femto

The DESY research magazine – Issue 01|15

ZOOM

BIG DATA

What use is the deluge of data?

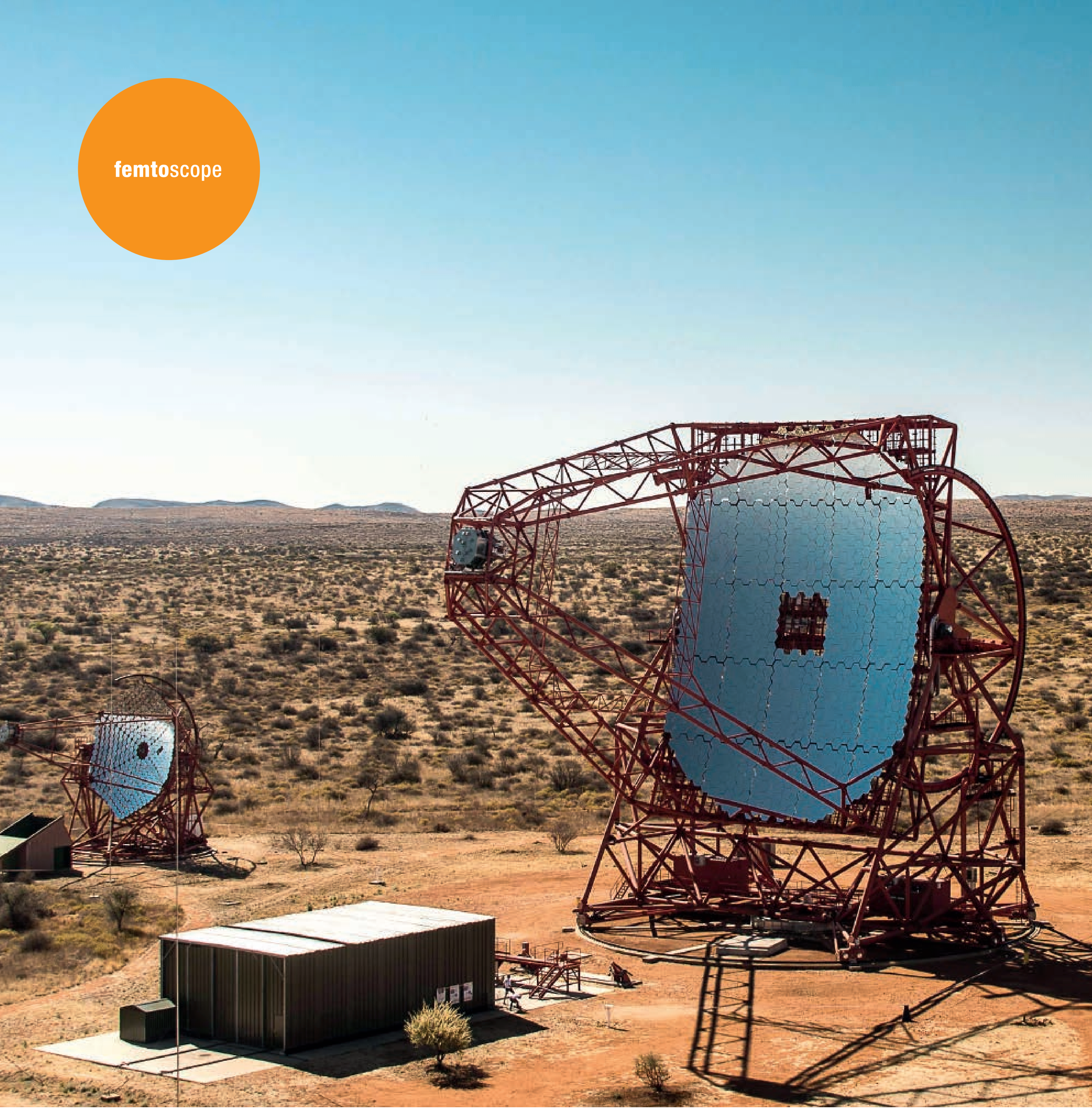
Moonstruck
Peering into the eyes of a bizarre marine midge

Superslow motion
Magnetic memory at work

Caught in a flash
Free nanoparticles shown for the first time in 3D

DESY

femtoscope



In the Namibian desert, the gamma-ray observatory H.E.S.S. (High Energy Stereoscopic System) is searching for cosmic sources of high-energy radiation. Gamma radiation is created when particles in cosmic objects undergo extreme acceleration and interact with gas or light in their surroundings. The H.E.S.S. team has now discovered a new class of gamma-ray source – the superbubbles. The superbubble known as 30 Dor C was probably formed by several supernovae and strong stellar winds.

Space seen from the desert

Picture: Christian Foehr / H.E.S.S. Collaboration

It's a shell-shaped structure, 270 light years across, and has now been shown to be a source of extremely high-energy and intense gamma radiation. Astrophysicists used the reflecting telescopes to observe the Large Magellanic Cloud – a galaxy neighbouring the Milky Way – for a total of 210 hours. The rate at which stars are formed and, at the end of their lives, explode as supernovae is especially high there. In addition to the superbubble, the astrophysicists also discovered two other extremely bright gamma-ray sources: the pulsar wind nebula of the most powerful pulsar seen to date and an intensely radiating supernova remnant. “This shows the huge discovery potential behind the use of combined reflecting telescopes,” says H.E.S.S. spokesperson Christian Stegmann of DESY. The scientists are already planning for the future with the Cherenkov Telescope Array (CTA), which should provide even more sensitive and higher-resolution images of gamma radiation starting in 2020.

Science, 2015; DOI: 10.1126/science.1261313

Contents

ZOOM

BIG DATA

What use is the deluge of data?

Page 14–31

When it comes to quantifying information, kilobytes and megabytes are small fry. In today's world of more and more data, gigabytes are standard, terabytes are nothing special, and even petabytes are soon racked up. The DESY computer centre alone holds some 11 petabytes – i.e. 11 million gigabytes or 11 quadrillion bytes – of scientific data. With more and more data being collected and analysed at an ever faster rate, a new buzzword has emerged to describe this trend: big data. In addition to providing us with new insights, big data also creates new challenges.

CAMPUS

ZOOM

SPECTRUM

Science in brief

Moonstruck

Researchers are peering into the eyes of a bizarre marine midge

At low water, soon after the full moon, an amazing spectacle plays out in Europe’s intertidal zones. In the evening hours, as if to a secret command, myriads of marine midges emerge in unison and begin a frenetic mating dance. In an hour or so, their life as adult insects will have come to an end. During that time, the male must find a female, help free her of her pupal case and fertilise her. His work completed, the male then dies, leaving the female to hastily lay a tube-like clutch of eggs before she, too, expires.

Despite the briefness of it all, such breeding habits make biological sense. In this inhospitable environment, synchronised reproduction makes it easier for the marine midge to find a mate. Full moon is the ideal moment for this ritual, when the united powers of the sun and moon produce the so-called spring tides. At this time, high water is exceptionally high, and low water exceptionally low, thereby exposing more of the seaweed and algae that provide a habitat for the marine midge larvae.

The question is: how do the larvae see the full moon?

But how do the larvae know when the spring tides are coming? “The marine midge has four different clocks: a circadian clock, a lunar clock, an annual clock and a tide clock,” explains Gerta Fleissner, a chronobiologist at the Goethe University in Frankfurt am Main. “These clocks are synchronised by various signals: the tide, moonlight, twilight and so on.” While the circadian rhythms have been very well investigated, the much longer lunar cycle is considerably more complex to analyse.

The marine midge *(Clunio marinus)* passes through a total of four larval stages. In the third of these stages, the larva needs to see the full moon to cast its skin and reach the fourth and final larval stage: the pupa. It is the light of the full moon that triggers this change and thus determines when the insect emerges from the pupa. “The basic question,” says Fleissner’s husband and research colleague, Günther Fleissner, “is: how do the larvae see the full moon?”

The eye as light meter

The Fleissners’ hypothesis is that the larva compares, from one night to the next, the moment of maximum brightness, i.e. it analyses the dynamics of the moonlight. “The exciting thing is that to do this, the marine midge must have eyes that register contrast successively rather than simultaneously,” explains Günther Fleissner. “We’ve discovered that in the run-up to the full moon, the marine midge’s eyes change in function. During this period, the eyes work like a light meter. In other words, the eyes change their biological function, and they become more sensitive. This alteration lasts about a week.” It results from changes to the pigments in the eye.

This is an unusual observation inasmuch as the normal shielding pigments, which are known as ommochromes, do not change at all. The change in function must therefore come from another pigment in the eye of the marine midge. Study of the literature led the Fleissners to the very common pigment melanin, which is also present in human skin, where it is responsible for, amongst other things, producing a sun tan. Unlike ommochromes, melanin contains various metals, which are visible under X-rays.

It was here that DESY’s ultrabright X-ray source, PETRA III, entered the picture. “We scanned the midge with a very fine X-ray beam and discovered an extremely high metal content at the location of the eyes,” explains DESY researcher Gerald Falkenberg, the head of the experimental station where the investigation was conducted. “This was the indication that it was indeed melanin.”

On the trail of melanin

Various types of melanin exist, including the darker pigment eumelanin and the lighter pigment pheomelanin. “It looks very much as if *Clunio marinus* has found a way to change melanin from one form into the other,” Gerta Fleissner explains. “If so, that’s very unusual, because the two proteins are produced on the basis of different processes. Normally, only one of the two forms of melanin is produced.” Metals seem to play a key role in this change in transparency.

Thanks to the high spatial resolution of PETRA III, the researchers from Frankfurt were able to investigate individual melanosomes in the marine midge’s eyes and observe details on a scale of half a millionth of a metre (0.5 micrometres). Melanosomes are collections of melanin enclosed in tiny

The marine midge has four different clocks: a circadian clock, a lunar clock, an annual clock and a tide clock

bubble-like membranes. The investigations showed that the concentrations and distribution of calcium, copper, zinc, nickel and other metals in the melanosomes changes with the phase of the moon. This observation supported the hypothesis that the two proteins occur alongside one another in the pigment granules and that it is solely the relative concentration of light and dark melanin that determines the transparency of the shielding pigment.

Pictures: Alfred Wegener Institute

Research applications for biomedicine

These changes in concentration go along with the observed change in function: whereas the larva's eyes are almost completely opaque during the darker nights when the moon is new, they become transparent around the time of the full moon. And thus, instead of functioning as an imaging device which, like a pinhole camera, allows the passage of light from one direction only, the eyes become a light meter.

A precise investigation of this process provides more than just an understanding of the bizarre lifecycle of the moonstruck marine midge. "Melanisation is a big topic for biomedicine, because the melanins are free-radical scavengers, and melanisation plays a key role in neurodegenerative diseases," says Gerta Fleissner. "And we think that *Clunio marinus* can tell us something about this."

Germany has only one rocky island: Heligoland. On the west and northeast coasts of the island, in particular, large shore platforms dry out at low tide. This is the habitat of a wide variety of algae, snails, crustaceans and other creatures that have adapted to the changing conditions of the rocky foreshore. The moonstruck marine midge lives in the feathery threads of algae and the tangled roots of larger types of seaweed.

At femtosecond intervals

It sounds rather matter-of-fact that the new X-ray free-electron lasers generate extremely intense flashes at a time scale measured in femtoseconds. But what's a femto? It is the factor that follows the milli, micro, nano and pico, and is completely beyond our ability to conceptualise. A femtosecond is a quadrillionth of a second – an unimaginably short period of time, especially if you consider that the reaction time of a human being is about one tenth of a second. During such a tenth of a second, 100 trillion femtoseconds will go by. Chemical reactions occur on the femtosecond scale, as does the movement of biomolecules. This is what makes femtoseconds so interesting for researchers. It also makes the new X-ray lasers so spectacular, because their extremely intense flashes enable scientists to achieve exposure times of a few femtoseconds. This will enable scientists to shoot entire films of biochemical processes or the behaviour of nanomaterials.

But before researchers can do that, they not only need X-ray lasers that emit light at femtosecond intervals – the entire complex facility also has to be synchronised to operate with femtosecond precision. That's why a group of developers headed by Holger Schlarb and Sebastian Schulz from DESY and Adrian Cavalieri from the Max Planck Institute for the Structure and Dynamics of Matter have created a purely optical system that coordinates all of the components of DESY's FLASH X-ray laser at a previously unprecedented level of precision. The system makes FLASH the most precisely synchronised X-ray free-electron laser in the world.

The optical system is approximately 10 times more precise than the electronic systems that are generally used for synchronisation. "It allows us to synchronise all of the independent accelerator subsystems and external experiment lasers at FLASH with a precision of 28 femtoseconds – a mere fraction of the X-ray pulse duration of 90 femtoseconds," says DESY physicist Holger Schlarb. The best synchronisation that electronic systems can achieve is about 100 femtoseconds – a restriction that limits the scientific potential of X-ray free-electron lasers.

Capturing chemical reactions on film

The new high-precision synchronisation delivers X-ray flashes with stable characteristics at the right time. "Doing that at the femtosecond scale is hard enough," says Schlarb. "However, the experimentalists need more than that – they require simultaneous, independent optical laser pulses, which we generate with laser systems in the FLASH experimental hall." Such laser flashes in the visible light region are needed to initiate the ultrafast processes that are to be filmed. The "filming" of chemical reactions or phase transitions

The reaction time of a human being is about one-tenth of a second. This amounts to 100 trillion femtoseconds

The FLASH free-electron laser is 300 metres long overall. The electrons are accelerated in the yellow modules.

Holger Schlarb helped to develop the optical system that coordinates all of the FLASH components with an extremely high level of precision.

At FLASH, all of the independent accelerator subsystems and external experiment lasers can be synchronised with a precision of 28 femtoseconds.

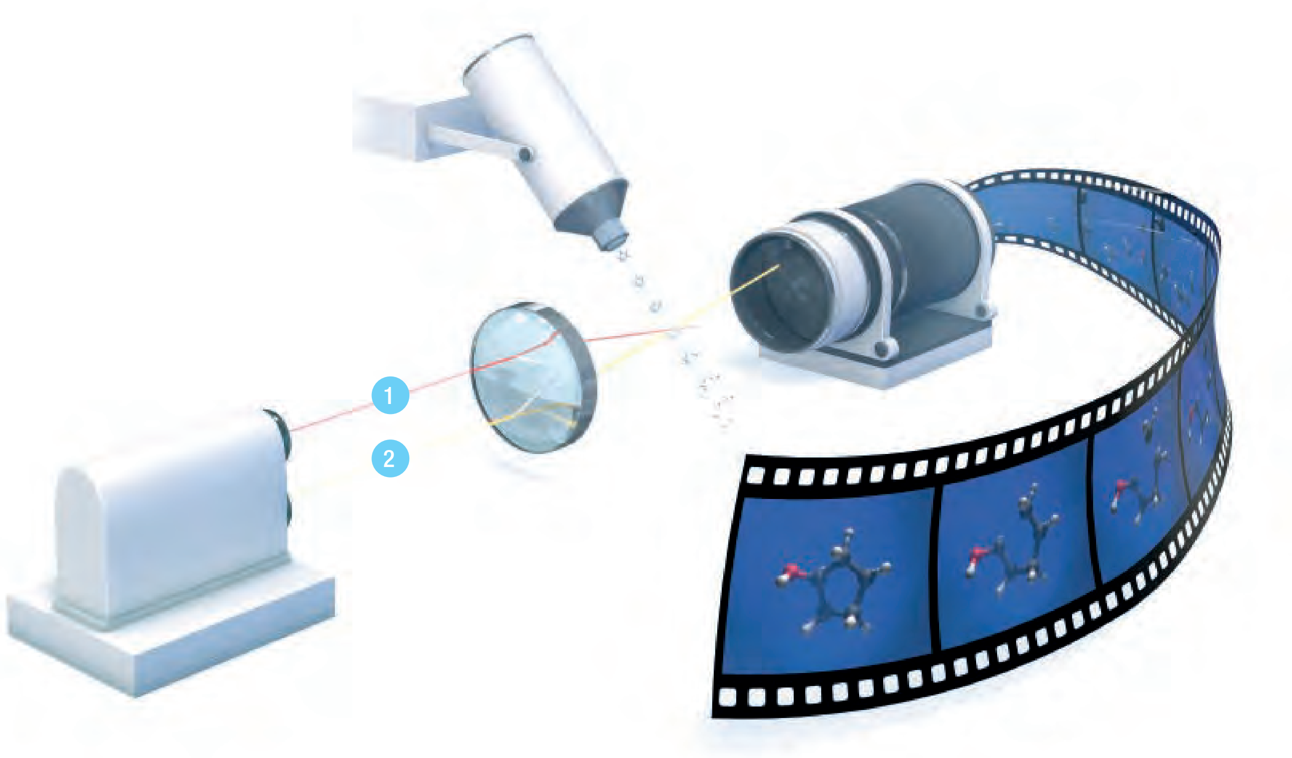
Picture: Heiner Müller-Elsner, © DESY

in materials consists of innumerable snapshots that are strung together like a flip book. The optical laser flash thus triggers a reaction of which the X-ray laser flash takes a snapshot. Because the sample is destroyed in the process, the reaction has to be restarted using a fresh sample. The X-ray laser then takes the next image at a slightly later moment in the reaction. This procedure is repeated until the “film” is completed. This complex process can only work if the optical and X-ray lasers are coordinated with extreme precision. Only then can the snapshots really document the reaction with an accuracy of a few femtoseconds.

To achieve the maximum possible temporal resolution in such studies, the optical laser pulse should be delivered with a timing precision that is only a fraction of the X-ray laser pulse duration. The optical synchronisation system demonstrated for the first time at FLASH does just that. As the quality of the optical synchronisation is limited primarily by the duration of the X-ray flashes, the precision should be even better with shorter X-ray pulses. Such precise synchronisation would not only allow ultrafast experiments with highest temporal resolution, but also open up new possibilities for the further development of X-ray free-electron lasers, and in particular of the future European XFEL X-ray laser, which is currently under construction from the DESY campus to the neighbouring town of Schenefeld.

Optical synchronisation enables ultrafast experiments

Nature Communications, 2015; DOI: 10.1038/ncomms6938



Molecular cinema – how to film a chemical reaction

Researchers use extremely short X-ray flashes to follow chemical reactions. Essentially, the process works like this: A light pulse (1) triggers the chemical reaction. Shortly thereafter, an X-ray flash (2) generates an image, and a detector then records what occurred during the reaction. A series of such snapshots at different intervals produces a “film” of the reaction.

Picture: DESY

femtopolis

Perfect sample holders

Why cactus needles are unbeatable in high-tech application

The large experimental hall of the cutting-edge X-ray source PETRA III contains row upon row of the latest generation of high-tech components and detection instruments. The offices are functionally furnished. They contain no superfluous knickknacks – except for a small green cactus on a window sill. Does that mean a scientist has tried his or her hand at interior decorating? Is the prickly ball perhaps there to create a feeling of cosiness? Not in the least, because even cacti have a very specific purpose at DESY. The researchers want the cactus for its needles, which make perfect sample holders. The longer and sharper the needles are, the better. Cactus needles aren't just inexpensive and always available – unless you manage to let even an undemanding desert plant die of thirst. Most importantly, they are also almost invisible in the X-ray light that the scientists use to examine samples. These samples generally have a crystalline structure that diffracts the X-rays in a characteristic way. Scientists can use such diffraction patterns to determine the internal structure of proteins or new materials down to the last atom.

By contrast, the organic carbon found in the needles of many cactus species isn't crystallised and hardly diffracts the intense X-rays produced by the PETRA accelerator at all. As a result, the sample's diffraction pattern doesn't get falsified. This is especially important when investigating biological samples, which produce only weak diffraction signals. To precisely analyse the structure of a biomolecule, the researchers first create a 10–20 micrometre wide crystal of the molecules in question. Under a microscope, the crystal is then attached to the tip of a cactus needle with a tiny amount of adhesive. Although the needle is very narrow at its tip, it is broad and stable at its base, so the sample can be firmly and securely positioned in front of the detector, where it is irradiated with X-rays.

Neither steel needles nor thin glass cannulas can be used as sample holders because the atoms in their crystal lattices scatter X-rays much more strongly than do biological materials. Glass is also too brittle. Commercially available carbon fibres are not suited to the task either. Although they don't scatter X-rays very much, they are too thin and flexible to hold the samples securely. That's why cactus needles have become the sample holders of choice for experimenters. But not all cactus needles are alike, and some are not suited for use as sample holders. Some are too thin or have a bent tip, while others contain silicate, which scatters X-rays too much.

The only way to solve this problem is to go to the local DIY centre, where a wide selection of inexpensive cacti is available to be tried out. Species that have sharp, long and rigid needles have proven to be especially effective. A small pillar-like cactus turned out to be particularly well suited, which is why it has already sacrificed half of its needles to research. Although this cactus is no longer attractive, the scientists didn't allow it to dry out even during the months in which no measurements were conducted at PETRA III. The cactus simply belongs here!

Cacti also appeal to DESY physicist Oliver Seeck.

Picture: Heiner Müller-Elsner, © DESY

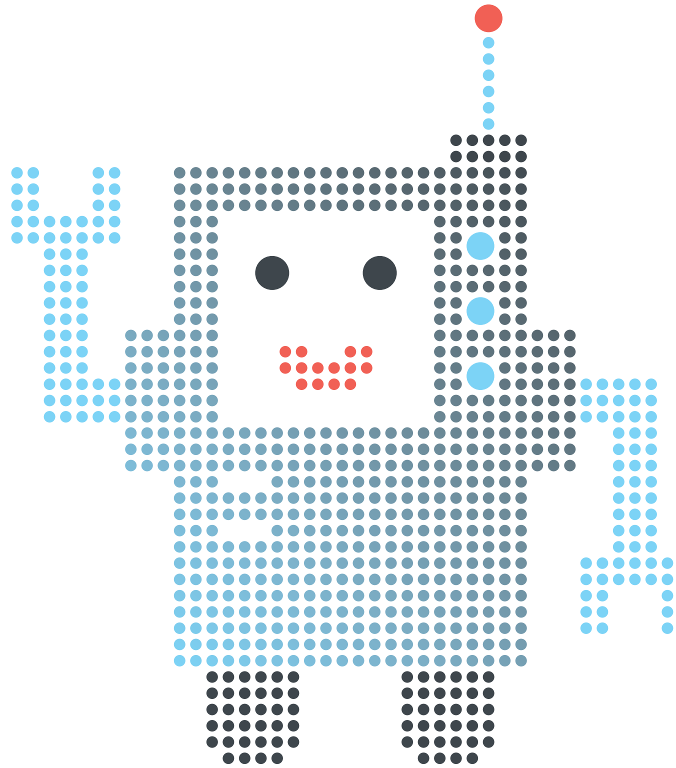
ZOOM

BIG DATA

What use is the deluge of data?

When it comes to quantifying information, kilobytes and megabytes are small fry. In today's world of more and more data, gigabytes are standard, terabytes are nothing special, and even petabytes are soon racked up. The DESY computer centre alone holds some 11 petabytes – i.e. 11 million gigabytes or 11 quadrillion bytes – of scientific data. The LHC, the world's largest accelerator, generates 25 petabytes of experimental data a year. And that's not to mention the floods of data that flow through social networks like Facebook and Twitter. With more and more data being collected and analysed at an ever faster rate, a new buzzword has emerged to describe this trend: big data. In the world of research, making intelligent use of such data flows can yield new insights. And, in many scientific disciplines, these enormous volumes of data have themselves become the object of new forms of research. Society, meanwhile, faces big challenges to ensure that sensitive personal information is properly protected.

TERA, PETA, EXA



Scaling the mountains of data to reach new insights

"I'm afraid it's rather loud in there," says Patrick Fuhrmann, as he opens the door to the DESY computer centre. He has to raise his voice above the penetrating drone of the air conditioning and the ventilator fans of countless computers and hard disks. The latter are racked in tall cabinets with banks of flashing LEDs and thick hanks of cable bulging out at the back.

DESY physicist Patrick Fuhrmann handles huge volumes of data.

Fuhrmann goes to the rear of the computer centre, where a huge container stands. "That's a robot that handles data on magnetic tape," he says, indicating the glass window in the wall of the container. Inside are thousands of tapes stacked on shelves, just like a library. Robotic arms whizz along special tracks, pluck tapes from the shelves and transport them to the tape drives. "It's our tape archive," Fuhrmann explains. "Storing our data on magnetic tape provides a safety backup, should one of our hard disks crash, and also gives us a long-term archive."

There is a huge amount of data stored in the DESY computer centre: all in all, 11 petabytes of scientific data – i.e. 11 million gigabytes – on several thousand hard disks, and around the same amount on magnetic tape. The amount of processing power is impressive, too: some 15 000 processor cores analyse data 24 hours a day, 365 days a year. "Over the years, the data volume here has literally gone through the roof," says Volker Gülzow, Head of IT at DESY. "Fifteen years ago we were dealing with just a couple dozen terabytes. That would now fit on a few of the latest generation of hard disks." Now, in 2015, the amount of data is almost one thousand times greater.

The flood of data at the DESY computer centre is symptomatic of a global trend. It is estimated that the volume of data worldwide currently doubles every two years. As computers, hard disks and data transmission become ever more powerful, they now have the ability to process and store huge amounts of bits and bytes. One result has been an exponential increase in the data traffic flowing through the Internet and, in particular, social networks such as Facebook and Twitter. Using new and sophisticated algorithms, companies like Google and Amazon are able to analyse this mass of data and target users with personalised advertising. At the same time, intelligence agencies sift through mountains of communications in the hope of tracking down terror suspects and enemies of the state. The catchphrase used to describe these increasingly diligent efforts to gather and analyse information is both snappy and vague: "big data".

The amount of data worldwide doubles around every two years

Mountains of data in Geneva

Science, too, has undergone its own data explosion. The prime reason for this is that the sensors, special cameras and detectors used to observe experiments have become ever more powerful. They provide greater resolution, sharper images and faster readings – and thus larger amounts of data in a shorter period of time. One result is that it becomes more and more difficult to transmit, store and analyse these data.

"The old technologies are no longer up to the job," Gülzow explains. "We need to strike out in new directions." This means, on the one hand, innovative hardware and software to gather and reliably store the data and, on the other, quick algorithms to analyse them. To crunch the mountains of raw data and sift out new scientific findings, innovative and sophisticated computer programs are required.

"Over the years, the amount of data here has literally gone through the roof". Volker Gülzow, Head of IT at DESY.

There are 11 petabytes – i.e. 11 million gigabytes – of scientific data in the DESY computer centre

A prime example is the world's largest accelerator, the Large Hadron Collider (LHC), at the European particle physics laboratory CERN near Geneva. Installed in a 27-kilometre ring-shaped tunnel, the LHC accelerates hydrogen nuclei to enormous energies. When these are made to collide, this can result in the creation of previously unknown elementary particles, such as the Higgs boson, detected in 2012. Such experiments are incredibly complex, however, and rely on enormous volumes of data.

For a start, these high-energy collisions of hydrogen nuclei rarely produce a Higgs particle – on average, only one is created in every 10 billion collisions. Moreover, this exotic particle can only be detected indirectly, as it is unstable and decays into countless other particles immediately after creation. It is these other particles that are registered and measured by the huge detectors. And it is only after exact analysis of extremely many such particle tracks that researchers can be fully sure they have generated a Higgs boson.

Rigorous selection

The challenge begins with the actual data taking. The LHC detectors comprise millions of individual sensors, which deliver signals several million times a second, thus producing a veritable deluge of data. Every second, ATLAS – one of the two

In the ATLAS detector, 150 million individual sensors take 40 million measurements a second. That amounts to 1000 terabytes of raw data a second. However, only around a millionth of this is stored.

large LHC detectors – generates some 1000 terabytes. That's enough to fill around 500 standard computer hard disks.

Under normal operating conditions, the LHC runs around the clock. In other words, it produces much more data than can feasibly be stored. Fortunately, a lot is irrelevant, as CERN physicist Michael Hauschild explains: "Most of the collisions produce particles we already know." There is therefore no need to store these data. Instead, they are eliminated by sophisticated electronics before they leave the detector.

"ATLAS has three filters," Hauschild explains. "The first comprises chips programmed to make an initial selection within a few millionths of a second and reject those data that are definitely of no interest." The signals that pass this first barrier must then go through two further filters. These are based on special

Only one collision in every million is analysed

software algorithms that run on server farms with hundreds of computers and eliminate all data that is likely of no interest for further analysis. At most, of the 1000 terabytes per second, only one gigabyte per second remains. In effect, this means that only one collision in every million is analysed. But how can physicists be sure that interesting events are not being inadvertently eliminated? To monitor for this eventuality, they always allow the data from a small number of collisions to pass through unfiltered. "These control events help us check whether our filter parameters are right," says Hauschild. "To date, we seem not to have missed out on anything."

Global computer network

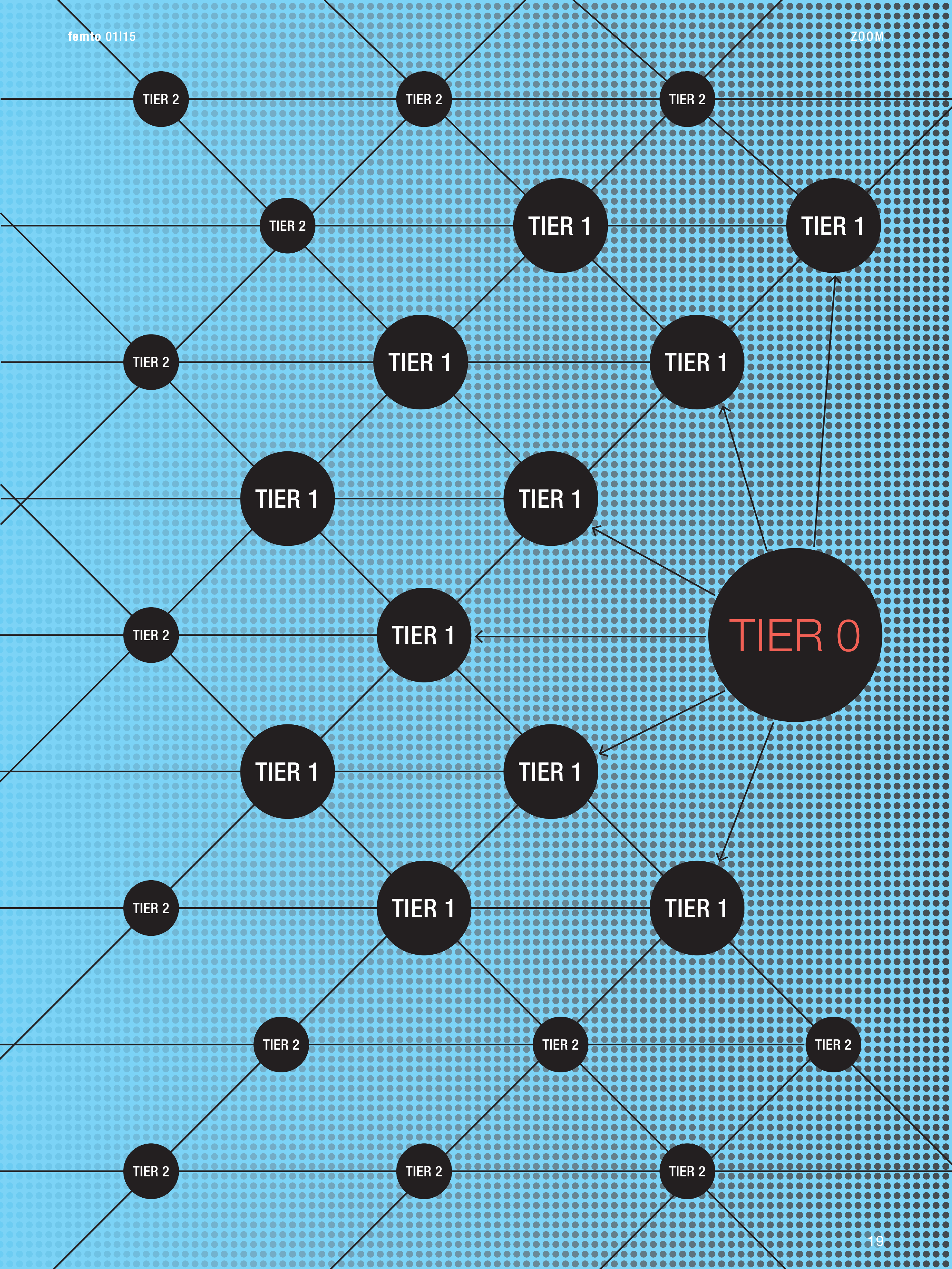
Yet even a reduced number of signals produces a downright flood of data. Each year, the LHC experiments generate around 25 petabytes of measurement data. To handle this huge volume, scientists have developed a special computing infrastructure known as the Worldwide LHC Computing Grid. Unlike the World Wide Web, this not only enables data transfer but also provides access to computing resources around the globe. "Whenever scientists want to evaluate data, they upload their analysis program to the grid," explains Joachim Mnich, Director of Particle and Astroparticle Physics at DESY. "And this program does the processing wherever there are computers with available capacity."

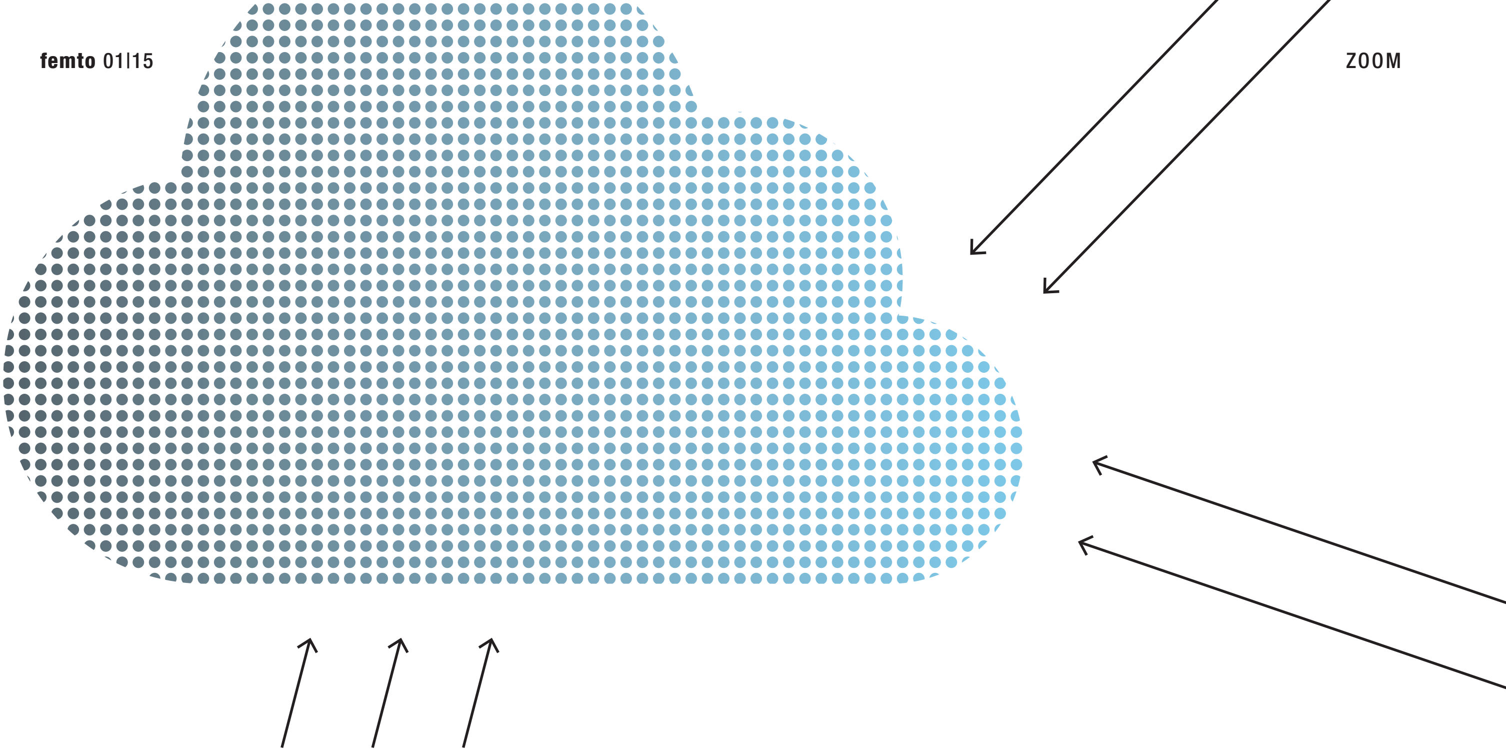
For example, whenever an LHC researcher at DESY wishes to analyse data, the complex algorithms required for this process are not necessarily run at the DESY computer centre. Instead, this operation is performed wherever there is available capacity. This might be at a German Helmholtz centre or university or at another institute in Europe, Asia or the USA. Special software determines where the relevant data are located and where these can be analysed as quickly and as inexpensively as possible.

Grid computing

To cope with the flood of data from the LHC, IT experts use a concept known as grid computing. This means that computers around the world are linked in such a way that they function together like a huge supercomputer. More than 170 computer centres in nearly 40 countries are involved in the Worldwide LHC Computing Grid (WLCG). The grid is made up of various levels, called tiers. The raw data generated by the LHC are first backed up on tape at CERN in Geneva, the Tier 0 centre. Following initial processing, these data are then distributed to a number of Tier 1 centres around the world for storage and reconstruction. These centres have the requisite storage capacity for large data volumes and are constantly connected to one another through the grid.

In turn, the Tier 1 centres make the data available to the Tier 2 facilities further down the grid hierarchy. These facilities, which comprise either a single computer centre or a number of interconnected ones, constitute the decisive level for the scientific analysis of the data using grid tools. DESY operates such a Tier 2 centre, spread between its Hamburg and Zeuthen sites, for the LHC experiments.


TIER 0
TIER 1
TIER 2



If possible, data are processed wherever they happen to be, so that they do not have to be sent back and forth around the globe.

It took almost 10 years to devise and set up this innovative network. “By the time the LHC went into operation, the grid was already up and running,” says Fuhrmann. “Back in March 2010, only a few hours after the LHC had delivered its first high-energy collisions, the data had already arrived for analysis at computer centres around the world.”

Whenever scientists want to evaluate data, they upload their analysis program to the grid

This method also enabled the discovery of the Higgs particle. Researchers combed through the two trillion or so collisions recorded by ATLAS and CMS, the second large LHC detector, up until the end of 2012. This required a huge amount of computing capacity, which was provided with the help of the grid plus small computer centres such as the National Analysis Facility (NAF) at DESY. This facility ensures particle physicists throughout Germany rapid access to LHC data and provides them with additional resources for data analysis.

Intelligent storage

Another crucial element of the grid is a technology developed by DESY together with the US research centre Fermilab and the Nordic Data Grid Facility (NDGF). The system, known as dCache, enables intelligent management of huge volumes of data. Not only must the grid’s storage facilities collect data reliably, they also have to ensure the information is accessible from anywhere in the world. “dCache automatically moves data back and forth between different storage media such as hard disks and magnetic tape, according to what’s best for the user,” explains Fuhrmann. “The system acts like a buffer to ensure that data are constantly available for processing, irrespective of where they come from.”

In addition, dCache reduces costs. As it saves multiple copies of a file in different locations, less expensive hardware can be used. If one of the storage media should fail, the system knows exactly where to find another copy of the file. Today, around half of all LHC data is stored in dCache systems around the world. Moreover, the technology is now used for other scientific projects – including the European radio telescope LOFAR and the Helmholtz project Supercomputing and Modelling for the Human Brain (SMHB) – and also by business organisations such as companies from the banking sector.

Research cloud

Notwithstanding the success of the LHC grid, a new development is already on the horizon – the transition to a research cloud. Cloud computing is already used in the commercial sector, where, in simplified terms, it offers companies the chance to rent – online and according to need – not only data storage but also processing time. Yet there is a downside to such services. “It’s not clear at all whether data security and privacy can be properly guaranteed,” says Fuhrmann.

As a result, these commercial services are not really suitable for use by the scientific community – which is therefore working on its own cloud infrastructure. In essence, this would work as follows: scientific computer centres throughout Europe, which to date have largely operated independently, join forces to form a shared computing and data cloud. The LHC grid would be either totally or partially integrated within this new cloud, the aim being to create a system that is significantly easier to use than the grid. This would then attract users not only from particle physics research – a community typically well versed in IT – but also from other disciplines such as biology, chemistry and materials science.

The future belongs to cloud computing

However, there are still several problems to be overcome until then. How, for example, can users be securely identified? And how can participating computing centres charge the others for the capacity they provide? These are questions to be addressed by the European Grid Initiative (EGI) and by a recently approved EU project by the name of INDIGO-DataCloud, in which DESY is involved. DESY has also developed a prototype for its very own internal data cloud based on dCache, which will soon be ready for rollout.

Huge amounts of data from photons

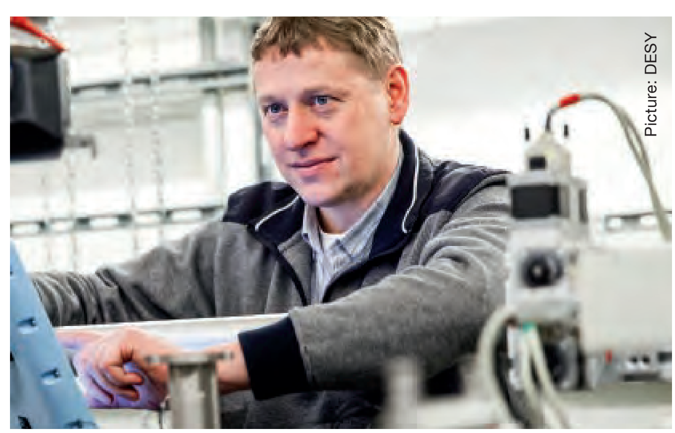
Aside from particle physics, there are other fields of research where big data is making its presence felt. These include research with particle accelerators used as sources of extremely bright light. These devices accelerate electrons to virtually the speed of light and then divert them through special magnetic fields. These cause the electrons to follow a zigzag course, which in turn induces them to emit extremely collimated high-intensity radiation – primarily X-rays, but also UV light. Researchers from a range of disciplines use this radiation to conduct detailed examinations of a whole range of samples. These include nanoparticles, semiconductor materials and protein molecules as well as rock samples, archaeological artefacts and even famous paintings.

Two such research light sources are operated in Hamburg, both of them among the world's best. PETRA III, a ring accelerator over two kilometres in circumference, produces brilliant flashes of high-intensity X-ray radiation. The free-electron laser FLASH generates ultrashort pulses of X-ray radiation with laser-like properties. Meanwhile, work continues to create an unparalleled new light source: the European XFEL X-ray laser, due to start operation in 2017. Almost three and a half kilometres in length, this facility will produce ultrashort high-intensity X-ray laser flashes.

All of these light sources generate ever increasing amounts of data. There is a simple reason for this: the detectors that record how the X-rays are scattered by the samples under examination are getting better and better. Just like the latest digital and mobile phone cameras, which always have more megapixels than their predecessors, the resolution provided by the X-ray detectors is constantly rising. The frame rate too is growing rapidly: while FLASH typically delivers 300 flashes a second, the European XFEL will produce 27 000 of them. In addition, robots are increasingly being used to exchange samples fully automatically, which enables measurements to be taken more quickly. And, last but not least, where researchers would once have been satisfied with taking a few still shots of a sample, today they increasingly tend to shoot actual films – which, naturally, makes data volumes skyrocket. "The European XFEL is due to start operation in 2017," says Gülzow. "And it could produce up to 100 petabytes of data a year."

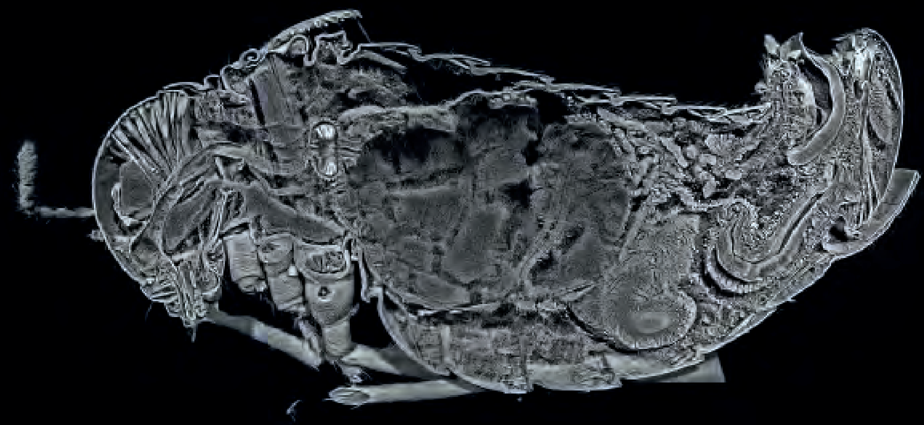
X-ray films in 3D

A prime example is the field of microtomography. Like a CT scanner in a hospital, this technique generates 3D X-ray images. Compared to medical uses, however, the researchers work with substantially more powerful radiation, which enables a much



Picture: DESY

"In real-time tomography, 100 images per second are recorded – that pushes data volumes through the roof". Felix Beckmann, expert in microtomography

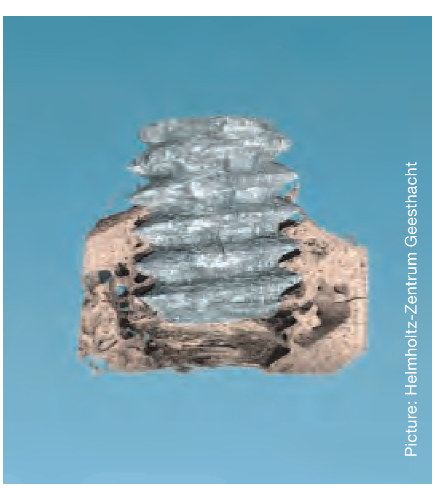


A microtomogram of a flea-like insect (Caurinus sp.) that plays a key role in the research of insect evolution. Length: two millimetres; resolution: less than 1.22 micrometres (thousandths of a millimetre).

higher resolution. This even provides images of microscopic details, such as the structure of the tiny ossicles in the human ear, the teeth of fossilised dinosaur embryos, or the pores and cracks in a high-tech alloy.

"Producing a tomogram typically requires us to store around 12 gigabytes of data per minute," explains Felix Beckmann from the Helmholtz-Zentrum Geesthacht, which operates a materials science outstation at DESY. "That's still manageable." The problems start when there's a need to "film" complex processes. "For such real-time tomography, we have to record 100 images a second," says Beckmann. "That pushes data volumes through the roof."

The technique provides fascinating insights. Researchers from Karlsruhe, for example, have been studying the movement of a beetle's hip joint, while experts from the Helmholtz-Zentrum Geesthacht are investigating how fast certain implants corrode – an important piece of information concerning their behaviour in the human body. Other scientists, meanwhile, are looking in detail at how shape-memory alloys change their form at certain temperatures.



Real-time tomography serves to investigate dynamic processes such as the degradation of a bone screw made of a new type of magnesium alloy.

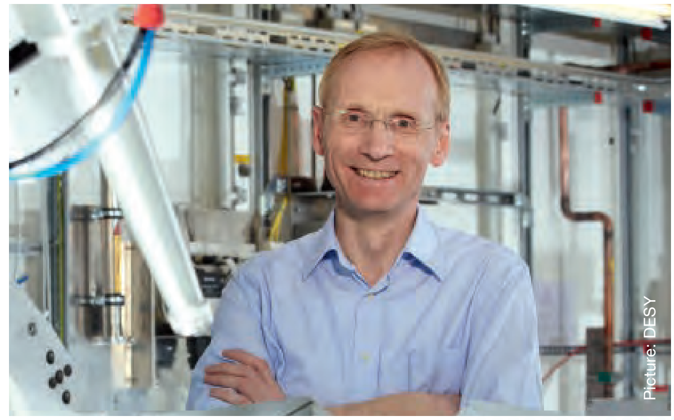
Yet the data-processing challenges are enormous. A single measurement can quickly generate terabytes of data. "In the past, we've been able to transfer the data directly to the DESY computer centre," Beckmann explains. "But, increasingly, we're having to install extra buffers just in order to be able to handle the flow of data."

Imminent storage crisis

Structural biologists are facing similar challenges. Using an X-ray beam from the accelerator, they can determine the precise atomic structure of biomolecules – a key factor in the development of new drugs. "At present, each of our experimental stations at PETRA III generates one terabyte a day," says Thomas Schneider, a scientist at the European Molecular Biology Laboratory (EMBL), which has an outstation in Hamburg. With the next generation of detectors, which could be installed beginning next year, this should rise to 20 terabytes a day, that is, around five petabytes a year.

"The present infrastructure for archiving and data transfer is not up to handling such huge volumes," Schneider explains. At the same time, another factor threatens to aggravate the situation. The Deutsche Forschungsgemeinschaft, Germany's largest independent research-funding body, now recommends that any raw data that have led to a scientific publication should be kept for a minimum of 10 years. That way, it will be possible to check whether a scientist later accused of fraudulent research was in fact guilty of such misconduct.

At present, however, it is unclear who should be responsible for this long-term data storage – major research institutes such as DESY or the scientists from the universities that make use of the accelerator facilities in Hamburg. “Presumably, institutes like DESY will be obliged to take responsibility, since the universities simply can’t afford to do so,” says Edgar Weckert, Director of Photon Science at DESY. “But then we’ll be in danger of drowning in all the data.”



“In the future, there’s going to have to be some pre-sorting of datasets so as to retain only those that are scientifically relevant”. Edgar Weckert, Director of Photon Science at DESY

To prevent that happening, experts are working on new approaches – merely the purchase of new computers and hard disks will not suffice. “One possibility is to use intelligent data reduction algorithms,” says Weckert. “At present, the experimental stations usually store all raw data. But in future, there’s going to have to be some pre-sorting of datasets so as to retain only those that are scientifically relevant.”

Many areas of science face the challenge of how to deal with enormous quantities of data

Slimming down data

In the world of digital photography, such data compression is common practice. Saving an image as a JPEG reduces the volume of data to a fraction with virtually no loss. Researchers could use similar algorithms to significantly curb the flow of data from X-ray detectors. Depending on the type of experiment, a reduction by a factor of 10 seems feasible. However, a substantial amount of development work is still required to make this possible. In particular, the compression algorithms need to become much faster so that the data can be written to hard disks quickly enough.

Yet even with such a reduction, researchers at PETRA III will still have to deal with a huge amount of data. DESY has therefore teamed up with the computer manufacturer IBM to develop a powerful new solution for the storage of large volumes of data. Known as “Elastic Storage”, this system is designed to handle and store up to 20 gigabytes a second – enough data to fill around five standard DVDs. Another aim is to provide researchers with immediate access to data after a measurement, so that they can directly view and evaluate them while at the experimental station.

Meanwhile, the challenges involved in data analysis continue to grow as well. Increasingly, supercomputers are used to evaluate the data from PETRA III. “Unlike computer farms, which are used to analyse data from the LHC, the processors in a supercomputer are closely networked with one another,” explains Gülzow, Head of IT at DESY. “That’s the only way to properly simulate and understand processes in which neighbouring cells actively interact.” Although DESY does not have its own supercomputer on site, it is a contract partner of the John von Neumann Institute in Jülich and, as such, has access to JUQUEEN, currently one of the world’s fastest supercomputers.

Whatever the future brings, one thing is clear: big data is set to remain a key issue for research. “Many areas of science now face the challenge of how to deal with enormous quantities of data,” says Gülzow.

Big data in science

Medicine

The use of CT, MRT and ultrasound systems generates huge volumes of medical data. These imaging techniques are increasingly precise and produce images containing ever greater volumes of data. Meanwhile, new techniques are also emerging. These include digital pathology, the creation of high-resolution scans of extremely fine tissue sections. These digital slides are easier to archive and can be analysed by a computer. Moreover, experts are now combining information from different imaging techniques, with the aim of providing greater diagnostic accuracy and improved treatment.

Climate research

To understand global climate change in greater detail, researchers are developing an extremely complex model of the Earth as a system, which maps the physical, chemical and biological interactions between the atmosphere, the oceans and the biosphere – including the influence of human behaviour. It is based on a fine-meshed lattice model of our planet, constructed by supercomputers. The requisite simulations often take several months and generate mountains of data. The German Climate Computing Centre (DKRZ) in Hamburg alone has to annually archive up to 10 petabytes of modelling data generated by supercomputers.

Astronomy

Just like particle accelerators, telescopes are now being built on a whole new scale. Future projects such as the Square Kilometre Array (SKA) and the Cherenkov Telescope Array (CTA) comprise several dozen or even hundreds of individual telescopes spread across a number of areas, each several square kilometres in size. CTA is on the lookout for high-energy gamma rays, SKA for radio waves. Both facilities will generate huge amounts of data: CTA between three and five petabytes a year, SKA quite possibly more than one petabyte a day.

Ecology

Researchers are busy trying to unravel the complex mysteries of a variety of ecosystems. Scientists from the University of Hamburg, for example, are investigating the interaction between marine organisms. How is climate change affecting ocean life, and what are the consequences for the fishing industry? To answer such questions, the scientists must handle huge volumes of data: during cruises on their research vessels, they amass several terabytes of image data; and they use computer models to predict what will happen in the future, which likewise generates a mass of data.

Data – then and now

The total amount of data amassed by humanity from the Stone Age to the year 2002 is now generated in the 10 minutes it takes to have a coffee break.

The telephone modem of old had a data transfer rate of 300 bits per second; the current world record for a single optical fibre is 42 terabits per second.

The good old floppy disk would store 160 kilobytes of data; today's hard disks hold 4 terabytes.

The first Intel processor from 1971 had a clock rate of 740 kilohertz and comprised 2300 transistors; today's Intel processors have a clock rate of up to 3.9 gigahertz and as many as 5.5 billion transistors.

The first commercial digital camera from 1976 had 0.01 megapixels; today's commercial cameras have up to 50 megapixels.

1024

kilobyte [kB]

thousand bytes

e-mail

megabyte [MB]

million bytes

photo

1048576

1073741824

gigabyte [GB]

billion bytes

HD film

HD

terabyte [TB]

hard disk

1099511627776

trillion bytes

petabyte [PB]

1125899906842624

quadrillion bytes

DESY

DESY computer centre

Big data for more dynamism

Leibniz prizewinner and DESY researcher Henry Chapman uses hundreds of thousands of X-ray images to determine the atomic structure of biomolecules, and investigates how biochemical processes can be filmed with molecular resolution. The vast quantities of data generated in this process provide a window on the moving world of molecules.

Henry Chapman studies nature at incredibly small scales using the very latest research light sources: X-ray free-electron lasers, which produce ultra-short X-ray pulses that enable exposure times in the femtosecond (quadrillionth of a second) range. He investigates the complex, three-dimensional structure of biomolecules at a spatial resolution in the angstrom (ten-millionths of a millimetre) range – on the scale of single atoms. He is also a pioneer in visualising the dynamics of biomolecular processes.

Chapman came to Hamburg after stations in Australia, New York and California. He is the head of the Coherent Imaging Division at the Center for Free-Electron Laser Science (CFEL) at DESY and professor of physics at the University of Hamburg. In recognition of his work, Henry Chapman has been awarded one of the prestigious Gottfried Wilhelm Leibniz Prizes 2015.

Chapman explains why his investigations into the world of atoms and molecular dynamics require such large volumes of data: “We use powerful X-ray lasers to obtain structural information about samples that are too small to be investigated using conventional X-ray sources. From biomolecules in particular, it’s often difficult to grow large crystals. Sometimes you only get minuscule nanocrystals. This type of sample scatters X-rays only weakly. We therefore need extremely intense laser pulses in order to record a diffraction image just before the radiation destroys the sample. Each of these images only contains part of the structural information we require. We therefore have to record hundreds of thousands of diffraction images. It’s only by putting them all together that we get the complete molecular structure.” The resulting dataset comprises between two or three terabytes (trillion bytes).

In one year, Henry Chapman and his team produce the same amount of data – one petabyte – as is currently stored in the international Protein Data Bank, which contains around 100 000 entries

For example, in order to analyse the enzyme Cathepsin B, which plays a role in sleeping sickness, researchers fed a fine jet of water containing millions of tiny enzyme crystals into the beam of the LCLS, an X-ray laser at the US research centre SLAC in California. The X-ray laser fired 120 flashes per second at the stream of enzyme crystals. On average, one in eleven flashes hit a crystal, thus generating a total of 293 195 diffraction images, which could only be processed by means of a large

parallel computer. "When combined, the images initially gave us a 3D map of the entire diffraction properties of the enzyme," says Chapman. "On that basis, we were able to determine its structure to a resolution of 2.1 angstroms."

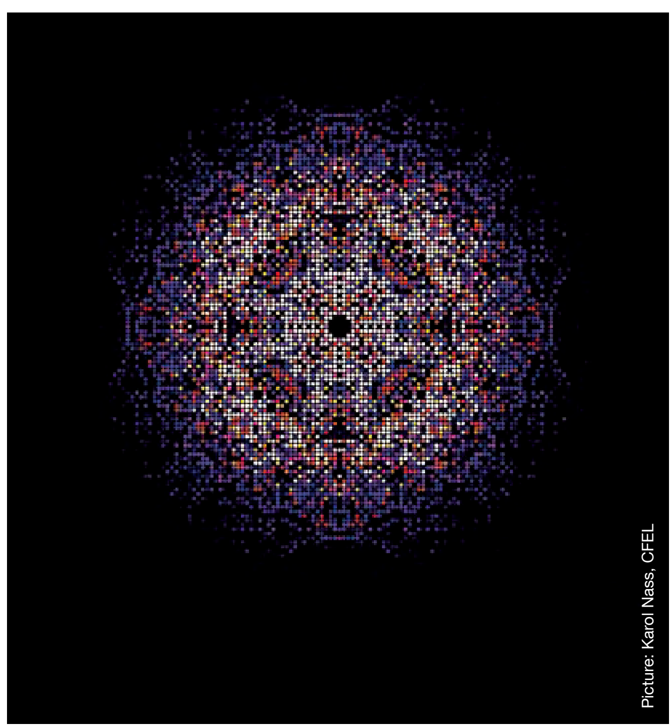
By way of comparison, structural analysis of an enzyme using a conventional X-ray source generates several hundred or thousand diffraction images and datasets in the region of a few gigabytes (billion bytes). In other words, an X-ray laser produces around 1000 times as much data. But the effort is worth it: innovative X-ray free-electron lasers such as the LCLS or the European XFEL, which is currently being built in Hamburg, generate pulses of light with a brilliance billions of times that of conventional X-ray research light sources – and, in the case of the European XFEL, even do so at a rate of 27 000 times a second. "This high pulse rate will also result in a much higher data rate, which in turn opens up new avenues for structural analysis," Chapman explains. "We'll be able not only to determine static structures but also to make dynamic, time-resolved measurements of molecular reactions."

In a groundbreaking study, Chapman and his research colleagues showed that the extremely short pulses of light from an X-ray laser can indeed be used to record the rapid dynamics of biomolecular processes in a kind of ultraslow-motion sequence. They were able to produce snapshots at a temporal resolution in the femtosecond (quadrillionth of a second) range. "By arranging these ultrafast snapshots into a film, we can produce a slow-motion sequence of the molecular dynamics," Chapman explains.

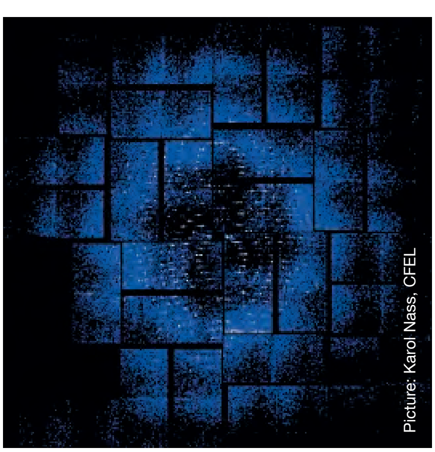
"Our work paves the way for the time-resolved investigation of dynamic processes with atomic resolution"

Henry Chapman, DESY

This process involves an enormous amount of data. If, for example, a dynamic process is recorded over a thousand time steps, this results in a thousandfold increase in the amount of data to several petabytes (quadrillion bytes) and several hundred million diffraction images per experiment.



The intensity map for the enzyme Cathepsin B is generated from almost 200 000 individual diffraction images of crystals. It was used to determine the three-dimensional molecular structure of the enzyme.



X-ray diffraction image of a single crystal of the enzyme Cathepsin B. The detector is made up of 64 individual elements, visible as a chequerboard of tiles. The diffraction by the crystal causes the regular pattern (white).

“There’s no sense in size alone”

In the world of science, huge amounts of data provide greater resolution, sharper images and faster readings. They reveal global patterns as well as minuscule details, and they bring progress and new knowledge. Yet these mountains of data arise in a whole variety of contexts. Social media networks, for example, collect in-depth information on the habits and personal preferences of their users. To discuss the opportunities and the risks presented by big data, femto talked with Volker Gülzow, Head of IT at DESY and a man well versed in handling huge volumes of data.

femto: What exactly is big data?

Volker Gülzow: Big data is, so to speak, the new kid on the block. The world of IT is always in need of a new buzzword. Big data, in the sense of large amounts of data, is not a new phenomenon. At DESY, we’ve always had to deal with large volumes of data from the accelerators, and that’s not about to change. What’s new is that business has become interested in “social” data and that lots of people now feel a powerful need to use digital technology to communicate with others or express themselves.

femto: What are the main sources of big data?

Volker Gülzow: Many classic areas of the natural sciences produce immense amounts of data. At the same time, the world’s appetite for data is constantly growing too. One of the reasons for this is that industry wants to improve its products and, under the buzzword “Industry 4.0”, aims to make manufacturing more efficient with, unfortunately, a reduced workforce. Meanwhile, the advertising sector wants to enhance product placement through the analysis of data from social media. At the same time, the security services are interested in gathering a host of data, which is certainly legitimate but can also raise certain issues. Amazingly, however, we ourselves also feed this trend by, for example, having an app on our mobile phone that monitors our physical condition while we’re out jogging and then just happens to transfer these data to a cloud.

femto: Can we say that “big is beautiful”, or is size alone a meaningless quality?

Volker Gülzow: There’s certainly no sense in size alone. And, in my opinion, not every dataset is vital for the advancement of society. We shouldn’t forget about the ecological consequences of Google, Facebook and all the rest. Behind them are a whole load of computers and storage media that eat up a monstrous amount of electricity. The question is: do we really need to post and share everything, right down to the latest holiday snaps, whatever the cost?

femto: What opportunities and possibilities does big data open up?

Volker Gülzow: Big data can certainly help us in a lot of applications. The secret is to combine these data in a meaningful way, so as to be able, for example, to identify complex interrelationships in environmental or medical issues. In the world of physics, we're used to combining data. But we're also in the fortunate situation of having well-structured and reliable data. That's different with social media. And that's where the problems begin. Unfortunately, it's often impossible to combine these data in any meaningful way. Something seemingly obvious soon proves to be baseless. Remember the decline in stork numbers and the parallel decline in the birth rate in Germany in the 1970s? And the data are not always reliable.

femto: Facebook and Google already know more about us than some people think is healthy. Should we be frightened of big data?

Volker Gülzow: There are those who say, "Where's the problem, I've got nothing to hide." I think that's pretty naïve. These people are not aware of how and for what purpose these personal data are being used. It's very easy to get around the purported anonymisation of data. For example, there was a story in the news about how data from taxi rides of the last year had been linked with the telephone book, which then revealed how often some prominent people had been hanging out in dubious bars. There are studies that show that in more than 90 percent of all cases, you only need four anonymised datasets in order to identify the person behind them.

femto: Are we ready for big data? Do we have the technical and social tools we need to exploit the opportunities offered by big data and to minimise the associated risks?

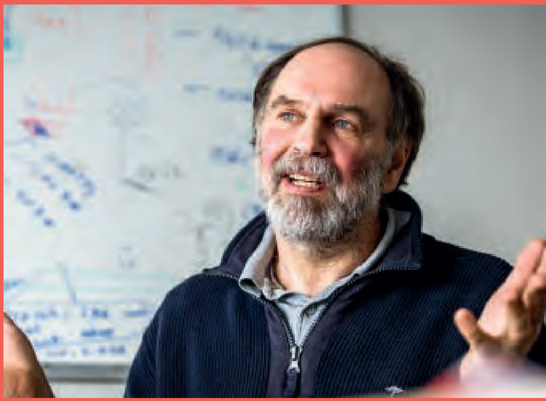
Volker Gülzow: No, not really. Most people don't have any idea of the risks. Either that, or they don't care about them. Just think about how opposed Germans once were to the census – and now look at what people post online today! That's why, when it comes to personal information, we at DESY are very keen on data minimisation. On the technical side, there are a lot of challenges, but we'll overcome these. And experience also shows that if something is technically feasible, it will get used.

femto: Where is the most pressing need for action?

Volker Gülzow: There are many areas where data are now combined and then used, for example, to produce assessments of creditworthiness that are completely false, because there's something wrong with the data quality or the algorithms. We need clear and properly regulated mechanisms to prevent or rectify things like this. And the kinds of institutions involved have to be made accountable. One big problem is that legislation can't keep up with the pace of IT development, so that you sometimes have the feeling that the lawmakers just haven't understood the problem we're facing.

femto: Which use of big data are you particularly looking forward to?

Volker Gülzow: I'm not interested in my refrigerator being able to do the shopping, and I don't care whether the person behind the cheese counter already knows what kind of cheese I like. I'm interested in classic scientific uses. And if I can help us take things a little step further, then that would be great!

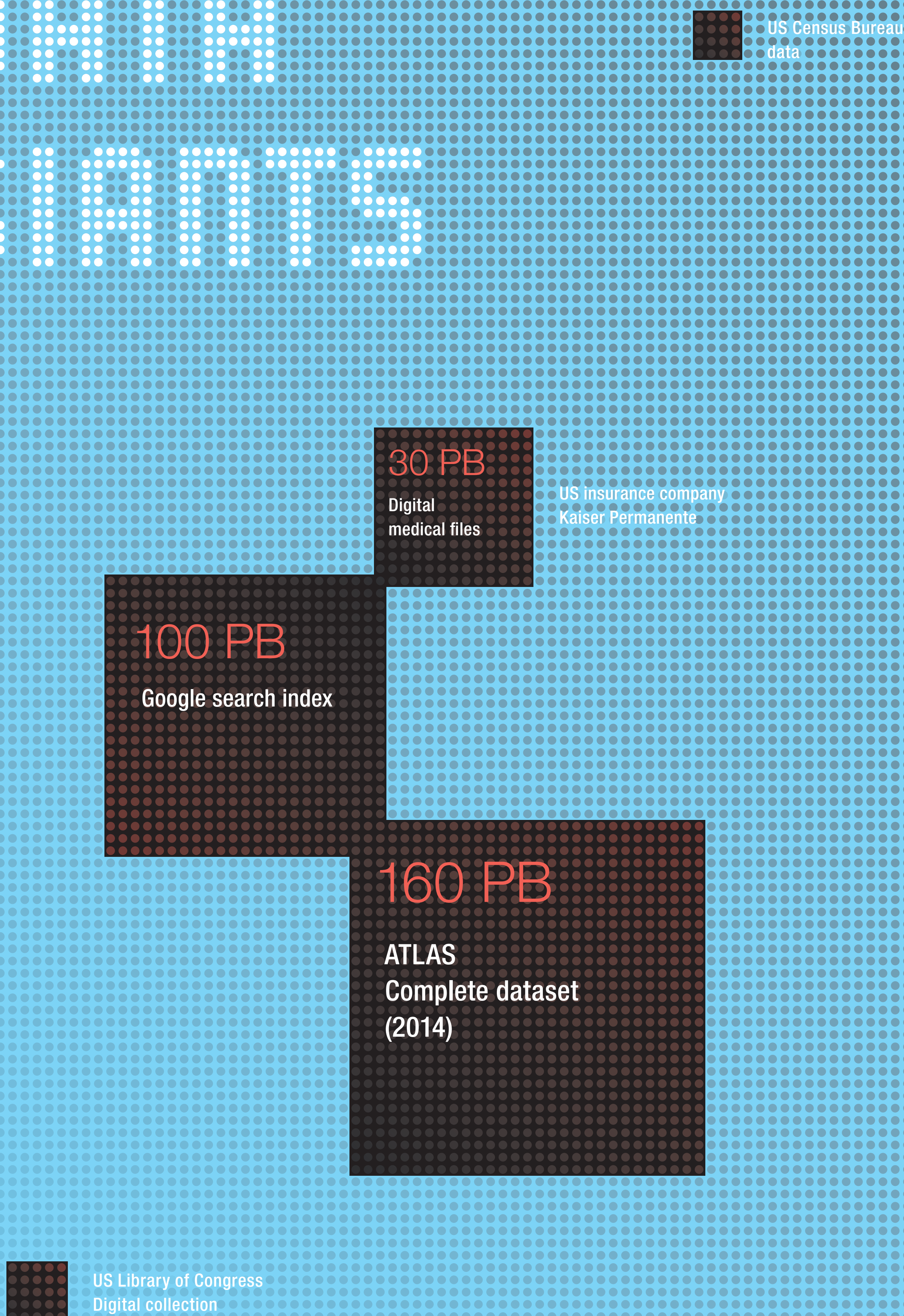


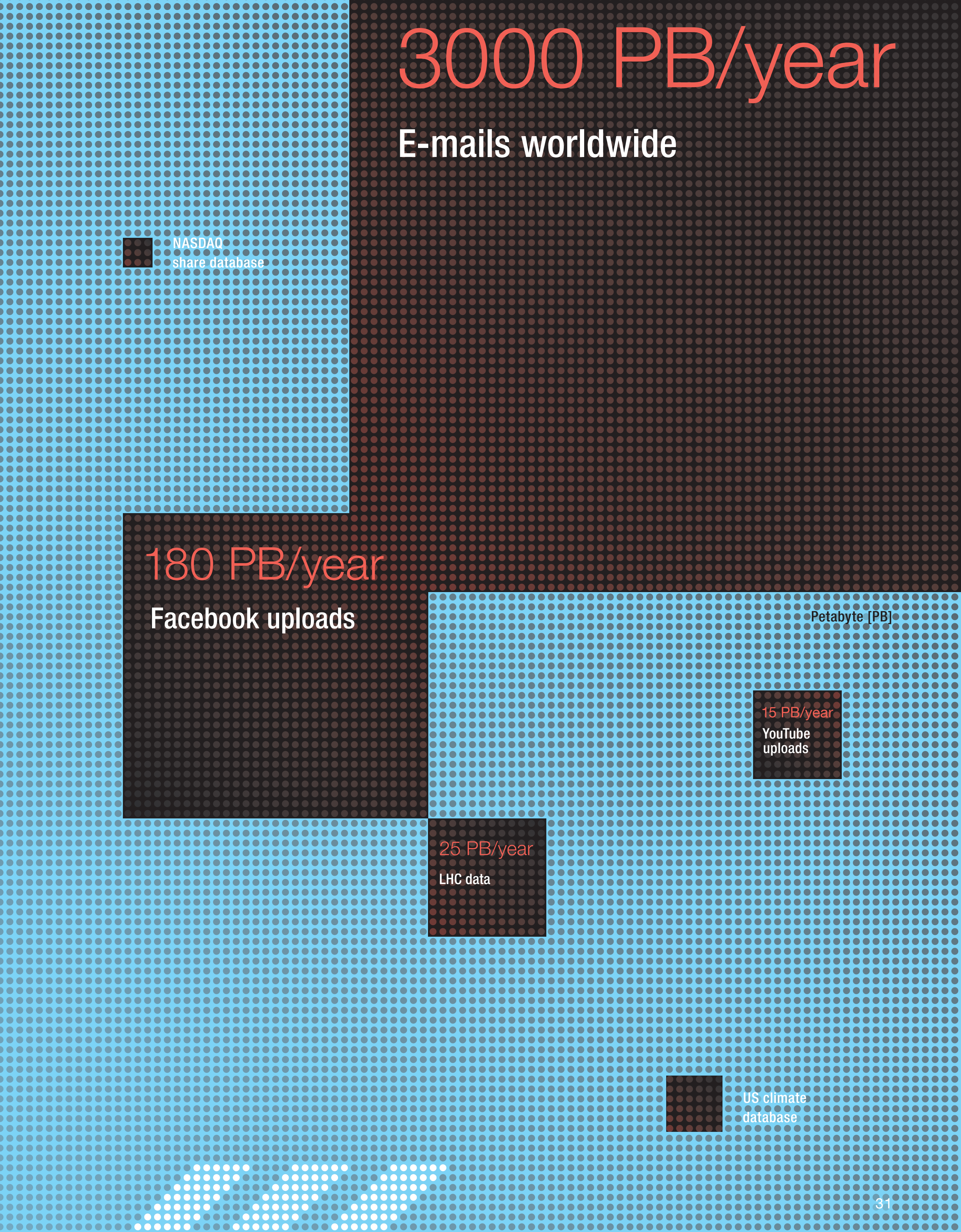
Pictures: DESY

femto: And which use are you frightened of?

Volker Gülzow: One of the major problems occurs when you combine data that are of mixed reliability. There's room here for a whole a lot of mischief. This can quickly have direct personal consequences or lead to undesirable social developments. We need to be much more vigilant here. And we also have to be careful that in all this rapid progress we don't lose older people along the way.

DATA GIANTS


US Census Bureau data
30 PB
Digital medical files
US insurance company Kaiser Permanente
100 PB
Google search index
160 PB
ATLAS
Complete dataset
(2014)
US Library of Congress
Digital collection


3000 PB/year
E-mails worldwide
NASDAQ
share database
180 PB/year
Facebook uploads
Petabyte [PB]
15 PB/year
YouTube
uploads
25 PB/year
LHC data
US climate
database

Close-up

Researchers observe transient state on catalyst surface

Picture: SLAC National Accelerator Laboratory

Artist's representation of an unprecedented snapshot made during the reaction of CO to form CO_2

An international team of researchers has made the first ever observations of the fleeting transient states that occur during the conversion of carbon monoxide into carbon dioxide at the surface of a catalyst. The team of scientists, which included researchers from the University of Hamburg and DESY, used ultrashort pulses from the X-ray laser LCLS at the SLAC research centre in California.

An optical laser pulse was first used to heat a ruthenium surface, which formed a simple catalyst. This activated the adsorbed carbon monoxide molecules (CO) and oxygen atoms. Using a special observation technique – X-ray absorption spectroscopy – the team was able to determine how the electronic structure of the participating oxygen atoms changed while carbon dioxide molecules (CO_2) formed. The process is similar to the way in which a car's catalytic converter works. The observed transient states agreed well with the results of quantum-chemical calculations.

Surprising, however, was the quantity of reaction partners that were activated into a transient state – and just as surprising was the discovery that only a small fraction of these went on to form stable CO_2 molecules. "It's as if you are rolling marbles up a hill, and most of the marbles that make it to the top roll back down on the same side again," says the head of the research team, Anders Nilsson of SLAC and Stockholm University.

Science, 2015; DOI: 10.1126/science.1261747

Infinitely cartwheeling molecules

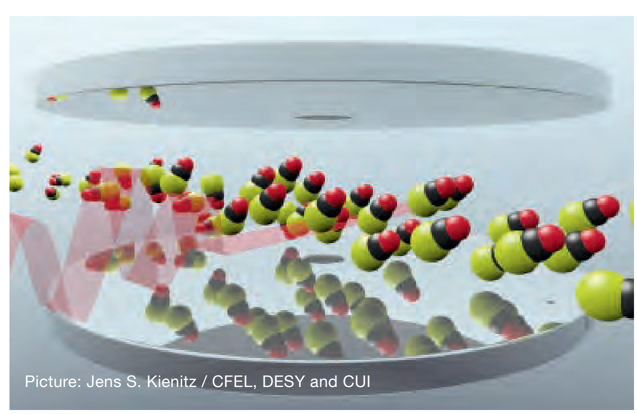
Synchronous rotations for new types of investigations

Using a physical trick, Hamburg-based researchers have caused whole groups of molecules to perform synchronised cartwheels almost infinitely. The technique opens up new possibilities for imaging molecules and the dynamics of their chemical reactions.

The team headed by Jochen Küpper at the Center for Free-Electron Laser Science (CFEL) at DESY fired an infrared laser pulse at a stream of carbonyl sulphide molecules, all of which were in their lowest energy state. The laser pulse mixed this state with the first excited quantum state. As a result, the molecules started to synchronously perform what is called an inversion, so that the sulphur atoms of the molecules all simultaneously alternated between pointing up or down. "In a sense, the laser forces the molecules to perform synchronous cartwheels," explains lead author Sebastian Trippel.

The researchers point out in their paper that the method works not only for the two lowest energy states, but also in principle for all states of a linear molecule. "This targeted molecular choreography opens up new possibilities for holding ensembles of free molecules in the X-ray beam of a free-electron laser in a controlled fashion, so that they can be investigated," says Küpper.

Physical Review Letters, 2015; DOI: 10.1103/PhysRevLett.114.103003



Picture: Jens S. Kienitz / CFEL, DESY and CUI

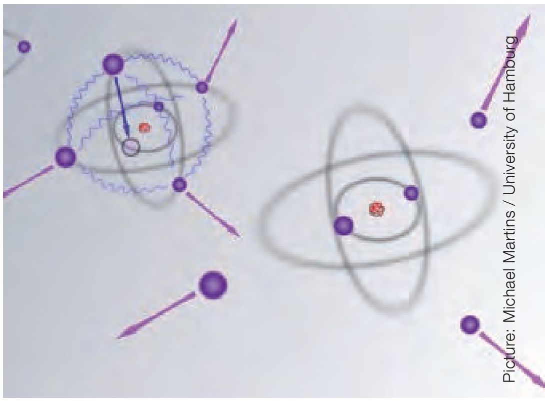
An infrared laser (red) causes the initially disordered molecules to perform synchronised cartwheels.

Electron billiards

Evidence of a direct four-particle interaction

Physicists at DESY's research light source PETRA III have observed a rare four-particle interaction. During this interaction, an electron ejects three other electrons from the electron shell of a carbon atom. The observations of the team of researchers headed by Alfred Müller and Stefan Schippers of the University of Gießen in Germany are relevant to many multielectron processes in physics and could help scientists better understand specific experiments using X-ray lasers.

In the experiment, an electron that had been excited by X-ray radiation jumped to a higher energy level in a carbon atom. The hole left by this electron was almost immediately filled by an electron from the higher level spontaneously dropping down to the lower level. In rare cases, the energy of this electron is transferred to two and sometimes even to three other electrons of the carbon atom, which are catapulted out of the atom's electron shell in a process known as Auger decay. In their experiment, the researchers were able to measure the frequency of this triple Auger decay in carbon for the first time. Just 0.01 percent of Auger decays result in three electrons being catapulted out of the atomic shell.



Triple Auger decay: An electron catapults three further electrons out of the atomic shell.

Physical Review Letters, 2015; DOI: 10.1103/PhysRevLett.114.013002

A matter of the mixture

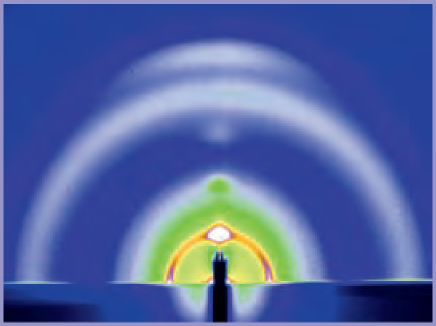
Magnetic nanoparticles boost solar cell performance

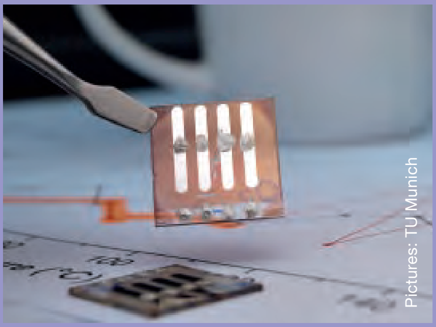
Magnetic nanoparticles can boost the performance of polymer solar cells – provided the mixture is right. This was demonstrated by an X-ray investigation at DESY's research light source PETRA III. Researchers headed by Peter Müller-Buschbaum from the Technical University of Munich in Germany have observed that an admixture of around one percent by weight of the nanoparticles increases the efficiency of certain organic solar cells.

The light creates pairs of electric charge carriers in the solar cells. To generate electricity, these charge carriers must be separated before they recombine. The scientists determined that nanoparticles of magnetite (Fe_3O_4) increase the lifetime of a charge carrier pair so that more such pairs can be separated – which increases the efficiency of the cell by 11 percent, from 3.05 to 3.37 percent.

However, adding too many of the nanoparticles is not a good idea, because they change the internal structure of the organic solar cells, as was shown by the X-ray studies. "The solar cells we looked at will tolerate magnetite nanoparticle doping levels of up to one percent by weight without changing their structure," explains co-author Stephan Roth from DESY. The researchers expect that doping with nanoparticles will further enhance the performance of other types of polymer solar cells.

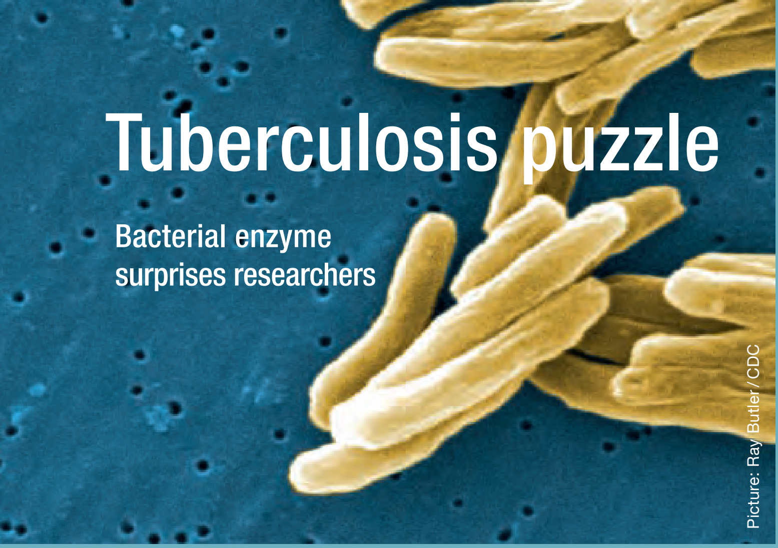
Advanced Energy Materials, 2015; DOI: 10.1002/aenm.201401770





Crystalline structures within polymer solar cells cause characteristic diffraction patterns in experiments with X-ray radiation (top).

The organic solar cells studied, shown here on a glass slide for research purposes (bottom).



Tuberculosis puzzle

Bacterial enzyme surprises researchers

Tuberculosis bacteria under the electron microscope

An international research partnership has discovered the previously unknown function of a tuberculosis enzyme complex. The group of scientists headed by Matthias Wilmanns of the Hamburg outstation of the European Molecular Biology Laboratory (EMBL) on the DESY campus were really looking for enzymes involved in the assembly of the unusually thick cell wall of the pathogen *Mycobacterium tuberculosis*. Together with colleagues from the Polish Academy of Sciences, the researchers had identified the complex AccD1-AccA1 as a worthwhile target.

The Hamburg-based researchers sent genetically modified bacteria that lacked this complex to French colleagues at the research centre CNRS who, surprisingly, determined that there were no differences between the cell walls of the modified bacteria and those of normal tuberculosis bacteria. Cooperation partners from Zürich then analysed which other molecules of the bacterium were connected with the enzyme complex under investigation, while EMBL researchers determined the exact architecture of the enzyme.

"We discovered that this enzyme is involved in a completely different process," reports Wilmanns. The results show that it plays a role in degrading the essential amino acid leucine, one of the building blocks of proteins. "As far as we are aware, this is the first indication that Mycobacteria possess their own degradation pathway for leucine," explains Wilmanns. "Even though this isn't the answer we expected, it's important to be able to fit another piece of the puzzle into place. It will help to find more specific inhibitors against other complexes of this protein group."

PLoS Pathogens, 2015; DOI: 10.1371/journal.ppat.1004623

Superstrong magnetic field in a very confined space

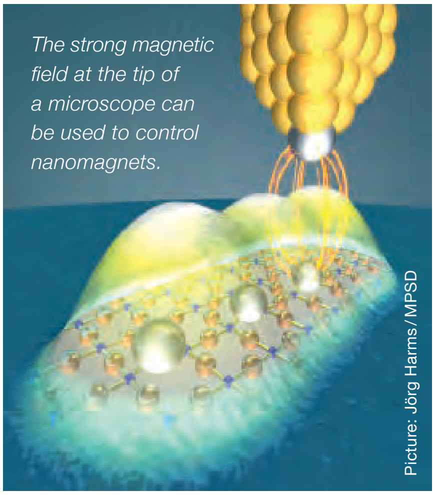
Max Planck physicists succeed in reversing the polarity of tiny nanomagnets

The force of the world's strongest permanent magnet in a speck the size of an atom – that's the record-breaking combination created by a team of researchers from the Max Planck Institute for the Structure and Dynamics of Matter at the DESY campus in Hamburg. The team used a fine microscope tip to control the magnetic polarity of a chain of three iron atoms. The new effect could spur the development of quantum computers and open up alternative concepts for future storage media.

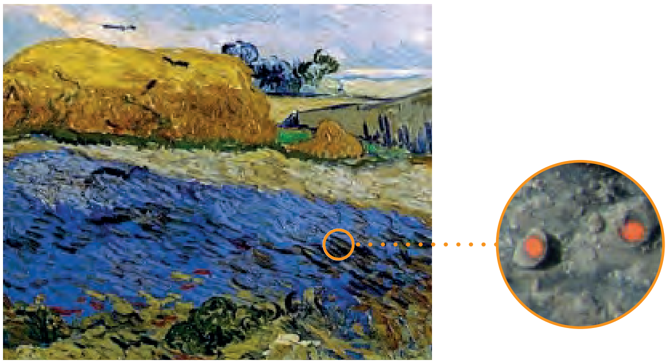
First, the scientists arranged three iron atoms into a short chain on a smooth copper surface using the fine point of a scanning tunnelling microscope. This created a tiny nanomagnet. Then they manoeuvred a magnetic tip very closely over the iron chain. Due to a quantum effect, a magnetic field with special properties formed between the microscope tip and the iron chain. "On one hand, it is extremely strong, on the other hand spatially very confined," says Shichao Yan, lead author of the study.

The field strength reaches a value of several tesla – more than the best permanent magnet. The extent of the field, however, is limited to an area about the size of an atom. "With this strong spatial confinement, we can control tiniest nanostructures individually," explains Sebastian Loth, the leader of the group. "A single small group of atoms can be made to reverse its magnetisation. Its neighbourhood, however, remains completely unaffected." That makes the new effect interesting as a control process for quantum bits – the switching units of a quantum computer. The phenomenon could also be relevant for data storage.

Nature Nanotechnology, 2014; DOI: 10.1038/nnano.2014.281



The strong magnetic field at the tip of a microscope can be used to control nanomagnets.



Wheat Stack under a Cloudy Sky (Korenschelf onder wolkenlucht; 1889, Kröller-Müller Museum). The bright orange-red can still be seen in and under the tiny paint sample (right).

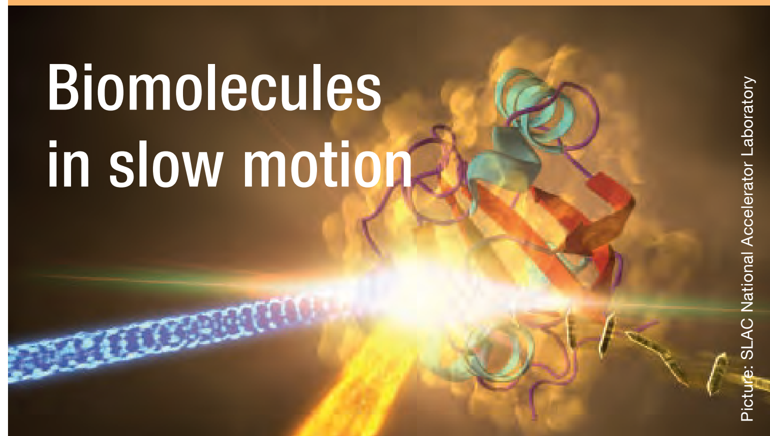
Van Gogh's fading orange

In studies at DESY, Belgian researchers have discovered traces of an extremely rare lead mineral in a painting by the Dutch artist Vincent van Gogh. The find sheds light on the decay process of the bright orange-red pigment red lead. Its bright orange-red colour is familiar to many as a corrosion protection primer for iron. Artists have been familiar with the pigment since antiquity. Its saturated colour fades with time, however, due to a variety of aging processes. For example, red lead (Pb_3O_4) reacts in the presence of light and air to form white lead (hydrocerussite) and white lead ore (cerussite).

The team headed by Koen Janssens of the University of Antwerp has now further clarified this process. The researchers studied a microscopically small paint sample from van Gogh's painting Wheat Stack under a Cloudy Sky using the bright X-rays from DESY's research light source PETRA III and determined the distribution of the different compounds in the sample with extremely high spatial resolution. In the process, they stumbled on the extremely rare lead carbonate mineral plumbonacrite.

"This is the first time that this substance has been found in a painting from before the mid-twentieth century," reports Frederik Vanmeert, lead author of the paper. "Our discovery sheds new light on the bleaching process of red lead." Based on their new insights, the scientists have proposed that the influence of light first initiates the reduction of the red lead to PbO. Carbon dioxide from the atmosphere or from the decomposition products of the binder of the oil paint then reacts with this PbO to form plumbonacrite as an important intermediate that absorbs additional CO_2 to form white lead and white lead ore.

Angewandte Chemie, 2015; DOI: 10.1002/ange.201411691



Picture: SLAC National Accelerator Laboratory

Biomolecules in slow motion

An international research team with participation from DESY has used the world's most powerful X-ray laser to watch a light-sensitive biomolecule at work. The study demonstrates that extremely short X-ray flashes can record the rapid dynamics of biomolecules in ultraslow motion.

The scientists led by Marius Schmidt of the University of Wisconsin in Milwaukee used photoactive yellow protein (PYP) as a model system for their studies. PYP is a receptor for blue light and part of the machinery of photosynthesis in certain bacteria. When it captures a photon of blue light, it goes through what is known as a photocycle, during which it harvests the photon's energy, before returning to its initial state. Most of the steps of the PYP photocycle have been well studied, which makes the molecule an excellent candidate for validating new investigative methods.

The researchers used the X-ray laser LCLS at the US accelerator centre SLAC in California for their ultrafast snapshots of the PYP dynamics. Thanks to the ultrashort and ultrabright flashes of X-rays from the LCLS, the scientists were able to observe how PYP changed its shape in the course of the photocycle. The resulting images have a resolution of 0.16 nanometres, making them the most detailed images of a biomolecule ever recorded using an X-ray laser.

The measurements showed the sequence of the PYP photocycle in finer detail than previous studies, thus demonstrating that the new technique works. "This is a real breakthrough," states Henry Chapman of DESY. "Our study is opening the door for time-resolved studies of dynamic processes with atomic resolution."

Science, 2014; DOI: 10.1126/science.1259357

Researchers film magnetic memory in superslow motion

The good old compact and video cassettes are examples, as are magnetic stripe cards and floppy and hard disks – they are all magnetic data storage media that are electromagnetically written and read. But magnetic memories are by no means history. "Although laptops and other mobile devices these days are increasingly using non-magnetic storage technologies, such as flash memory, the price of magnetic data storage devices means that there is no substitute for them when it comes to storing large amounts of data," says Philipp Wessels, who is studying magnetic memories as a member of Markus Drescher's group at the Hamburg Centre for Ultrafast Imaging (CUI). "The trend is towards storing data in the cloud, and the cloud is magnetic."

The trend is towards storing data in the cloud, and the cloud is magnetic

To better understand the dynamics of magnetic storage media, Wessels and his researcher colleagues have been filming a candidate for tomorrow's data storage media in superslow motion at DESY using an X-ray microscope. The resulting film shows how magnetic vortices form in ultrafast memory cells. "Our exposures show for the first time how exactly the magnetisation proceeds," says Wessels. "For the first time, we can observe the switching of these magnetic cells in detail." For their investigation, the researchers chose a memory cell made of an alloy of nickel and iron, which can be magnetised in less than a billionth of a second.

Using a custom-built X-ray microscope developed by the University of Hamburg in association with the University of Applied Sciences in Koblenz, the scientists watched a memory cell being erased and then having new data written to it. The extremely short flashes of X-ray light from DESY's research light source PETRA III enabled them to observe the process with a temporal resolution of 0.2 billionths of a second (200 picoseconds). The degree of magnetisation can be measured via the extent to which the polarised X-ray radiation is absorbed by different regions of the sample. In the process, the X-ray microscope is able to resolve minute details, down to 60 millionths of a millimetre (60 nanometres).

Magnetic vortices in the memory cell

In their studies, the scientists used tiny square memory cells, each side of which was two thousandths of a millimetre (2 micrometres) long. Four magnetic regions – so-called domains, whose magnetic field varies either in a clockwise or anti-clockwise direction – form in each of the memory cells. These magnetic domains are triangular, with their apexes meeting in the centre of the memory cell, where they give rise to a magnetic vortex core.

When the contents of a memory cell are erased by an external magnetic field, the magnetic vortex core is driven out of the cell. "In our experiments, we were for the first time able to measure the speed at which the vortex cores are expelled from the material," explains Jens Viefhaus from DESY. It turned out that the core shoots out of the memory cell at a speed of more than 3600 kilometres per hour. "Because this process is reproducible, we were able to make a reliable measurement of the velocity," adds CUI researcher

Philipp Wessels of the University of Hamburg is studying magnetic memories.

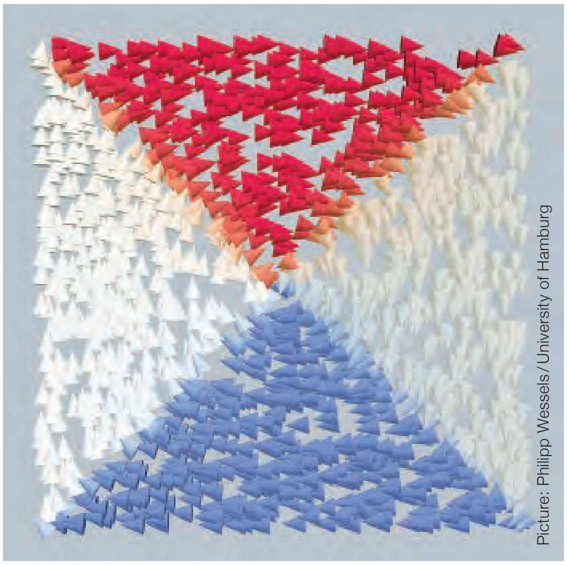


Illustration of the magnetic field direction within the memory cell

Guido Meier. “In order to make these measurements, we had to use very intense and stable magnetic excitation pulses.”

A better understanding of the magnetisation dynamics can lead to faster and better storage media.

Zigzag in superslow motion

The external magnetic field forces the entire memory cell into a state of uniform magnetisation. When the field is switched off, the cell once again forms four magnetic domains with a central vortex – whose direction depends on that of the applied magnetic field; in other words, new data is written to the cell. The process is complex, however. “The four-domain state develops via a complicated zigzag pattern, and for the first time we were able to watch ‘live’ as this state was formed,” reports Wessels. This behaviour is in good agreement with the results of simulations. Thanks to thesuperslow motion, the high-speed dynamics of the process can now be observed in greater detail. “The same method can be used to study the dynamics of any other magnetic material,” notes Wessels. “Our experiments can help us to understand how quickly data can, in principle, be written to a magnetic storage medium, coded in the form of domains.” A better understanding of the magnetisation dynamics can lead to faster and better storage media.

Physical Review B, 2014; DOI: 10.1103/PhysRevB.90.184417

femtomenal

Colder than outer space

Nowhere and nothing in the universe can be colder than minus 273.15 degrees Celsius. The British physicist Lord Kelvin therefore chose this point as the absolute zero of the temperature scale that bears his name. Here on Earth, we are a long way from such temperatures. German headlines start announcing record cold snaps when the temperature drops below minus 30 degrees, and Antarctica’s record low was minus 93 degrees. Outer space is considerably colder, at minus 270.42 degrees Celsius or 2.73 Kelvin.

The accelerator experts on the DESY campus in Hamburg, however, handle temperatures even colder than outer space. The accelerator modules of the free-electron lasers FLASH and European XFEL have to be operated at minus 271 degrees to bring the electrons up to speed. That’s the temperature at which the accelerator elements, which are made out of the metal niobium, become superconducting.

Things can get even colder if cooling down individual atoms rather than kilometre-long accelerators. Such ultracold atoms can reach temperatures within billionths of a Kelvin – or even less – of absolute zero.

X-ray flashes show free nanoparticles for the first time in 3D

Nanoparticles are conquering our everyday world. The applications of these tiny particles – which are far too small to be visible to the naked eye – range from sunscreens to effect coatings, from colour filters to electronic components and even to medical applications such as cancer treatment. The dimensions of these particles are measured in nanometres (millionths of a millimetre) – hundreds of times thinner than a human hair.

For the first time ever, a German–American research team using DESY's X-ray laser FLASH has determined the three-dimensional shape of free-flying silver nanoparticles. “The functionality of nanoparticles is linked to their geometric form, which is often very difficult to determine experimentally,” explains Ingo Barke from the University of Rostock in Germany. “This is particularly true when they are present as free particles – that is, in the absence of contact with a surface or a liquid.”

The form of a nanoparticle can be calculated from the characteristic way in which it scatters X-ray radiation. X-ray sources such as FLASH thus serve as a kind of super microscope for the nanoworld. Previously, researchers typically reconstructed the spatial structure of nanoparticles from multiple two-dimensional images taken from different angles. This procedure is unproblematic for particles on solid substrates, as images can be made from many different directions so that the particles' three-dimensional form can be unambiguously reconstructed.

“Bringing nanoparticles into contact with a surface or a liquid can significantly alter the particles, so that you can no longer see their actual form,” says Daniela Rupp from the Technical University (TU) of Berlin. A free particle in flight, however, can only be imaged once before it either escapes the imaging area or is destroyed by the intense X-ray radiation. So a method is required to create, with a single laser pulse, a diffraction pattern that contains the entire structural information of a particle.

The functionality of nanoparticles is linked to their geometric form

A complete picture with a single flash

The physicists led by Thomas Möller of the TU Berlin and Karl-Heinz Meiwes-Broer and Thomas Fennel of the University of Rostock, working with colleagues from SLAC and DESY at the X-ray laser FLASH, have now accomplished this using a trick. Instead of recording the diffraction pattern in a small angle around the direction of the incident X-ray pulse as is usually the case, the researchers recorded it in a wide angular range around the nanoparticle. “This approach virtually captures the structure from many different angles simultaneously, without the need to make multiple exposures of the particle,” explains Fennel.

The researchers tested this method on silver nanoparticles with diameters of 50 to 250 nanometres (0.00005 to 0.00025 millimetres) that were guided through the X-ray beam in a carrier gas. The test not only demonstrated that the method works but also produced some surprising results. It revealed that comparatively large nanoparticles come in a much greater variety of shapes than had been expected.

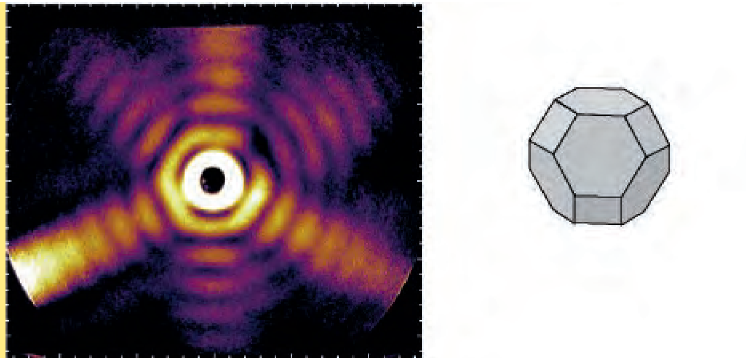
The external form of free nanoparticles is a result of various physical principles, particularly the particles' efforts to minimise their energy. As a result, large particles composed of thousands or millions of atoms often yield predictable shapes, because these atoms can only be arranged in a particular way to achieve an especially low-energy state.

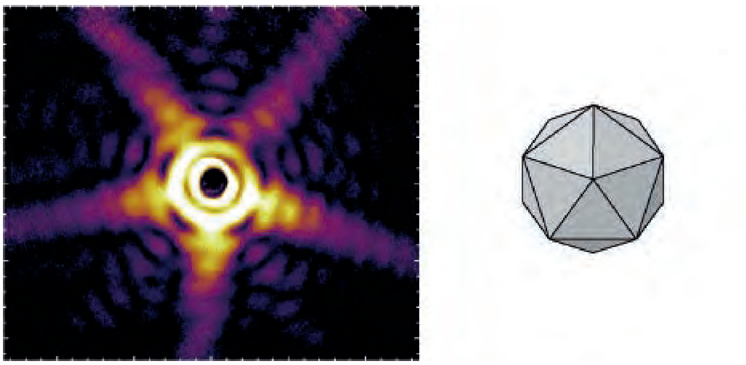
Silver in 3D

In their experiment, however, the researchers observed numerous highly symmetrical three-dimensional shapes, including several types known as Platonic and Archimedean bodies. Examples include the truncated octahedron (a body with eight regular hexagonal and six square faces) and the icosahedron (a body with twenty faces, each an equilateral triangle). The latter is actually only especially stable for extremely small particles consisting of a few atoms, and its occurrence in free particles of this size was previously unknown. “The results show that metallic nanoparticles retain a type of memory of their structure, from the early stages of growth to a yet unexplored size range,” says Barke.

Due in particular to the large variety of shapes, it was very important that the scientists use a fast computational method to map the shape of each individual particle. To this end, the researchers used a two-stage process. After first determining the general shape, the scientists refined

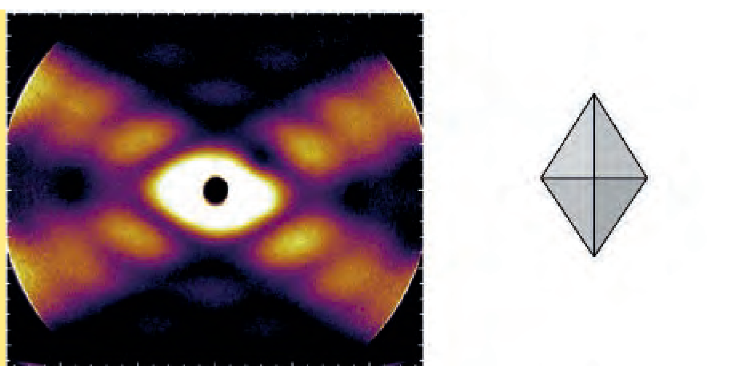
X-ray diffraction image (left) of a nanoparticle in the shape of a truncated octahedron (right) with a diameter of 200 nanometres

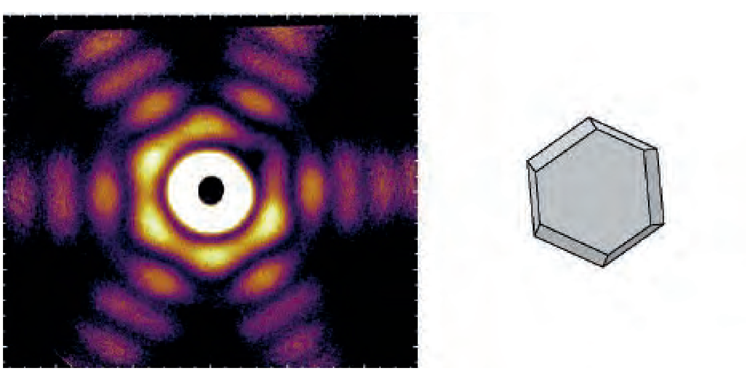




X-ray diffraction image (left) of a nanoparticle in the shape of an icosahedron (right) with a diameter of 240 nanometres

X-ray diffraction image (left) of a nanoparticle in the shape of a decahedron (right) with a diameter of 180 nanometres





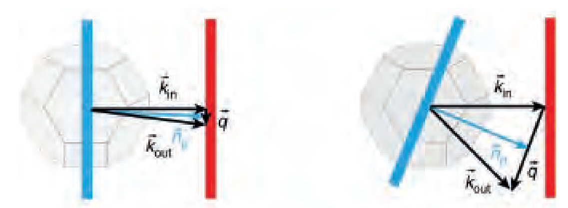
X-ray diffraction image (left) of a nanoparticle in the shape of a truncated twinned tetrahedron (right) with a diameter of 150 nanometres

The structure is captured from many different angles simultaneously, without the need to make multiple exposures of the particle

Pictures: Hannes Hartmann / University of Rostock

In the experiment, silver nanoparticles are generated in a particle source (left in the picture) and directed into the focus of the FLASH laser (blue). The light scattered from the individual particles is recorded using a detector (right).

When the scattered light is recorded only in a small angle (left), it contains information solely about the projection of the particle. Images taken in a wide angle (right) comprise similar information to many images taken in a small angle while the particle is being rotated.
In this way, wide-angle images provide information about the three-dimensional structure.



In future projects, particles could be directly "filmed" in three dimensions during growth or during phase transitions

Pictures: Hannes Hartmann/ University of Rostock

its details using more complex simulations on a supercomputer. This approach proved to be so efficient that it could not only reliably determine a wide variety of shapes, but also distinguish between varying orientations of the same shape.

This new method for determining the three-dimensional shape and orientation of nanoparticles with a single X-ray laser pulse opens up a wide range of new research directions. In future projects, particles could be directly "filmed" in three dimensions during growth or during phase transitions. "The ability to directly film the reaction of a nanoparticle to an intense flash of X-ray light has long been a dream for many physicists – this dream could now come true, even in 3D!" says Rupp.

Nature Communications, 2015; DOI: 10.1038/ ncomms7187

Shedding light on pathogens

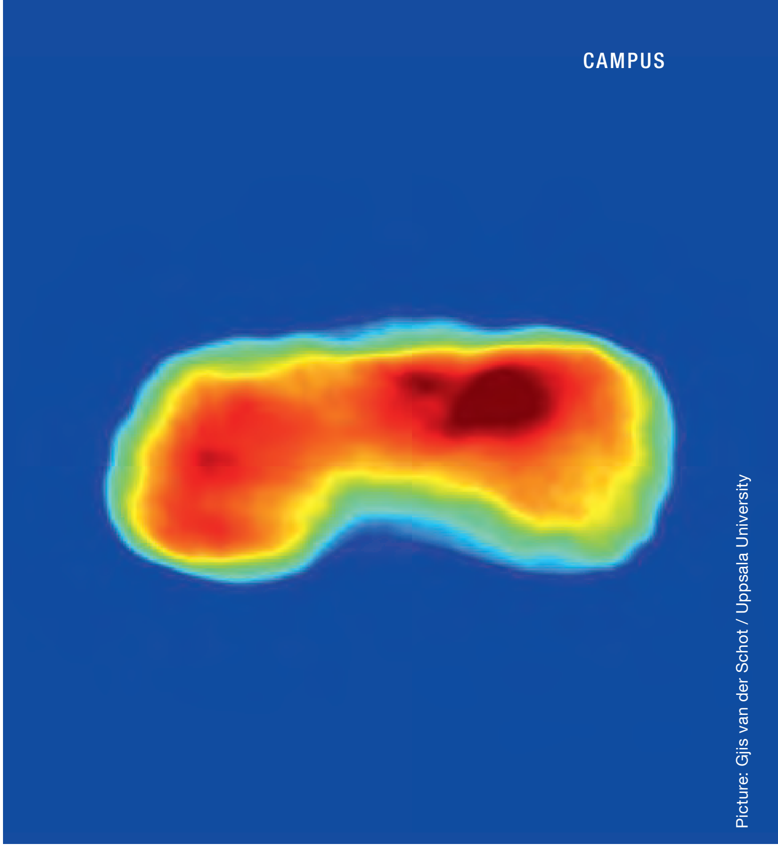
Researchers study viruses and living bacteria

Researchers have used the world's most powerful X-ray laser to penetrate deep into the biological microcosm. The scientists led by Janos Hajdu from Uppsala University in Sweden succeeded in imaging live bacteria and even individual cell organelles. They were also able to obtain a three-dimensional image of a virus. The experiments, in which DESY scientists also participated, were carried out using the X-ray laser LCLS at the US research centre SLAC. They illustrate the enormous potential of such investigations.

Biological cells are normally dead and have been chemically fixed when they are studied using X-rays. "If you want to completely understand the details of how a cell functions, you need the cell alive," says Hajdu. To demonstrate the imaging of living cells, the researchers sprayed cyanobacteria through the X-ray beam of the LCLS. When the X-ray beam hits such a bacterium, a characteristic diffraction pattern is produced, from which the structure of the cell under investigation can be calculated.

"Although the cells are destroyed by the X-ray beam, the ultrashort and extremely bright flashes from an X-ray laser enable diffraction patterns to be recorded sufficiently quickly that a correct image of the sample can be obtained before it is vaporised," says DESY researcher Anton Barty, one of the scientists involved in the experiments. "The flashes outrun the damage." In this way, the scientists were able to reconstruct individual bacterial cells with a resolution of 76 nanometres. One nanometre is one millionth of a millimetre. "The data collected suggests that we could get down to 4 nanometres, which is the size of a protein molecule," says the leader of the experiment, Tomas Ekeberg of Uppsala University.

In a further experiment, the researchers sprayed individual cell organelles into the X-ray beam. Organelles are the functional units within cells. The scientists selected carboxysomes – organelles measuring around 100 nanometres in diameter that are responsible for carbon fixation – from cyanobacteria. Molecules normally have to be crystallised for conventional



Picture: Gijs van der Schot / Uppsala University

Reconstruction of a cyanobacterium from an X-ray diffraction pattern

X-ray investigations. However, many biomolecules are extremely difficult or even impossible to crystallise. That's why researchers are searching for methods of studying individual non-crystalline objects using X-ray diffraction.

"With the carboxysomes, we have reconstructed the smallest single non-crystallised biological particles ever imaged with an X-ray laser," says the leader of the experiments, Max Hantke of Uppsala University. The calculated images show details down to a size of 18 nanometres.

The researchers demonstrated the potential for investigating viruses in a third study, in which the three-dimensional structure of the mimivirus was determined using the X-ray laser. The mimivirus is classed as a large virus and has a diameter of around 750 nanometres. The researchers sprayed intact viruses through the X-ray beam and combined the single portraits of the individual viruses into a three-dimensional model of a mimivirus. "Our investigations have only given us a fleeting glimpse of the potential of future studies of three-dimensional biological structures using X-ray lasers such as the LCLS or the European XFEL, which is currently being built in Hamburg," emphasises DESY researcher and co-author Henry Chapman.

Nature Communications, 2015; DOI: 10.1038/ncomms6704
Nature Photonics, 2014; DOI: 10.1038/NPHOTON.2014.270
Physical Review Letters, 2015; DOI: 10.1103/PhysRevLett.114.098102

by Johannes Kretzschmar

Imprint

femto is published by
Deutsches Elektronen-Synchrotron DESY,
a research centre of the Helmholtz Association

Editorial board address
Notkestraße 85, D-22607 Hamburg, Germany
Tel.: +49 40 8998-3613, Fax: +49 40 8998-4307
e-mail: femto@desy.de
Internet: www.desy.de/femto
ISSN 2199-5192

Editorial board
Till Mundzeck (responsible under press law), Ute Wilhelmsen

Contributors to this issue
Frank Grotelüschen

Design and production
Diana von Ilsemann
Diana Schröder

Translation
TransForm, Cologne
Ilka Flegel

Printing and image processing
reset, Hamburg

Copy deadline
March 2015

FSC
www.fsc.org
MIX
Paper from responsible sources
FSC® C008151

The DESY research centre

DESY is one of the world's leading particle accelerator centres. Researchers use the large-scale facilities at DESY to explore the microcosm in all its variety – ranging from the interaction of tiny elementary particles to the behaviour of innovative nanomaterials and the vital processes that take place between biomolecules. The accelerators and detectors that DESY develops and builds at its locations in Hamburg and Zeuthen are unique research tools. The DESY facilities generate the most intense X-ray radiation in the world, accelerate particles to record energies and open up completely new windows onto the universe.

HELMHOLTZ
ASSOCIATION

DESY

DESY is a member of the Helmholtz Association,
Germany's largest scientific organisation.