Security – Now and in the Future

Information is long established as a vital business asset – and protecting this information is a key challenge for the 21st century. Prof. Antonia Wachter-Zeh wants to use mathematical codes to develop new encryption techniques able to withstand attacks from even the most powerful quantum computers. Her methods could also help enable long-term data storage using DNA.

Link
www.ei.tum.de/en/lnt

Gesamter Artikel (PDF, DE): www.tum.de/faszination-forschung-27

Sichere Daten – auch in der Zukunft

D

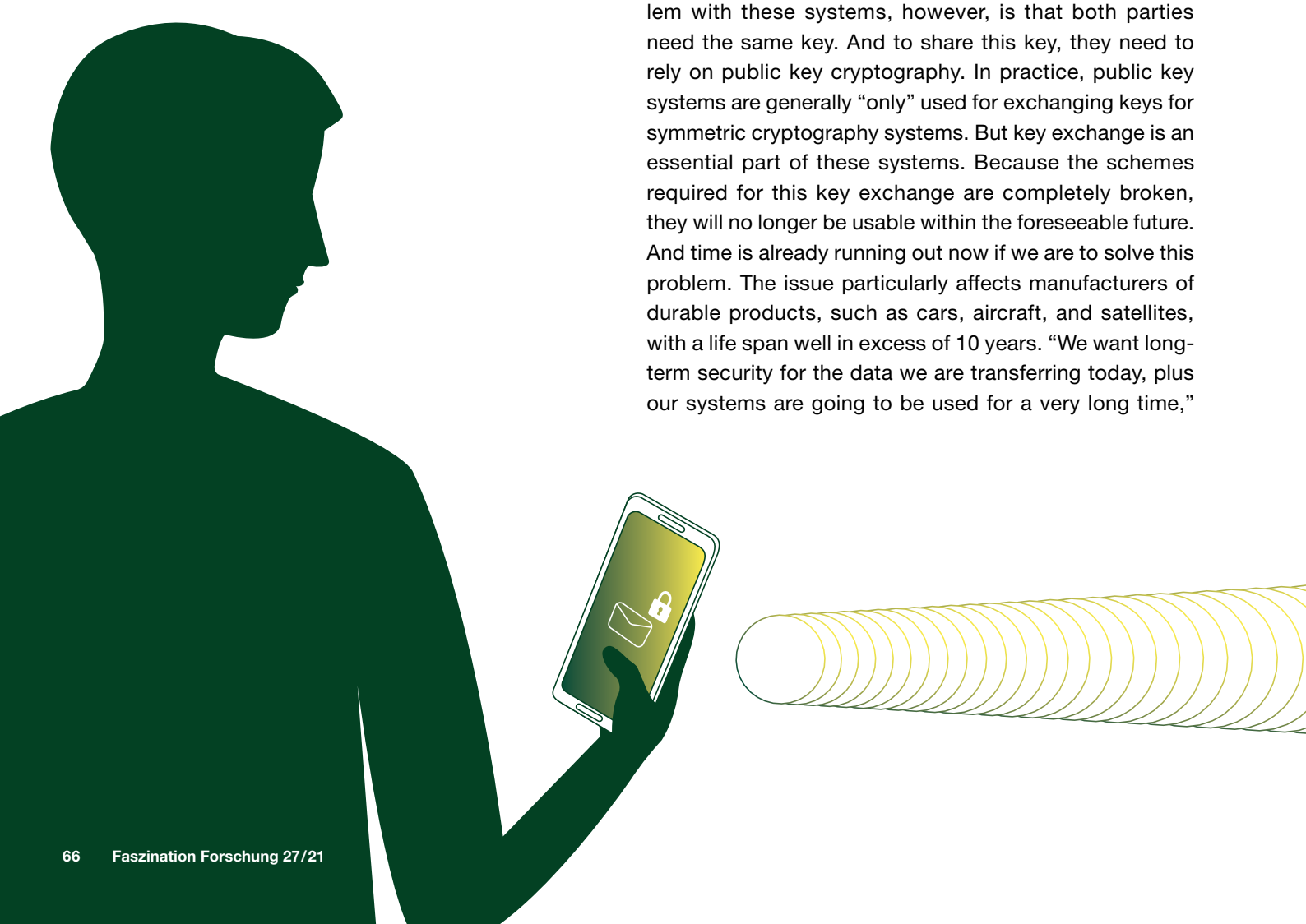
Das Verschlüsseln von Daten verhindert das Auslesen und den Missbrauch sensibler Informationen. Mit dem Bau von Quantencomputern ist die Sicherheit gängiger Verschlüsselungsmethoden in Gefahr. Die auf der Quantenphysik beruhenden Rechner können aktuelle Kryptoverfahren knacken – was die gesamte Kommunikationssicherheit in der IT in Frage stellt. Zwar sind leistungsfähige Quantenrechner erst im Stadium von Prototypen, doch die Zeit für die Lösung des Problems ist schon heute kritisch. Antonia Wachter-Zeh, Professorin für Codierung und Kryptographie an der TUM, entwickelt mit algebraischen Codes neue Verschlüsselungsmethoden, die auch Quantencomputern standhalten sollen. Dieses Konzept der algebraischen Codes wendet Wachter-Zeh auch auf die neue Methode der DNA-Speicherung an. Bei diesem Speicherverfahren werden Daten direkt in DNA-Strängen abgelegt. Die dabei auftretenden Lese- und Schreibfehler will die TUM Professorin mit algebraischen Codes korrigieren. □

Graphics: ediundsepp (source: TUM)

In an ever more digital world, data security is hugely important. To protect data and guarantee the security of data transfer, data for applications such as home banking, online shopping, and sharing business documents is encrypted. A number of different encryption schemes are used, depending on the application.
Asymmetric cryptography is one widely-used cryptography system. Asymmetric cryptography algorithms feature a public key and a private key and form the foundation of many crypto-procedures at work in today's IT world. We rely on public key systems when we send an email, do our internet banking, or communicate with a web server. From a mathematical perspective, public key encryption is predicated on the fact that calculating the private key from the public key is extremely difficult – in fact it's beyond even the most powerful modern supercomputers.
But that may not be the case for much longer. The advent of quantum computers means that widely-used public key algorithms such as RSA and elliptic curve cryptography are no longer secure. It's not yet clear when quantum computers powerful enough to crack these algorithms will be available, but it's clear that they're coming. Google, IBM and other IT companies have already developed early prototypes. The German Federal Office for Information Security expects machines able to crack public key encryption to be available by the early 2030s.

The quantum menace

TUM Professor of Coding and Cryptography and DFG Heinz Maier-Leibnitz Prize winner Antonia Wachter-Zeh is very clear about the significance of this threat. "As soon as powerful quantum computers with sufficient qubits become available, widely-used public key cryptosystems have a big problem," she explains. As part of the ERC Starting Grant project inCREASE, her research group is seeking to develop new cryptographic systems that are also resistant to attack from quantum computers.
It's worth noting that quantum computers are not a threat to all cryptography systems. For very efficient symmetric cryptography algorithms, such as AES, quantum computers pose relatively little threat. Increased computing power can be counteracted by using longer keys. The problem with these systems, however, is that both parties need the same key. And to share this key, they need to rely on public key cryptography. In practice, public key systems are generally "only" used for exchanging keys for symmetric cryptography systems. But key exchange is an essential part of these systems. Because the schemes required for this key exchange are completely broken, they will no longer be usable within the foreseeable future. And time is already running out now if we are to solve this problem. The issue particularly affects manufacturers of durable products, such as cars, aircraft, and satellites, with a life span well in excess of 10 years. "We want long-term security for the data we are transferring today, plus our systems are going to be used for a very long time,"



says Wachter-Zeh. “Satellites in orbit, for example, are very difficult to update. But we still want the data transferred via those satellites to be secure in 20 years’ time.” Another reason Wachter-Zeh believes we need to deal with this problem now is that today’s communications are often stored. “We need to make sure that even in 10 or 20 years’ time, no-one will be able to decrypt data transferred today.”

Post-quantum cryptography

Moving to quantum computer-resilient cryptographic systems is therefore something we need to do now. Faced with this problem, the US National Institute of Standards and Technology (NIST) launched a quantum-resilient cryptography competition, which has now entered its final round. The goal is to standardize “post-quantum cryptography”, i.e. cryptographic systems that can’t be cracked even by quantum computers.

Development of post-quantum cryptography is based on mathematical concepts which can be divided into a number of groups. Antonia Wachter-Zeh’s group is researching secure encryption techniques that use error correction codes. This promising approach is based on correcting errors which arise during data transfer or storage. Error correction codes enable error-tolerant data encoding. How many errors can be corrected depends on the type of code.

“Codes like this are used in conventional communications technology. In the simplest case they involve adding redundancy to a message,” explains Wachter-Zeh. “In communications engineering, if you transfer a codeword and the communication channel introduces an error, then you want the recipient to be able to resolve the error and compute the correct message. This principle can also be used for encryption. You deliberately introduce a specific error, so that it can’t be decrypted by anyone else.”

Wachter-Zeh is an expert on a specific class of error correction codes called rank metric codes. The NIST classes rank metric-based techniques as extremely promising, but thinks a lot more research is needed. Wachter-Zeh’s research group has developed some of the world’s most efficient rank metric decoding techniques and has proposed a new public key system based on rank metric codes (LIGA).

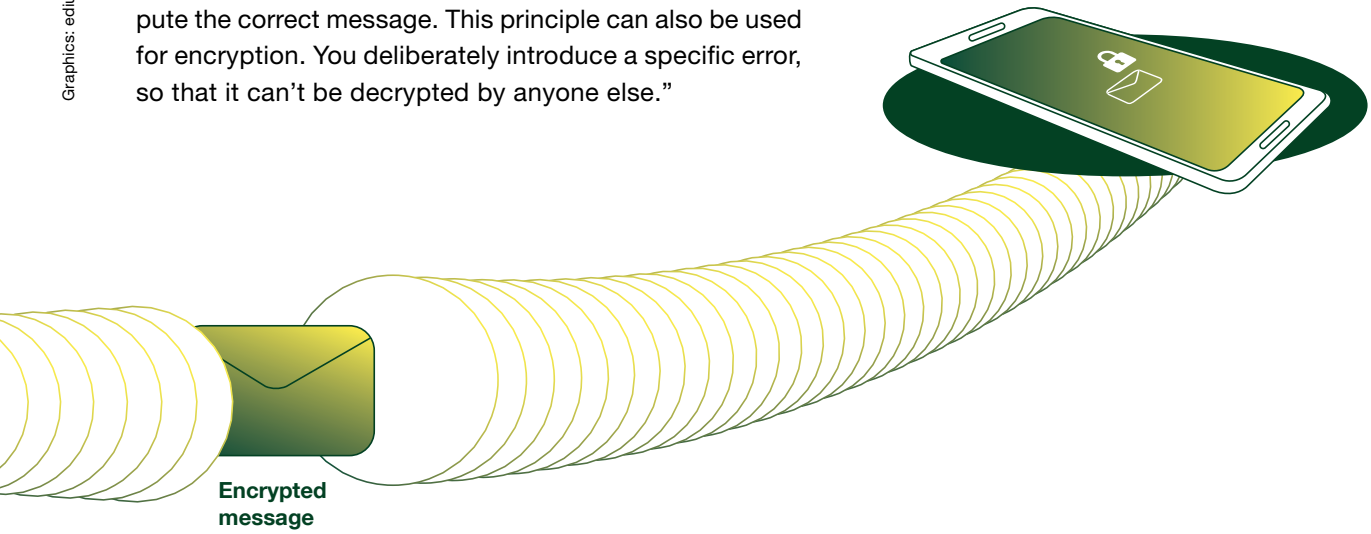
DNA-based storage

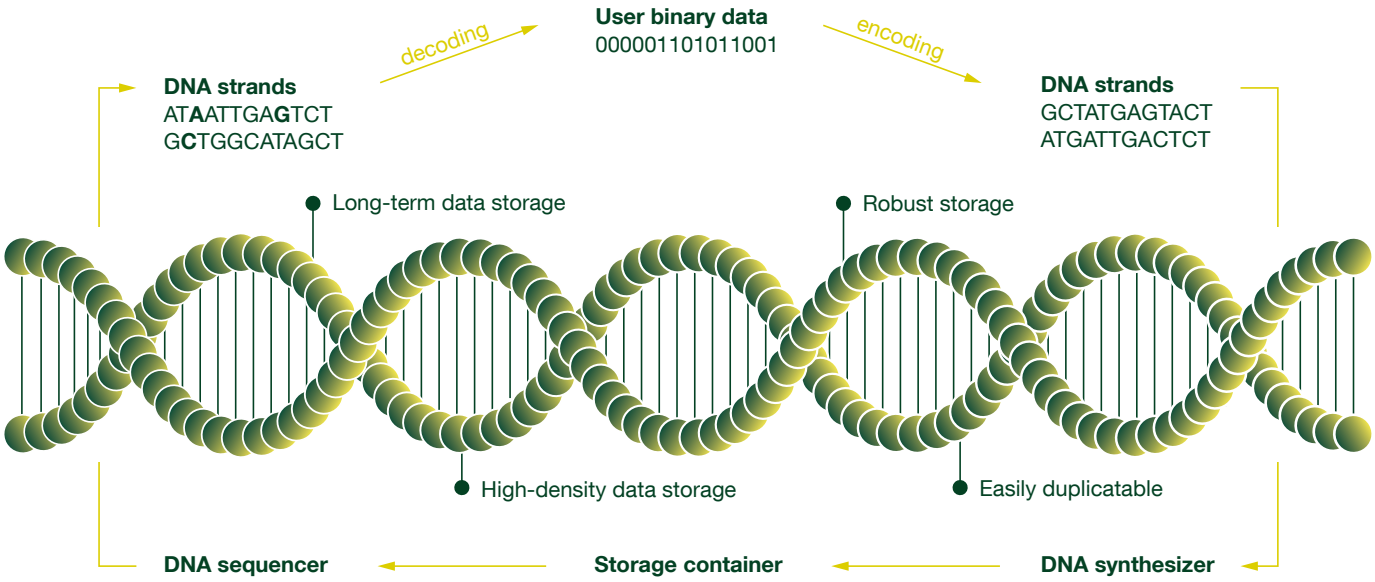
Wachter-Zeh hopes that error correction codes will also provide solutions to another highly topical problem – data storage using DNA. This new storage technique involves storing data in strands of DNA. The zeros and ones of digital data are converted into the four bases found in DNA: adenine (A), cytosine (C), guanine (G) and thymine (T). In the simplest case, two digital bits are combined into one DNA base.

Storing data in DNA strands has the potential to solve an increasingly pressing problem – conventional storage media such as DVDs and hard drives have a very limited life span. Where data needs to be stored for longer periods – several decades for example – there is a high risk of data loss.

With DNA storage, data keeps for much longer – as demonstrated by the fact that we are able to extract DNA from ancient mammoth bones. “The ability of fossils to store DNA implies that DNA could be used for archival data storage. If we want to store data for generations, DNA-based storage is an excellent option.” ▷

Graphics: ediundsepp (source: TUM)





Storing data in strands of DNA involves converting the zeros and ones of digital data into the four bases of DNA: adenine (A), cytosine (C), guanine (G) and thymine (T). DNA storage has the potential to solve several problems such as life span and capacity of conventional storage media.

Prof. Antonia Wachter-Zeh

is a professor in the TUM's Department of Electrical and Computer Engineering. She completed her MSc in communications technology at Ulm University in 2009. Wachter-Zeh completed her PhD at Ulm and the University of Rennes in France in 2013. From 2013 to 2016 she worked as a postdoctoral researcher at the Technion-Israel Institute of Technology in Haifa, Israel, and from 2016 to 2020 was Tenure Track Assistant Professor at TUM. Wachter-Zeh has been awarded the DFG's Heinz Maier-Leibnitz Prize and is funded by an ERC Starting Grant. Her research interests are coding theory, cryptography, and information theory and their application to storage, communication, privacy, and security.

Another advantage of DNA-based storage is that gene sequences are very dense and enable the storage of a great deal of data in a very compact space. “The maximum data density for the medium currently used for archival storage – tape – is 100 GB per cubic millimeter. For DNA it’s 10^9 GB per cubic millimeter,” explains Wachter-Zeh.

Eliminating errors

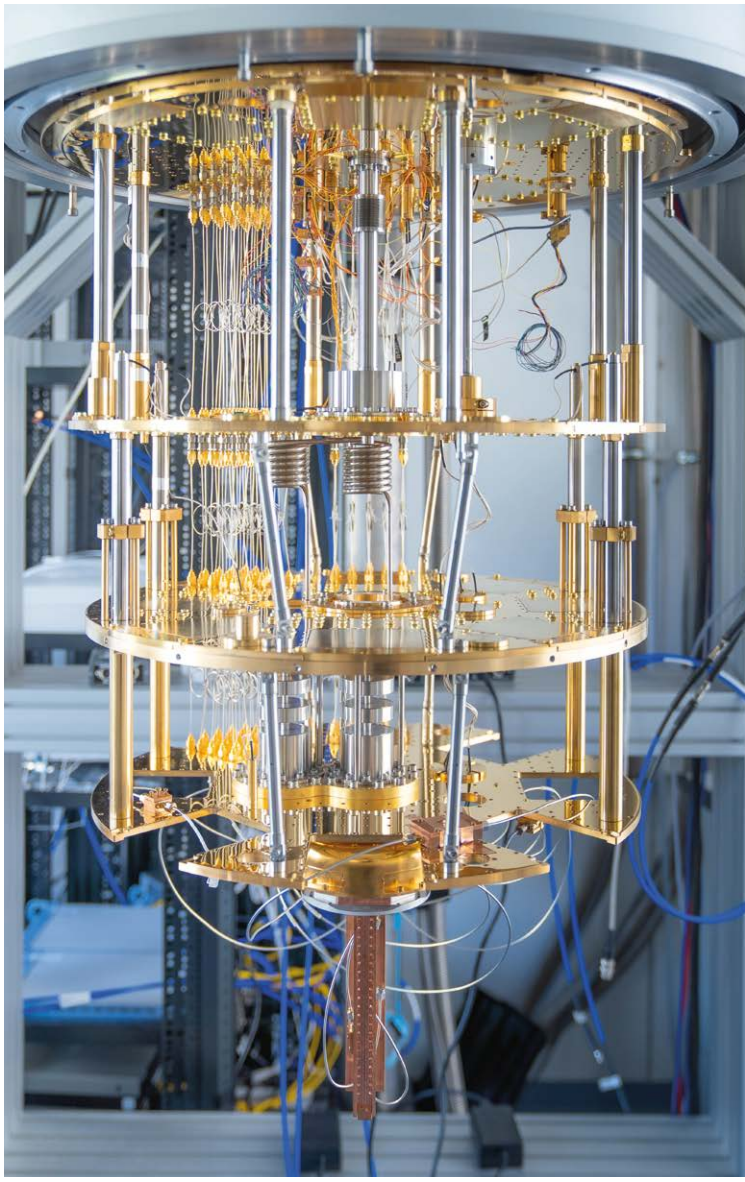
Before DNA-based storage can be used in practice, however, a number of obstacles need to be overcome, particularly in the areas of reading and writing the data. Writing data to DNA using DNA synthesis is currently the costliest part of DNA storage. Wachter-Zeh’s research group is trying to make this process more efficient. “We want to write multiple sequences simultaneously and as rapidly as possible,” explains the TUM researcher. “If a machine can do this in parallel, the question is what is the theoretically best way of doing this?” She is therefore examining synthesis techniques in which multiple strands are constructed step by step and in parallel using a fixed supersequence.

The second focus for her research group is reading the data. This is currently the biggest challenge for DNA-based storage. Basically, data is read by sequencing the DNA. Because each DNA strand is replicated many times, the read sequences are always clustered around the written data. The sequencing process sequences a large number of DNA strands which then need to be reassembled in order.

In addition, various errors can occur. Sequences can go missing, because they have not been read. The order of the sequences can also be lost. Within sequences, it can also occur that bases are inserted, deleted, or duplicated. When reading the data, all of these errors need to be corrected. “Researchers have found that without error correction it is almost impossible to reconstruct data from DNA systems,” explains Wachter-Zeh. “There are simply too many errors. Error correction is therefore essential – for which we can use known communications engineering techniques.”

Here again, Wachter-Zeh is relying on error correction codes. The existence of base deletions and other errors mean, however, that she is having to develop completely new techniques. For her research group, her methodological expertise is invaluable. Her group has developed a general principle to enable coding using unordered sequences. Wachter-Zeh is also a world leader in the development of codes for correcting insertions and deletions. ■

Klaus Manhart



Cryogenic infrastructure for operating a superconducting quantum computer at the Walther Meissner Institute, funded by the Munich Cluster of Excellence MCQST. Many surfaces and electrical contacts are coated with gold to achieve minimum electrical resistance and maximum thermal conduction.

Picture credit: Astrid Eckert/TUM, C. Hohmann/MCQST; Graphics: ediundsepp (source: TUM)