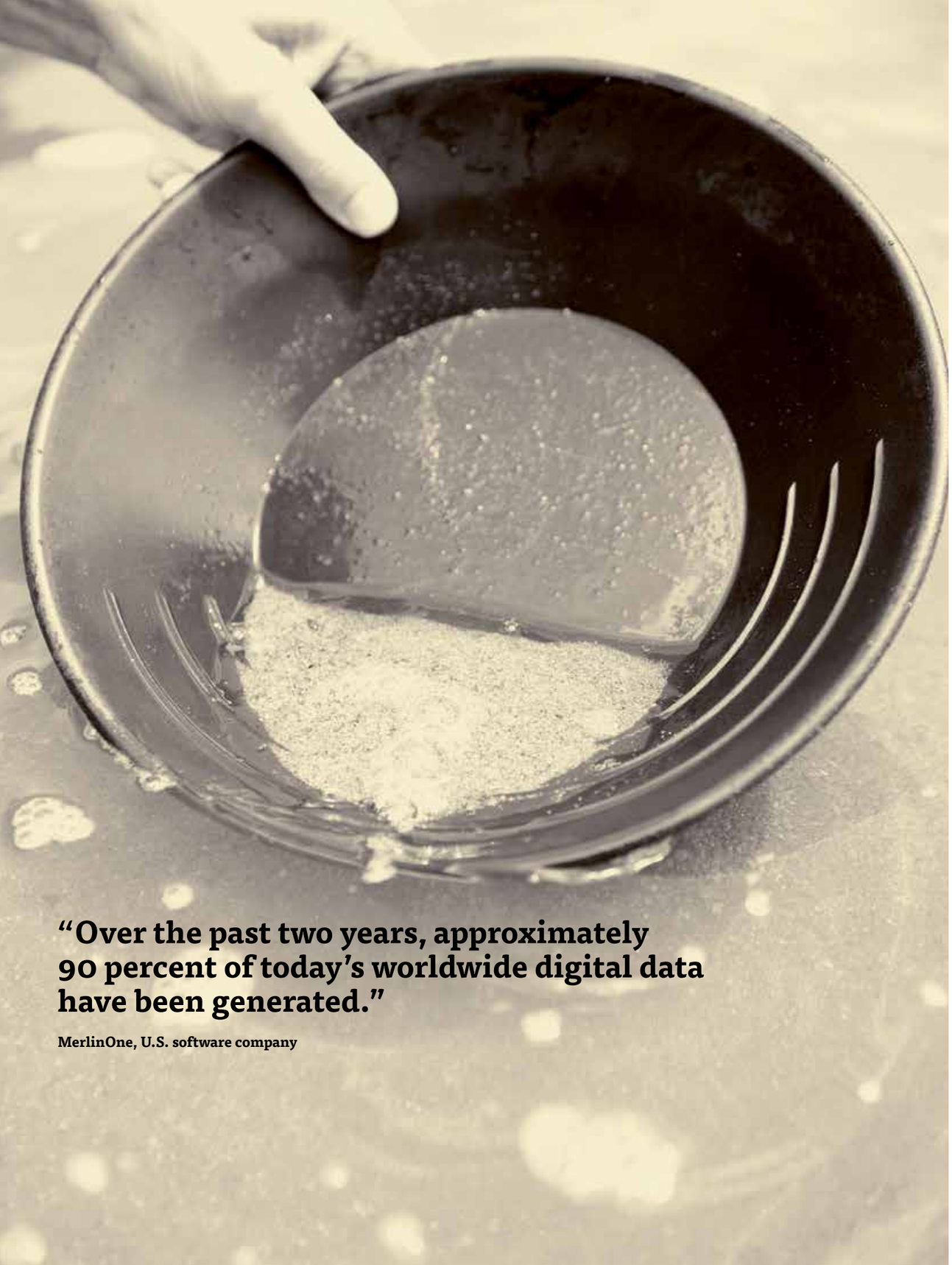
INFORMATION IS THE NEW GOLD

How will we save our data in the future?

ALL PHOTOS: ISTOCKPHOTO



“Over the past two years, approximately 90 percent of today’s worldwide digital data have been generated.”

MerlinOne, U.S. software company

Data are the new oil. As a commodity, they are fuelling tomorrow’s economy, they increase efficiency and sustainability and they personalize shopping or medicine. They provide the foundation for the new revolution of the world of work by means of artificial intelligence (AI). Data are incessantly generated – from the ground, from the air or the orbit, by people as well as by machines. Dozens of terabytes are added every second. With an upward trend! To store and administer this treasure requires the latest technologies, innovative concepts and an abundance of energy.

Over the past two years, approximately 90 percent of today’s worldwide digital data have been generated. At least, this is the overall percentage aggregated by the U.S. software company MerlinOne. And this data pile is growing incessantly. Every day, about two and a half quintillion bytes are added. Quintillions! Every day! You would need almost 300 million DVDs to store this amount of data. If you stacked them up, they would reach into space. An end of this trend is not in sight. On the contrary: technologies such as autonomous driving, the Internet of Things (IoT) or the constantly growing number of eyes and ears in space will even accelerate the trend in the future. However, this means that the demands upon technologies and concepts that are necessary for the storage and administration of the enormous amounts of data are increasing simultanesouly.

Flashback and outlook: from kilobytes to yottabytes

To be able to better understand current developments, it is worthwhile to take a look at physics and the early days of computer science. The basic unit of all data is bit. This so-called binary digit can only take two values – 0 and 1, true and false, on and off. A sequence of eight bits equals one byte, which can express 256 different states. The first alphabet for computers was developed from this starting point. One byte represented a letter, a number or a special character. A text with the length of one A4 page equals a few thousand bytes, thus a few kilobytes. A music track in MP3 format takes a few megabytes (106 bytes) and a movie on a DVD a few gigabytes (109 bytes). Hard disks with a storage capacity in the terabyte range (1012 bytes) are a standard today. But even larger quantities of data are possible. For example, Google is processing data volumes of several petabytes (1015 bytes) on a daily basis. From a quintillion upwards, the unit exabytes (1018 bytes) is used and zettabytes (1021 bytes) and yottabytes (1024 bytes) are more and more becoming standard vocabulary, particularly regarding IoT, autonomous driving and digital personalized medicine.

And the winner is: magnetic data storage

As the base unit can only have two states, there are some options for data storage (see Box 2). For example, holes punched into paper sheets can later be scanned by pins. These so-called punch cards or punch tapes were essential for data inputs at the beginning of the computer age. But scanning can also be performed by means of lasers. For example, nano-scale pits in the carrier material of a CD, DVD or BluRay disc. Of course, electric charges can also be used for data storage. The volatile RAM modules are based on this principle, as well as the non-volatile flash memories, which are for example used in USB sticks and Solid State Drives (SSD). However, one specific technology evolved into the workhorse of computer technology –magnetic data storage.

When IBM released the first Hard Disk Drive (HDD) in 1956, a designated room was required to store it. IBM 350 was 1.73 metres in height, 1.52 metres in width and almost three-quarters of a metre in depth. 50 aluminum plates rotated 1,200 times per minute in its interior and had a storage capacity of 3.75 megabytes. Approximately 650 U.S. dollars of monthly rent were due for this storage. The capacity of modern HDDs has now reached 15 terabytes. “Since the development of the first hard drive, storage density has increased by nine orders of magnitude,” Olav Hellwig states. “This means that one billion times more data can be stored in the same space today.” Hellwig is a physicist. For almost 14 years, he worked for major hard disk manufacturers

in Silicon Valley. Today, he is leading the research group "Functional Magnetic Materials" at Helmholtz Center Dresden-Rossendorf (HZDR) and holds a professorship at Chemnitz University of Technology. According to the expert's assessment, "conventional hard disk technologies can still be optimized, but with the conventional surface storage density provided on a hard disk, we can at best increase capacities by one order of magnitude."

What is sufficient for private users, has to be larger for companies …

For the foreseeable future, there may certainly be sufficient capacity for home PCs. In contrast, companies longe since have to resort to professional storage solutions. "On the one hand, there are dedicated storage networks," Martin Falk explains. "These are central storage systems, basically large cabinets full of hard drives. They then distribute data between individual servers." The IT expert works as a consultant in Tim AG's Professional Service and is in close contact with the producers of computer centres. Another possible variant are so-called hyper-convergent systems, which are currently up and coming. Instead of connecting individual hard disks, they link entire servers with CPU, RAM, hard disk space and obviously network connectivity by means of a virtualization software. Thus, a single virtual machine is created from many servers, which can be accessed via a unified administration interface. The advantage: "These systems can be scaled. This means that every time a new server is added, the overall system receives more RAM, more CPU power and more storage, of course."

A company that specializes in such computer centres is Cloud& Heat Technologies GmbH, based in Dresden. "Flexibility is a central aspect in computer centres," says Robert Pawlik and explains why he is relying on modular, expandable systems. "IT systems have to be economically efficient. Thus, it is of course counterproductive to set up an overhead, i.e. unused computing capacity. Therefore, the optimal utilization of the systems is key here, not the largest possible capacities." The Business Development Manager knows that there are various ways to optimize the cost-effectiveness of the systems. The availability of cooling capacity is one of them, for example. Another important point is limited space. "Often, the server rooms available in commercial property are not sufficient and newly contructed buildings are very expensive. In these cases, flexible, modular options must be created that are suitable to scale the capacities as required." For some time now, container computer centres therefore have become a trend. Cloud&Heat also offer such expandable systems as "computer centres in a box". The containers are set up outside of buildings, and therefore they do not need any additional server rooms.

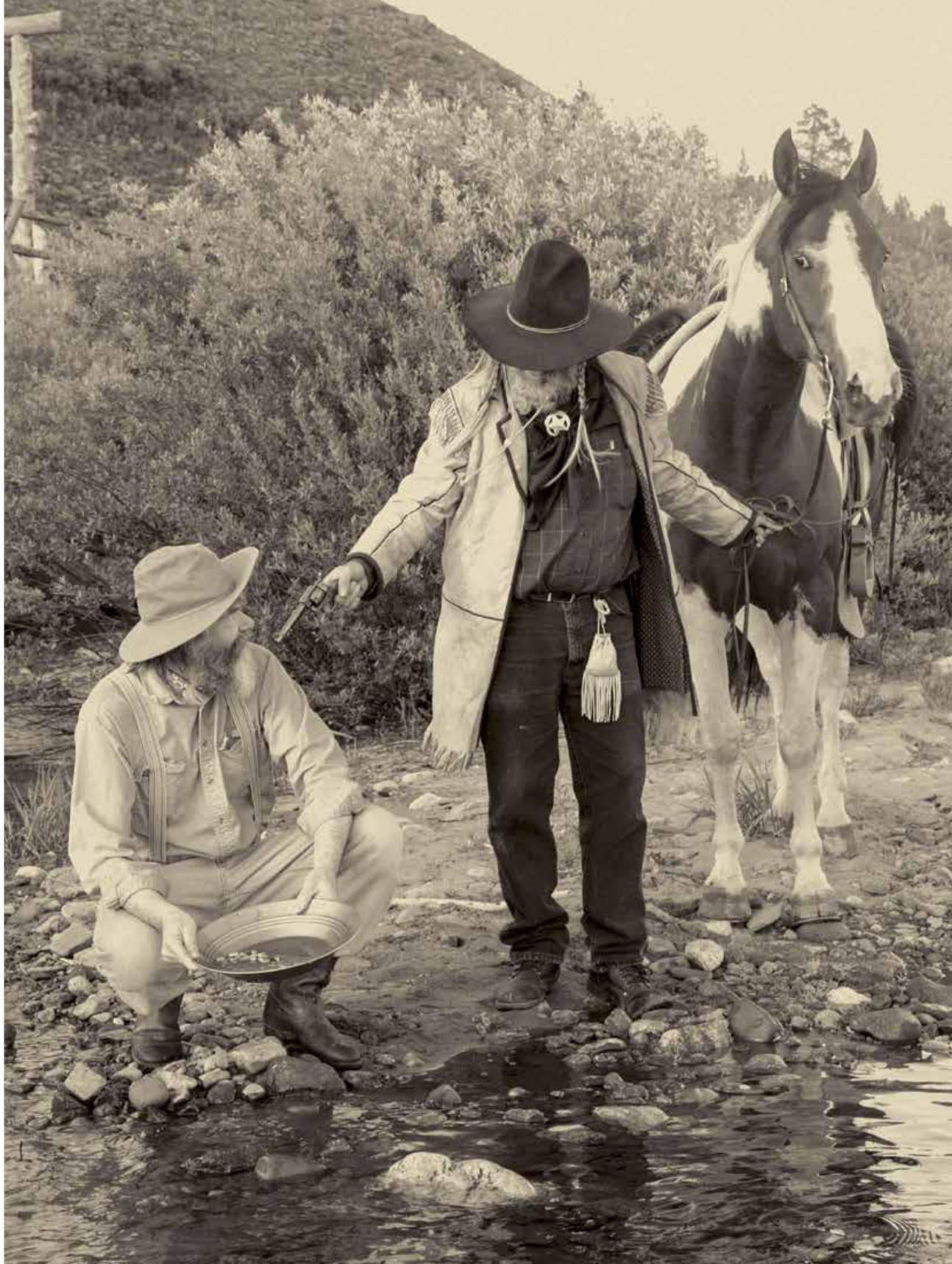
From "tiers" to flash modules

But change is not only emerging for server systems, but also for data carrier technologies. "Currently, so-called tiered systems are still state of the art," explains Falk. Super fast SSDs, quickly rotating RAS disks and high capacity HDD represent the different tiers. The data are first transferred to the SSD and are available from there very quickly. However, if they are retrieved rarely, the system gradually shifts them to the next slower tier. "This technology is on the downgrade," Falk states. In his opinion, the trend is away from rotating disks and towards pure all-flash systems that are based on SSDs or on flash modules only.

"Today, this is the case simply because modern SSDs already have a larger capacity than rotating disks." In fact, capacities of 16 terabytes are readily available, and SSDs with as many as 100 terabytes of

"Since the development of the first hard drive, storage density has increased by nine orders of magnitude."

Dr Olav Hellwig, professor of Functional Magnetic Materials at Chemnitz University of Technology



“The trend is away from rotating disks and towards pure all-flash systems.”

Martin Falk, Consultant in the Professional Service at Tim AG

storage capacity have been announced in the consumer sector. The largest rotating hard drives today have a capacity of 15 terabytes. “The storage density is also considerably higher. This means that the stored information requires significantly less space, which in turn saves assembly space, and above all, energy.” However, SSDs have a major disadvantage. Currently, they are considerably more expensive than HDDs. “But this is depending on supply and demand. The more they are used in computer centres, the cheaper they will become.”

A look at statistics shows that there is sufficient demand on the market and it is also growing steadily. In 2017, a study commissioned by the German Association for Information Technology, Telecommunications and New Media (Bitkom) counted 53,320 “computer centres” in Germany alone. On the one hand, these include regular server cabinets, whose number has remained stable at around 30,500 over the past four years. On the other hand, there are large computer centres with more than 5,000 square metres of room area. Their number has increased from 70 to 90 in the same period of time, which equals almost 30 percent.

More performance for less energy

But how does the choice of storage technologies affect the operation and economic efficiency of small and large computer centres? Falk is sure that it is not crucial at all, how much data can be stored on a medium. “Size, power consumption and heat development play a much bigger role,” he says. “Above all, operators of large computer centres employ extensive server infrastructures. Their energy costs are major items. And in turn, storage technology has a large impact on energy consumption.”

Hellwig explains why this is the case: “IT hardware is entirely based on the flow of electrons. But every time charges are moved, they generate Ohmic heat. The higher the density of memory cells or processor units is, the warmer it gets. This is not only problematic regarding heat dissipation, but also requires more and more energy for cooling.”

Microsoft sinks a computer centre in the sea for cooling purposes

As this effect does not only have an impact on the balance sheets of operators of computer centres, but is also increasingly polluting the environment, solutions must be found. Microsoft tested a drastic solution off the coast of Scotland in June 2018. The software giant sunk a container computer centre with a length of twelve-metres in the rough North Sea. The 864 servers with a storage capacity of 27.6 petabytes are cooled with the cold sea water that has a temperature of only a few degrees. This reduces power consumption significantly. Both the data transfer and the power supply take place via submarine cables. To make the project even more environmentally friendly, the company relies on regenerative energy from tidal power, sun and wind. It seems as if the first test was successful for Microsoft, because the company only recently announced that they plan to sink more computer centres into the sea.

A residential building in Dresden is heated with recovered energy

Even though this “American Way” seems to be pay off economically: it is probably not the be-all and end-all to discharge waste heat into the sea without further use. Therefore, Cloud&Heat go one step further. No matter whether their hardware is stored in usual server rooms or in containers, it is water-cooled. Pawlik explains the underlying technology as follows: “The water is led into the server cabinets and cools down the components by means of heat exchangers.” According to him, 90 percent

“With the waste heat produced by a computer centre, you can heat the company swimmingpool.”

Robert Pawlik, Business Development Manager at Cloud&Heat

What is the difference between data, information and knowledge?

Of oil, gold, power and a new economic branch

“Data are the oil of the 21st century.” Or: “Information is the new gold.” When talking about the future in the cloud, pithy statements are not long in coming. Of course, we have known one thing for a very long time: “Knowledge is power.” But what makes data different from information, and how does information in turn differ from knowledge? And what makes them valuable?

In the first instance, data are just an accumulation of individual numerical values. They can be gained by a measuring device as well as by pressing a button at your home PC. An example of this: “51499406–01632344231220181015”. At first glance, this makes little sense. But read on…

Data are the basis of information society. They are created incessantly, they are stored, processed and put in correlations. In the right context, information is derived from them. For example, our string of characters translates into the coordinates “51.499406, -0.163244” (or 51° 29’57.9”N 0° 09’47.7” W) and a time specification. In particular, 23 December 2018 at 10:15. This is more useful, but something is still missing. For example, more information such as the photograph of a red double-decker bus shared on an Instagram account.

Information can be combined and from this combination, knowledge is created. Coordinates reveal my location: the luxury department store Harrods in London. The photograph of the bus confirms the location. The time stamp suggests that I am on a Christmas shopping tour. This knowledge can now be utilized in many different ways. My friends could now be jealous that I am on a spontaneous short holiday. My employer’s human resources department could be suspicious, because they have just filed my certificate of illness. And I could suddenly receive a popup advertisement on my phone. The nice electronic toy that I have been searching for on the internet for such a long time is on offer in a shop just three streets away from my current location …

However, data and information can not only be used to find out locations and personalize advertisements. Tools from the field of AI, such as machine learning or predictive analytics, particularly draw knowledge from data and information, which was unimaginable until recently. Satellite data and drone images are used to calculate the ideal time for agricultural irrigation, fertilization or harvesting. Failure of important components can be predicted with data provided by sensors on machines and modern electric cars know exactly when it is time for a visit in a workshop.

Data and information have long since become a commodity of equal importance as crude oil or iron ore. And thus they are the foundation for a completely new way of earning money, which is particularly in the focus of the big players of digital economy. Amazon, for example: the company started business as an internet bookstore almost 25 years ago and today is a synonym for online shopping. However, they have earned most of their money with leasing out computing power and storage space for a long time now. While this division made profits of 4.3 billion U.S. dollars, the company recorded approximately 200 million U.S. dollars in mail order losses last year. But not only Amazon have discovered the cloud for themselves. Microsoft has also been investing strongly with significant success. As recently as in October 2018, the company knocked its competitor off the throne in the cloud market. They created turnovers of approximately 26.7 billion U.S. dollars there, which is 3.3 billions more than Amazon. Google and IBM have long since secured their pieces of this lucrative pie, in the Far East Alibaba is gaining some ground and the German company SAP is also heavily involved.

of the resulting heat are transferred into the water cycle. And then? "We are working at very high temperature levels", he outlines the special features of his plants. "The water exits the system at 60 degrees Celsius. It can then supply a heating system via a buffer tank and heat drinking water or the company swimmingpool." Pawlik reveals that one of the company's computer centres is located in a residential building with 56 accomodation units. The recovered heat equals the base load of the building. "But water cooling by means of heat exchangers has yet another advantage," he states. "For example graphics cards no longer require their very large fans. Thus, considerably more of them fit in the same space." If you know that graphics cards can process data 100 times faster than normal CPUs and therefore meet a great demand particularly in the fields of artificial intelligence and blockchain, you will appreciate this advantage.

Utilizing the heat produced by computer centres as a commodity is certainly a good step towards sustainability. But what if modern IT produced less, or ideally no heat at all? This is where Hellwig's research comes into the picture. "Today, the density of processed and stored data on magnetic hard disks is so high that the underyling structures have long reached the nanoscale level," the physicist explains. "And in nanomagnetism, we encounter considerably different phenomena than at macroscopic level." For example, there are spin-polarized currents. These are electrical currents that carry a magnetic momentum. Thus, they are also magnetic and can switch the bit cells. Unlike the external magnetic fields that are normally used, spin-polarized currents act very locally inside the conducting paths. Therefore, the switching process can be controlled easier. A positive side effect: as no external magnetic field is required, the energy requirements are often reduced, too.

Pure spin waves could become the solution to the heat problem

In fact, there are already storage elements today, which make use of spin-polarized currents. The technology is called Magnetoresistive Random Access Memory, abbreviated with MRAM. It combines the advantages of very fast, but volatile RAM modules with non-volatile, but limitedly rewritable flash memories. MRAM modules are already available today. However, their high price considerably limits their fields of application. For example, they are used in the aerospace industry, as data buffers in server systems or for industrial facilities, where data losses caused by power failures are less tolerable than extra costs for storage capacities.

But nanomagnetism offers yet another possibility: Hellwig and his colleagues at HZDR investigate the use of pure spin waves. Unlike spin-polarized currents, these are not based on electron flows. "Magnetic interactions that take place without a transfer of charges have an advantage. Without Ohmic heat, the energy demand decreases and the problem of heating is avoided. Although this technology is still in its infancy, it is a very exciting field of research for physicists."

The next few years will show which of the technologies and concepts will eventually catch on – no matter whether they are electrical or magnetical, rotating or solid, with heat utilization or sea cooling. "Apart from technical feasibility, costs are the decisive factor whether technologies catch on at all and if they do so, how quick and how sustainable this process is," Hellwig knows from experience. "In any case, I do not believe that a new technology will wipe everything that has been established so far off the table. I rather think that different solutions will co-exist in parallel for different use cases."

"Apart from technical feasibility, costs are the decisive factor whether technologies catch on at all and if they do so, how quick and how sustainable this process is."

Dr Olav Hellwig, professor at Chemnitz University of Technology

INNOVATION & FUTURE > DATA – THE NEW GOLD

Interview

Gold diggers? Information is the gold of the 21st century. This is how to protect it.

High-tech hard disks, energy-efficient systems and innovative storage concepts are undoubtedly important pillars of a data-based world. But how can data be managed not only efficiently, but also securely? How can they be protected against prying eyes and defended against criminal energies? And why are data security and data protection a great deal more than just buzz words? All these are questions that Benedikt Fischer is confronted with every day. The expert for cloud services is working for the IT service provider aConTech and in the following interview, he explains why every company should take part in the data security marathon and which obstacles have to be overcome on the way.

Data are often referred to as the oil of the 21st century and information as the new gold. Of course, such a treasure deserves first-class protection. What is the actual situation of data security in German companies?
I think all facets from white to black are represented. There are companies that have been strongly committed to data security for years and also invest a lot in this topic because of their professional responsibility. But there are also companies that have slackened the reigns regarding this issue due to a lack of know-how or budget. In the course of the General Data Protection Regulation, an upswing in security and data protection can actually be observed. But I am worried that some of it will fizzle out and we will quickly revert to old routines: the typical "everything will be OK" patterns of thought. This is why I believe that data security should be one of the top priorities in companies, at least for the next one or two years. Unfortunately, this is currently often not the case yet.

Trojans, data leaks, ransomware: extorting companies with their own data seems to have become a lucrative source of income for hackers. How great can damages caused by data theft be today?
It is unpleasant when data are encrypted to extort companies for ransom, of course. In these cases, the data are lost for the moment. But this can be resolved with reasonable effort. If data are changed or tapped permanently and this remains undetected, the impact is much worse. This hurts considerably more, because it is usually noticed only with a significant time delay. Of course, money is the largest motivation, particularly in criminal industrial hacking. The hackers want to get away with as much loot as possible. And they achieve this aim best if they cause a lot of damage and then undo their changes in return for blackmail payments. But there are two further topics that play a key role: theft of professional know-how from the company and the reputational damage dealt after data losses have become known to the public.

How can hackers intrude a company's system? What are the most important loopholes at the moment?
Phishing emails, which lead to hacked websites, are still among

Short glossary on storage technology from the perspective of cleanrooms

There can be no modern data storage without cleanrooms

Data is stored magnetically on hard disks, abbreviated with **HDD**, which stands for hard disk drive. In their interior, there are rotating disks of glass, aluminium or magnesium alloys with multiple wafer-thin magnetizable layers made of cobalt, chromium and platinum, which are often laminated with atomically thin non-magnetic coupling layers. In a distance of about three nanometers, a read / write head floats above the surface and defines the so-called bit cells of a data track during the writing process. Because even the tiniest impurities result in an early death of the hard disk, they are coated by means of the so-called sputtering technology under the purest conditions (we reported on this topic in issue 18, 03-2018). And also the read / write heads are produced in cleanrooms. Around 100,000 of them can be placed on a silicon wafer.

Flash memory chips of modern Solid State Drives (SSD) pass the same processing steps of microelectronics production, too. On them, data are not stored magnetically but electronically in the form of charges. SSDs are nonvolatile memories and thus data can be kept on them without power. Their advantage over HDDs: they are lighter, faster and require less energy. Also, they have no moving parts and are therefore quite robust. Their disadvantages: each initiated deletion process causes a bit of damage to the SSD. Consequently, they are not indefinitely rewritable. Also, they currently cost almost ten times more than HDDs of the same size. While some experts predict declining prices as a result of rising demand, others doubt this in view of the raw materials needed for their production. For example, HDDs in which glass is used as a substrate, require silicon dioxide, which is both abundantly available and easily accessible. In contrast, the chips in SSD require monocrystalline and high-quality silicon.

Random Access Memory, abbreviated with **RAM**, also is a semiconductor memory that is produced in the cleanrooms of electronics manufacturers. But unlike flash memories, it needs constant replenishment. Without electricity, it "forgets" the data stored on it and therefore ranks among the volatile memories. But in return, RAM modules are considerably faster than Flash technology. Therefore, their main application is the working memory in computers, smartphones and tablets.

the most important loopholes. However, the so-called social hacking is currently also on the rise. That means that hackers do not even target technical components in this case. Instead, they try to reach a person via fake social media contacts, take over their identities and thus penetrate their companies' security measures. Typical hacker attacks, as known from former times, are recently ocurring less and less. It requires enormous effort to break through firewalls today. It is much easier to be on the streets, talk to people and give manipulated USB sticks to them. This is hard to believe, but it still works best.

Does this mean that people are the greatest weakness in data security?
This formulation is a bit drastic. But how people act is indeed one of the greatest weaknesses in this area. Therefore, knowledge, enlightenment and training are very important components to increase data security.

For example also the hint that "1 2 3 4" is not a safe password?
I generally have reservations about passwords, because you always encounter the same problem with them. Either they are easy to remember and consequently easy to guess. Or they are too long to remember. In this case they are at best stored in so-called password safes with their own passwords, which then open all the gates at once if they fall into the wrong hands. Or they are noted somewhere and are thus no longer safe. But there are many other ways to prove somebody's identity 100 percent secure – face ID, fingerprints, voice and video combined and many more. Anyway, personally I am a big fan of two-factor authentication. For example, if your notebook registers your phone in the proximity and you simultanesouly authenticate yourself by face ID, you will no longer need a password. This is a completely different level of security. And such methods are technologically definitely possible today.

This is certainly particularly true in case of accounts of system administrators.
Correct. And there is another weakness. Administrators often have the habit of permanently running all applications with administrative rights. You actually rarely need these rights, for example, when you configure and change data as part of a typical change management routine. For daily business, for example to check emails, create documents, perform monitoring tasks or introduce new products, no permanent administrator rights are necessary. If I was a hacker, I would soon realize who is an administrator and who is not. If I then got access to the administrator's account, I would suddenly be allowed to access everything, I could read every email and could get all data. For hackers, these accounts are therefore very valuable and are attacked targetedly.

How can companies react to this threat?
For a company this means two things: firstly, administrators need to sensitize their perception of these circumstances so that they are not permanently running all applications with administrator rights. And secondly, accounts with these rights should be specifically protected by means of privileged user access concepts.

"Data security is like a marathon. It starts with registering for the run, not with discussing your approach to the finish line"

Benedikt Fischer, expert for cloud services, aConTech

An in private life? There, everbody is a "privileged user" and should focus on safety. But this is often different in reality. Do we handle our data too carefree at home?
Yes, absolutely. But that is also a question of knowledge and enlightenment. Raising awareness of data security is a permanent process and sometimes means tilting at windmills. This is mainly caused by the fact that that the technology is still relatively new. Ten years ago, smartphones were not as popular as they are today. Consequently, there currently is much too little experience which consequences the disclosure of personal data has in the long term. I think that this is the reason for the careless use of this technology. We simply do not know what can happen to us. Undoubtedly, many data profiles have already been created today. Only the methodologies to protect them comprehensively are far from sophisticated. But this is about to change quickly.

Does the sensitization for data security and data protection consequently have to start much earlier in our lives, probably as a subject at school?
This would be favourable. The question of how people can handle their data and protect themselves from identity theft should definitely be a part of the curriculum throughout Europe, or even better, around the world. Only then, people would be able to responsibly decide on their own data. This is another reason why many people today act like ostriches. They put their heads in the sand and are then sure that nothing will happen. However, I think that most people know that something is happening here. But they do not deal with it, because the effects are not visible to them.

This is really hard to grasp. For example, take the news of 50 million Facebook accounts that were recently hacked. Suddenly, someone can read a post that I perhaps have already shared with the whole world anyway. What harm can this do?
First of all, you lose your identity. And this certainly has far-reaching consequences. Facebook is an example of so-called identity providers. You can authenticate yourself on many portals with your Facebook account. Anyone who has your identity, at the same time has a carte blanche. Today, everything runs fully automated and for example, credit card data or spicy photos can fall into the clutches of hackers which in turn can be used for the next attempt of extortion.

If one side has to be arduously sensitized to the topic of data security, but the other side exploits every weakness with a high degree of criminal energy, does this mean that the battle for data is actually lost?
I would not put it like this. Yes, the world has changed. Today, hacking is automated, bot-based and supported by highly mafia-like or even government-sponsored structures. And certainly there are people who have done little for the security of their data so far. If you wake up today, you might actually think: Oh God, we are at war and I have almost lost it. And even people who insist on one hundred percent protection could easily despair. But in between these two extremes, every company has many options to create hurdles for hackers that are as high as possible. Because hackers quickly realize whether they are facing professionals as their opposition or whether they can break through security with justifiable effort. I usually compare data security with a marathon. It starts with registering for the run, not with discussing your approach to the finish line. There are thousands of steps between the start and the finish line. Data security is a Herculean task, but nothing is lost. ●