Comparison of UDO Medium Generations**

UDO MEDIUM	Generation 1	Generation 2	Generation 3
Capacity	30 GB	60 GB	120 GB
Transfer Rate	Up to 8 MB/s	Up to 12 MB/s	Up to 18 MB/s
RPM	2000 RPM	3000 RPM	3600 RPM
Avg Seek Time	25 msec	25 msec	25 msec
Numerical Aperture	0.7	0.7	0.85
Medium Layers	1	2	2
Encoding	1.7	1.7	ML
Sector Size	8 KB	8 KB	8 KB
SCSI Transfer Rate	80 MB/s	80 MB/s	80 MB/s
Load Time	5 sec	5 sec	5 sec
Unload Time	3 sec	3 sec	3 sec
MSBF	750,000	750,000	750,000

Plasmon advances UDO with the release of UDO2 60GB drives and mediums. The Archive Appliance is further developed to include UDO2 drives and mediums having the backwards compatibility of reading UDO1 medium in UDO2 drives (Fig. 5.3). The G Series libraries have even greater backwards compatibility with mixed drive and medium capability [98].

An alternative method for long-term data storage is to combine reflective and recording layers in a single layer, which should have a good adhesion to the substrate material. This may improve the overall reliability of the optical recording medium and reduce the probability of separation during its usage. The information on disk is recorded directly on the reflective metal film. The information on the positive photoresist layer is recorded according to already existing CD standard using laser-beam recorder. Then, the mask is created for following chemical etching of chromium film. Analysis of morphology of recorded pits has proved suitability of the long-term data storage disc to existing standards [92-98, 185]. Chromium optical disc created in this way is successfully read in a standard DVD player. Thus, the Cr film can be used to create medium for long-term storage. The use of the protective layer in this case is only necessary for protection against impurities [95-98].

Longer lifetime of data storage is projected by using optical medium with cermet recording layer (branded name is M-Discs). Structure comparison of standard DVD and M-Disc is shown in Fig. 5.4.

One of the reasons of longer lifetime is using the thermosensitive layer with high reflectance that allows one does not use light refractive layers. Thus, the expected longevity of the M-Disc is over a thousand years of storage at 25 °C and a relative humidity environment of 50%, but at a temperature of 40 °C and 70% relative humidity expected longevity is 53 years [96, 98]. National Institute of Standards and Technology (NIST) of USA defined that M-disc shelf life is equal to 1000 years (letter “M” in the title means “Millennium”. However, it should be noted that polycarbonate withstands severe temperature changes only for a short time.

Project of the company Millenniata is devoid of the biggest disadvantage of a standard DVD-R, because its safety depends on metallic reflecting layer and the condition of the organic layer to which the information is recorded. Danger to standard medium is the gradual oxidation of the reflective layer that usually consists of

Table 5.3. **Readability of M-Disc DVDs in common usual drivers**

Manufacturer and model	Read M-Disc DVD	Read M-Disc BD-R	Manufacturer and model	Read M-Disc DVD	Read M-Disc BD-R
LG WH16N540	Yes	Yes	Matshita UJ8B0AW	Yes	NA
Asus DRW-24B1ST	No	NA	TSST SN-208FB	Yes	NA
Teac DV-W516C	No	NA	Teac DV-W28S-V	Yes	NA
Matshita BD-MLT UJ272	Yes	Yes	Plextor PX-B320SA	Yes	Yes

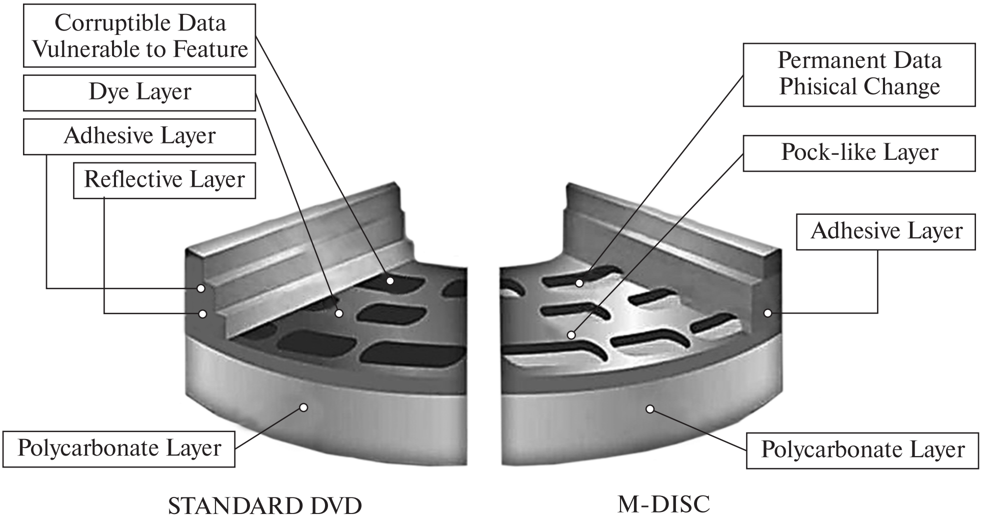


Fig. 5.4. Structure comparison of standard DVD and M-Disc [98]

aluminum or silver. Some manufacturers (for example, Verbatim) use as "blanks" a shelf gold or silver-gold alloy. However, these engineering solutions are only a very small part of the problem. When it comes to the impact of the environmental conditions the dye is the most vulnerable place: bright sunlight, high humidity and temperature promote dissolution of the dye layer and disappear of saved files. New M-Disc offer a capability of full data read out in existing DVD players. However, a special drive is required to write the special medium. It should be noted that M-Disc has exactly the same care and handling limitations as a conventional CD/DVD recordable medium, but does create a far more firmed physical data record with at least the potential for long term storage.

Verbatim M-Disc is a new standard in digital data storage. This standard has been designed to preserve and protect files by engraving your information into a patented "rock-like" layer, which is resistant to light, temperature, and humidity. Industry standard ISO/IEC 10995 tests carried out by Millenniata showed the expected mean lifetime of an M-Disc to be 1,332 years, with just 5% of discs showing signs of data loss after 667 years. Therefore, the projected lifetime is expected to be several hundreds of years. BDXL 100GB capacity disc makes archiving essential data easier for home users as well as business & enterprises users.

M-Disc BDXL discs incorporate titanium for added longevity. Combined with the M-Disc patented "rock-like" recording layer this provides ultimate protection for you precious memories & data. M-Disc recordable DVDs should be readable in 90 percent of the DVD drives installed (see the table below).

M-Disc released 4.7 GB DVD discs, which are suitable for archiving documents and perhaps most treasured photos. Users of long-term data storage systems have mentioned that devices like the M-disc have impressive durability [169].

5.2. Creation of optical mediums for long-term data storage based on highly stable materials

Mediums for long-term data storage based on highly stable materials must have features such as immunity to technology obsolescence, high temperature tolerance, immunity to water damage, unaffected by electromagnetic radiation, highly durable over long periods of time [92-98].

The variety of technical solutions is proposed for technology of long term data storage which is based on using highly stable materials for both substrate and recording medium. First designed and manufactured optical discs had a substrate of silica glass. Their data retention time could not exceed of 20-30 years because of degradation with time of recording layer based on chalcogenide glasses [96]. Materials for the manufacture of optical discs having long term data storage time must be chemically, thermally and mechanically stable. The main element determining parameters of the optical disc is a transparent layer (substrate) through which the information is read and a recording layer on which information is recorded. Disc substrate is an external unprotected layer and it defines the types of materials and technologies that can be applied in the deposition of the recording layer, so the choice of the substrate medium is the key element in development of technology of long-term data storage. Table 5.4 shows characteristics of the materials that can be used for manufacture of optical disc substrate.

Table 5.4. **Physical and chemical properties of materials suitable for the manufacture of optical discs for long-term data storage**

Parameters	Material					
	Sapphire (Al_2O_3)	Quartz (SiO_2)	Fused quartz (SiO_2)	Yttrium aluminum garnet ($Y_3A_{15}O_{12}$)	Magnesium aluminate (MgA_2O_2)	Diamond (C)
State of matter	Crystalline	Crystalline	Amorphous	Crystalline	Crystalline	Crystalline
Optical type of crystal	Negative	Positive	—	—	—	—
Mohs hardness	9	7	5.3-6.5	8.5	8	10
Fusing temperature (K)	2300	1960	1350 (softening)	2210	2400	1100 (bum down)
Thermal conductivity coefficient k (W/(mK))	~34	3	1.3	14	15	3200
Coefficient of linear thermal expansion $\alpha \times 10^{-6}$ (1/K)	5.6	0.55	0.55	8.0	7.5	1.0
Chemoresistance	1*	2*	3*	1*	3*	1*
Resistance to UV radiation	Not degraded	Not degraded	Degraded	Not degraded	Not degraded	Not degraded

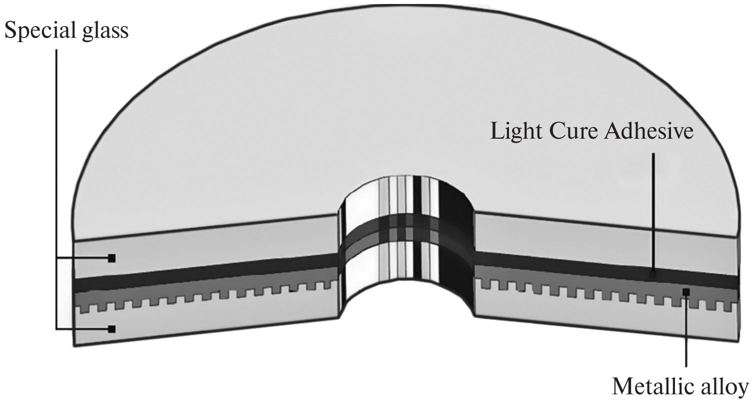


Fig. 5.5. Glass disc for long-term data storage [98]

In Table 5.4 the following designations are used: 1* — soluble in concentrated solutions of fluoride salts (such as, BaF_2, MgF_2, PbF_2, and oxides) at very high temperatures (>1000 °C); 2* — soluble in alkaline aqueous solutions at high temperatures (>300 °C) and solution of hydrofluoric acid; 3* — soluble in aqueous alkaline solutions and solution of hydrofluoric acid. From the Table 5.4 one can see that for long term data storage it is appropriate to apply a substrate from highly stable single crystal materials.

Much attention is paid to creating technology of ROM type optical mediums which will be able to storage data during a few hundreds of years under severe climate conditions. We proposed a technology of long-term optical medium creation based on the microrelief information writing in the silica glass substrates [99, 91]. Optical discs prototypes created by this technology have been demonstrated on the International exhibition of computer technologies CeBIT 2008 [92]. This technology was implemented on an industrial level by German company Syylex [98]. Structure of the glass disc is shown in Fig. 5.5.

The key point of this technology is a choice of special glass as disc substrates. Glass from which discs are made has superior resistance to chemicals. Thus, in order to dissolve it in liquid acids or alkalis an additional high temperature impact is needed (about 1000 °C). This glass is plastic only at temperatures above 850 °C. Discs from special glass "GlassMasterDisc" can store recorded information during 1000 years. Term "GlassMasterDisc" has been defined by French Laboratoire national de metrologie et d'essais.

During testing different types of discs were subject of prolonged exposure of temperature of 90 °C at humidity of 85%. Under these conditions information of DVD-ROM became inaccessible after 375 hours, and WORM type medium failed much earlier. Only glass disc has stood the test drive that lasted 1,500 hours. This glass disc is able to retain recorded information for 1000 years being on Earth as well as in space. According to this technology, information media are made to order on special technological equipment. This equipment is similar to ones which are used to

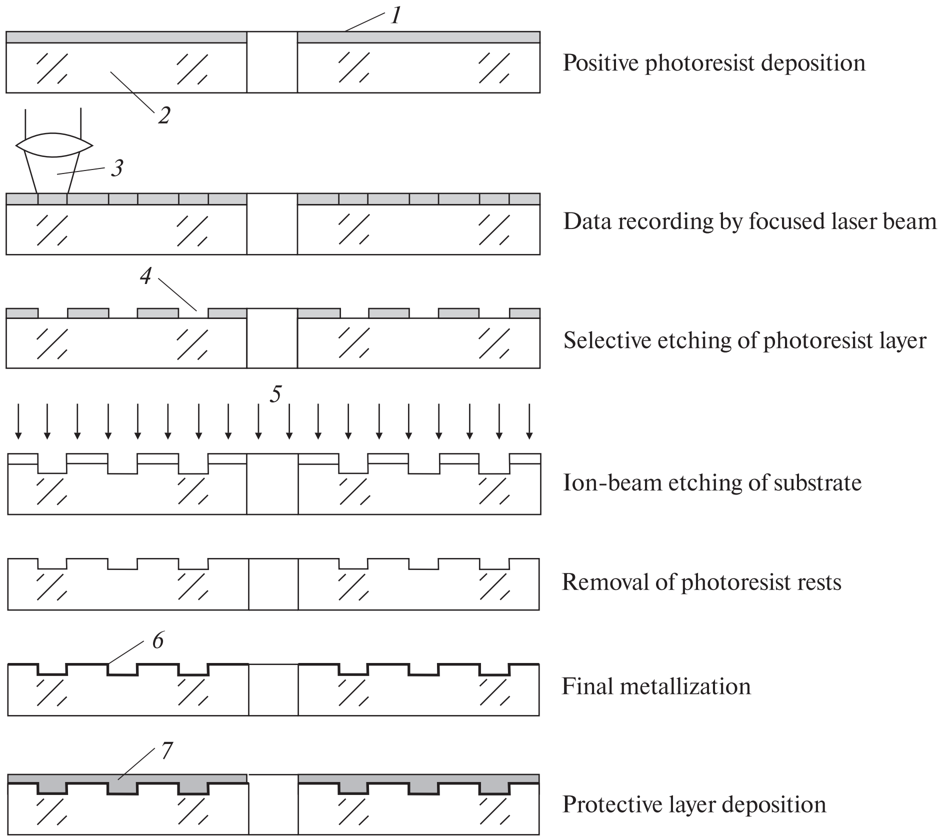


Fig. 5.6. Scheme of data recording process on the optical disc for long-term data storage: *1* — sapphire substrate; *2* — Cr film; *3* — positive photoresist layer; *4* — focused laser beam; *5* — etching mask; *6* — ion beam; *7* — reflective metallic layer [92]

produce master-disc during compact-discs manufacturing [99]. Recording layer is consists of pits, which have been etched in glass substrate (Fig. 5.6). Light-reflective layer provides the ability to read out information in any DVD-BD driver. Another glass substrate is attached to the first one via UV cure adhesive. It provides additional mechanical strength of medium [90, 98].

In 2007 Ajay Pasupuleti came up with the idea of using the same technology used in semiconductor fabrication to create the tiny images while he was pursuing a doctorate in microsystems. Silicon wafers can keep data safe for centuries. NanoArk's process starts with documents or images being digitized and then etched onto the wafers. Meanwhile, data from the documents such as certain key words are put into a searchable directory. Potential customers include banks, real estate companies, law firms and educational institutions [105].

The images are stored by utilizing semiconductor fabrication techniques. These images are organized and managed using metadata in the form of a barcode. Each

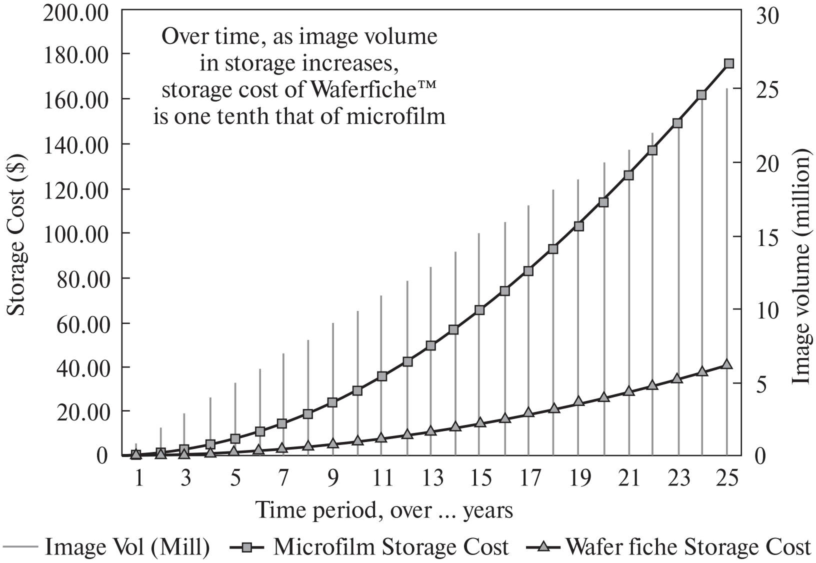


Fig. 5.7. Comparison of storage costs – microfilm and waferfitche [111]

barcode is a unique identifier that contains the location information for each specific image on the silicon wafer substrate. The system further provides an identifier in an electronic database that references the appropriate barcode and describes the contents of the image. The images and barcodes are transferred to specific predetermined locations on the silicon wafer. The stored images are retrieved by use of a software program that searches for a user's queries in the electronic database and outputs the specific barcode to the image reader. The image reader translates the barcode information for the desired image and drives the optics or the silicon wafer to the appropriate location [113].

In the future technology Waferfiche for long-term storage of data was created, in which recording is performed on the surface of a silicon wafer in the form of visual codes. The technology Waferfiche was originally developed for NASA and is now available and affordable for archiving your documents. With the remarkable ability of the medium to withstand environmental elements, its small size, and the ability to see the documents with a simple magnifying glass, this proves itself to be the best way to ensure your documents are protected. Waferfiche is resistant to high temperatures and large lengths of time in contact with water. Cost is about 10X less to store on the Waferfiche medium than microfilm (see a comparison in Fig. 5.7).

Silicon wafer can hold up to 2,000 documents depending on the document size and resolution. Waferfiche is compact and easier to store than microfilm. Documents can be read with light and magnification when technology is not available. In the Waferfiche™ technology, data is stored in such a way that it is visible to the human eye with or without magnification. The data from print, digital or any other media is converted to images as a first step. These images, with the help of photolithography tools and fabrication techniques are then imprinted and etched on silicon wafers. The use of silicon makes the information temporarily resistant to high temperatures (up to 400 °C) and water, ensuring longevity, which is very useful in preserving documents. In this technique, since the stored data is not processed or digitized before storing, the data is stored for long periods without any loss of data over time.

The added advantage of the new technique is retrieval of the data can be as simple and straightforward as magnifying the image on the silicon wafers thereby eliminating the need for a computer. This feature enables archival of data in a technology free environment. Also depending upon the semiconductor fabrication technique used (smallest feature size in the order of 200 nm, 100nm, 90 nm or below), nano-scale images can be imprinted thereby making it possible to store large amount of data on a single silicon wafer. NanoArk's Waferfiche technology employs a photolithographic process to inscribe minute copies of documents onto thin silicon discs. The Waferfiche's inherent material properties render it resistant to fire over a limited duration and, more importantly, practically impervious to water damage. NanoArk notes that the biggest threat to any kind of archived information is moisture [110, 111].

Modification of the medium for long-term storage has been proposed, which consists in the fact that the data is stored in a metal layer on a substrate of a highly stable material. The storage and archival of digital and analog data representing images and documents is accomplished by the imaging the top metal layer on a wafer or substrate that represents the data that can be extracted by optical methods. The storage is accomplished by memory device that are embedded either in the same wafer or substrate that contains the archival data or in the same package that contains the said wafer or substrate. The storage data can be extracted by means of a connector as digital data that can be interfaced to a computer. The stored data accurately represents the archived data so that the stored and archived data are reasonably similar in nature. The system in a package may be either a multi-chip module in a single package, or a wafer or substrate and a printed circuit board assembly that are inside a single package [112].

A team of Dutch and German physicists have developed a technology that can preserve data for up to 1 billion years. The team chose elemental tungsten because it has a very high melting point of 3,422 degrees Celsius and low thermal expansion. Basically, if you build something out of tungsten, it will remain mostly unchanged over time. Tungsten is somewhat malleable, though, so the researchers encapsulated the metal in silicon nitride. This inert solid is durable and is transparent to light, which allows the tungsten pattern to be visualized [118].

The suggested data storage system consists of a medium where the data is represented by one material embedded within a second, different material as schematically described in Fig. 5.8. The base materials were selected - tungsten for the data and Si_3N_4 for the encapsulating material. The Si_3N_4 has a high fracture toughness and low thermal expansion coefficient. Another important feature of the Si_3N_4 is its transparency to light. A very thin film would also be transparent to electron beams. These materials are readily available and generally used in microfabrication.

An alternative to the transparent medium is a medium with high contrast in reflection. Contrast can be enhanced by using an optical beam of a single wavelength. If the layer thickness is tuned correctly, constructive interference can occur in the parts where there is no metal and destructive interference in the parts where metal is present or vice versa, as shown schematically in Fig. 5.9.

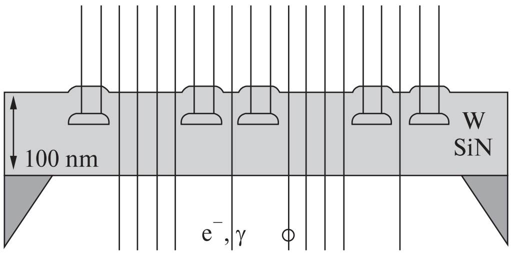


Fig. 5.8. W-Si_3N_4 WORM medium, which is transparent to electron or photon beams [106]

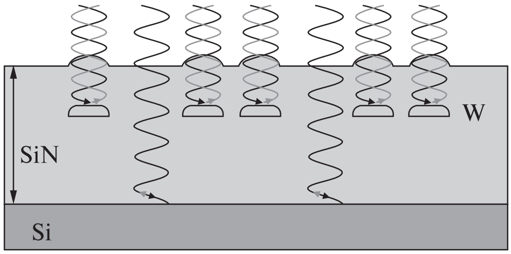


Fig. 5.9. W-Si_3N_4 WORM medium, which is readable by constructive interference on the silicon base and destructive interference on the tungsten islands [107]

This makes it possible to have a much thicker base because the sample does not need to be optically transparent. It should however still be ensured that not all the light gets absorbed by the nitride layer. With light in the visible spectrum, this method can be used for low density data.

The medium consists of a 338 nm layer of LPCVD Si_3N_4 on a bare silicon wafer. The tungsten is patterned using optical lithography and a mask containing the QR codes. The pattern is etched using Ar ion beam etching and a top layer of PECVD nitride of 225 nm is deposited on top of the tungsten patterns. The process steps are schematically shown in Fig. 5.10. Because readout by optical microscope means that the data density is low, it is also necessary to have a higher density storage method. The higher data density storage can be achieved by embedding the data in the nitride for readout by electron beam. For the high data density sample, tungsten lines are used instead of islands. With this sample it is possible to simulate high density data with a linewidth below 100 nm. The lines will make it easier to create very small structures and inspect the sample after thermal exposure by means of an SEM. A drawback of these lines is the variation in stress in the sample in the direction along the lines compared to across the lines.

A scanning electron micrograph of the developed sample is shown in Fig. 5.11. A short O_2 reactive ion beam etching (RIBE) step is used to transfer the pattern into the BARC layer. The BARC pattern is transferred into the tungsten layer by argon ion beam milling as can be seen in Fig. 5.11 (*b*). The entire sample is subsequently covered with Si_3N_4 by a PECVD process to encapsulate the tungsten lines. The result is shown in Fig. 5.11 (*c*).

In the cross-section image from bottom to top, the silicon, the LPCVD silicon nitride, the tungsten lines and the PECVD silicon nitride can be seen. The Si_3N_4 in the final product is much thicker than the thickness schematically shown in Fig. 5.12 to

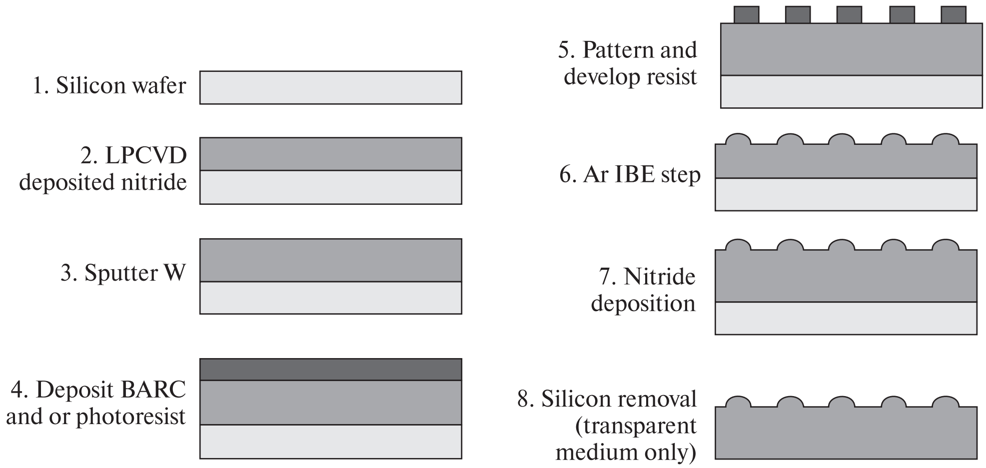


Fig. 5.10. Design of the HDP disk. Each pixel in the large QR codes consists of a micro-QR code [105]

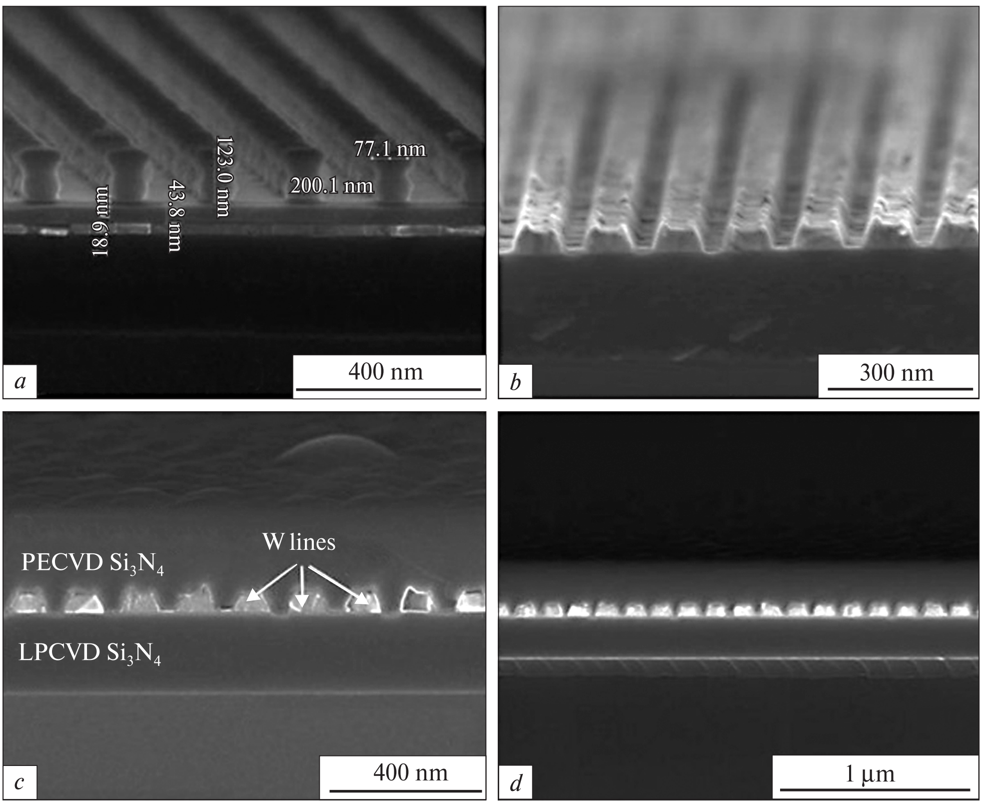


Fig. 5.11. Scanning electron micrograph of the test sample before etching (*a*). Scanning electron micrograph of the test sample after etching containing W lines (*b*). Scanning electron micrograph of the cross-section of the encapsulated lines in the test sample (*c*), scanning electron micrograph of the sample after 1 h. at 473 K [99] (*d*)

observe possible spreading of the tungsten clearly. For the optical transparent sample, the silicon needs to be removed from the bottom of the sample and a medium with tungsten lines encapsulated in a Si_3N_4 matrix remains. The silicon removal of step 8 shown in Fig. 5.13 is not performed to ensure mechanical stability of the sample. The data on the medium for long term data storage is written in two-dimensional (matrix) barcodes, which can be read back by a camera and computer. These two-dimensional barcodes were introduced for cases where more information needs to be stored than can be accommodated by their one-dimensional predecessors, but are now becoming increasingly popular.

Fig. 5.12. Design of the HDP disk. Each pixel in the large QR codes consists of a micro-QR code [114]

The implementation was the quick response (QR) code [114], which can be easily decoded by today's smartphones. The level of QR code containing the largest amount of information can lose up to 7 % of the data before the code becomes unreadable. For the encoding of the final disk, it is likely that a coding scheme would be required which focuses on easy decodability. By keeping the size of the QR code low, it is possible to read out the disk by an optical microscope. For the demonstration, the entire disk was covered with a centimeter sized QR code. Each pixel of the code consists of a set of much smaller QR codes with pixels of only a few micrometers in size as shown in Fig. 5.12. The initial attempt to create a medium containing embedded data which is able to survive for 1 million years is promising. The optical readable data in the form of QR codes was able to survive the temperature up to 713 K (see Fig. 5.13).

Exposure of the medium to higher temperatures shows degradation due to the difference in thermal expansion coefficients between the two different types of Si_3N_4 and the W lines. The top layer of Si_3N_4 starts to shows cracks and the W is exposed to the environment. This leads to the "whiskers" being grown under the influence of oxygen and high temperature. The amount of readable QR codes decreases at higher temperatures, but this seems to be largely due to the detection scheme as the tungsten is still present. The medium can survive high temperatures, up to 1 hour at 848 K, without visible degradation of the medium but at higher temperatures the medium degrades rapidly. A model for a high density recording medium, consisting of W lines with a width below 100 nm embedded in a Si_3N_4 matrix, has been successfully fabricated. An accelerated ageing test was performed by storing the sample at 473 K for one hour. There was no visible degradation of the sample or the tungsten lines. If we only take diffusion into account this proves that the sample will survive for well over 1 million years when stored at 300 K. The tungsten discs could end up being less stable in real life than the initial testing indicates due to the elements or exposure to chemical agents [20].

In order to provide a durable storage of data on optical media it is appropriate to use highly stable single crystal materials as a substrate. One of these materials is single crystal quartz [28, 62-65]. Numerous experiments have showed that oxide glasses are materials which are able to provide a high-speed recording of giant amount of data. Retentivity of data may be millions of years at room temperature. In addition, discs of quartz glass can withstand a thermal shock with temperature gradient of 1,000 °C [119-126].

The main benefit of silica discs is that they can easily withstand external influence and remain unscathed. For example, such medium with information can easily survive the fire or even the impact of electromagnetic radiation. Scientists expect that such media will be stored for more than a hundred years.

Several types of quartz data media have been reported. In the first developments of Hitachi, the emphasis was on achieving a long shelf life of information, rather than on creating high-capacity media. In Fig. 5.14 a general view of quartz data media is shown. The use of fused silica is expected to provide new technology for semi-perpetual data storage of valuable data such as historically important legacies and public documents as well personal data that individuals may wish to preserve data for future generations [132].

In Fig. 5.14 company researcher displays the storage unit, consisting of a sliver of glass 2cm square and 2mm thick, which can hold 40MB of data per square inch, about the same as a standard CD. The data is written in binary format by lasering dots on the glass in four layers, but the researchers say adding more layers to increase storage density isn't a problem. The solution of printing binary code onto quartz glass is not too dissimilar to records. There may not be commercialized technology available to get the data, but the simplicity of the format makes creating a retrieval system very straightforward. Just as adding a needle to a spinning record will produce sound, placing the binary code into any computer with a simple program allows users access

Fig. 5.13. Optical microscope images of the same QR code; left — after fabrication, center — after 2 hours at 613 K and right — 2 hours at 763 K [115]

Fig. 5.14. Hitachi quartz data media [132]

to the data — all you need is an optical microscope to get to the information. It sounds simple enough, and that is exactly the point — the data can be easily accessed no matter what the future technologies of the digital age bring. Even better, the data is safe from fire, chemicals, and water — almost anything, except perhaps a hammer [119-130]. The new technology will be suitable for storing historically important items such as cultural artifacts and public documents, as well as data that individuals want to leave for posterity.

Improvement of quartz media parameters is conducted in the direction of creating multilayer ones. It is proposed to realize a multilayer (WORM) data recording on some types of fumed silica mediums using a femtosecond laser [150-154]. The idea of the optical memory based on femtosecond laser writing in the bulk of transparent material [153, 154]. More recently self-assembled nanogratings produced by ultrafast laser writing in glass were proposed for the polarization multiplexed optical memory, where the information encoding would be realized by means of two birefringence parameters, i.e. the slow axis orientation (4th dimension) and strength of retardance (5th dimension), in addition to three spatial coordinates [154]. A distinctive feature of such media is that the information is recorded by the laser inside the media rather than on its surface. It guarantees protection and long shelf life of recorded information.

It has successfully achieved read/write of digital data in 100 layers of fused silica glass, a recording density comparable to Blu-ray Disc. One hundred multi-layer data recording was verified by the application of newly developed noise reduction technology to overcome interference from data recorded on other layers while trying to access data written in deeper layers within the fused silica glass [127]. Data was recorded with a femtosecond pulse laser. Read/write in 100 layers was verified using the following technologies.

(1) High-quality data recording and reading via application of a spherical aberration correction lens. When light is focused on a recording layer deep within fused silica glass, a phenomenon known as spherical aberrations which degrades the quality of the spot occurs. To then form a dot at the degraded spot, laser intensity needs to be increased but in fused silica glass, the focal points of powerful lasers are distorted depth-wise causing dots to be formed in other layers close to the target layer, resulting in noise in reading the data. Employing a spherical aberration correction lens, spot quality degradation was controlled without the need to increase laser intensity, and the formation of dots in the deeper recording layers was verified. Further, by also employing an aberration correction lens to the optical microscope used to read the data, it was confirmed that high-quality images could be attained.

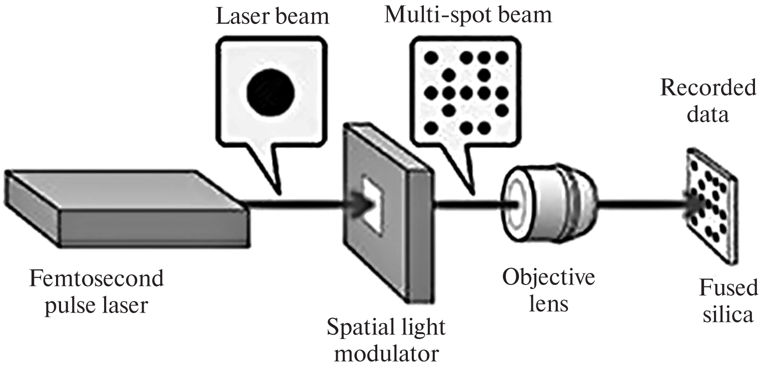


Fig. 5.15. The concept of storing data optically in the bulk of non-photosensitive transparent materials [162]

(2) Reduced read error by applying a noise-cancelling read algorithm. Images captured by the optical microscope are converted to digital signals in reading data: spots where dots are formed are converted to “1”, and those without dots are converted to “0”. As the layers become deeper within the fused silica glass, it was found that in capturing the images, noise due to crosstalk from other dots in other recording layers increased. To overcome this noise, an image-processing algorithm was applied which detects dot area (size) to distinguish dots from noise. By determining and eliminating image signals below a certain size as noise, it was confirmed that a read signal error rate below 10^{-3}, a standard for practical applications, could be achieved [112].

The concept of storing data optically in the bulk of non-photosensitive transparent materials (such as fused quartz, which is renowned for its high chemical stability and resistance) via femtosecond-laser allows for high-capacity optical recording by multiplexing new degrees of freedom (e.g., intensity, polarization, and wavelength).

Method of data storage that makes use of three spatial and two optical dimensions has been developed (Fig. 5.15). On the macroscopic scale, the self-assembled nanostructures behave as uniaxial optical crystals with negative birefringence. The alignment of the nanogratings gives rise to optical anisotropy (a form of birefringence) of the same order of magnitude as positive birefringence in crystalline quartz [162].

In conventional optical storage such as DVDs, data is stored by burning tiny pits in one or more layers of the plastic disc, thereby making use of three spatial dimensions. In fused quartz media two additional (optical) dimensions have also exploited. When the data-recording femtosecond laser marks the glass, it makes a pit with a nanograting. This nanograting produces birefringence that is characterized by two additional parameters. The slow-axis orientation introduces a fourth dimension, and the strength of retardance — defined as a product of the birefringence and the length of the structure — forms a fifth dimension. These two parameters are controlled during recording by the polarization and light intensity, respectively. By adding these additional optical dimensions to the three spatial coordinates, we achieve 5D optical data storage (see Fig. 5.16) [163].

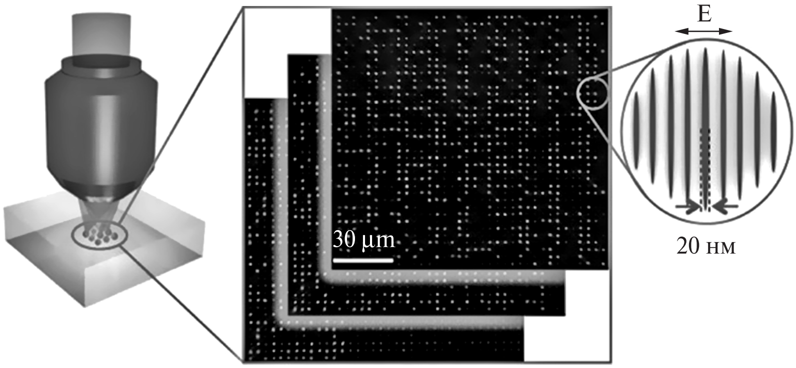


Fig. 5.16. 5D optical data storage, written in fused quartz using a femtosecond laser. Three spatial dimensions and two optical ones (the slow-axis orientation and the retardance) are exploited. Each voxel contains a self-assembled nanograting which is oriented in a direction perpendicular to the light polarization [163]

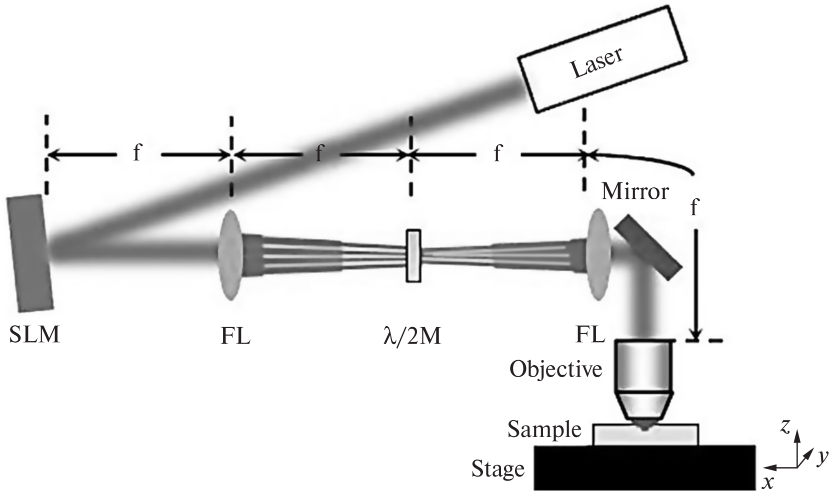


Fig. 5.17. 5D optical storage writing setup: femtosecond laser, spatial light modulator (SLM), Fourier lens (FL), half-wave plate matrix (L/2 M), dichroic mirror, 1.2 NA water immersion objective, silica glass sample, translation stage [163]

The experiments were performed with a femtosecond laser system Pharos (Light Conversion Ltd.) operating at 1030 nm and delivering 8 pJ pulses of 280 fs at 200 kHz repetition rate. The intensity distribution at the focal plane was modulated via a spatial light modulator (SLM), which split the incident light into 256 beams. The hologram generated on the SLM was reimaged via a 4-f optical system on the back pupil of the objective (Fig. 5.17). In addition, a half-wave plate matrix, imprinted by the laser nanostructuring of fused silica, was added to the 4-f optical system for the motion free polarization control. The laser beam was focused with a 1.2 NA water immersion microscope objective at the depth of 140 pm below the surface

Fig. 5.18. Appearance of fused silica media for long-term storage [172]

of the silica glass sample mounted on a three-axial translation stage (ABL1000, Aerotech Ltd.) [163]. The combination of the SLM and a half-wave plate matrix allowed the removal of relatively slow rotating and moving components for retardance and slow axis orientation control (Fig. 5.17).

An adapted weighted Gerchberg-Saxton algorithm was used to set the split beam energy at several levels at the back focal plane of the objective [133]. Combined with a phase distribution of Fresnel lens, various levels of intensity at different depths of the focal plane could be achieved (Fig. 5.18). The polarization direction was controlled by the half-wave plates matrix, where beams passing through the selected segment can generate the targeted polarization state [163].

The first digital documents have recorded (including copies of the Universal Declaration of Human Rights, Newton's *Opticks*, the Magna Carta, and the King James Bible) across up to 18 layers using optimized parameters (light pulses with energies of 0.2 μJ and a duration of 600 fs at a repetition rate of 500 kHz) [164].

At the same time, more than 200,000 text files and archive documents or more than 500 films can be written in one quartz multilayer disc. Thus, it is allow mankind to safely storage all accumulated important data for next generations [172].

Implementation of 5D storage needs a solution of complex problems that related to necessity of using high-precision equipment. That equipment must guarantee ultra-high accuracy of laser beam hit at the required point in the volume of rotating and moving media. Besides, it is necessary to provide a high speed of parameter modulation of femtosecond laser pulse. Significant engineering obstacles are also appeared during data retrieval at the high speed from medium volume using parameter recognition of reading beam modified by the nanostructures.

It should be noted that important part of the project is a solution of material problem of revealing optimal composition of oxide glass and optimal regimes of laser beam interaction with them. All this is necessary to achieve maximal speed of data recording.

Diamond is a unique platform material whose extreme properties and multifunctionality are enabling an ever-growing set of applications. Diamond typically contains impurities and other defects whose varying concentration and composition give

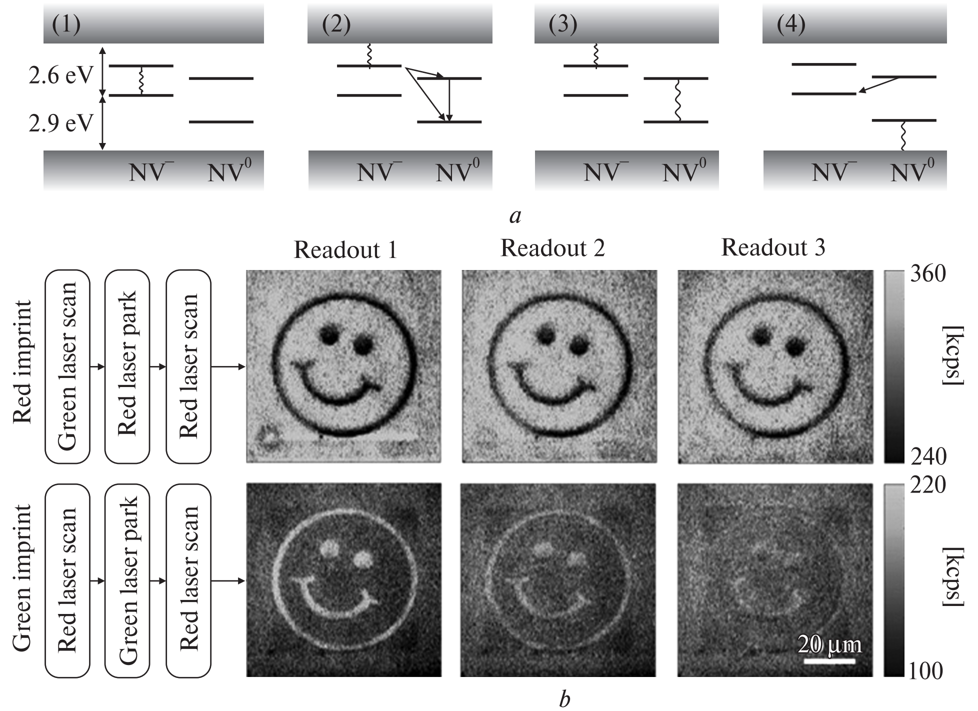


Fig. 5.18. Charge manipulation and readout in diamond

gems their signature colors. The paramagnetic centers can be located individually using confocal microscopy, initialized via optical pumping, and read out through spin-dependent photoluminescence measurements [145, 146]. Diagram of data recording using impurities in diamond is shown in Fig. 5.19.

Fig. 5.19 shows process of charge manipulation and readout in diamond: (**A**) Energy diagram for NV^- and NV^0. In (1) and (2), the successive absorption of two photons (wavy arrows) of energy equal or greater than 1.95 eV (637 nm) propels the excess electron of an NV^- into the conduction band, leaving the defect in the neutral ground state (solid arrows). In (3) and (4), an NV^0consecutively absorbs two photons of energy equal or greater than 2.16 eV (575 nm) transforming into NV^-. CB, conduction band; VB, valence bands. (**B**) Top: A binary pattern on an NV^--rich background is imprinted via spatially selective red illumination (632 nm, 350 μW, 100 ms per pixel). Bottom: Starting from an NV^--depleted background, the pattern results from selective illumination with green laser light (532 nm, 30 μW, 5 ms per pixel). From left to right, images are the result of three successive readouts of the same original imprint via a red scan (200 and 150 μW for the upper and lower rows, respectively). In all cases, the image size is 100 × 100 pixels, and the integration time is 1 ms per pixel. kcps, kilocounts/s [146].

The proposed diagram of data recording is a version of photoluminescent reverse memory. This memory based on the principle of electron capture and it is close to the development of recording in the rare earth-based media (CaS-EuSm, for example). Diamond media have to be stored in the dark. Disadvantage of such medium is obvious; information is erased during read out. Thus, permanent rewriting of information is required. One possible way to capacity increasing is usage of spintronics, i.e. additional data encoding through nucleus and electrons spins at the impurities areas. This will allow controlling the polarization of luminescence. So, this type of memory is very similar to the 5D-DVD, only technologically more difficult and single-layer recording.

A multilevel recording (data encoding through intensity levels of luminescence with multiple data reliability decreasing), sub-diffraction recording method (photoactivation which is used in photoluminescence microscopy), and use of additional layers for service information are also considered.

The main problem related to the diamond memory is a low speed of recording: time of one pixel recording is about 1 ms. One pixel for one-level recording is correspond to 1 bit of information. Thus, in spite of high mechanical strength the proposed diamond media will hardly be used for the long-term storage of data.

5.3. Sapphire optical discs with graphical information presentation for long-term data storage

The use of sapphire as a base allowed creating unique carriers which are characterized by high strength to mechanical damage, chemicals and can withstand temperature of above 1000 °C. None of the existing digital media can provide storage under these conditions. Mass production of single-crystals sapphire has mastered today in many countries. Sapphire is used for producing LEDs, protective smartphone screens, windows for submarines and spacecraft. The analysis of existing highly stable uniaxial single crystal materials has demonstrated that the best material for the optical disc substrate is sapphire [91-95], on the internal surface of which the information micro relief structure is formed [89, 91-95]. It has a high chemical stability, its wear resistance is 8 times greater than that of steel, it is thermally stable up to 1600 °C and it is optically transparent in the range from 0.17 microns to 5.5 microns. Sapphire is the hardest material being made of cheap components, and technology for growing single-crystal sapphire is simple and affordable process.

A few different optical carriers for long-term storage have been proposed in which sapphire substrates are used. For several years there have been experiments with using sapphire discs for permanent storage. This seems similar to the Norsam HD-Rosetta archival preservation technology, which has been around for decades. HD-Rosetta™ archival preservation technology utilizes unique microscopic processes to provide analog and/or digital data, including information as texts, line illustrations or photos on nickel plates. Density can be as high as over 10,000 pages per plate.

One of these optical carriers for long-term storage is media in which a layer of platinum with recorded information is protected by sapphire wafers (Fig. 5.20). It is well-known The Nanoform: Sapphire Disk System of Records Preservation used by ANDRA (Fig. 5.20).

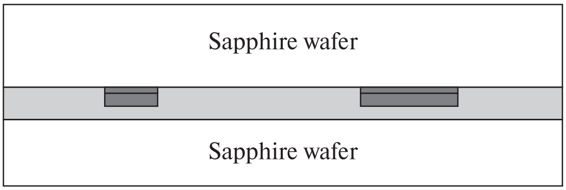


Fig. 5.20. Schematic representations of a nanoform [115]

The ultimate long-term storage solution: two 20cm (8in) sapphire disks, molecularly fused together, with a thin layer of platinum in between, and inscribe up to 40,000 miniaturized pages of text or images on the platinum. The information would be read with microscope. The disk is expected to have a lifetime of 10 million years [115].

The manufacturing process involves lithography already used for microscopic text writing. The two main technological advances rely on the choice of the materials and the use of a specific process. The precious and refractory metal layer, on which the information is engraved, is buried between two sapphire wafers stuck together by molecular bonding. These are both protected by patents [166, 167]. A team of scientists made a hard disk from sapphire which it claims will last 1 million years when looking to preserve records of where nuclear waste repositories were buried, not just in the near-future, but for tens of thousands of years. A key application would be as a solution for how future societies will be able to identify areas of buried nuclear waste. Nuclear reactors produce radioactive waste that needs to be safely stored for up to one million years. Once a disposal method is determined, future societies will need to know where the waste is buried. According to *Science* magazine Finland, France, and Sweden are the furthest advanced in the process of finding a geologically suitable site. While designers of such repositories are confident the waste can be buried safely, the fear is that future archaeologists may dig in the wrong places. Markers would be a way to allow them to know the sites where they should not dig [165].

With a sapphire disk, the warning message could be encoded into varied forms of written human communication, including words, pictograms, and diagrams, and in turn linguists and artists are involved in the project. The researchers say thus far they have no idea what language to use [165]. Currently, there's no digital format that we know can last that long with certainty. Two problems: First, the prototype disk cost $30,000 dollars. Second, what language to use when writing on the disk [115].

Sapphire disc with recording on a titanium nitride film has been proposed. The disc is covered with titanium nitride and the laser writes, literally on this "ink". A second disc, perfectly smooth, is applied thereto, and, under the effect of heat, merges with the first [117].

The Fahrenheit 2451 project has the ability for people to order their own nanoforms or jewelry with content. It appears that the nanoform (composed of two slices of synthetic sapphire which can withstand 2451 degrees Fahrenheit) currently etches analog letters and images on a disk. At the smallest size, every character has a size of 10 μm, and the disk can be read with a digital microscope or similar enlarging device. Disks must be created and inscribed in a clean room environment. The cur-

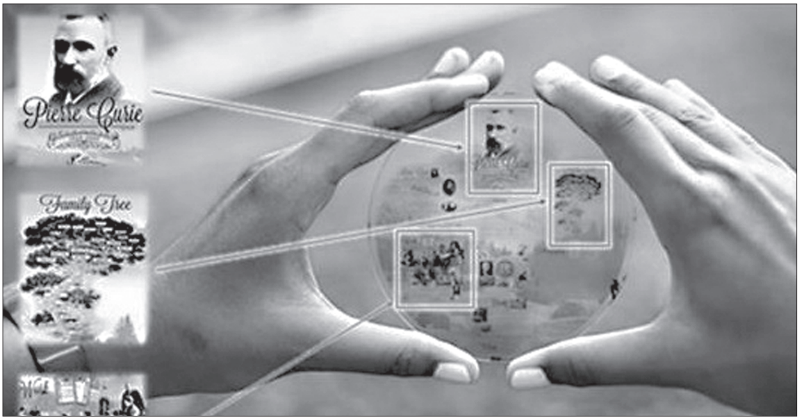


Fig. 5.21. A hard disk from sapphire of the Fahrenheit 2451 project

rent facility is Grenoble, France. The 8'' nanoform can store up to 10,000 letter pages at 150 dpi or 650 × 850 pictures (Fig. 5.21). The FAQ indicates it is possible to digitize the content again. It requires a professional scanner with 25,600 dpi precision, which you can find in most of labs and printing press, and the "quality of your pictures and documents will be approximately the same as the original." There is a viewing platform online to view the images [119].

The new In-volume Selective Laser Etching (ISLE) technique provide the first process for producing microchannels and shaped holes and cuts in transparent parts made of fused silica, borosilicate glasses, sapphire, and ruby. ISLE is a two-step process. In the first step, material transparent to the laser radiation is modified locally inside the volume (Fig. 5.22). For this purpose, ultra short laser radiation is focused inside the material. In the focus volume, the intensity of the laser radiation must be large enough to enable the absorption by multi-photon processes because the material is transparent at lower intensities. Therefore, laser radiation with pulse duration of picoseconds or shorter and focusing to a few microns is necessary. The material is heated by the absorbed energy during a short time to high temperatures, resulting in large local pressures. Because only a volume of several cubic microns is heated, the material cools rapidly and is quenched in a modified state. For example, single

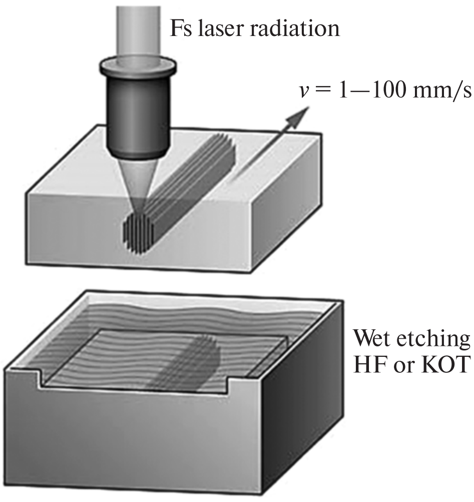


Fig. 5.22. Processing steps of ISLE: Writing of the modification with fs-laser radiation and subsequent selective wet chemical etching of the modified material

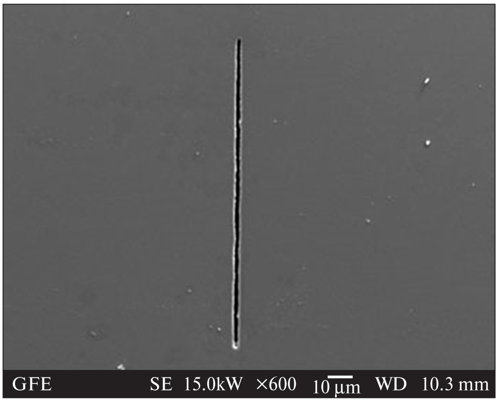


Fig. 5.23. Electron micrograph of a cross section of a 1 cm long microchannel in sapphire with 1 µm widths and 125 µm heights

crystalline sapphire is modified to an amorphous form. A connected volume inside the material is modified by scanning of repetitive pulses, which must have at least one contact to the surface of the work piece.

In the second step, the work piece is exposed to a strong acid or lye for some time. Because the material, modified by the laser radiation, is etched much faster than the unmodified material, the modified volume is selectively removed, resulting in hollow structures inside the work piece. The selectivity, that is the ratio of the etching rate of the modified material to the etching rate of the unmodified material, can be as high as 1,000. Therefore, very tight channels, cutting kerfs, and structures with a precision in the range of a few microns can be produced (Fig. 5.23) [118].

The ISLE process is characterized by:

- high energy efficiency (modification instead of evaporation);
- material efficiency (cutting kerf of a few micro meters);
- high precision in three dimensions (1-µm focus, no debris);
- scalability to high velocity using lasers with large repetition rates.

These advantages come about because the material is not removed during laser processing but modified locally. Thus, 3D shapes and complex structures can be produced with high efficiency and precision. The main challenge of the process development has its origin in the same fact: Since the modified material is not removed during processing, stress is accumulated in the work piece, which may result in cracking of the material. For its main applications in precision mechanics and medical engineering, the ISLE technique is used to cut out parts made of sapphire and glass [118].

With your eyes you will discern the pages, not the text (Fig. 5.24, 5.25). As these are analogical data, methods to retrieve information are numerous:

- A digital microscope will display the images and texts directly on computer.
- A powerful magnifying glass.
- A precise scanner (will not render a readable text when written at the smallest size).

The sapphire disc represents a dramatic breakthrough in long-term data storage domain, offering the possibility of media with lifetimes longer than several thousand years and resistant to fire, flooding, rats, etc. This innovation introduces a new era for the conservation of the memory of our human heritage and history as well as national or company's archives, and even family records. From this perspective, Arnano can be compared with the Arctic "doomsday vault" initiative, inaugurated in Febru-

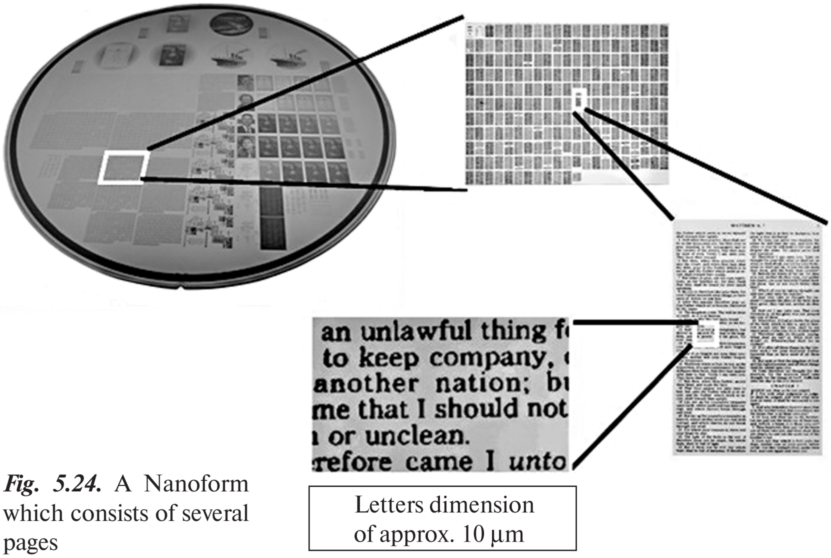


Fig. 5.24. A Nanoform which consists of several pages

Fig. 5.25. Digital devices to retrieve information from a Nanoform

ary 2008 by European Commission President Jose Manuel Barroso and Nobel Peace Prize winning environmentalist Wangari Matai, filled with samples of the world's most important information, and aimed at providing mankind with a Noah's Ark of data in the event of a global catastrophe. The nanoform offers the possibility of reproducing any document in various grey levels, and even in colors, by separately storing the images of each fundamental color (for example: Red, Green, and Blue). The information storage density is 20 times higher than that of a standard A6 microform. It is possible to store the equivalent of 10,000 A4 pages on a single 200mm sapphire disk. The reproduction is not forgeable: indeed, the data embedded within sapphire cannot be changed mechanically (scratching) or chemically. It could in principle be destroyed by a high power laser, but proof of the intervention and of the forgery would then be perceptible [165].

The sapphire disc represents a dramatic advance in long-term data storage technologies; with several specific advantages in comparison with numerical or physical solutions that are used today:

• Durability of the access to information: the information is maintained in an analog form, it is easily read thanks to simple optical techniques, based on unchanging physical principles; the information will never be impacted by the triple obsolescence of numerical techniques, i.e. the supports (very low lifetime of the CDs, DVDs disks, etc.), the software, and the equipments (new systems and peripherals);

• Preservation of the information stored with a high security level, and very low storage cost: the nanoform is resistant to biological, chemical, physical (light, radiation), mechanical, and heat and water attack due to the innovative use of sapphire substrates. For example, sapphire has a very high chemical immunity and can resist immersion of more than 500 years in a chemical mixture comprising hydrofluoric and nitric acids. Therefore, in comparison with current classical approaches, the nanoform will not need any particular environmental storage condition (except for mechanical protection against risk of a buildings collapse due to earth quake etc.) [165].

5.4. Sapphire optical discs for long-term storage of information with digital data representation

Creation of manufacturing technology of sapphire media on which information is recording in widely used formats (such as DVD and Blu-ray) is of great interest. It should be noted that in these formats information can be recorded with higher area density in comparison with analog representation of data. Actual problem is a creation of methods of information survivability assessment on optical carriers, substrates of which are made of single-crystal sapphire. The idea of use the sapphire substrates in long-term storage media was proposed more than a decade ago [30]. However, realization of potential opportunities of such carriers was difficult due to significant distortions of laser beam during its passing through sapphire substrate. The reason of these distortions is high birefringence value of sapphire. Birefringence also does influence on data read-out from usual compact-discs, but birefringence value of polycarbonate is 20-30 times less [118]. Significant reduction of polycarbonate substrate birefringence is possible by increasing the time of injection molding and special heat treatment. Birefringence decreasing of sapphire substrates using similar methods is not possible, and it is necessary an additional changing of the laser beam shape.

The sapphire substrate on which information is recorded in the form of a microrelief, and the data are read through the sapphire substrate can be a key element of technology for optical data storage with high reliability and long shelf life [139, 140]. Thickness of the substrate may be 0.4-1.2 mm, which provides sufficient mechanical strength to the carrier.

However, sapphire substrates have significant birefringence Δn_{spf} (n_o = 1.780, n_e = 1.772, $\Delta n_{spf} = n_e - n_o = -8 \cdot 10^{-3}$) at λ = 442 nm), which leads to significant aberrations when focusing the laser radiation through the sapphire substrate of the carrier. Phase distortions with an unusual light polarization (p-polarization) occur when laser beam which forms a spherical wave front, falls on a single crystal sapphire substrate. Phase distortions are superposition of astigmatism and spherical aberrations of various orders. In particular, the phase distortions lead to the fact that the s-and p-polarized light will be focused at a different depth and the distance between focal points ΔF is defined as follows [94]:

$$\Delta F = 2H\Delta n / n_o , \tag{5.1}$$

where H is the thickness of optical carrier substrate. Thus, phase distortions lead to the fact that distance between two spots (s- and p-polarized beams) several times higher than the spot size.

There is a possibility to compensate the aberrations caused by the anisotropy of the sapphire substrate using an additional compensating plate with a positive uniaxial material (inverse to the sapphire). Schematic representation of the method of aberration compensation is shown in Fig. 5.26.

Condition of spherical aberration compensation in general form can be written as:

$$\frac{n_{plc}^2 - 1}{8n_{plc}^3} NA_1^4 \frac{H_{plc}}{\lambda} = \frac{n_{spf}^2 - 1}{8n_{spf}^3} NA_1^4 \frac{H_{spf}}{\lambda} + \frac{n_{com}^2 - 1}{8n_{com}^3} NA_2^4 \frac{H_{com}}{\lambda}, \tag{5.2}$$

where n_{plc} is refractive index of polycarbonate, H_{plc} is thickness of polycarbonate substrate, n_{spf} is average value of the refractive index of sapphire between the ordinary and extraordinary rays, H_{spf} is thickness of sapphire substrate, n_{com} is average value of the refractive index between ordinary and extraordinary rays for the material used for the manufacture of the compensating plate, H_{com} is thickness of compensating plate, NA_1 is numerical aperture in region behind the objective, NA_2 is numerical aperture in the region where compensating plate is placed, λ is wavelength of laser radiation.

Condition of astigmatism compensation in general form can be written as:

$$\Delta\Phi = \Delta\Phi_1 + \Delta\Phi_2 = NA_1^2 \frac{H_{spf}}{n_{spf}^2} \Delta n_{spf} + NA_2^2 \frac{H_{com}}{n_{com}^2} \Delta n_{com} = 0, \tag{5.3}$$

where Δn_{spf} and Δn_{com} are differences of the refractive indices for the extraordinary and ordinary rays of sapphire and material of compensating plate, respectively. Fig. 5.27 shows a modified optical scheme of readout system. From the standard scheme, the modified optical scheme differs by additional compensating plate 6, which is placed between the objective lens and the carrier. During data reproducing light generated by a laser diode (1) through the diffraction grating (2), a quarter-wave plate (4) and the beam splitter cube (3) is directed to a focusing objective lens (5). The focusing lens focuses laser beam through the compensating plate (6) and media substrate (7) to the relief structure of information media. Presence of single crystal compensating plate leads to the fact that ordinary and extraordinary beams are focused by focusing depth in one plane.

Birefringence value Δn_{com} of compensating plate material must have inverse value to the birefringence value of sapphire Δn_{spf}. Defocusing effect of the laser beam

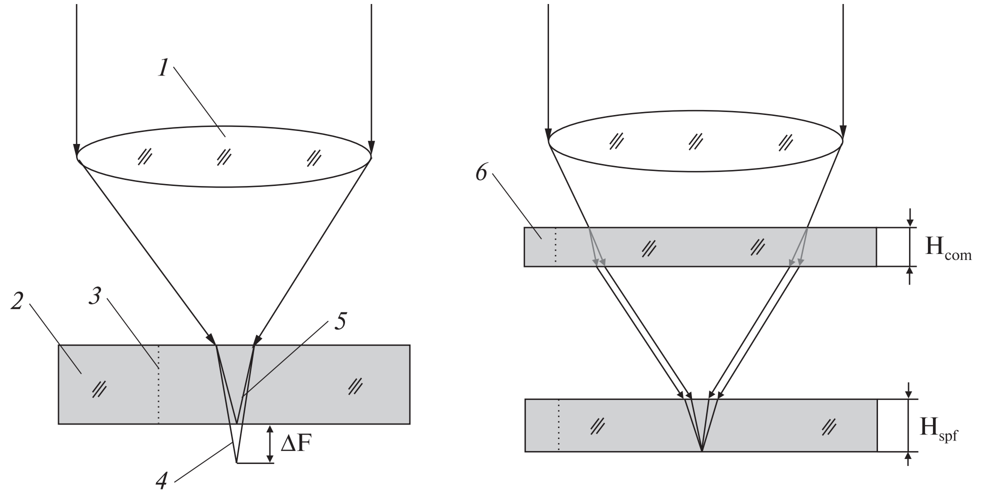


Fig. 5.26. Schematic representation of aberration compensation: *1* — objective lens, *2* — sapphire substrate, *3* — optical axis, *4* — extraordinary ray, *5* — ordinary ray, *6* — compensating plate

with different polarizations is eliminated by passing the beam through an additional compensating plate.

Table 5.5 shows the examples of single crystal materials which can be used to manufacture of the optical media substrate and the compensating plates. As a material of the compensating plate one can use positive single crystal quartz ($n_o = 1.544$, $n_e = 1.553$, $\Delta n_{kvr} = n_e - n_o\ 8 \cdot 10^{-3}$). The condition of astigmatism aberration compensation in this case can be written as:

$$\frac{H_{spf}}{n_{spf}^2}\Delta n_{spf} + \frac{H_{kvr}}{n_{kvr}^2}\Delta n_{kvr} = 0, \quad (5.4)$$

where H_{kvr} is the thickness of quartz compensating plate. Higher orders of aberrations will be fully compensated when the materials of optical disc substrate and the compensating plate have the same refractive indices for the ordinary ray. Otherwise, the compensation is partial. Accordingly, the second main condition for choosing the material from which the compensating plate of optical readout system must be manufactured is that the refractive index of the compensating plate material n_{com} should be close to the refractive index of the substrate of the optical carrier n_{sub}.

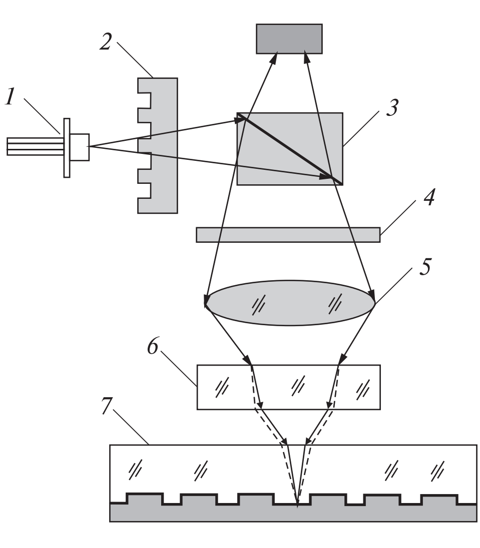


Fig. 5.27. Modified optical scheme of readout system

Conditions for obtaining optical system with minimal residual aberrations of the optical disc with a high-stable single crystal sapphire substrate and single crystal quartz compensating plate can be written as [93]:

$$1)\ H_{kvr}/H_{spf} = 0.62 \div 0.72,$$
$$2)\ H_{kvr} + H_{spf} = H_{sum} \pm 5\%, \qquad (5.5)$$
$$3)\ |n_{com}/n_{sub}| < 0.3$$

First condition is a condition of the astigmatism compensation, the second is a condition for compensation of spherical aberration of plane-parallel substrate, the third is condition for sufficient for optical recording compensation of higher-order aberrations, H_{sum} is thickness of the polycarbonate substrate. Table 5.6 shows examples of constructional parameters (thickness) of the disc with sapphire substrate and quartz compensating plate for different formats of optical data storage.

It should be noted that recording of information on sapphire discs can be done using technology which is used during recording of information on a master disc in the technological process of CDs production. The conventional scheme of data recording process on master disc consists of following steps [150]. The positive photoresist layer is deposited on the substrate surface. The information is recorded on photoresist layer via laser recording system by focused laser beam. Parameters of laser recording system define the format of optical disc for long-term data storage. The

Table 5.5. **Examples of single crystal materials which can be used to manufacture the optical media substrate and the compensating plates**

Material	Parameters		
	n_o	n_e	$\Delta n = n_e - n_o$
Sapphire (Al_2O_3)	1.765	1.757	–0.008
Single crystal quartz (SiO_2)	1.542	1.551	+0.009
Orthovanadate yttrium (YVO_4)	1.993	2.212	+0.219
Rutile (TiO_4)	2.574	2.86	+0.286
Zircon ($ZrSiO_4$)	1.96	2.015	+0.055

Table 5.6. **The thickness of sapphire and quartz layers for various formats optical discs**

Type of media	The thickness of sapphire substrate (mm)	The thickness of quartz plate (mm)	Uncompensated aberration $\Delta\Phi / \Phi_0$, 100%
CD	0.714	0.486	1%
DVD	0.357	0.243	1.6%
CD (400 nm readout)	0.714	0.486	0.3%

etching mask is performed by selective etching of recorded photoresist layer. The ratio of photoresist and substrate materials etching rate defines the thickness of photoresist layer. There must be a 150 nm photoresist layer thickness for performing information later with 110-130 nm depth.

There are a few methods of forming the microrelief on the sapphire substrates. Let us show the possibilities of their application.

Plasma dry etching is very effective to form fine structures on the substrate, but it is accompanied by several problems such as a low etching rate and poor etch mask selectivity [149]. Technologies used for improving luminous efficiency of GaN-LEDs (Gallium Nitride-Light Emitting Diodes), pillar or cone shaped structures are periodically generated on sapphire substrates, which is called PSS (Patterned Sapphire Substrates) are used in order to format a microrelief structure on the sapphire substrates. Nano-PSS fabricated on the sapphire substrates have the size: height 100 – 750nm, spacing of 230nm [149].

Conventional photolithography and etching are used for PSS fabrication. Either dry etching or wet etching can be used as the etching process. Dry etching is more commonly used than wet etching for the mass production of PSS, since it allow one to control the pattern shape easily. Wet etching has recently attracted much attention from the view point of its low fabrication cost compared with dry etching. Patterned sapphire substrates (PSS) can be fabricated by wet etching solutions with different mixture ratios of H_2SO_4 to H_3PO_4 and different temperatures. It was found that the mixture ratio and temperature of the etching solutions affect the ratio of the pattern diameter to the pattern depth. In addition, the observed pattern shape was strongly affected by the mixture ratio [151]. The main disadvantage of wet etching is application of high toxic reagents and high temperature of etching (520-550 K).

The most effective way to format the nano-PSS is dry etching using ICP-RIE (Inductively Coupled Plasma-Reactive Ion Etching) and employing photo resist (PR) as a mask. However, there have been challenges related to mask lithography and sapphire etching. First of all, to prevent deformation of the resist mask due to heat from plasma, the mask was UV-cured and hard-baked to increase its heat resistance and it was confirmed that deformation didn't take place with processing carried out at 250 °C. Heat generated by the plasma will deform the PR, and the heat resistance of the PR mask has to be improved. Typically, the highest post-development process temperature for PR is around 120 °C, but since high RF power is required for the etching of sapphire, heat from the plasma to the substrate increases, and the PR mask is deformed (or burned) due to the increased substrate temperature as shown. There are two ways to prevent this heat-induced problem: (i) directly cool the sapphire substrates to keep the PR mask below its deformation temperature; (ii) increase the maximum processing temperature of the PR mask. Direct cooling is carried out using mechanical or electrostatic clamping. A significant influence on the process of ion etching makes the magnitude of the bias voltage. To adjust the height, selectivity against PR mask was crucial, and it was confirmed that selectivity can be controlled by adjusting bias RF Power, process pressure, and CHF_3 flow rate. As bias power was

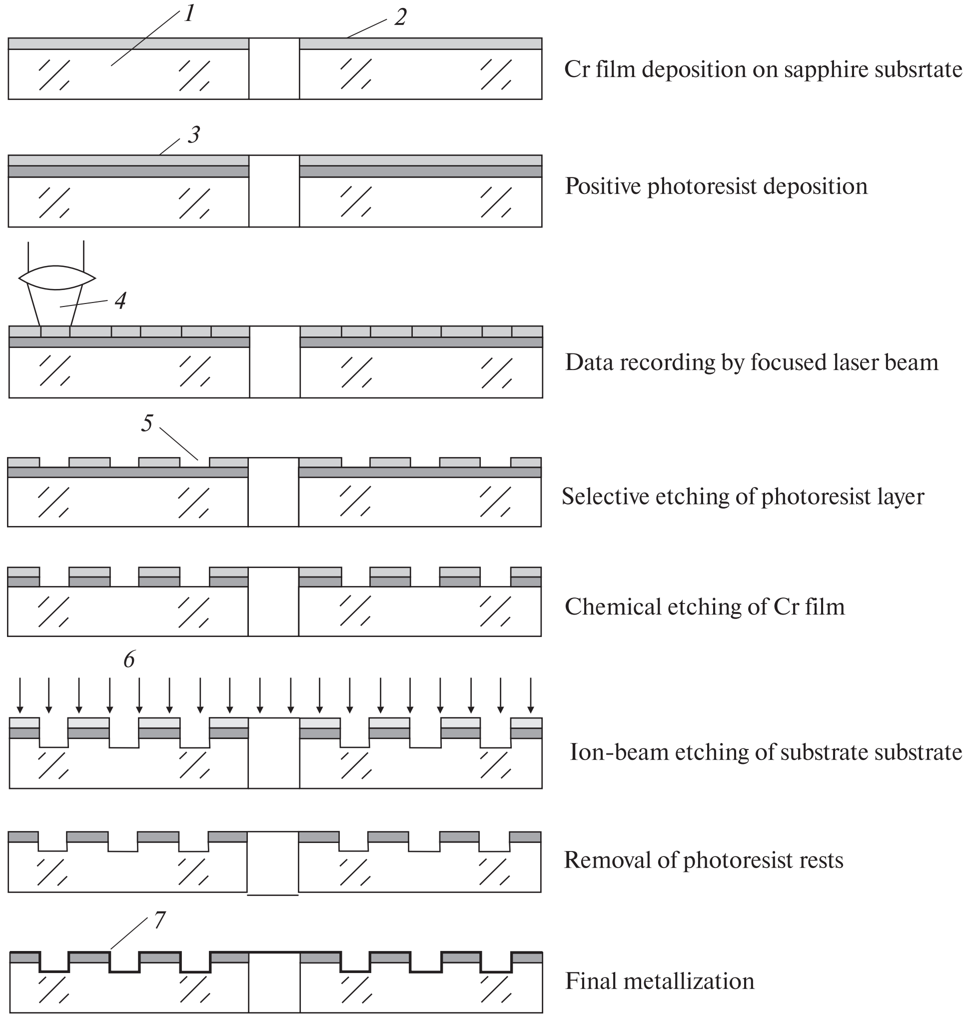


Fig. 5.28. Improved scheme of data recording process on the optical disc for long-term data storage: *1* — sapphire substrate; *2* — Cr film; *3* — positive photoresist layer; *4* — focused laser beam; 5 — etching mask, *6* — ion beam, *7* — reflective metallic layer

increased, the selectivity decreased, and this means that the change of PR mask etching rate is larger than that of sapphire, indicating that selectivity can be controlled with bias power [149].

The technology of making diffractive optical elements on sapphire substrates using the double mask Cr-PR is known. Etching of sapphire substrate is performed by ion-chemical method in CF_4 medium through double mask of Cr-PR. Etching rate of sapphire of 130 E per minute was obtained in a high-frequency discharge $f = 13.56$ MHz with the voltage of automatic displacement on a table with a sample of $U \approx$ 300 V [148]. This technology can be applied during long-term sapphire discs manu-

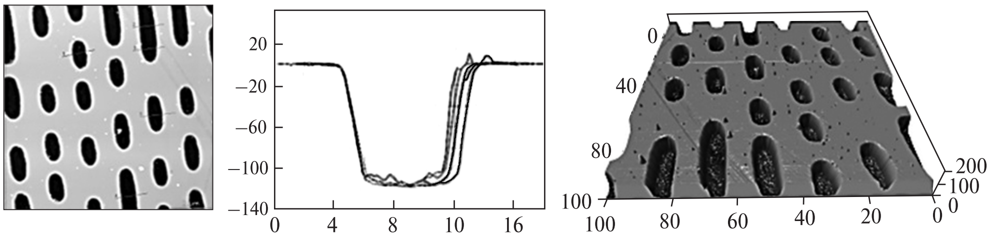


Fig. 5.29. AFM image of sapphire substrate after ion-beam etching

facturing. Ion-beam (and reactive ion-beam) etching is one of the most suitable methods for microrelief creation on the surface of high-stable materials [71]· In our experiments the information relief on the substrate surface was performed by reactive ion beam etching. It was shown that maximum depth of microrelief on sapphire substrate is equal to 95 nm and limited by initiation of surface charge induced by positive ions bombardment in the case direct (zero frequency) ion-beam etching. Experimental results have shown that in this case the speed of etching equals 3 nm per minute for sapphire and 10 nm per minute for photoresist. In order to increase microrelief depth the using of additional metallic masks has been proposed. In further investigations it was suggested to use an additional layer of chromium, since it has a high adhesion to substrate (Fig. 5.28) [138,139].

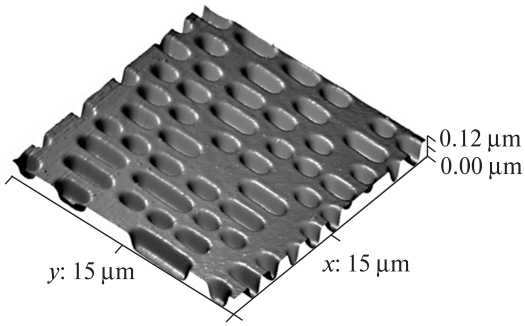


Fig. 5.30. General view of sapphire substrate of long-term data storage media

Chromium film was deposited by thermal evaporation. Thickness of Cr film was equal to 50 nm. The positive photoresist layer is deposited on the Cr film by centrifugation and subsequent annealing in air atmosphere with T = 363.15 K for 1 hour. Thickness of photoresist layer was 150 nm. Chemical etching of Cr film was performed by 20% $CeSO_4$ solution for 2 minutes. Further, the ion-beam etching of sapphire substrate was performed through windows in Cr and photoresist layers. AFM image of sapphire substrate after ion-beam etching is shown in Fig. 5.28.

From Fig. 5.29 one can see that depth and width of pits are about 115 nm and 600 nm, respectively, which is sufficient for the production of sapphire discs for long-term data storage at modern optical information formats. General view of sapphire substrate of long-term data storage media is shown in Fig. 5.30.

Also note that not only the physical storage medium needs to last, but also the entire surrounding infrastructure, like data centers where data is stored, needs to be designed taking a long term view on its survivability. There is no gain in storing costly sapphire disks in a building that may be crippled by earthquakes or floods [129].

5.5. Metallic media for long-term information storage with optical data read-out

Historical experience of document storage showed that metallic media allow us to store information for a long time under certain conditions. One of the biggest and best preserved hoards of ancient sealed books, which had been secretly hidden for centuries, has been discovered in Jordan. Early indications are that some of the books could date from the first century CE/AD and may be among the earliest Christian documents, predating the writings of St. Paul. The books, or "codices", were apparently cast in lead, before being bound by lead rings. Their leaves — which are mostly about the size of a credit card — contain text in Ancient Hebrew, most of which is in code. If the relics are of early Christian origin rather than Jewish, then they are of a huge significance. His collection consists of more than 20 codices (early books), cast mostly in lead and containing cryptic messages in Hebrew and Greek along with symbols such as the menorah. A discovery in the Middle East of more than 20 ancient lead plate "books" — each with 5 to 15 pages — is being hailed by some as one of the most important religious discoveries of the past. Others are calling it ridiculous [122, 123].

Now about 100 thousand pages of text in 2500 languages are collected in the archive. Upon completion of work, the accumulated material is supposed to be recorded on a nickel disk ("Rosetta disc") by engraving. According to the developers, it will allow to store the archive from 2 to 10 thousand years.

The Rosetta Disk contains over 13,000 pages of information on over 1,500 human languages (Fig. 5.31). In 1998 the idea was put forward that a modern Rosetta Disk might be made using micro-etching technology.

The disk would have as many translations of the same text as possible, all etched into metal that could last for thousands of years [116]. The pages are microscopically etched and then electroformed in solid nickel, a process that raises the text very slightly — about 100 nanometers — off of the surface of the disk. Each page is 400 microns across and can be read through a microscope at 650X as clearly as you would from print in a book. Individual pages are visible at a much lower magnification of 100X. The outer ring of text reads "Languages of the World" in eight major world languages. Proceeds support the Rosetta Project and to build the largest open, publicly accessible collection of resources on the world's languages [144, 155].

The group that implements the Rosetta project believes that one of the key problems is not the technique of the carrier creation, but the way of encoding information. A disadvantage of the disk is that its surface is fragile and can easily be scratched so an encasing for the disk is essential for long time survival and repeated readout might damage the surface of the disk. The Rosetta disk is a 3-inch nickel disk with an estimated lifespan of 2,000—10,000 years. The disk is mounted beneath a glass hemisphere. This graphic side of the disk is pure titanium. A black oxide coating has been added to the surface. The text is etched into that, revealing the whiter titanium. This bold sign board is needed because the pages of genesis which are etched on the mirror-like opposite side of the disk are nearly invisible. This business side of the disk is pure nickel. Picking it up you would not be aware there were 13,500 pages of linguistic gold hiding on it. The nickel is deposited on an etched silicon disk. In effect the Rosetta disk is a nickel cast of a micro-etch silicon mold. When the disk is held at the right angle the grid array of the pages forms a slight diffraction rainbow. You need a 750-power optical microscope to read the pages. The Rosetta disk is not digital. The pages are analog "human-readable" scans of scripts, text, and diagrams. Among the 13,500 scanned pages are 1,500 different language versions of Genesis 1-3, a universal list of the words common for each language, pronunciation guides and so on [116].

The technology for micro-etching was adapted by Los Alamos Labs from gallium-ion beam micro-circuitry FIB (focused ion beam) machines to etching text instead of circuits. Diagram of the recording process is shown in Fig. 5.32. The etching rate in the irradiated areas is proportional to the radiation dose.

They specifically developed this in order to store data for the nuclear waste storage program that had a congressional mandate of 10,000 years. This technology had been licensed exclusively by a company called Norsam to make high density archival materials; in fact they called it HD Rosetta. The technology theoretically could put around 300,000 pages on a 2.8" metal disk, with the caveat however that would need a very specialized electron microscope to read it. They estimated that at 10X that scale, or 30,000 pages per disk, the content would be readable with optical microscopes in the 1000X range (17th century technology). The human eye readable side of the disc would have an outer ring of text, in eight major world languages, spiraling down to about 0.2mm high. Inside of that was a listing of all the languages represented on the disk in order, and grouped by major geographic region. That text was all at 0.18mm high, and was written in a special "engravers font". This font used as few curve points as possible, and approximated a single stroke instead of the usual outline information. This was important because if the etching machine outlined each charac-

Fig. 5.31. Rosetta disc is a Physical component of "The Rosetta Project"

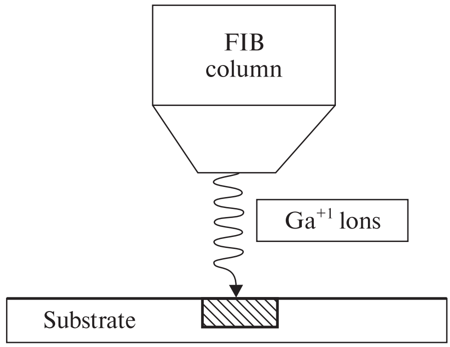


Fig. 5.32. Diagram of the recording process [124]

ter, which is the standard True Type font method, the characters would get muddled by the necessary thickness of the outline. A single stroke font is more like how you write a character with a pen, utilizing the thickness of the line itself, instead of outlining each character. In the center of all of that is an image of the world. This disk is made out of commercially pure (CP) 0.035" thick titanium cut into a 2.8" circle by water jet, and then coated in a matte black oxide. The material and cutting cost around $100, and having 15 of them coated cost $600 (we made plenty of extras for testing with). The side of the disk with all the data pages on it was also etched by Norsam. The pages written so that ended up in a circle to maximize the page size on the disk. We had also learned along the way that we should not expect to be able to etch more than 15,000 pages, and still have them be easily readable by optics. This meant that we had to edit the archive. We ended up taking the first page of each element of the archive, adding all our phonetic word lists, as well as a few pages at the beginning and end about the project. This left us with about 13,500 pages to be etched. The pages from the archive organized into folders, one folder to a row to be etched on the disk. These were also grouped by geographical region and named with a special file name that had meta data about each file (its name, region, three digit code, page number etc.). I then fed these through a program, ironically called Debabelizer, to automatically "stamp" each page at the top with the meta data in the file name. The digital stamping was done in another special font called OCR-b which allows for Optical Character Recognition. This would ostensibly allow for a computer to catalog all the data on the disk if it were scanned back in with a microscope. These files "padded" with blank files so that if they were organized in a grid, we would end up with all our files in a circle in the center of that grid. These folders were then sent by hard drive to Norsam who etched them into silicon in their machine. Unfortunately they could not maximize it to the disk size as they had said, and we ended up with the pages written smaller than hoped in a 1.9" circle. This silicon master was then transferred to another substrate, and then used as a mold for successive layers of nickel plating to create a pure nickel disk [116].

Metal matrixes are used when manufacturing ROM compact discs. Serial production has not changed in its core since the release of records. First, a metal matrix-stamp is made, with which the records of molten plastic are stamped. Using metal media will allow solving a number of problems regarding the long-term storage: avoid multiple rewrites and changes forms of information; solve the issue of storing metadata and so on [113, 92].

Nickel digital discs function exactly like CDs, and allow for easy retrieval of information with the use of a nickel disc reader [131]. Norsam Company also produces metal CDs and DVDs. Customers send Norsam the data they want placed on the disk via a digital file. Norsam then creates a master disc, writing the data onto the disk [136]. Norsam's data writing system (called NORSAM HD-ROM for High-Density Read-only Memory), which is currently under development, uses a highly focused charged particle beam instead of the much larger laser beams used by today's optical data storage devices. Norsam's initial charged particle beam can be focused to a

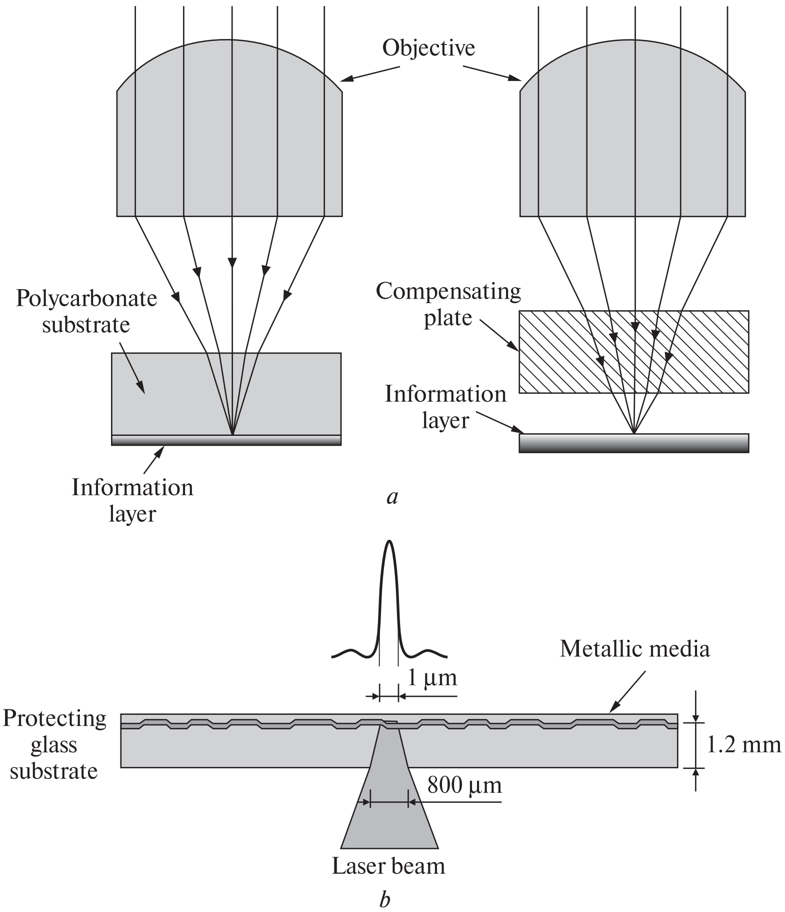


Fig. 5.33. Modification of optical system for information read-out from metallic media (*a*); metallic media with protecting glass substrate (*b*)

50-nanometer spot size, which could correspond to as much as 165 gigabytes of data storage on a CD-sized (120mm) disc [140]. This is due to the fact that the size of the minimal elements on the surface of the metal data carrier recorded in the CD format is (0.8-0.6) μm, and in the DVD format is (0.4-0.35) μm. Table 5.7 shows the causes of damage to metal carriers and possible methods to eliminate their effects.

A more reliable protection of the metal carrier against dust, organic and inorganic solution contamination can be achieved in the case of carrying out the medium in such a way that the information is reproduced by focused laser radiation through a transparent substrate 1.2 mm thick for carriers recorded in CD format or 0.6 mm for media recorded in DVD format. The connection of the metal carrier and the transparent protective substrate can be done by photopolymer lacquers which provide a high optical homogeneity of the connecting layer. A general view of the carrier with a protective layer in the form of a transparent substrate is shown in Fig. 5.33. Thickness

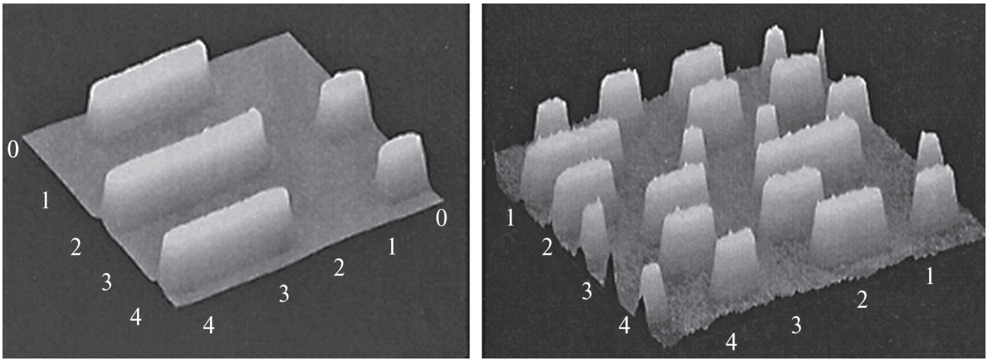


Fig. 5.34. Information recording area with different recording density: CD format (track pitch of 1.6μm) (*a*); DVD format (track pitch of 0.74μm) (*b*)

Table 5.7. **Causes of damage to metal carriers and possible methods to eliminate their effects**

The main causes of damage	The consequences of damage	Troubleshooting methods
Dust pollution	Increase in the number of corrected errors, the appearance of uncorrected errors	Washing with deionized water, drying on a centrifuge
Influence of water	Increase in the number of errors due to the appearance of spots on the surface of the carrier	Electrochemical cleaning of the carrier, washing with deionized water, drying on a centrifuge
Abrupt temperature changes, the impact of fire	Deformation of the carrier, which leads to an increase in the beating of the carrier	Using special adapters to reduce the beating of the storage medium
The impact of air pollution, including sulfur oxides, nitrogen, carbon	Minor increase in the number of errors due to the appearance of layers on the surface of the carrier	Electrochemical cleaning of the carrier, washing with deionized water, drying on a centrifuge

of the metal layer can be reduced to 30-100 microns to reduce mass of the storage media. One of the variants of the protective layer can be a polymer film with a thickness of 0.2-0.3mm, which is coated on the information surface of the carrier [120].

This method is used to protect the information surface of metal stamps used for copying CDs from surface contamination. To ensure a uniform film of a certain thickness, the protective layer is applied by centrifugation. Most often, thermoplastic materials (varnishes) are used for the protective layer, which require annealing at a temperature of 60-90 °C for 10-20 minutes.

The advantages of metal matrices are obvious: manufacturing technology of such matrices is serial, tested for a long time; metal matrixes are made of high-strength nickel-vanadium alloy (withstand stamping up to 80,000 polycarbonate discs at temperatures above 200 °C in a corrosive atmosphere); cost of the basis for manufacturing the metal matrix does not exceed 50 euros; the manufactured metal matrix can be reproduced on a device like all known CD-ROMs directly connected to a personal computer. The experience of metal stamps storing proved the possibility of long-term storage of recordings. The biggest impact on the reproduction of the information form metal media is the contamination of the media surface with dust and traces of evaporated water.

A characteristic feature of these varnishes is the possibility of easy separation from the metallic surface. The conducted experimental researches have shown the possibility of information reproduction through a transparent protective layer 0.1÷0.2 mm thick without modernization of the optical reading system. Protective layer can be removed and replaced after a period of operation of the metal recording medium. If necessary, chemical cleaning of the surface of the metallic carrier from contamination and residual material can be done during this operation. Reproduction of information from a metal carrier protected by a polymer layer can be carried out on standard CD players in which the optical head is modified by adding to the objective lens a compensating plane-parallel plate of 1.2 mm thick. A general view of the information recording area in digital form on the modern metal media with different recording density is shown in Fig. 5.34. On metal media recorded in CD format, the minimum length of the projections is 0.8 μm, the height is 150 nm, and in DVD format — 0.4 μm and 120 nm, respectively. And unlike other storage mediums (such as tape and plastic-based CDs and DVDs), thin metal disks can withstand exposure to electromagnetic radiation, extreme temperatures and saltwater, keeping the data stored (technically, etched) on them safe and legible for, according to some estimates, thousands of years.

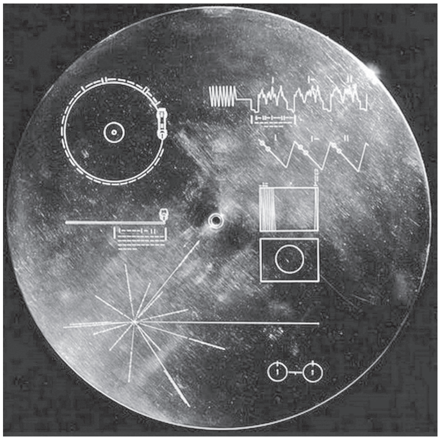


Fig. 5.35. Voyager's NASA's golden disc [92]

Metal carriers have already been used for secure storage of information. The Voyager message is carried by a phonograph record — a 12-inch gold-plated copper disk containing sounds and images selected to portray the diversity of life and culture on Earth (Fig. 5.35).

Scheme of radiation of the hydrogen atom for obtaining metric and temporal units, and a map of pulsars on which the position of the Sun in the Galaxy are also reproduced on the plate [92].

5.6. Conclusions

1. The use of recording media with greater thermodynamic stability makes it possible to increase the storage time on recordable optical media. The use of such media became possible due to a significant increase in the power of lasers used to information recording.

2. It is possible to substantially increase the guaranteed storage time of data on carriers with optical reading of the recorded information by using carrier substrates made of highly stable materials. New media with metallic (nickel and tungsten), silicon, high-strength silicate glass and synthetic sapphire substrates have been developed and are being created. Media with such substrates use the ROM mode. The historical experience of document storage showed that such carriers have kept recorded information for centuries. The media produced at the present time from such materials have a much higher density and recording speed.

3. The greatest guaranteed periods of data storage can be provided on sapphire carriers. Such carriers withstand heating to 1000 °C, the effect of water vapor and organic solvents; have a high mechanical strength, which excludes the appearance of scratches. Visual coding methods are used for recording information on most long-term data storage carriers with substrates from highly stable materials. To write information to sapphire substrates, it is intended to use codes used to represent data on compact discs.

4. Development of special recording layers is the main direction of data storage reliability improving on write-once optical mediums. Voids in polymers and glass materials which are permanent laser-induced physical changes can provide an approach to long-lifetime storage without data degradation.

5. The variety of technical solutions is proposed for technology of long term data storage which is based on using highly stable materials for both substrate and recording medium. Materials for the manufacture of optical discs having long term data storage time must be chemically, thermally and mechanically stable.

6. The sapphire disc represents a dramatic advance in long-term data storage technologies; with several specific advantages in comparison with numerical or physical solutions that are used today.

7. Not only the physical storage medium needs to last, but also the entire surrounding infrastructure, like data centers where data is stored, needs to be designed taking a long term view on its survivability. There is no gain in storing costly sapphire disks in a building that may be crippled by earthquakes or floods.

CHAPTER 6

OPTICAL LONG-TERM DATA STORAGE PROGNOSTICATION AND FAULT DETECTION

6.1. Optical media longevity testing procedures

Accelerated lifetime testing assumes a stable aging process. Changes in temperature and other environmental conditions do not introduce new chemical reactions to the aging process. The Arrhenius and Eyring methods are based on this assumption. These methods require: Multiple test conditions ((combinations of heat & humidity). The product aging behavior is log-linear under the test condition. An acceleration of the aging reaction rate through elevated temperature and/or humidity [172].

Lifetime estimation method for optical disks is based on the Eyring acceleration model and statistical analysis. The statistical distribution of life data for the optical disk assumed a lognormal distribution. Analysis of statistical techniques based on the ISO standard is proposed as a life expectancy prediction method for recordable disks. The standardized life expectancy of an optical disk is defined as the minimum lifetime of 95% survival probability at 25/50% RH with a 95% confidence level. An acceleration test was conducted using a high-density optical disk under stress conditions of temperature and relative humidity. The statistical distribution function of optical disk lifetime data must be clarified experimentally to apply this technique to lifetime estimation of optical disks in the future [94]. A physical lifetime factor of the optical disk degrades in the recording layer. Therefore, recording-layer degradation can be accelerated using a high temperature and high humidity environment. The acceleration test model is applied to the Eyring acceleration model, which incorporates temperature and relative humidity (RH). The Eyring model equation is an expression based on the laws of thermodynamics as [94]:

$$t = AT^d exp\left(\frac{E_s}{kT}\right) exp\left\{R\left(B\frac{C}{T}\right)\right\}, \quad (6.1)$$

where t — lifetime time, R — relative humidity, E_s — activation energy, k — Boltzmann's constant, T — absolute temperature (Kelvin), and

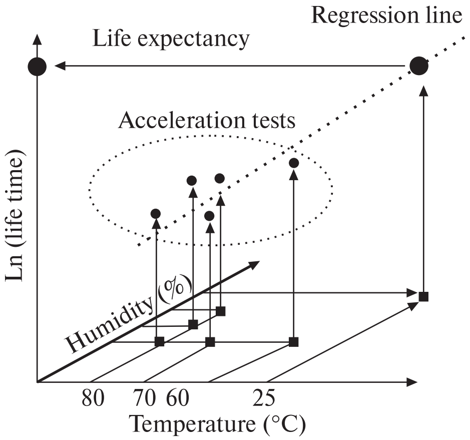


Fig. 6.1. Outline of the acceleration Eyring model [157]

A, B, C, d — constants. When the temperature range is small, as in this case, the Eyring equation is modified. The simplified Eyring's model equation is shown as the following [157]:

$$t = A \cdot exp\left(\frac{E_s}{kT}\right) exp(B \cdot R). \quad (6.2)$$

In that equation, taking the natural logarithm of both sides produces

$$\ln(t) = \ln(A) + \left(\frac{E_s}{kT}\right) T^{-1} + B \cdot R \quad (6.3)$$

Fig. 6.1 shows the lifetime estimate model using Eyring acceleration tests. An acceleration test of the stress condition was executed with different temperatures and relative humidity. Lifetime data were measured under respective conditions.

Constants A, B and E_s were calculated using multiple linear regression analysis of measurement results; results were used to determine the Eyring equation of the evaluated sample disk [157].

The measurement item of the life time criteria used the byte error rate (BER) of the read-out digital signal (Fig. 6.2). In the experiments the time of end-of-life (life time) was assumed as the time at which BER reached 9×10^{-4}: BER $=9 \times 10^{-4}$ corresponds to the maximum PI error number 280 of eight error correction code (ECC) blocks in the digital versatile disk (DVD) system.

The standardized life expectancy of an optical disk is defined as the minimum lifetime of 95% survival probability at 25/50%RH with a 95% confidence level. The following two items were assumed to apply a statistical analysis.

Statistical distribution of the lifetime of the high-density PC optical disk applies lognormal distribution:

$$\begin{cases} f(t) = \dfrac{1}{\sqrt{2\pi}\sigma t} exp\left\{-\dfrac{1}{2}\left(\dfrac{\ln(t)-\mu}{\sigma}\right)^2\right\}, t > 0, \\ f(t) = 0, t \le 0 \end{cases} \quad (6.4)$$

where μ — log mean, σ — log standard deviation.

The lognormal cumulative distribution function F(t) and reliability function R(t) are shown as the following equations

$$\begin{cases} F(t) = \dfrac{1}{\sqrt{2\pi}} \displaystyle\int_0^t \dfrac{1}{\sigma x} exp\left\{-\dfrac{1}{2}\left(\dfrac{\ln(t)-\mu}{\sigma}\right)^2\right\} dx, \\ R(t) = 2 - F(t). \end{cases} \quad (6.5)$$

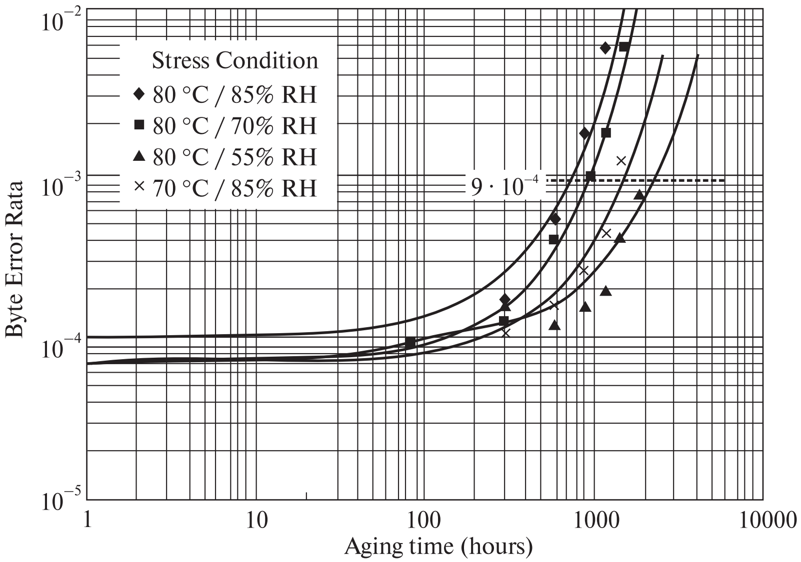


Fig. 6.2. Typical experimental result of BER measurement [144]

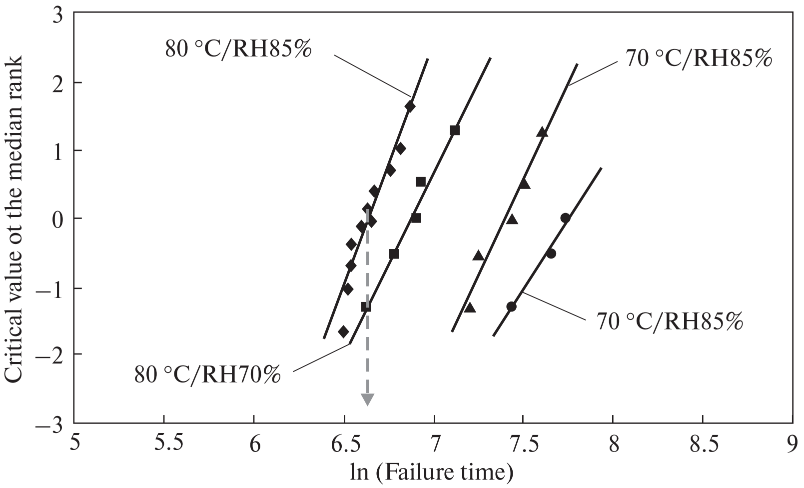


Fig. 6.3. Relation between the limit value of F(t) for the median rank of lifetime data and the natural logarithm of the lifetime data [94]

The failure mode does not change according to acceleration test stress conditions. The relation between the limit value of function F(t) for the median rank of lifetime data obtained for measurement and the natural logarithm of the lifetime data are shown in Fig. 6.3.

The log mean and the log standard deviation at each stress condition were calculated with data from this figure using least-squares regression. The time of the log

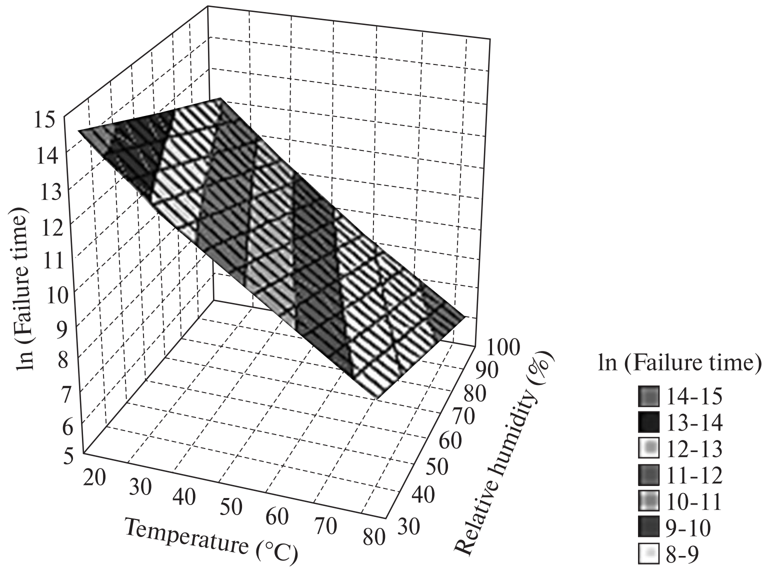


Fig. 6.4. Natural logarithm of failure time (log mean) at temperature range 20 °C-80 °C and RH range 30%-100% [84]

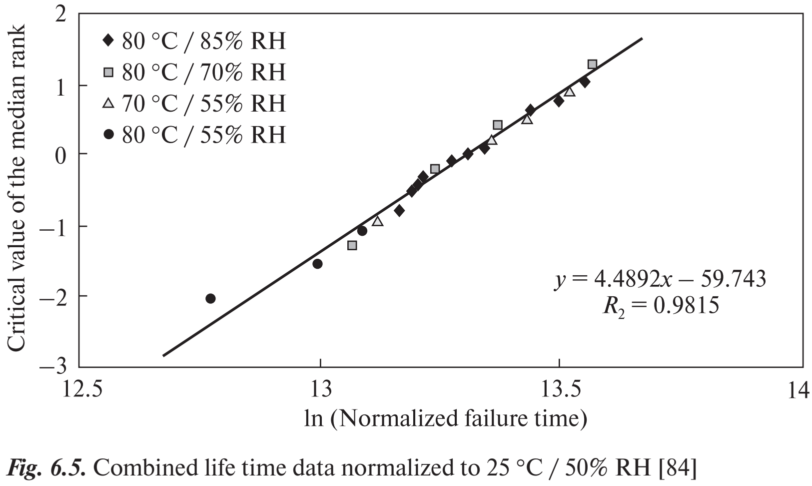


Fig. 6.5. Combined life time data normalized to 25 °C / 50% RH [84]

mean shows the natural logarithm of lifetime data for an optical disk with 50% survival probability. The values of A, B and (Es/k) in the Eyring eq. (3) determined from experimental results. Multiple linear regression analyses made using lifetime data (log mean) and each stress condition of the temperature and relative humidity.

This equation easily calculates the failure time (log mean) at any combination of temperature and relative humidity. For cases within the ranges of temperature 20 ˚C-80 ˚C and relative humidity 30%-100%, the natural logarithms of failure time (log mean) are shown in Fig. 6.4.

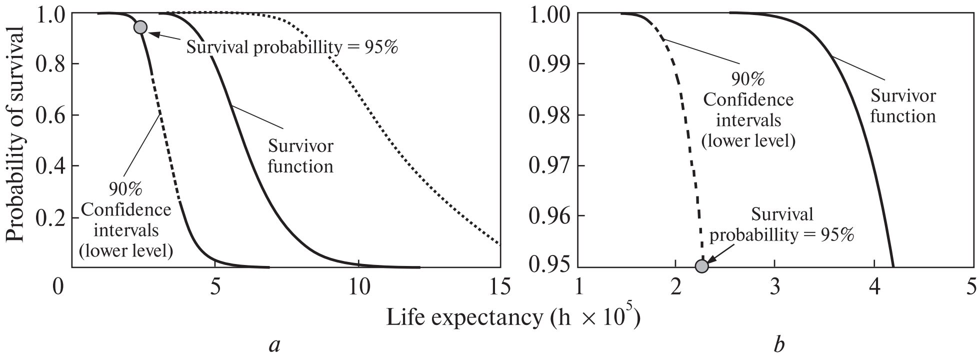


Fig. 6.6. Probability-of-survival function at 25 °C / 50%RH (*a*), expanded scale (*b*)

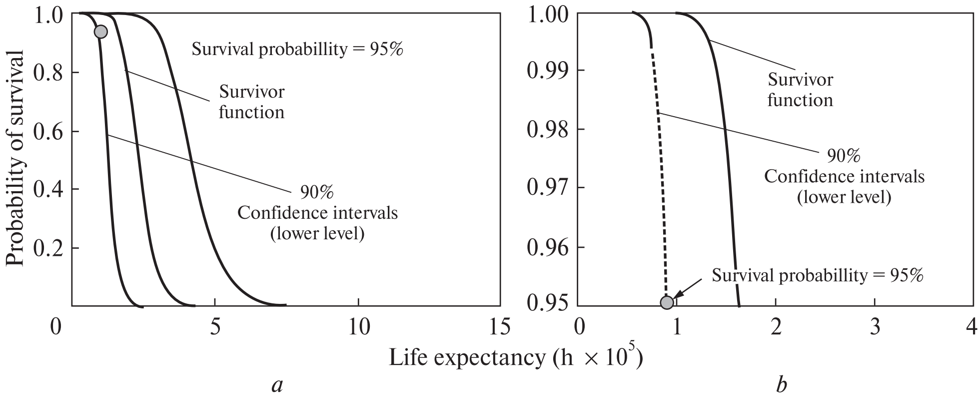


Fig. 6.7. Probability-of-survival function at 30 °C / 60%RH (*a*), expanded scale (*b*)

The acceleration factor was used for normalization of lifetime data at each stress test. The relation between the limit value of the function *F(t)* for the median rank of the lifetime data and natural logarithm of the normalized lifetime data is shown in Fig. 6.5.

All lifetime data in this figure show linearity. The statistical distribution of lifetime data is inferred to be a lognormal distribution. A standardized life expectancy, $t_{50}\% = 612 \times 10^3$ (h) is calculated from the lognormal distribution of all lifetime data in Fig. 6.6.

Fig. 6.6 (a) shows the reliability function (*R(t)* = 1-F(t)) under the 90% confidence interval. The reliability level of 95% defines the life expectancy. Therefore, Fig. 6.6 (*b*) can be obtained finally by graphing the range of the reliability function as 95% or more. The life expectancy by the 95% confidence interval can be presumed as 228×10^3 (h), when reliability (probability of survival) is assumed as 95% at 25°C / 50%RH.

The statistical distribution function of optical disk lifetime data must be clarified experimentally for future application of this technique to lifetime estimation of optical disks [84].

Fig. 6.7 shows the life expectancy for another combination condition of temperature and RH, 30°C / 60%RH.

The Arrhenius method applies to high temperature stress conditions only. The Eyring method applies to multiple simultaneous stress conditions (heat & humidity). Both methods require the definition of a failure point. The application of these test methods to optical discs is standardized under ISO 10995 and ISO 16963 Both ISO standards specify: Multiple Test Conditions Verification of linearity of behavior under test DVD Failure Point is PISum8 = 280. The Arrhenius and Eyring methods are the standard. Other failure modes need to be considered but are not evaluated by these methods. Test results obtained at only one temperature condition do not provide sufficient information to evaluate lifetime under normal storage conditions [123].

Standard accelerated life testing only addresses Data Layer Failure chemical reactions such as oxidation of materials [123]. The only way to know the condition of a digital collection is constant and comprehensive testing. This cannot be stated too strongly; no collection using CD-R DVD-R/+R as an archival carrier should be without a reliable tester. The error correction capability of most replay equipment will mask the effects of degradation until the errors are well into the uncorrectable region. When this point is reached, all subsequent copies are irreversibly flawed. On the other hand, a comprehensive testing regime allows for best possible planning of preservation strategies by acting on the known, objective and measurable parameters that digital archiving make possible. In the well-documented digital archive, metadata will record the history of all objects, including a record of error measurements and any significant corrections [124].

6.2. Methods and systems for the rapid evaluation of optical media reliability

The lifetime of optical storage media is conventionally estimated using an accelerated ageing test method based on the Arrhenius or Eyring methods, which consider the effect of temperature and humidity. In the process of testing is controlled by the number of errors when data is reproduced from optical media. Fig. 6.8 shows the change in the maximum PI sum 8 valuemeasured at 250-hour intervals during the accelerated agingtest. As max PI sum 8 value is generally considered as aprimary factor in determining the playback stability, theaverage values calculated from 20 disks were used in thisstudy. The results show that the average PI sum 8 value increases rapidly after 750 h and reaches the failure pointof over 280 after 1500 h of the accelerated aging test [137].

The following standards and studies are available on this subject:

• ISO Standard 18921:2008 “Imaging materials — Compact discs (CD-ROM) — Method for estimating the life expectancy based on the effects of temperature and relative humidity”;

• ISO-Standard 18927:2008 "Imaging materials — Recordable compact disc systems — Method for estimating the life expectancy based on the effects of temperature and relative humidity";

• standard ECMA-379 (equivalent to ISO/IEC 10995:2008) "Test Method for the Estimation of the Archival Lifetime of Optical Media";

• tests conducted by the National Institute of Standards and Technology in the USA as part of the "Optical Media Longevity Study";

• tests conducted by CDs21 Solutions and the Archive Disc Test Center in Japan [138]

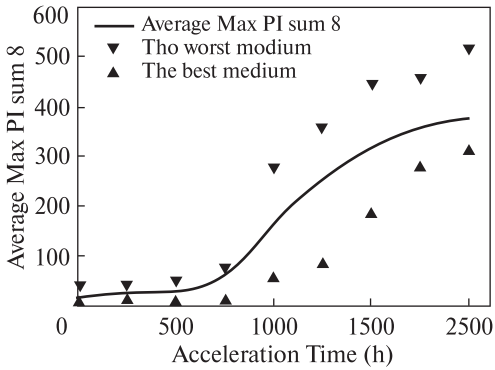


Fig. 6.8. Average max PI sum 8 value with aging time. The PI sum 8 values were averaged from 20 DVD + R test disks [137]

A widely-accepted key signature of the health of the data on a disc is the PI8(max) value.The ISO/IEC-10995 standard6, and later the ECMA-379 standard7, give one measure that is widely accepted as the key parameter for measuring the quality of the written data: PI Sum 8 (PI8). The PI8 statistic is a measure of the number of times there has been a Parity Inner (PI) error in 8 consecutive ECC blocks. The specified maximum allowable value is 280, meaning there must be ≤280 PI errors in any 8 consecutive ECC blocks.

This PI8 parameter and its specified maximum value [PI8(max)], are particularly meaningful in light of the structure of the data on DVDs and the way the error correction coding operates. Accordingly, the data on a recordable DVD is at risk when this parameter exceeds the specified value of 280.

Some DVDs start out with a measured PI8(max) value as low as 20. This would give a degradation headroom value of 260, where degradation headroom = PI8(max — spec value) — PI8(measured value), or 280 — 20. Thus a DVD with a degradation headroom value of 260 is superior to a DVD with a degradation headroom value of 100.

While we do not believe that the PI8(max) parameter gives a full picture of the quality of the data on a recordable DVD, it is the accepted standard, and we do believe that this parameter has merit. Accordingly, one major way of determining the initial quality of the data written on a recordable DVD is to look at the initial value of PI8(max). If this value approaches the spec value of 280, the degradation headroom of the DVD is very poor, and the disc cannot degrade much before the PI8(max) value is exceeded and the data would be at risk [168].

At best, the quality of blank recordable CD and DVD media can be described as

variable. The recordable CD and DVD-manufacturing industry has become a market place driven by narrow profit margins and large quantities. Recordable CD and DVD manufacturing equipment has become smaller, cheaper and more self-contained. As a consequence, the production of reliable data carriers for the quality market has largely been replaced by manufacturers of recordable CD and DVD, producing recordable CD and DVD for the low cost market [164].

To test the durability of this data-storage mechanism used accelerated aging measurements. These tests reveal that the decay time of the nanogratings is $3 \times 1020 \pm \pm 1$ years at room temperature (303K), showing the unprecedentedly high stability of nanostructures imprinted in fused quartz. Even at elevated temperatures of 462K, the extrapolated decay time is comparable to the age of the Universe (13.8 billion years). Based on the tests, we believe that these copies could survive the human race [148]. Best results for optical media will be found in properly cased, dark stored media, at between 4-20C, and 20 to 50 percent relative humidity. Also be careful of the label side, which people think is the tough side, as damage here can also destroy recovery [121].

6.3. Long term data storage application area

Heritage institutions — libraries, archives, and museums — traditionally bear the responsibility of preserving the intellectual and cultural resources produced by all of society. This important mission is now in jeopardy around the world due to the sheer volume of information which is created and shared every day in digital form. Digital technology, in dramatically easing the creation and distribution of content, has generated exponential growth in the production of digital information. The digital universe is doubling in size every two years and will grow tenfold between 2013 and 2020. Preserving this vast output is difficult, not just for its extent, but because much of it is ephemeral. Digital information does not have the same longevity as physical objects, documents, and books, which often will survive for centuries. Digital file formats, storage media, and systems are ever evolving, jeopardizing the future readability and integrity of digital heritage over much shorter timeframes than does the deterioration of paper and physical objects, and its availability for capture is fleeting. The survival of digital heritage is much less assured than its traditional counterparts in our collections. Identification of significant digital heritage and early intervention are essential to ensuring its long-term preservation [164].

The challenge of long-term preservation in the digital age requires a rethinking of how heritage institutions identify significance and assess value. The proliferation and abundance of digital heritage and information, and the ephemeral nature of much of it, means that heritage institutions must be proactive to identify digital heritage and information for long-term preservation before it is lost. Traditional forms of cultural heritage — books, periodicals, government records, private correspondence, personal diaries, maps, photographs, film, sound recordings, artefacts, and works of art, to name a few — now have digital equivalents, which often fit well within our existing practices and mandates. But the digital environment has created new forms of expression, ranging from web pages and interactive social media sites to private research databases and online gaming environments that blur boundaries and lines of responsibility and challenge past approaches to collecting [164].

The long-term preservation of digital heritage is perhaps the most daunting challenge facing heritage institutions today. Developing and implementing selection criteria and collecting policies is the first step to ensuring that vital heritage material is preserved for the benefit of current and future generations.

The amount of digital assets, whether born digital or digitized objects from analog and paper artifacts, is growing rapidly. Unlike companies which are required to retain their records for a relatively short period of time to comply with the Sarbanes-Oxley Act, national archives and digital libraries have to face daunting challenges of long-term preservation. Indeed, in order to fulfill the mission to provide discovery and access to digital assets over a long period of time, institutions must develop strategies and mechanisms to effectively preserving these assets. Besides the volume issue, another complicating aspect of digital preservation is data heterogeneity due to the fact that data might originate from various software products specific to diverse application domains. Moreover, organizations have increased their portfolios to disseminate a wide range of file formats from textual documents, geospatial images, audio visual records, web pages, and database files [131].

The amount of content in our digital archive is increasing rapidly. What are the most important objects or collections to preserve? Should we periodically revisit what is in the archive and deaccession content that is less important? In a world of finite resources we decided that we need to determine our preservation priorities in order to better preserve the important content. The low system complexity, easy availability of technology, and inexpensive hardware and media makes the recordable and rewritable CD and DVD a popular option with archives. The cost of a more reliable data storage system, however, could be less even for quite small holdings, if averaged across the entire collection. There are, however, occasions where a collection manager may still choose to use a discrete optical media such as recordable CD or DVD [164].

Hitachi Digital Preservation Platform is an optical storage solution that uses existing Blu-ray and eventually M-DISC technology to preserve the nation's most valuable resource, its information. HDPP lets federal agencies access their data whenever they want, no matter how long it is stored, at lower total cost of ownership [134].

Medical information is an important resource used for the clinical care of patients, health service planning, monitoring, improvement and medical research. This information is vital to patient safety and wellbeing and must be managed with consideration for its confidentiality and sensitivity. Health creates and collects a vast amount of information, much of which is confidential personal health information. All information needs to be managed, stored and disposed according to its classification, business requirements and retention period. Suitable retention periods, storage conditions and the use of recommended disposal methods will ensure that information is managed, protected and accessible. Electronic information must remain available, accessible, retrievable and usable for as long as a business need exists or as long as legislative, policy and archival provisions and procedures require them to be kept. Electronic information including databases must only be disposed in accordance with an approved Retention and Disposal Authority. Long Term Management of Electronic Records Policy ensures that electronic information of continuing value maintains its functionality and is migrated forward in accordance with the relevant retention and disposal authorities when hardware and software changes [132]. Because patient health records (even old ones) could potentially need to be accessed at any time, simply copying old records off to tape and storing the tapes in a vault simply is not an option. Patient health data needs to be able retrievable within a reasonable amount of time. Not only must patient health data be retrievable, but in many cases it must also be updatable. If a patient who has not been seen in quite some time suddenly comes into the facility requiring treatment, the facility must be able to retrieve the patient's health records in a timely manner and to be able to make updates to those records [155].

The patient information collected by clinicians across the nation constitutes a vast, distributed library of incalculable importance to patients, to relatives, and to public health: the human phenome. Personalized medicine will involve correlations over a lifetime between the patient's genome and phenome. (Created by the interaction of genome with the environment, the phenome consists of all the expressed characteristics of an individual, from microcellular chemistry to behavior; in health care the phenome is more loosely defined as the information gathered about the patient.) Applied genomics is also an essential element of much translational medicine and clinical research, and will inherently require access to clinical information.How long to store depends on the intended use. Because some patients will encounter a clinical problem for whose management old clinical information can be crucial, a case can be made for storing clinical information at least for the lifetime of every patient. The working assumption would be that the cost of storing more information than will actually be useful is outweighed by the value to those who will benefit at time of need. As genomics begins to suffuse clinical medicine, the argument for lifetime storage strengthens because genome/phenome correlations and the effect of environmental factors will probably appear over time in complex ways. Lifetime may not be long enough: information about genomic expression could have implications for blood relatives, including descendants, so that storage of a patient's clinical information beyond death may continue to have value for patient care. The enormous potential of genome/phenome information for biomedical research also pleads for storage of clinical information for generations.Some period between a century and eternity is probably best. Storage for such long periods, much longer than current retention policies, will doubtless require specific provisions for archiving systems in which information is stored separately from the records maintained for quotidian care. The information generated during patient care constitutes a valuable information resource, the phenome, of value to patients, the public, and biomedical science. Preserving clinical information for generations rather than the ephemeral periods now required will entail complex revision or creation de novo of information retention policies for both print and digital data [149].

The EU stipulates that plans for the final disposal of radioactive waste must include the means to preserve knowledge about the facility "for the longer term". The

IAEA states that "long term safety" cannot rely on active institutional control, but passive controls could reduce the risk of intrusion over a "longer time scale than the one envisaged for active controls. Clarification is needed as to whether the meaning of these terms is consistent with the project definition, being the time with no repository oversight. The EU, Switzerland, the UK and the USA all mention the preservation of knowledge. The EU demands that the required concepts or plans for post-closure should include "the means to beemployed to preserve knowledge of that facility in the longer term". The guidance in Switzerlandspecifies that "documentation has to be prepared on long-term securing of knowledge of thegeological repository", and the UK guidance consistently refers to the preservation of "knowledge andadequate records". The UK guidance also provides information on what meta data to preserve. In the USA, the EPA regulations for WIPP require measures that will be used to "preserve knowledge" about the repository [165].

Some the types of the information need high reliability long-term data storage were determined. For example, it is necessary to provide long-term storage of results of tests (acoustic emission data) of complex engineering buildings, the term of exploitation of which makes over one hundred years, medical information, codes of electronic digital signatures, information about activity of enterprises with a high level of regulation and activity on the part of public authorities.

Important digital heritage, including master files with associated metadata, should exist in multiple copies that are stored in at least two different physical locations. Heritage institutions can use a mix of on-site, off-site, and distributed cloud-based storage, but digital originals should be backed-up in at least one other location. Storage sites should be chosen to diminish the risk of loss due to natural or man-made disasters and economic or political crisis

UNESCO assumed a leadership role and decided to establish a global mechanism to safeguard all forms of documentary heritage. It therefore created the Memory of the World Programme with the stated objective of preserving and protecting the world's memory, in the form of its documentary heritage, for the benefit of all. This memory is a shared, common legacy that is important for both current and future generations. Under the Programme, UNESCO implements preservation projects, conducts training workshops and provides technical advice on preservation and access issues as it seeks to build awareness of the contribution of documentary heritage to efficient and accountable management and governance, and to education and sustainable development. To achieve its objectives, the Programme operates at the international, regional and local levels to ensure that no matter the means of recording memory, universal and permanent access to it can be assured [165].

Heritage institutions are traditionally entitled to preserve the intellectual and cultural production of their own nation, whatever their form or their medium may be. This long-standing mission is now challenged by the sheer amount of content which is produced and published every day in digital form, especially on the internet. National web domains range from thousands to millions of websites; on which millions to billions of files are posted, updated or deleted every day. Government administrations and private companies produce an unprecedented mass of digital records. What is not saved today by heritage or other institutions is at risk of being lost in the near future. Digital technologies have indeed dramatically eased the creation and production of original content and reduced their costs.Heritage institutions are thus facing a conundrum: selection is necessary, as it is legally, economically and technically impossible to gather all current digital production. The main goal of memory institutions in the field of digital preservation is to reduce the risk of loss and to retain cultural heritage in digital form and ensure their legibility, interoperability, availability, and authenticity over a long period of time. Preservation of electronic data is done not by one single technique, but by regarding all possible scenarios which may cause loss of the material. An efficient system for long-term storage and archiving should solve the problems which might cause losses or damages [146].

Digital audio archives are not peculiar to a single branch of knowledge. On the contrary, they appear to be a virtual space in which different kinds of expertise convene and deal with unusual, original research questions concerning audio preservation, cataloguing, transcription, analysis, data re-using, and access rights management. Oral historians, linguists, and anthropologists have often underlined the urgent need to protect analogue and born-digital audio archives collected by professional scholars and ordinary people interested in languages, dialects, tradition, popular music, and ethnology. In every respect, audio archives are a precious resource: linguists, anthropologists, ethnographers, oral historians have spent years collecting materials that deserve safeguarding and circulation. However thousands of hours of speech recordings collected for different purposes, despite having been digitally preserved, are still inaccessible to the communities for which they have been produced, not to speak of the wider audience. In most cases, audio archives collected in the humanities and social sciences are still in the hands of the original researchers. It can even be very difficult to get the basic datasets documentation and even more difficult to persuade researchers and private citizens to provide open information about their data. Crucially, the UNESCO Convention for the Safeguarding of the Intangible Cultural Heritage, Article 2 defines this material as belonging to Intangible Cultural Heritage domains, which include:

- oral traditions and expressions, including language as a vehicle of intangible cultural heritage;
- performing arts;
- social practices, rituals and festive events;
- knowledge and practices concerning nature and the universe;
- traditional craftsmanship [155].

Moving image and sound content is at great risk. 74 per cent of professional collections are small: 5,000 hours or less. Such collections have a huge challenge if their holdings are to be digitized. About 85 per cent of sound and moving image content is still analogue, and in 2005 almost 100 per cent was still on shelves rather than being in files on mass storage. Finally, there is a major problem of material that is scattered, unidentified, undocumented and not under any form of preservation plan [256].

6.4. Conclusions

1. The lifetime of optical storage media is conventionally estimated using an accelerated ageing test method based on the Arrhenius or Eyring methods, which consider the effect of temperature and humidity.

2. Lifetime estimation method for optical disks is based on the Eyring and the Arrhenius acceleration model and statistical analysis. The statistical distribution of life data for the optical disk assumed a lognormal distribution. In the process of testing is controlled by the number of errors when data is reproduced from optical media.

3. Important digital heritage, including master files with associated metadata, should exist in multiple copies that are stored in at least two different physical locations. Heritage institutions can use a mix of on-site, off-site, and distributed cloud-based storage, but digital originals should be backed-up in at least one other location. Storage sites should be chosen to diminish the risk of loss due to natural or man-made disasters and economic or political crisis.

SUMMARY

Keeping information for a long time has always been a challenge. Deciding how to create a long-term archive involves choosing the right storage system with the right technology under the proper environmental conditions. Determining how long something will last has long been an important area of study for science and technology, particularly in materials science.

Many advances have been made, and much is known today about how to reliably predict the life expectancy of a product, based on the materials used to make it and the conditions of its use. These advances are readily applied to the field of data storage. The most common failure mechanisms for materials include oxidation, corrosion, breaking of chemical bonds and mechanical wear. Each of these failure mechanisms is exacerbated by elevated temperature, humidity, and exposure to light (including UV). That is the reason that any controlled environment that is intended for archival storage always includes controlled temperature, humidity and light. There are three basic technologies available for storing digital data: magnetic (magnetic tape and hard-disk drives), solid-state (flash memory), and optical (CD, DVD and BD). Each of these technologies uses wellknown processes and materials to manufacture the storage media, and each of these technologies has known failure mechanisms, which have been studied. Creation of manufacturing technology of long-term optical media is of great interest. Data on stamped optical discs is recorded at the time the discs are manufactured, and cannot be altered by the optical disc drives. They are nearly impervious to change, except by extreme conditions. Actual problem is a creation of methods of information survivability assessment on high-stable monocrystal substrates discs. Realization of potential opportunities of such optical discs was difficult due to significant distortions of laser beam during its passing through sapphire substrate. It was shown that the problem of data reading through a substrate of negative single crystal sapphire can be solved by using for reading a special optical system with a plate of positive single crystal materials. The experimental results confirm the efficiency of the proposed technical solution.

ACRONYMS

BD	Blu-ray disc
CD	Compact disc
CD-R	Compact disc-recordable
CD-ROM	Compact disc read-only memory
CD-RW	Compact disc-rewritable
COM	Computer output microfilmers
DPN	Digital preservation network
DNA	Deoxyribonucleic acid
DVD	Digital versatile disc
IDC	International data corporation
IoT	Internet of things
IFLA	International federation of library associations
HDD	Hard disk drive
MAID	Massive array of idle disks
MPEG	Moving picture experts group
NAS	Network attached storage
LE	Life expectancy
LTO	Linear tape open
LTFS	Linear tape file system
MTBF	Mean times between failure
NAND	Non-volatile storage technology
OAIS	Open archival information systems
QR	Quick response code
ODF	Open document format
PDF	Portable document format
SD	Secure digital memory card
SNIA	Storage networking industry association
SSD	Strong state drives
TIFF	Tagged image file format
UDO	Ultra density optical

REFERENCES

1. Voos, H. (1972). *Handbook of Data Processing for Libraries (Book Review)*. College & Research Libraries, 33(1), 55-56.
2. Grigorescu, A., Cerchia, A. E., & Jeflea, F. V. (2016). *Modern technologies for data storage, organization and managing in CRM systems*. Proceedings of the ICA, June 2016. IIT.
3. McMillan, G., Schultz, M., & Skinner, K. (2011). *Digital preservation*. Washington, D.C.: Association of Research Libraries.
4. Adkins, Lesley; Adkins, R. A. (2000). *The keys of Egypt: the obsession to decipher Egyptian hieroglyphs*. HarperCollins.
5. Wayleith, P. R. (2008). *Data security: laws and safeguards*. New York: Nova Science.
6. Yaranal, M. A., & R. (2015). *Comprehensiveness of Digital Information in Digital Age — A Survey on Research Trends in National Law School of India University*. SRELS Journal of Information Management, 52(2), 119.
7. Lee, G. (2014). *Storage Networks*. Cloud Networking, 139-161.
8. *Permanent Digital Data Storage: A Materials Approach*. (n.d.). Retrieved October 5, 2017, from http://www.bing.com.
9. Scaling the Facebook data warehouse to 300 PB. (n.d.). Retrieved October 05, 2017, from https://code.facebook.com.
10. Graham, W. (2012). *Introducing the Facebook Platform*. Beginning Facebook Game Apps Development, 171-200.
11. Song, J., & He, B. (2013). *NAND FLASH based three-tier fault-tolerant storage strategy*. IEEE Conference Anthology.
12. Graham, W. (2014). *The Cultural Imperative*. Big Data, Big Innovation, 47-78.
13. Lueck, T. (2014). *Internet Archive: Digital Library of Free Books, Movies, Music, and Wayback Machine*. The Internet Archive Companion. American Journalism, 31(2), 299-301.
14. Lueck, T. (2015). *Long-Term Preservation of Digital Documents*. (n.d.). Long-Term Preservation of Digital Documents, 3-20.
15. *Sony and Panasonic announce the Archival Disc: is it the answer to the long-term digital storage conundrum?* (n.d.). Retrieved October 05, 2017, from http://www.imaging-resource.com.
16. *Millenniata and LG to Release M-Disc Archival DVD Burner*. (n.d.). Retrieved October 05, 2017, from http://blog.cdrom2go.com.
17. *Hitachi Storage Performance Monitoring*. (n.d.). Retrieved October 05, 2017, from http://www.solarwinds.com.
18. *Hard Disk in Andhra Pradesh*. (n.d.). Retrieved October 05, 2017, from https://www.olx.in/andhrapradesh/q-hard-disk.

19. ArXiv, E. T. (2013, October 31). *Million-Year Data Storage Disk Unveiled*. Retrieved October 05, 2017, from https://www.technologyreview.com.
20. *Eternal 5D data storage could record the history of humankind*. (n.d.). Retrieved October 05, 2017, from https://www.southampton.ac.uk.
21. Bancroft, C. (2001). *Long-Term Storage of Information in DNA*. Science, 293 (5536), 1763-1765.
22. Wilson, A. (2010). *How Much Is Enough: Metadata for Preserving Digital Data*. Journal of Library Metadata, 10(2-3), 205-217.
23. G.A., Kartha (2015). *Study on the Environmental Stability of Nickel Ion Doped Photopolymer Material for Data Storage Applications*. Washington, D.C.: Association of Research Libraries.
24. Schweizer, T., Rosenthaler, L., & Fornaro, P. (2017). *Content-based Interoperability: Beyond Technical Specifications of Interfaces*. Archiving Conference, 2017 (1), 34-38.
25. Coughlin, T. M. (2008). *Fundamentals of Optical Storage*. Digital Storage in Consumer Electronics, 53-71.
26. Minor, D., & Kozbial, A. (2013). *The Chronopolis digital network: the economics of long-term preservation*. A Handbook of Digital Library Economics, 115-123.
27. *Amazon Glacier — Cloud Archive*. (n.d.). Retrieved October 05, 2017, from https://aws.amazon.com/glacier.
28. *Backblaze Online Backup*. (n.d.). Retrieved October 05, 2017, from https://www.backblaze.com.
29. *Google Drive — Cloud Storage & File Backup for Photos, Docs & More*. (n.d.). Retrieved October 05, 2017, from https://www.google.com/drive.
30. *Welcome to Microsoft OneDrive*. (n.d.). Retrieved October 05, 2017, from https://onedrive.live.com/about/en-us/plans.
31. *National Archives of the Netherlands*. (n.d.). Retrieved October 07, 2017, from http://en.nationaalarchief.nl.
32. *Open Yearbook*. (n.d.). Retrieved October 07, 2017, from http://uia.org/s/or/en/1100014774.
33. Portree, D. S. *Viking on the Moons of Mars*. Retrieved October 07, 2017, from https://www.wired.com/2013/04/viking-on-the-moons-of-mars-1972.
34. Gutenberg, P. (n.d.). *Digital dark age*. Retrieved October 07, 2017, from http://self.gutenberg.org/articles/eng/Digital_Dark_Age.
35. Mann, M. (n.d.). *MarexMG Project*. Retrieved October 07, 2017, from http://www.marexmg.org/ fileshtml/howtoSSTV.html.
36. Borghi, M., & Karapapa, S. (2013). *Copyright and Mass Digitization*. Oxford: OUP Oxford.
37. Lenk, B. (2012). *QR-Code*. Kirchheim unter Teck: Lenk Fachbuchverl.
38. Roebuck, K. (2011). *QR Code: High-impact strategies — what you need to know: definitions, adoptions, impact, benefits, maturity, vendors*. Dayboro: Emereo Pty Limited.
39. Honeyson J. (2011). *Shotcode*. Retrieved October 07, 2017, from https://www.slideshare.net/honey211/shotcode.
40. *DATABASES: Library of Congress E-Resources Online Catalog*. (n.d.). Retrieved October 07, 2017, from http://eresources.loc.gov.
41. Zomaya, A. Y., Sakr, S., & Sahni, S. (2017). *Handbook of big data technologies*. Cham, Switzerland: Springer.
42. *Preserving the scholarly record historical record cultural record for this and future generations*. (n.d.). Retrieved October 07, 2017, from http://dpn.org.
43. Bailey, C. W. (2006). *Institutional repositories*. Washington, DC: Association of Research Libraries, Office of Management Services.
44. Jeflea, F. V. (1999). *Information-based access to storage: the foundation of information systems*. Proceedings of the IEEE, June 1999, IEEE Computer Society Press.
45. *Infotoday Conferences*. (n.d.). Retrieved October 07, 2017, from http://www.infotoday.com/cilmag/sep12/Huwe--Digital-Migration-Strategies-Old-and-New.shtml.

46. *Information Systems: Definitions and Components*. (n.d.). Retrieved October 7, 2017, from http://www.bing.com.
47. Yang, J. (2015). *Summary Report of ISO/IEC 10995 Test Program. RITEK: Global Home*. Retrieved October 17, 2015, from http://www.ritek.com/m-disc/eng/download/001.pdf.
48. Vries, J., Schellenberg, D., Abelmann, L. (2013). *Towards Gigayear Storage Using a Silicon-Nitride/Tungsten Based Medium*. ArXiv. Retrieved May 11, 2013, from http://arxiv.org/ abs/1310.2961.
49. Clery, D. (2012) *A Million-Year Hard Disk. Science*. Retrieved May 21, 2012, from: www.sciencemag.org/ news/2012/07/million-year-hard-disk.
50. Petrov, V.V., Krychyn, A.A., Shanoylo, S.M,, etc. (2009). *High-density optical information recording*. NAS of Ukraine, Institute for information recording, Kyiv: NASU, 282.
51. Office, U. G. (2016, May 25). *Information Technology: Federal Agencies Need to Address Aging Legacy Systems*. Retrieved October 07, 2017, from http://www.gao.gov/products/GAO-16-696T.
52. *Self-assembling polymer arrays improve data storage potential*. (n.d.). Retrieved October 07, 2017, from http://news.wisc.edu/self-assembling-polymer-arrays-improve-data-storage-potential.
53. Domingo, J. S. (2016, January 21). *Seagate Backup Plus Portable Drive*. Retrieved October 07, 2017, from https://www.pcmag.com/article2/0,2817, 2497199,00.asp.
54. *MTBF and power supply reliability*. (n.d.). Retrieved October 07, 2017, from https://www.electronicproducts.com/Power_Products/AC_DC_Power_Supplies/MTBF_and_power_supply_reliability.aspx.
55. *S.M.A.R.T. technology*. (n.d.). Retrieved October 07, 2017, from http://www.ariolic.com/activesmart/smart-technology.html.
56. *IBM Tape Library Slot and Capacity Calculator 2017 v5.5*. (n.d.). Retrieved October 07, 2017, from https://www.ibm.com.
57. *Home Altosystems*. (n.d.). Retrieved October 07, 2017, from https://www.altosystems.com.
58. *Magnetic tape*. (n.d.). Retrieved October 07, 2017, from https://www.alibaba.com/showroom/magnetic-tape.html.
59. *What is magnetic tape used for?* (n.d.). Retrieved October 07, 2017, from https://www.reference.com.
60. *The Role of Future Magnetic Tape Technology for Digital*. (n.d.). Retrieved October 7, 2017, from http://www.bing.com.
61. *LTO Archiving Solutions & LTFS Asset Managment*. (n.d.). Retrieved October 07, 2017, from http://1beyond.com/archiving.
62. *LTO-6 SAS External Tape Drive*. (n.d.). Retrieved October 07, 2017, from https://www.bhphotovideo.com.
63. *IBM Tape Library Guide for Open Systems*. (2016, September 30). Retrieved October 07, 2017, from https://www.redbooks.ibm.com.
64. Office, U. G. (2000) *Imaging materials — Polyester-base magnetic tape — Storage practices: ISO 18923*. Genf: ISO.
65. *Oracle StorageTek SL8500 Modular Library System*. (n.d.). Retrieved May 17, 2017, from https://www.oracle.com/storage/tape-storage/sl8500-modular-library-system/index.html.
66. Hidaka, H. (2017). *Embedded flash memory for mcu/soc applications*. S.l.: Springer international Pu.
67. Liu, D. (2012). *Write-activity-aware NAND flash memory management for PCM-based embedded systems*. Hong Kong: Dept. of Computing, The Hong Kong Polytechnic University.
68. Hidaka, H. (2009) *SSD: Solid State Drive 2010*. Tōkyō: Nikkei BP Sha.
69. Park, K. (2016). *Data processing using flash storage: some opportunities and limitations*. Ann Arbor, MI: ProQuest LLC.
70. Sand S. (n.d.). *A better computing experience*. Retrieved October 27, 2016, from https://www.sandisk.com/home/ssd.

71. Servers, S. (2013, August 15). *Facebook to rely on cold flash technology for cold data storage needs*. Retrieved October 07, 2017, from https://storageservers.wordpress.com/2013/08/15/facebook-to-rely-on-cold-flash-technology -for-cold-data.
72. DAlleyrand, M. R. (1989). *Image storage and retrieval systems: a new approach to records management*. New York: Intertext Publications.
73. *What is Microfilm?* (n.d.). Retrieved October 07, 2017, from http://www.microfilmworld.com/whatismicrofilm.aspx
74. *Computer-output microfilming*. (2007, January 01). Retrieved October 07, 2017, from https://link.springer.com/referenceworkentry/10.1007/1-4020-0613-6_3393.
75. *Digital Optical Tape: Technology and Standardization* (n.d.). Retrieved October 7, 2017, from https://www.bing.com.
76. *PA.Gov*. (n.d.). Retrieved October 07, 2017, from http://www.phmc.pa.gov.
77. *Systems Archives — Laser Photonics*. (n.d.). Retrieved October 7, 2017, from https://www.bing.com
78. Yaniv Erlich, Dina Zielinski. (Sep. 9, 2016) *Capacity-approaching DNA storage*. BioRxiv preprint: Association of Research Libraries..
79. Extance A. (2016) *How DNA could store all the world's data*. Nature, 7618 v., 537 p..
80. Bornhol, J., Ceze, L. (2016) *A DNA-Based Archival Storage System*. Proceedings of the Asplos, June 2016. ASPQ.
81. Lopez, R , Carmean D. *DNA Data Storage*. Proceedings of the Asplos, June 2016. ASPQ.
82. Clelland, T., Risca, V. and Bancroft, C. *Hiding messages in DNA microdots*. Nature, 399:533-534, 1999.
83. Church, G., Gao, Y. and Kosuri, S. *Next-generation digital information storage in DNA*. Science, 337(6102):1628, 2012.
84. Goldman, N., Bertone, P., Chen, S. *Towards practical, high-capacity, low- maintenance information storage in synthesized DNA*. Nature, 494:77-80, 2013.
85. Liang, P., Yongshik, P., Yi, X. (2011). *Maskless Plasmonic Lithography at 22 nm Resolution. Scientific Reports*. Retrieved October 25, 2011, from: www.nature.com/articles/srep00175.
86. Inoue, M., Kosuda, A., Mishima, K., Ushida, T., Kikukawa, T. (2010) *512 Gb recording on 16 layer optical disc with Blu-Ray Disk based optics*. Proc. SPIE, 7730, D1-D6.
87. Gu, M., Li, X., Cao, Ya. (2014). *Optical storage arrays: a perspective for future big data storage*. Light: Science & Applications, 3, 71-77.
88. Gu, M. (2013). *Optical data storage with diffraction-unlimited resolution lasers and electro-optics Europe*. Conference on and International Quantum Electronics Conference. Munich (Germany), 93-99.
89. Gu, M, Li, X, Lan, Th., Tien, Ch. (2012). *Plasmonic keys for ultra-secure information encryption. SPIE-Newsroom*. Retrieved October 3, 2012, from: http://spie.org/newsroom/4538-plasmonic-keys-for-ultra-secure-information-encryption.
90. Kudryavtsev, A.A., Moskalenko, N.L. (2013). *Is there any future of optical discs?* Semiconductor Physics, Quantum Electronics & Optoelectronics, 16(4), 362-365.
91. Nikles, D. E., Wiest, J. M. (1999). *Accelerated aging studies and the prediction of the archival lifetime of optical disc media,* Proc. SPIE, 3806, 30-36.
92. Petrov, V.V., Kryuchin, A.A., Gorbov, I.V., Kossko, I.O., Kostyukevych, S.O. (2009) *Analysis of properties of optical carriers after long-term storage*. Semiconductor Physics, Quantum Electronics and Optoelectronics 12(4), 399-402.
93. *Everspan*. (n.d.). Retrieved October 07, 2017, from http://www.everspan.com.
94. Kryuchyn, A. A., Petrov, V. V., Shanoilo, S. M., Lapchuk, A. S., & Morozov, Y. M. (2014). *Sapphire optical discs for long term data storage*. Optical Data Storage 3(2), 25-32 .
95. Petrov, V.V., Semynozhenko, V.P., Puzikov, V.M., Kryuchyn, A.A., Lapchuk, A.S., Shanoilo, S.M., Morozov, Ye.M., Kosyak, I.V., Borodin, Yu.O., Gorbov I.V. (2014). *Readout optical system of sapphire disks intended for long-term data storage*. arXiv, 1403.3119, 10.

96. Dobrovinskaya, E.R., Lytvynov, L.A., Pishchik, V., (2009). *Sapphire: Material, Manufacturing, Applications*. Retrieved May 5, 2009, from: http://www.springer.com/us/book/9780387856940.
97. Petrov, V.V., Semynozhenko, V.P., Puzikov, V.M., Kryuchyn, A.A., Lapchuk, A.S., Morozov, Ye.M., Borodin, Yu.O., Shyhovets, O.V., Shanoylo, S.M. (2014). *Method of aberration compensation in sapphire optical disks for the long term data storage*. Functional Materials, 21(1), 105-111.
98. *Archive Appliance Libraries*. (n.d.). Retrieved October 07, 2017, from http://www.plasmon.com/products/udo.
99. Gan, Z., Cao, Ya., Evans, R.A., Gu, M. (2011) *Three-dimensional deep sub-diffraction optical beam lithography with 9 nm feature size*. Retrieved October 11, 2011 from: http://www.nature.com.
100. LaFratta, C.N., Fourkas, J.T., Baldacehini, T. (2007) *Multiphoton Fabrication*. Angewandte Chemie, 46 (33), 6201-6379.
101. Cumpston, B.H. (1999). *Two-photon polymerization initiators for threedimensional optical data storage and microfabrication*. Nature, 398, 51-54.
102. Xue, J., Zhao, Y., Wu, J., Wu, F. (2008). *Novel benzylidene cyclopentanone dyes for two-photon photopolymerization. J. Photochem. Photobiol.* A Chem., 195, 261-266.
103. Cao, Y., Gan, Z., Jia, B., Evans, R. A., Gu, M. (2011). *High-photosensitive resin for super-resolution direct-laser-writing based on photoinhibited polymerization*. Optics Express, 19 (20), 19486-19494.
104. Dodson B. (2013*). New technique would allow a petabyte of data on a single disc*. *Gizmag*. Retrieved May 23, 2013 from: http://www.gizmag.com/petabyte-dvd-data-storage/28181.
105. Jinga, P.E., Andronescu, S. (2013). *2 nm Quantum Optical Lithography*. Optics Communications, 291, 259–263.
106. Stocker M.P.,, Li, L., Gattass, R.R., Fourkas, J.T. (2011). *Multiphoton photoresists giving nanoscale resolution that is inversely dependent on exposure time*. *Nat Chem.*, 3 (3), 223 227.
107. Li, L., Gattass, R.R., Gershgoren, E., Hwang, H., Fourkas, J.T. (2009). *Achieving λ/20 Resolution by One-Color Initiation and Deactivation of Polymerization*. Science 324 (15) 910-913.
108. Li, L., Gattass, R.R., Fourkas, J. (2009). *Dual-beam, 3D photolithography provides exceptional resolution*. SPIE Newsroom. Retrieved October 21, 2009 from: http://spie.org/newsroom/1690-dual-beam-3d-photolithography-provides-exceptional-resolution.
109. Li, L., Gattass, R.R., Gershgoren, E., Fourkas, J.T. (2009). *Achieving Resolution Far beyond the Diffraction Limit with RAPID Photolithography*. Retrieved October 19, 2009 from: http://spie.org/newsroom/1690-dual-beam-3d-photolithography-provides-exceptional-resolution.
110. Stocker, M.P., Li, L., Gattass, R.R., Fourkas, J.T. (2011). *Multiphoton photoresists giving nanoscale resolution that is inversely dependent on exposure time*. Nature Chemistry, 3, 223-227.
111. Fourkas, J.T. (2010). *Nanoscale photolithography with visible light*. J. Phys. Chem. Lett., 1(8), 1221-1227.
112. Li, L.J., Gattass, R.R., Gershgoren, E., Hwang, H., Fourkas, J.T. (2009). *Achieving lambda/20 resolution by one-color initiation and deactivation of polymerization*. Science, 324 (5929), 910-913.
113. Forman, D.L., Cole, M.C., McLeod, R.R. (2013). *Radical diffusion limits to photoinhibited superresolution lithography*. Phys. Chem. Chem. Phys., 15 (36), 14862-14867.
114. Scott, T.F., Kowalski, B.A., Sullivan, A.C., Bowman, C.N., McLeod, R.R. (2009). *Two-Color Single-Photon Photoinitiation and Photoinhibition for Subdiffraction Photolithography*. Science, 324 (5929), 913-917.
115. Tsuujioka T., Kume M., Horikawa Y., Ishikawa A., Irie M. (1997). *Super-resolution disk with a photochromic mask layer*. Jpn. J. Appl. Phys. 36 (1/1B), 526-529.
116. Tsuujioka T., Kume M., Irie M. (1997). *Theoretical analysis of super-resolution optical disk mastering using a photoreactive dye mask layer*. Opt. Rev., 4(3), 385-389.

117. Chen, Q., Tominaga, J., Men, L., Fukaya, T., Atoda N. (2001). *Superresolution optical disk with a thermoreversible organic thin film*. Optics Letters, 26 (5), 274-279.
118. Andrew, T.L., Tsai, H.Y., Menon, R. (2009) *Confining light to deep sub-wavelength dimensions to enable optical nanopatterning*. Science, 324 (5929), 917-921.
119. Masid, F., Andrew, T.L., Menon, R. (2013) *Optical patterning of features with spacing below the far-field diffraction limit using absorbance modulation*. Optics Express, 21(4), 5209-5214.
120. Ma, X., Wei, J. (2011). *Nanoscale lithography with visible light: optical nonlinear saturable absorption effect induced nanobump pattern structures*. Shanghai Institute of Optics and Fine Mechanics, 3, 1489-1492.
121. Zhang , Ch., Wang, K., Bai, J., Wang, Sh., Zhao, W., Yang F., Gu, Ch., Wang, G. (2013). *Nanopillar array with a $\lambda/11$ diameter fabricated by a kind of visible CW laser direct lithography system*. Nanoscale Res Lett., 8 (1), 280-285.
122. Coufal, H., Burr, G.W., Sincerbox, G.T. (2004). *Handbook of Lasers and Optics*. Springer-Verlag: New York.
123. Pham V.T., Lee S.K., Trinh M.T., Lim K.S., Hamilton D.S. (2006). *Nonvolatile two-color holographic recording in Tm-doped near stoichiometric $LiNbO_3$*. Korean Phys. Soc. 49, 533.
124. Петров, В.В., Крючин, А.А., Токарь, А.П. (1992). *Оптико-механические запоминающие устройства*. Киев.: Наукова думка.
125. Milster, T., Upton, R.S., Luo, H. (1999). *Objective lens design for multiple-layer optical data storage*. Opt. Eng., 38, 295-299.
126. Eichler, H.J., Kuemmel, P., Orlic, S., Wappelt, A. (1998). *Resolution-limited optical recording in 3D*. IEEE J. Sel. Top. Quantum Electron, 19 (17), 16096-16105.
127. Yan, A., Tao, Sh., Wang, D. (2005). *Multiplexing holograms in the photopolymer with equal diffraction efficiency*. Proc. of SPIE, 5643, 109-117.
128. Curtis, K., Psaltis, D. (1992). *Recording of multiple holograms in photopolymer films*. Appl. Opt., 31, 7425-7428.
129. Pu, A., Psaltis, D. (1996). *High-density recording in photopolymer-based holographic three-dimensional disks*. Appl. Opt., 35, 2389-2397.
130. Pu, A., Curtis, K. (1996). *Exposure schedule for multiplexing holograms in photopolymer films*. Opt. Eng., 35, 2824-2829.
131. Dhar L., Curtis K. (1998) *Holographic storage of multiple high-capacity digital data pages in thick photopolymer systems*. Opt. Lett., 23, 1710-1712.
132. Van De Nes, A.S. (2006). *High-density optical data storage. Reports on Progress in Physics,* 69, 53-63.
133. Walker, E., Rentzepis, P.M. (2008). *Two-photon technology: A new dimension. Nat. Photonics,* 2, 406-408.
134. Zhang, Yu., Dvornikov, A.S., Walker, E.P., Kim, N.H., McCormick, F.B. (2000) *Single Beam Two-Photon-Recorded Monolithic Multi-Layer Optical Disks*. Optical Data Storage. Proceedings of SPIE, 4090, 174-178.
135. Orlic, S. (2001). *Microholographic storage in photopolymers*. J. Opt. A: Pure Appl. Opt., 3 (72), 10-18.
136. Kikukawa, T., Inoue, M., Mishima, K., Ushida, T. (2010). *Recording characteristics of 10-layers recodable optical disc and a prospect for over 500G-byte recording*. Jpn. J. Appl. Phys., 49, 10-20.
137. Shirashi, J., Kobayashi, S., Miyashita, H., Hino, H. (2009). *New Signal Quality Evaluation Method for GB/Layer BDs*. ISOM 2009 Tech. Dig., 9, 74-75.
138. Tanaka, H., Takahashi, K., Ogasawara, M., Taniguchi S. (2013). *Advanced Radial Position Control of a Recording Beam for Super Multilayer Disc with Separated Guide Layer*. Japanese Journal of Applied Physics, 52 (9S2), 13-20.
139. Lapchuk, A.S., Kryuchin, A.A., Klimenko, V.A., Kolesnikov, M.U., Petrov, V.V. (1996) *Diffraction of Gaussian laser beam by three-dimensional grating of dielectric spheres*. Proc. SPIE, 3055, 160-169.

140. Shylo, S.A., Lapchuk, A.S., Song, J.S., Kim, K.S. (2005). *Optical Parameters of Light Beam in Multilayer Nano-Structures.* J. of the Korean Physical Society, 47, 18-22.
141. Glushko, B.A., Levich, E.B. (2008). *Fluorescent optical memory.* USA patent. G11B 007/24. № 6071671; published 10.02.2008.
142. Wang, M., Esener, S. (2008). *Three-dimensional optical data storage in fluorescent dye-doped photopolymer.* Appl Opt. 2000 Apr 10, 39(11), 1826-34.
143. Зубарева, Т.С. (2014) *Флуоресцентная микроскопия полного внутреннего отражения. TIRF – микроскопия.* Retrieved January 31, 2014, from: http://www.stormoff.ru/articles_565_101.html.
144. Vasara, A., Taghizadeh, M.R., Turunen, J., Westerholm, J., Noponen, E., Ichikawa, H., Miller, J.M., Jaakkola, T., Kuisma, S. (1992). *Binary surface-relief gratings for array illumination in digital optics.* Applied Optics, 31 (17), 3320-3336.
145. Soifer, V.A. (2002). *Methods for Computer Design of Diffractive Optical Elements.* Retrieved January 23, 2002, from http://eu.wiley.com/WileyCDA/ WileyTitle/ productCd-0471095338.html.
146. Korolkov, V.P., Nasyrov, R.K., Shimansky, R.V. (2006). *Zone-boundary optimization for direct laser writing of continu¬ous-relief diffractive optical elements.* Appl. Opt., 45 (1), 53 62.
147. Yan, A., Tao, Sh., Wang, D., Shi, M., Wu, M. (2005). *Multiplexing holograms in the photopolymer with equal diffraction efficiency.* Proc. SPIE, Advances in Optical Data Storage Technology, 5643, 10-19.
148. Nam, K. (2005). *Holographic applications based on photopolymer materials. International Workshop on Photonics and Applications.* Proceedings of the ICA, June 2005. Hanoi.
149. Sun, H.-B., Kawatal, S. (2004). *Two-Photon Photopolymerization and 3D Lithographic Microfabrication.* APS, 170, 169-273.
150. Хонина, С.Н., Волотовский, С.Г. (2009) *Фраксикон – дифракционный оптический элемент с конической фокальной областью.* Компьютерная оптика, 33(4), 401-411.
151. Хонина, С.Н., Волотовский, С.Г. (2010) *Исследование применения аксиконов в высокоапертурной фокусирующей системе.* Компьютерная оптика, 34(1), 35-51.
152. Yang, A.A., Wrigley, Ch.Y., Lindmayer, J. (1996). *Optical storage medium utilizing electron trapping film layers sandwiched with electrodes.* USA patent. G11C13/04. № US5502706 A; published 26.03.1996.
153. Akselrod, M. (2010). *Aluminum oxide material for optical data storage.* USA patent. G11B7/243. № US 6846434 B2. Published 15.12.2010.
154. Goldsmith, P., Lindmayer, J., Wrigley, C. (1990). *Electron trapping. A new approach to rewritable optical data storage.* Proceedings of SPIE, 1316, 312-320.
155. Zhang, Yu., Dvornikov, A.S., Walker, E.P., Kim, N.H., McCormick, F.B. (2000) *Single Beam Two-Photon-Recorded Monolithic Multi-Layer Optical Disks.* Proceedings of SPIE, 4090, 174-178.
156. Zhang, Yu., Milster, T.D., Butz, J., Bletcher, W., Erwin, K.J., Walker, E. (2002) *Signal, Cross Talk and Signal to Noise Ratio in Bit-Wise Volumetric Optical Data Storage.* IEEE Catalog, 02EX552, 246-248.
157. Lindmayer, J., Goldsmith, P., Wrigley, C. (1989). *Electronic Optical-Storage Technology Approaches Development Phase.* Laser Focus World, 11, 109-119.
158. Yang, X., Wrigley, Ch.Y., Lindmayer, J. (1993). *Three-dimensional optical memory based on transparent electron-trapping thin films.* Proc. SPIE, 1773, 41-43.
159. Yang, X., Wrigley, Ch.Y., Lindmayer, J. (1993). *Three-dimensional optical storage system based on electron-trapping thin films.* SPIE, 560, 60-66.
160. Driggers R.G. (2000). *Encyclopedia of Optical Engineering* Retrieved October 5, 2000, from: www.crcpress.com/Encyclopedia-of-Optical-Engineering-Print/Driggers-Hoffman-Driggers/p/book/9780824709402

161. Beliak, Ie.V., Kravets, V.G., Kryuchyn, A.A. (2007). *Luminescence of the pyrazoline dye in nanostructured zeolite matrix*. Semiconductor Physics, Quantum Electronics & Optoelectronics, 10 (1), 33-35.
162. Beliak, I., & Butenko, L. (2011). *Development of fluorescent multilayer disc structure*. 22nd Congress of the International Commission for Optics: Light for the Development of the World..
163. Gu, M., Li, X. (2010). *The road to multi-dimensional bit-by-bit optical data storage*. Opt. Photon. News, 21, 29-33.
164. Li, X., Lan T.-H., Tien, Ch.-H., Gu, M. (2012). *Three-dimensional orientation-unlimited polarisation encryption by a single optically-configured vectorial beam*. Nature Communications, 998, 10-18.
165. Chang, S.S., Shih, C.W., Chen, C.D., Lai, W.C., Wang, C.R.C. (1999).*The shape transition of gold nanorods*. Langmuir, 15, 701-709.
166. Link, S., Burda, C., Nikoobakht, B., El-Sayed, M. A. (2000). *Laser-induced shape changes of colloidal gold nanorods using femtosecond and nanosecond laser pulses*. J. Phys. Chem., B104, 6152-6163.
167. Link, S., Burda, C., Mohamed, M.B., Nikoobakht, B., El-Sayed, M.A. (1999) *Laser Photothermal Melting and Fragmentation of Gold Nanorods: Energy and Laser Pulse-Width Dependence*. J. Phys. Chem. A, 103 (9), 1165-1170.
168. Ditlbacher, H., Krenn, J. R., Lamprecht, B., Leitner, A. & Aussenegg, F. R. (2000). *Spectrally coded optical data storage by metal nanoparticles*. Opt. Lett., 25, 563-565
169. Chon, J.W.M., Bullen, C., Zijlstra, P., Gu, M. (2007). *Spectral encoding on gold nanorods doped in a silica sol-gel matrix and its application to high density optical data storage*. Adv. Funct. Mater., 17, 875-880.
170. Zhang, J., Gecevičius, M., Beresna, M., Kazansky, P.G. (2013). *5D Data Storageby Ultrafast Laser Nanostructuring in Glass. OSA Postdeadline Paper Digest*. Retrieved August 12, 2013, from: https://www.osapublishing.org.
171. Shimotsuma, Ya., Sakakura, M., Kazansky, P.G., Beresna, M., Qiu, J., Miura, K., Hirao, K. (2010). *Ultrafast Manipulation of Self-Assembled Form Birefringence in Glass*. Adv. Mater., 22, 4039-4043
172. Shimotsuma, Z.Y., Sakakura, M., Kazansky, P.G., Beresna, M., Qiu, J., Miura, K., Hirao, K. (2010). Ultrafastmanipulation of self-assembled form birefringence in glass. Advanced Materials, 22, 4039-4043.
173. Beresna, M., Gecevičius, M., Kazansky, P.G., Tailor, T., Kavokin, A.V. (2012). *Exitation mediated self-organization in glass driven by ultrashort light pulses*. Applied Physics Letters, 101, 20-31.
174. Zhang, J. (2014). Seemingly Unlimited Lifetime Data Storage in Nanostructured Glass. Phys. Rev. Lett., 112, 33-39. Petrov, V. V., Kryuchyn, A.A., Shanoylo, S. M., Kryuchyna, L.I., Kossko, I.O. (2005). *The Metal Carriers for Long-Term Storage of the Information*. NAS of Ukraine, Institute for Information Recording, Naukova dumka, Kiev.
175. Gorbov, I.V., Petrov, V.V, Kryuchyn, A.A. (2007). *Using ion beams for creation of nanostructureson the surface of high-stable materials*. Semiconductor Physics, Quantum Electronics & Optoelectronics, 10 (1), 27-29.
176. Kryuchyn A.A. (2008). *Un disque optique en verre gravé pour l'archivage longue durée. Le magazine du stockage et de la gestion d'informations*, 247/248, 3-4.
177. *Petrov, V.V, Kryuchyn, A.A., Lapchuk , A.S., Gorbov, I.V., Manko, D.Yu., Fu, M., Shanoylo S.M., Morozov, Ye.M.* (2016). Long-term data preservation on sapphire optical discs. SPIE Proc. 9818(02), 4-10.
178. Fu, M., Xu, W., Le, Z., Gorbov, I., Manko D. (2016). *Data recording in digital form on sapphire optical disk*. Optics and Presision Engeeniring, 24(10), 110-115.

179. Petrov, V.V, Kryuchyn, Shanoylo, S.M., Kossko, I.O., Kravets , V.G. (2003). *Methods of Solving the Problem of Long-Term Information Storage Recorded in a Digital Form*. Rep. of the Nat. Acad. Sci. of Ukraine, 4, 52-58.
180. Petrov, V., Gorbulin, V., Kryuchyn, A. (2009). *Information security and problems of long-term storage of electronic documents.* Pomiary, Automatyka, Komputery w Gospodarce i Ochronie Srodowiska, 1 (Mar.), 6-8.
181. Petrov, V., Kryuchyn, A., & Gorbov, I. (2011). *High-density optical disks for long-term information storage*. 22nd Congress of the International Commission for Optics: Light for the Development of the World.
182. Petrov, V.V., Kryuchyn, A.A., Shanoylo, S.M., Kravets, V.G., Kossko, I.O., Belyak, Ie.V., Lapchuk, A. S., Kostyukevych, S. O. (2009). *Super-Dense Optical Information Recording*. Institute for Information Recording, NAS of Ukraine: Kyiv.
183. Petrov, V.V., Kryuchyn, A.A., Beliak, Ie. V., Lapchuk, A.S. (2016). *Multi-photon microscopy and optical recording*. Nat. acad. of sciences of Ukraine, Inst. for information recording, Akademperiodyka: Kyiv.

CONTENT

CHAPTER 4 LONG-TERM USABILITY OF OPTICAL MEDIA

CHAPTER 5 TECHNOLOGIES OF OPTICAL LONG-TERM DATA STORAGE DEVELOPMENT

CHAPTER 6 OPTICAL LONG-TERM DATA STORAGE PROGNOSTICATION AND FAULT DETECTION

Довготривале зберігання цифрової інформації є важливим науково-технічним завданням в умовах швидкого зростання обсягів інформації, представленої у цифровому вигляді. Значущим елементом вирішення цієї проблеми є створення спеціальних носіїв для довготермінового зберігання стратегічно важливої інформації, науково-технічної інформації та інформації, яка становить національне культурне надбання. Наведено дані про створення спеціальних оптичних носіїв, виготовлених із високостабільних матеріалів для довготермінового зберігання інформації.

Наукове видання

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ
ІНСТИТУТ ПРОБЛЕМ РЕЄСТРАЦІЇ ІНФОРМАЦІЇ

ПЕТРОВ Вячеслав Васильович, ЛЕ Тзичунь,
КРЮЧИН Андрій Андрійович, ШАНОЙЛО Семен Михайлович,
ФУ Мінлей, БЕЛЯК Євген Вячеславович,
МАНЬКО Дмитро Юрійович, ЛАПЧУК Анатолій Степанович,
МОРОЗОВ Євгеній Михайлович

ДОВГОТРИВАЛЕ ЗБЕРІГАННЯ ЦИФРОВОЇ ІНФОРМАЦІЇ

Англійською мовою

Видається в авторській редакції

Художнє оформлення *Є.О. Ільницького*
Технічний редактор *Т.М. Шендерович*
Комп’ютерна верстка *Н.М. Коваленко*

Підписано до друку 24.07.2018. Формат 70 × 100 /16. Гарн. Ньютон.
Ум. друк. арк. 12,03. Обл.-вид. арк. 13,75. Тираж 250 прим. Зам. № 5335.

Видавець і виготовлювач Видавничий дім «Академперіодика» НАН України
01004, Київ, вул. Терещенківська, 4

Свідоцтво про внесення до Державного реєстру суб’єктів
видавничої справи серії ДК № 544 від 27.07.2001