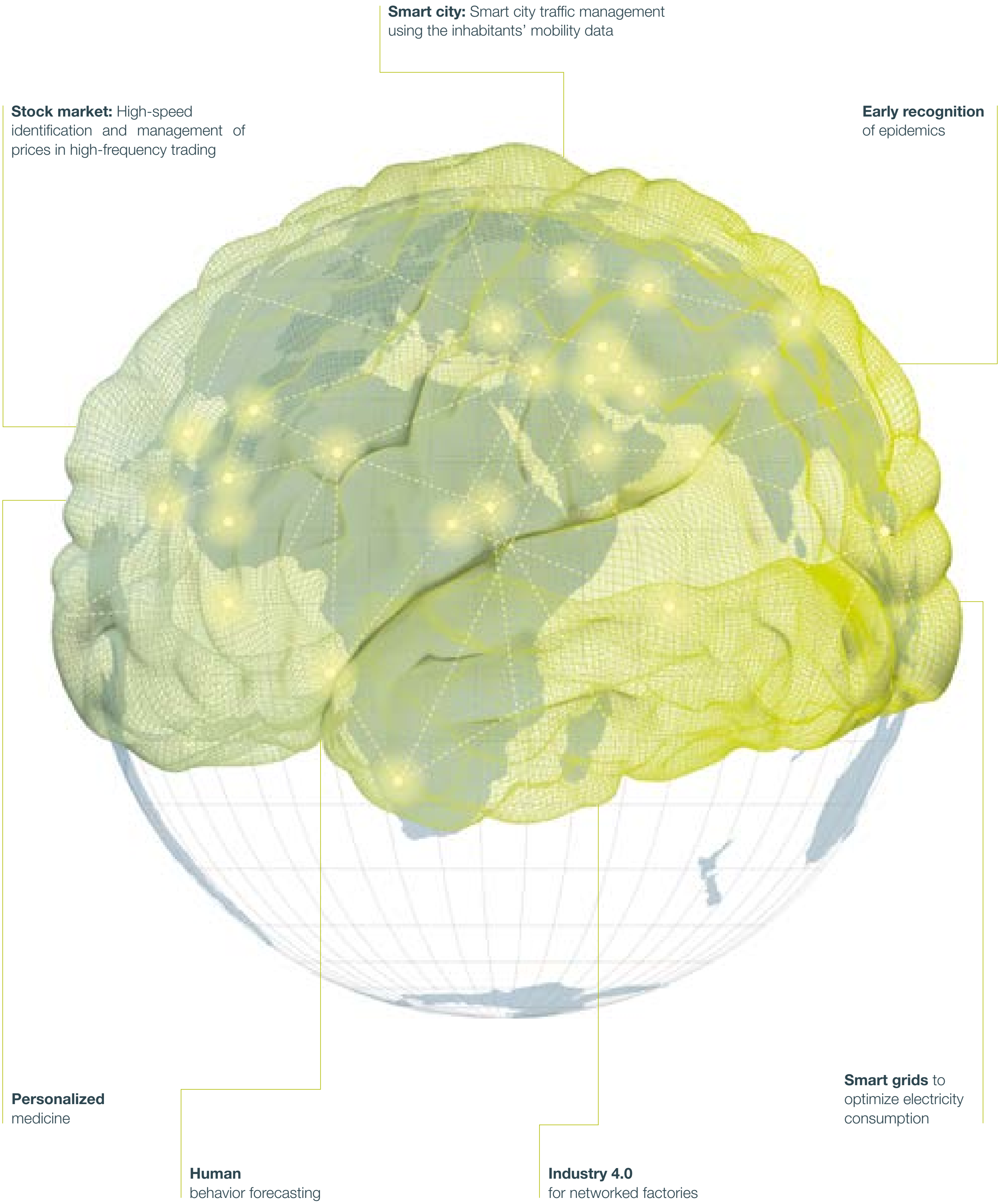


The network structure of the World Wide Web is similar to the connections between the nerve cells and the different areas of the human brain. According to Klaus Mainzer, the modern IT world can be viewed as the next stage in the biological evolution of information. The end result would be a kind of superorganism consisting of our planet and its IT networks as the nervous system.

“Algorithms are powerful and useful. **But by themselves they are blind**.”

Algorithms, until recently strictly the preserve of computer enthusiasts and mathematicians, are now pushing their way center stage – quite simply because of the key role they play in the supercomputers that are digitalizing our lives. The dawn of the big data world is welcomed by some as a golden era, but feared by others as the realization of Orwell’s worst nightmare. But what are the economic, social and legal implications of big data? What are the potential opportunities and risks? We interviewed philosopher of science Prof. Klaus Mainzer, Chair of Philosophy of Science at TUM and Director of the Carl von Linde Academy, about his latest book on computing the world, “Die Berechnung der Welt. Von der Weltformel zu Big Data,” to find out more.

Link
www.mcts.tum.de/philwiss

Picture credit: ediundsepp

The cascade-like spreading of an epidemic disease is represented as a cluster and a complex pattern in a network. Social infection works in a very similar way. An innovation catches on because people are influenced by the behavior of their neighbors.

Picture credit: Jooss/graphics: ediundsepp

“We deplete natural resources by using them – but this is not the case with data. Data can always be reused in different contexts, creating new ways to make a profit.”

Klaus Mainzer

Prof. Mainzer, the subtitle of your new book encompasses both "the theory of everything" and "big data." Basically these are two different ways of understanding the complex world we live in. How do they differ?

Mainzer: The search for a theory of everything is a very old idea. The great philosopher Plato was a mathematician who was convinced that everything in the world was made up of the accepted basic elements of the time: fire, water, air, earth and the celestial spheres. He believed that these elements were identical to the five regular solids of Euclidean geometry. These are the perfect symmetries of the Euclidean space. So he believed that the world was governed by mathematical laws. Even today, cosmologists are seeking to unite the known fundamental forces of physics and explain them using mathematical symmetry laws. But now, instead of Euclidean geometry, they are applying quantum physics and the differential geometry of the theory of relativity. There is one common idea behind all of this theory-led research: We need a good theory to know what it is we are looking for. Only then can we understand the complexity of the world. Big data, on the other hand, entails collecting data and writing algorithms to process this data. The algorithms recognize patterns, so this is a quicker method than the time-consuming search for laws. In essence, this is a data-driven science.

What are algorithms?

Mainzer: Algorithms are clear formal procedures for solving problems, just like the math we learned at school and the instructions in computer programs. Philosopher and mathematician Gottfried Wilhelm von Leibniz – after whom the supercomputing center in Garching, Germany, was named and who first used the binary numbers 0 and 1 for calculation, i.e., our modern bits – actually believed that all problems could be solved with algorithms if they were suitably encoded with numbers. In the 20th century, logician and mathematician Alan Turing defined a universal formal program allowing the simulation of all algorithms. Such a "Turing machine" could, in principle, simulate the supercomputer in Garching, as well as the program on your smartphone. With Turing's definition of computability, however, it is also possible to decide whether problems can even be solved algorithmically, how complex the solutions are and how long they will take. This has practical cost-benefit implications, because computing time costs money. Turing also considered the philosophical question of whether human thinking can be reduced to algorithms and computer programs. Logician Gerhard Gentzen, a contemporary of Turing, introduced formalisms that allowed the accuracy of algorithms to be verified. But what are the practical implications of all this? Well, take a computer program that controls a production process at BMW. The company would naturally like to rule out programming errors in advance and so avoid accidents and additional costs. What these examples show is that the problems of our ▷

"Algorithmen sind mächtig und hilfreich. Aber alleine sind sie blind."

Algorithmen – vor nicht allzu langer Zeit eher Thema für Nerds und Mathematiker – sind inzwischen in den Feuilletons angekommen. Hintergrund ist die umfassende Digitalisierung mithilfe gigantischer Rechnerkapazitäten, die Grundlage einer neuen Welt ist: der Big Data-Welt, die gerade dabei ist, unser Leben zu revolutionieren.

Von den einen wird sie als ein Eldorado gefeiert, von den anderen als Realisation Orwell'scher Phantasien gefürchtet. Doch welche Perspektiven eröffnen sich tatsächlich damit für die Wirtschaft, das Recht und die Gesellschaft? Welche Chancen und Risiken bestehen? Prof. Klaus Mainzer, Inhaber des Lehrstuhls für Philosophie und Wissenschaftstheorie an der TUM und Direktor der Carl von Linde-Akademie befasst sich in seinem aktuellen Buch "Die Berechnung der Welt – Von der Weltformel zu Big Data" mit diesen Fragestellungen. Bei allen Vorteilen, die Big Data für Wirtschaft, Gesellschaft, aber auch für die Wissenschaften mit sich bringt, hält er es für riskant, sich allein auf eine datengetriebene Methode zu verlassen. Wer mit Tunnelblick auf Effizienz und schnelle Ergebnisse auf die Frage nach Ursache und Wirkung verzichtet, gerät auf dünnes Eis. "Wenn wir das nicht begreifen oder vergessen, fliegen uns am Ende die Algorithmen um die Ohren." Letztlich geht es nicht allein darum, teure Programmierfehler in der automatisierten Produktion zu vermeiden, sondern um größere Gefahren.

Wie er eindrucksvoll zeigt, bringen Big Data-Methoden in der Medizin wenig nachhaltigen Nutzen, wenn Kausalzusammenhänge nicht geklärt werden. Ähnliches gilt für das "Precriming" und andere Versuche, mithilfe der Mustererkennung Kriminalität und Terroranschläge vorherzusehen, um präventiv eingreifen zu können. Das weckt nicht nur ungute Assoziationen mit dem Film "Minority Report", sondern bedroht für ihn ganz klar unsere Gesellschaft in ihren Grundfesten. "Wenn staatliche Sicherheit in totale Überwachung umschlägt, ist die Demokratie verloren."

Birgit Fenzel

Picture credits: Jooss

modern big data world are rooted in fundamental questions of logic, mathematics and philosophy. If we fail to grasp this or forget its importance, then ultimately, algorithms could wreak havoc. This is what my book is all about, and this is what has interested me ever since my student days.

Algorithms are really nothing more than procedures or rules for solving mathematical problems – so nothing to get excited about. Yet now they are being spoken of in revolutionary terms. What is so special about big data algorithms?
Mainzer: Simply put: the sheer volume of data. The Internet and the World Wide Web heralded the first digital revolution: now people can communicate with each other anywhere in the world with small-scale technology like cellphones, smartphones and apps. The first digital revolution was the Internet of people – where people communicate with each other. A computer is no longer a mere computation device. While it continues to process its bits, its real function is as a means of communication. Right now we are embarking on the second digital revolution – the Internet of Things. This will go beyond the interpersonal level to encompass communication between "things." Modern technology in the form of sensors, RFID chips and software interfaces will allow various objects to communicate with each other, independently of human interaction.

That sounds like science fiction...
Mainzer: But we have been living with the Internet of Things for some time already. We experience the Internet of Things through the enormous volume of data and signals produced. To take one example: Google processes 24 petabytes of data per day – that's 6,000 times more than the content of the American Library of Congress. Just one company processing the entire memory of a nation in a single day. Special algorithms are needed to handle such huge quantities of data. They break down the massive volume into subpackages, process these in parallel and search for data correlations and patterns at lightning speed. So in order to find the needle, you need to go through a giant haystack with a fine-toothed comb.

In your book, you refer to the masses of data as the "crude of the future." But in some ways, the hype around big data also recalls the gold rush of the 19th century. But instead of panning for gold in the rivers of the Yukon, the fortune-hunters of the big-data age are now mining for data with algorithms. How do you explain the hype in the business world?
Mainzer: Mining big data and finding correlations can help to predict customer and product profiles with much greater efficiency than ever before. But the gold rush or oil metaphor fits only to some extent.

Picture credits: Jooss

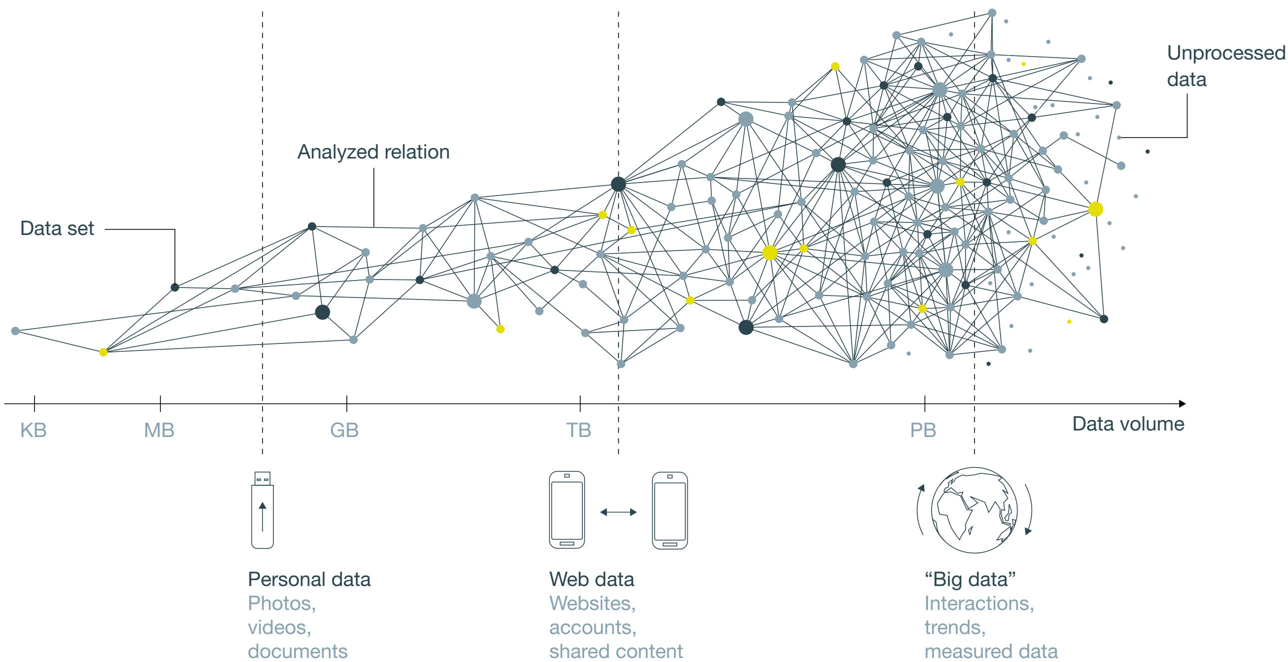
It certainly explains the upbeat mood in the business world – so where does it not sit comfortably?
Mainzer: We deplete natural resources by using them – but this is not the case with data. Data can always be reused in different contexts, creating new ways to make a profit. We will see new business models and value chains emerging. Data owners will earn money by licensing rights of use to their data, as will knowledge workers with skills in mass data management and entrepreneurs with new big data business ideas. The value chain will thus be completely transformed by big data.

How will this affect markets and consumers?
Mainzer: The impact will be huge. We are setting our sights on Industry 4.0, that is, the Internet of Things in the world of industry. Conventional mass production will be replaced by a new system of manufacture “on demand.” Tailor-made instead of off-the-shelf. The ultimate vision of Industry 4.0 is made-to-measure suits as standard. In the automobile industry, many manufacturers are already operating on an on-demand basis, with their suppliers abandoning stockpiling. Another example of where the Internet of Things has been embraced is container ports, where many operations have been automated using robotic vehicles that engage with each other. This world is coming into being at an exponentially fast rate, but hardly anyone is aware of it. Some people are talking about exponential technology and exponential companies that are created with information technology.

The idea of a customer getting exactly what they want sounds great – what’s the catch?
Mainzer: Well, there is a price to pay. An automated process requires trillions of sensors, and in some cases drones and cameras. Tons of data have to be collected, because the machines have to organize themselves. They interact with each other. This also means that the people who are integrated in this process are subjected to intensive scrutiny throughout the entire day. The controller computers know them better than they know themselves. And this is apart from the fact that the customer’s data is also collected. In the end, you will have transparent employees and transparent customers. In this respect, this way of producing data presents a huge social challenge.

Your book also deals with the introduction of big data to science. You use medicine as an interesting example of how these new technologies can benefit the scientific community. What is the state of play here?
Mainzer: Big data will have a massive impact on medical science. Consider that 400,000 papers have already been written on diabetes alone. No one could read all of these in their lifetime. Even while they are reading, the knowledge is growing at an exponential rate. That is why we need intelligent software to filter the information and find the most suitable treatment for the patient.

But you also use medicine to illustrate the potential pitfalls. Where does the problem lie?
Mainzer: If I rely solely on finding efficient data correlations through data mining, then I have not necessarily understood the causes of the disease. Even Steve Jobs, once the ▷



The growing diversity and complexity of Web services is leading to a petabyte-scale data explosion. The term "big data" describes the rapidly growing volume of data on the one hand, but also the systems used to analyze this data.

icon of effective and smart computing, died of cancer, despite using his fortune to exploit all the computing capacity and big data analysis available at the time. Jobs had his cancer cells sequenced at frequent intervals because of the constant mutation of the tumors. His physicians could then continuously adjust his treatment. Ultimately, though, as long as we remain in the dark about the causes of cancer and the mechanism of cancer cells, the analysis of data on a massive scale and the calculation of correlations will be of limited use.

Your book is intended as an argument against the promises of big data and for the continued relevance of basic research and philosophical reflection. But given the growing complexity of data, isn't the search for explanations, causes, theories and laws completely outdated?

Mainzer: It is true that some of the biggest proponents of big data, like Chris Anderson, are talking about the "end of theory." Back in 2002, American computer scientist and software entrepreneur Stephen Wolfram proclaimed a "new kind of science," in which computer experiments would replace mathematical proofs and theories. Wolfram had simulated extensive pattern formations of cellular automata on high-performance computers, discovered some remarkable correlations, and classified the patterns based on his observations. Together with my UC Berkeley colleague Leon Chua, I wrote a book refuting this theory entitled "The Universe as Automaton," in which we proved that it is only the fundamental mathematical laws of cellular automata that allow accurate forecasting and classification of patterns. In our follow-up book "Local Activity Principle," we expanded this argument to examine the emergence of patterns in nature, taking in physics, chemistry, biology and brain research. Here too, we found that it was only when the basic equations were known that accurate declarations and forecasts could be made on the emergence of structure and patterns. What we can generally say about science is that theory is often the best way to solve a problem. How will a mountain of data help me if I don't know what I am looking for? At CERN, for example, a huge amount of data is produced during proton collisions, but only a fraction can be analyzed by the computers now in existence. I have to know what I am looking for. The best example of this is theoretical physicist Peter Higgs, who predicted the particle named after him on the basis of his mathematical model in the framework of quantum field theory. Once experiments had confirmed his model, he was duly awarded the Nobel Prize last year. The Higgs boson is key to explaining how the universe began. But it took Higgs' model to know exactly where to look in the enormous number of events and volumes of data.

Seen in that light, is big data no more than shadow puppetry, like in Plato's cave?

Mainzer: Yes, of course. That is the data landscape of big data. Plato asked what was actually behind the shadows. His answer was the ideas of truth, goodness and beauty. ▷

Picture credit: Jooss/graphics: ediundsepp

“Pythagoras’ theorem is correct not because it has been confirmed in countless surveys, but because it ensues from the axioms of Euclidean geometry. Logically consistent.”

Klaus Mainzer

Picture credit: Jooss

“Theory is often the best way to solve a problem. How will a mountain of data help me if I don’t know what I am looking for?”

Klaus Mainzer

An inscription was said to hang above the entrance to the Platonic Academy: “Let no one ignorant of geometry enter here.” Geometry stood for mathematics in those days. Why? Because Plato believed that, before reflecting on the eternal ideas of truth, goodness and beauty, he must first learn truths that do not rely on observation or changes in perception. And in his view, that is the world of mathematics. Pythagoras’ theorem is correct not because it has been confirmed in countless surveys, but because it ensues from the axioms of Euclidean geometry. Logically consistent. For Plato, that was the first step into the world of unchangeable ideas. But one does not have to be a die-hard Platonist to understand his basic idea: that behind observed data lie structures and laws that must be recognized to understand certain facts. Finding patterns in data using algorithms is useful, but algorithms are blind without theory and laws. On the other hand, it must also be stressed that mathematical theories in the natural and social sciences remain empty without empirical data.

One of the chapters in your book deals with big data in the humanities and cultural studies. In the digital humanities, new automation methods are already in use. For example, metadata from old manuscripts is being algorithmically created to draw conclusions on the source, production conditions and context. This sounds like a very efficient and helpful approach at first, but is there a risk that it will take the “human” out of the humanities, as critics fear?

Mainzer: Here too, the critics have a point. Supercomputers cannot replace the appraisal and interpretation of a literary scholar. But software programs that automatically write standardized texts are already out there. I could easily foresee the same happening for scientific and engineering papers. They would follow certain standards. The software would write the article in the standard English of the particular discipline, and the human authors would add their findings. The citations would be perfectly executed. There would be no room for plagiarism, either, because the software could automatically trawl the Internet to see whether the findings existed elsewhere. There are already writing programs available that adapt themselves to the style of the author. We’ll just have to wait and see if they’re capable of producing Faust, Part 3.

As you have noted, it is practically impossible to predict social behavior, because there are no known equations of motion and evolution for individuals. But early indications from big data show that such predictions are indeed possible thanks to data analysis on a massive scale and the calculation of correlations. Why is that a double-edged sword?

Mainzer: With big data algorithms, we can predict not only product profiles, but also criminal profiles. In “pre-criming,” profiles of criminals are created to prevent them from offending – along the lines of the film “Minority Report.” But where big data differs from the movie is that you don’t need to read minds – you just need the metadata, such as where and when an e-mail was sent, the names of the sender and recipient in conjunction with the criminal profiles, and other data.

The ethical and legal challenge is to protect democratic rights in the age of big data and digitalization. This was made very apparent through the recent controversy surrounding the NSA.

Mainzer: That episode provided further proof that whoever has the biggest search engines, the most storage capacity and the best algorithms rules the world. After the fall of the Iron Curtain, some people got the impression that the US was on the retreat – militarily, economically and politically. But the Snowden revelations showed us that they had reckoned without the NSA and the power of American IT.

One consequence of this is the parallel development of data protection and security technology. It begs the question: Can we safeguard and strengthen personal rights and rights of self-determination on the Internet without compromising freedom through over-regulation?

Mainzer: Europe is not just the birthplace of industry, technology, research and science. Europe is also where the concepts of state under the rule of law and democracy were born. These philosophical ideas originated in ancient Greece, as well as in England, France, Germany and Switzerland – the birthplaces of Kant, Locke, Montesquieu and Rousseau, among others. If we can develop an information technology package that incorporates our rule-of-law and democratic standards, and that is on a par with the Americans’, then we could provide an example to the world along the same lines as we are currently endeavoring with our energy transition policy. It would be great if Germany could accomplish this in a good-hearted, respectworthy way, just like the national team at the World Cup. Also acting in partnership with the US, and not in opposition. Important negotiations on the trade agreement between the US and Europe are forthcoming. We must fight for our legal standards just as strongly as for our food standards, while at the same time demonstrating that we, too, are masters of technology and science. That would be my dream scenario at any rate.

Interview by Birgit Fenzel