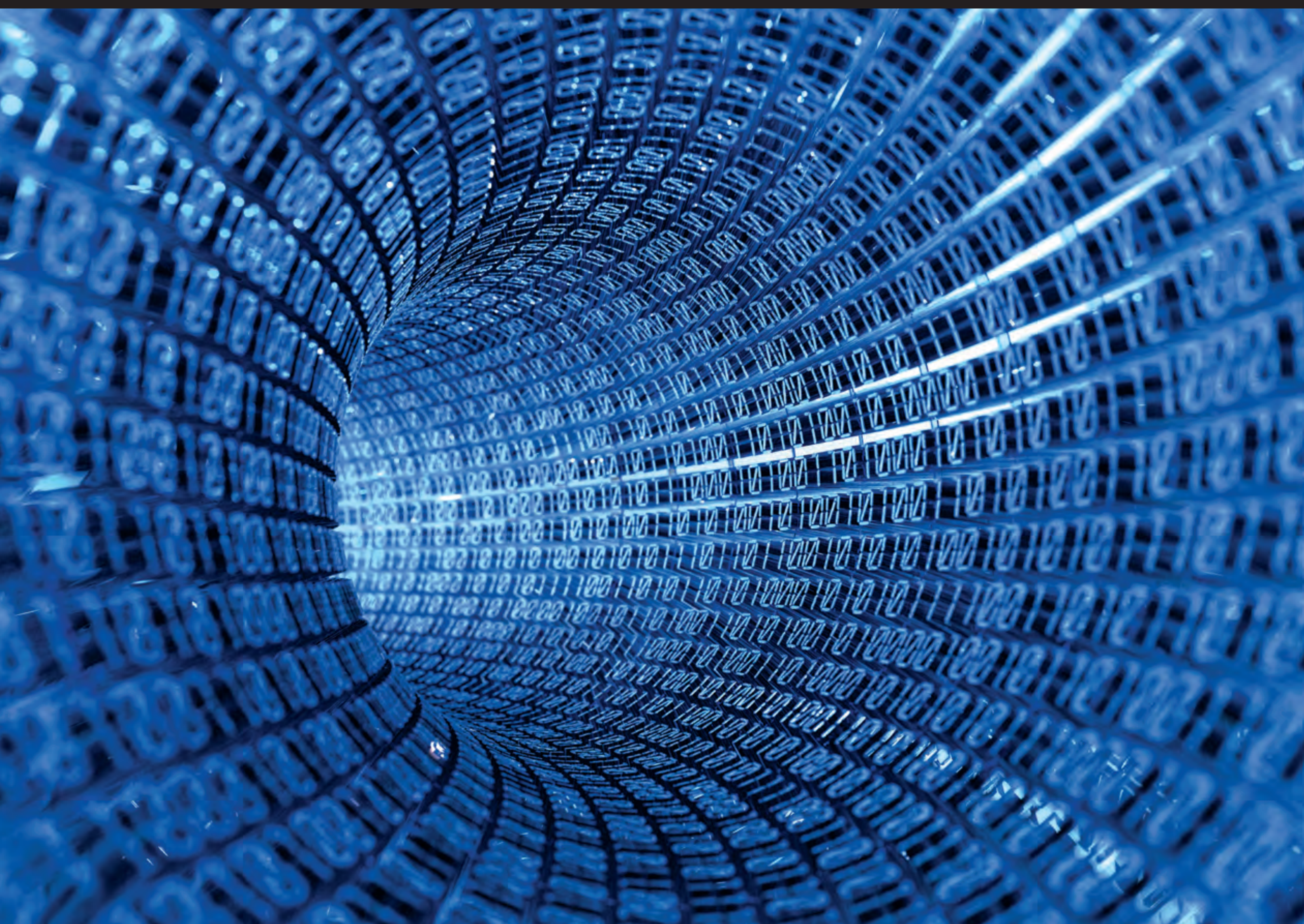
A Light for Science

ESRF news

Number 65 December 2013

Big data in focus



New lattice shapes light-source scene

Superconductor predicted from first principles

ESRF

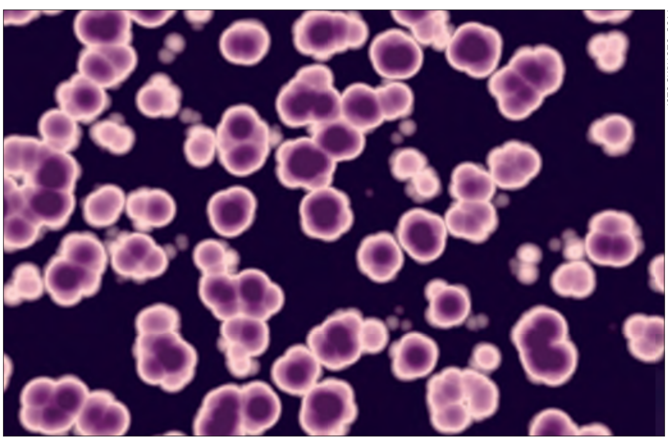
A light for science


B LEBAYLE

Fourth paradigm: the ESRF gears up for the big-data deluge, pp14–21.


FOJAB

Lattice pioneers: Sweden's MAX-IV to set new standards, p11.


ISTOCKPHOTO.COM

Novel project: bacteria targeted by a GlaxoSmithKline collaboration, p24.

On the cover: Synchrotrons face up to big data, pp14–21.

ESRFnews Number 65 December 2013


Ready to jump in at any stage of your Control System Project...
We build the complete
CS cost-efficiently from tried
and proven components
NEW
EPICS
TANGO
Control System Studio (CSS)
NATIONAL INSTRUMENTS Gold Alliance Partner
LABVIEW
ORACLE, mySQL, etc.
We master all control system
technologies for Big Physics
cPCI, PXIe, VME, xTCA, MTCA.4
FPGA programming
PVSS, FESA, DOOCS, TINE, ACS...
Timing system, MPS
Expert staff to cover peak loads and beyond*
*yes, we do maintenance and support, even write up the documentation!
Choose the arrangement that suits you best:
outsourcing (on-site or off-site), development services,
or a customized turnkey solution.
COSYLAB
Your TRUSTED Control System Partner
www.cosylab.com

ESRFnews

Editor
Matthew Chalmers
Tel +44 (0)7857 866 457
E-mail mdkchalmers@gmail.com

Editorial committee
Nick Brookes
Dominique Cornuéjols
Bauke Dijkstra
Andy Fitch
Claus Habfast
Axel Kaprolat
Joanne McCarthy
Harald Reichert
Francesco Sette

Subscription
Chantal Argoud
For subscriptions to the digital version of *ESRFnews* (print subscriptions are no longer possible), visit http://go.esrf.eu/esrfnews

ESRFnews is produced for the ESRF by
IOP Publishing
Temple Circus
Temple Way
Bristol BS1 6HG, UK
Tel +44 (0)117 929 7481
www.iop.org

Publisher
Susan Curtis
Group editor
Joe McEntee
Art director
Andrew Giaquinto
Production
Alison Gardiner
Technical illustrator
Alison Tovey
Display-advertisement manager
Edward Jost
Advertisement production
Mark Trimnell
Marketing and circulation
Angela Gage

ISSN 1011-9310

Big data: a blessing or a curse?

There is a lot of hype around the term "big data" these days. It touches many areas of science and industry but, truth be told, nobody is really able to define what big data means. My own definition, applied in the context of the ESRF, would be very pragmatic: a volume of data that is impossible to process by simple means, hence requiring significant investments in IT infrastructure to capture, store, transfer, analyse and visualise datasets.

Big data at the ESRF is driven by the availability of new, high-resolution pixel detectors capable of generating an image of several megapixels in a few milliseconds. We, the IT people, have to build the IT infrastructure of tomorrow to cope with data rates of several gigabytes per second from a single detector, knowing that there will be many more such detectors in the future on the ESRF's beamlines.

But IT infrastructure is only one piece of the puzzle. It is also vital that we structure the data, develop powerful algorithms that allow users to analyse data "on the fly" and provide a data analysis service to users who come to the ESRF with a sample but who are not IT literate. It will be equally important to define a data policy that states how long data are stored at the facility and who has access rights if and when data become accessible to others.

"Soon there will be a paradigm shift at facilities such as the ESRF"

Soon there will be a paradigm shift at facilities such as the ESRF because it will be almost impossible, and certainly impractical, for users to carry away the raw data. Consequently, the role of the facilities will have to be adapted to offer a data analysis service for big-data experiments.

User feedback

All this will not be possible without the strong involvement of all stakeholders, in particular the scientific user community. Big data means big money for the necessary software developments and IT infrastructure, and many of the big-data challenges cannot be addressed by a single facility. This has long been recognised, leading to collaborative IT projects such as PaNdata and CRISP, co-funded under the 7th EU Framework Programme, which are aimed at developing common tools. Many basic tools already exist, but they need to be implemented and improved based on user feedback. Other issues, such as automated metadata capture, massively parallel data analysis algorithms and data curation and archiving need to be addressed urgently.

In the frame of Phase II of the ESRF Upgrade Programme, new software and hardware are required to address the manifold challenges of big data. The active involvement of the user community and other laboratories will be vital in allowing us to keep pace with detector developments and to join forces so that we can define and enhance the tools for data processing. Failure to address these issues now will be painful for everybody.

Rudolf Dimper, *Head of the ESRF Technical Infrastructure Division*

In brief

Student program enters second year

The ESRF and ILL have teamed up with Université Joseph Fourier in Grenoble for the second year running to host a Summer Bachelor Program from 3 June to 14 July 2014.

The program is open to English-speaking science undergraduates and includes a two-week research placement at the ESRF or ILL, which are offering grants to help students meet tuition costs.

This year's successful students hailed from Canada, Poland, Spain, the US and the UK, and each found the exposure to research highly rewarding. "The ESRF has been the highlight for me," said Edurne Sagasta from Spain who worked in the ESRF's surface science laboratory. "This place is full of people wanting to understand how things work. It's what I've always wanted to do."

The application deadline for next year's program is 22 March 2014 (see www.ujf-grenoble.fr/international/bachelor-summer-program/ for more information).

Users invited to shape techniques

The plenary session of the 24th ESRF Users' Meeting will take place on 4-5 February 2014, giving users the opportunity to exchange scientific ideas with each other and with local staff. Three workshops will also be held, each highly relevant to the scientific potential of the new storage ring envisaged for Phase II of the ESRF Upgrade Programme.

The first, "functional materials for electronics", aims to bring together scientists working in micromechanics, microelectronics, spintronics and magnetic or ferroelectric storage and 2D electronic systems. It will address the open problems in materials research and potential applications in new electronic devices, focusing on the synchrotron techniques that are required.

Plenary lectures will be a high point of the 2014 Users' Meeting. Per Ahlberg of Uppsala University, pictured at this year's meeting.

The second workshop, devoted to structural biology, will combine a symposium celebrating the achievements of the recently closed ID14 beamline (see p13) with two half-day sessions concerning the status of the ESRF's new MX facilities. The workshop will also focus on the potential of synchrotron-based serial crystallography.

The third workshop, "high-pressure science at third-generation synchrotron facilities: state-of-the-art and future prospects", will concentrate on investigations of materials under extreme pressures and temperatures. The workshop will delineate this wide-ranging field in condensed matter research and establish research directions.

Registration is open until 21 January 2014, and more information can be found at: www.esrf.fr/home/events/conferences/um2014.html.

Users' corner

At the last proposal submission deadline on 1 September 2013, 918 new proposals were received and reviewed during the Beam Time Allocation Panel meetings on 24 and 25 October. The next deadline for submission of standard proposals is 1 March 2014, while for long-term proposals the deadline is 15 January 2014. Proposers with ongoing long-term projects must submit progress or final reports by 31 January 2014.

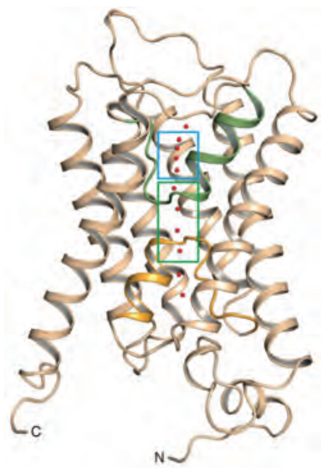
News from the beamlines

- **ID01** will be closed for upgrade from December 2013 until September 2014. User mode is planned from November 2014 on the diffractometer in the new nano end-station.
- The soft X-ray-dependent spectroscopy beamline **ID08** was closed on 9 October. The new soft X-ray beamline at ID32 will start commissioning early in 2014 and will be open for proposals for the 1 March 2014 deadline for both the XMCD and RIXS branches. Proposals for the third "open" experimental area will also be possible. In all cases, users should contact the beamline staff before submitting a proposal. More details can be found on p12.
- A new experimental station for magnetic scattering experiments with circularly polarised X-rays has been installed on **ID12**. This versatile instrument comprises a high-vacuum two-circle diffractometer operating in horizontal scattering geometry, with the sample and detector rotations driven from outside the vacuum by Huber goniometers with a very high accuracy and a reproducibility of approximately 5 μrad. The temperature of samples can be varied from 15–300 K, and an applied magnetic field of up to 4000 Oe can be reversed at a frequency of 1 Hz.
- An experimental setup for XMCD at high pressure, high field and low temperature is also open for users at **ID12**. Complementary to the ID24 setup, the photon energy range is extended down to about 3 keV using fully perforated diamond anvils that provide access to M-edges of actinides and L-edges of 4d-metals. A 6T split-coil room-temperature superconducting magnet has been adapted and a special cryostat has been built to cool both the cell and sample to 2.7 K. Spectra will be detected using a total fluorescence yield technique, allowing measurements both on bulk samples and dilute systems.
- **ID14-4**, the last remaining operating beamline on ID14, closed for user operation at the end of November 2013. The ID14 suite of beamlines will be replaced by ID30 and BM29. More details can be found on p13.
- At the nuclear resonance beamline **ID18**, the synchrotron Mössbauer source has become available for users, providing a tunable (typically ± 1 μeV) X-ray beam at 14.4 keV with an energy resolution of about 10 neV and 10 kHz "resonant" gamma-quanta (V Potapkin *et al.* 2012 *J. Synchrotron Rad.* **19** 559).
- The inelastic hard X-ray scattering beamline for electronic spectroscopy, **ID20**, is fully operational with 100% user operation at target performance.
- After many years of operation, the MAR 225 CCD detector at **ID23-2** will be replaced by a Dectris Pilatus3 2M detector. The new detector will allow more rapid data collection and will significantly improve the quality of data from small crystals thanks to reduced readout noise. A high-speed detector translation table will also be ready for operation in early 2014.
- The high-resolution powder diffraction beamline **ID31** will close in December 2013 and be moved to ID22. The beamline will re-enter user service in April 2014 with an energy range extended beyond the current limit of 63 keV.
- New experiment control software, **MxCuBE2**, has been released on the structural biology beamlines, providing a completely redesigned GUI that simplifies user interaction while allowing more complex experiments to be performed in an automated fashion (see p20 for more information).

How cells allow water to flow

Ultra-high resolution diffraction studies of crystals of a yeast aquaporin molecule carried out at the ESRF's ID29 beamline have revealed the detailed atomic mechanism that enables water to be transported across cell membranes – a crucial process for maintaining the stability of the body's internal environment.

While it has long been known that cells contain numerous channels that allow certain molecules and ions to cross biological membranes, it is less clear how membrane proteins permit some ions to pass through easily while blocking others that are structurally similar. Aquaporins must be highly selective in order to exclude hydroxide (OH^-) and hydronium (H_3O^+) as well as preventing the transport of protons (H^+), which can often be rapidly exchanged between water molecules linked by hydrogen bonds.



An aquaporin monomer comprises six transmembrane helices and one pseudo-transmembrane helix (orange and green) arranged to form a pore.

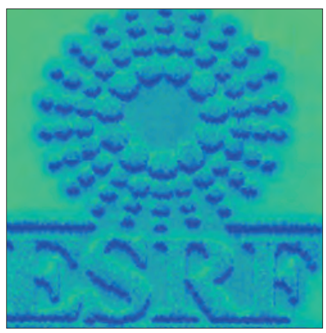
By mapping the structure of an aquaporin protein at a record resolution of 0.88 Å, researchers from the University of Gothenburg in Sweden and the University of Illinois in the US have revealed how the hydrogen-bond geometry of water molecules prevents H^+ conductance without compromising water transport. Electron density maps in conjunction with molecular dynamics simulations also explain how the transport of H_3O^+ ions is inhibited.

The results suggest a pairwise movement of water molecules through the aquaporin selectivity filter, analogous to processes that allow the cellular transport of potassium ions (*Science* **340** 1346). "This functional similarity points to underlying principles concerning how evolution has balanced the trade off of achieving highly selective filters while not binding the transported molecule so strongly that transport is inhibited," said Richard Neutze of the University of Gothenburg.

Surface probe aids chips

A powerful nano-diffraction technique developed at the ESRF's ID01 beamline can detect the slightest imperfections in the crystalline structure of a material, even when buried under a stack of different layers. In the demonstration image on the right, the imperfections take the form of a 90 μm-wide ESRF logo that was etched into a 40 nm-thick silicon germanium film on top of a silicon substrate using focused ion beams. The image shows the distribution of the integrated diffracted intensity obtained in the vicinity of the silicon [004] reflection caused by variations due to strain and tilts induced at the interface of the silicon substrate.

The method is generating significant interest in the microelectronics industry because it can rapidly reveal imperfections in thin films and heterogeneous structures such as semiconductor devices with a strain resolution of a few parts per million and a spatial resolution of 300 nm. "Solder on small contacts is one of the main causes of defects in microelectronic devices, while the technique can also help researchers understand how



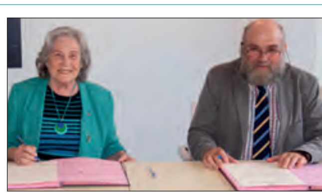
The new technique can reveal the slightest inhomogeneity in thin-layer depositions, as this demonstration image reveals.

copper rods used for 3D chip integration affect semiconductor performance," says Tobias Schülli, scientist in charge at ID01.

Israel boosts participation

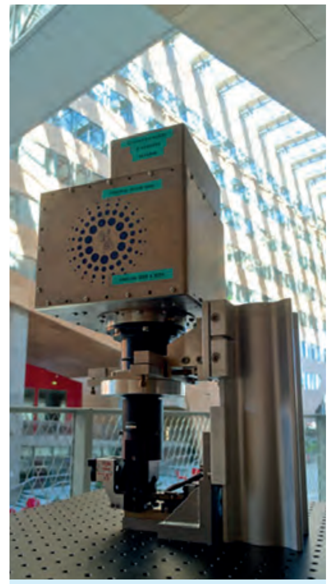
At a ceremony on 27 August, Israel representatives signed an agreement strengthening the country's links to the ESRF. A scientific associate since 1999, Israel has now renewed its membership for a fourth term of five years and raised its financial contribution by 0.5% to 1.5%.

Over 40 Israeli research groups use the ESRF, with excellent


Ruth Arnon, president of the Academy of Sciences and Humanities of Israel, and ESRF Council chairman Jean Moulin.

success rates for proposals that helped motivate the increased financial contribution. In line with overall ESRF use, about 50% of the proposals come from the protein crystallography community.

In 2009, ESRF user Ada Yonath from the Weizmann Institute in Israel shared the Nobel Prize in Chemistry for her studies on the structure and function of the ribosome, while users from Israel have made numerous other scientific breakthroughs – including the elucidation of the crystal structure of vaterite by a group from Israel's Technion institute last year.



Inspired by the ESRF, the exhibition is part of a trilogy dubbed *Thinkrotron.*

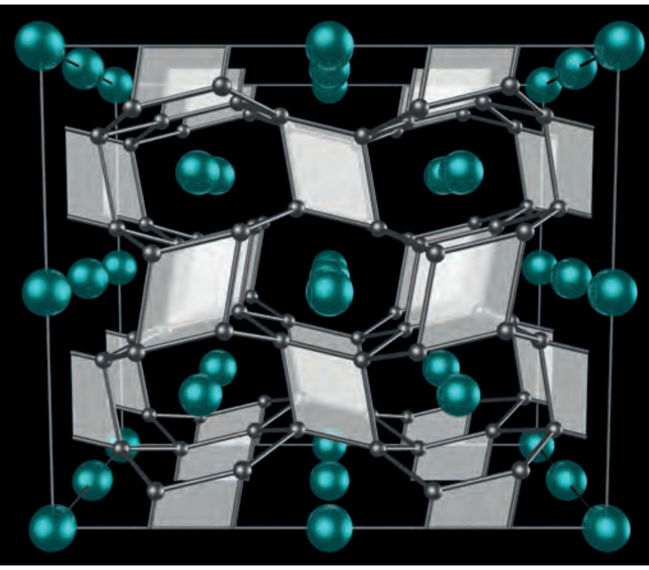
Instrumentation turned into art

After more than ten years of service, the X-ray camera on the ESRF's ID19 imaging beamline had been earmarked for retirement. Instead, it has been put on display at the Hôtel de Région Rhone-Alpes as part of the Lyon Biennial of Contemporary Art. The camera is a centerpiece of an exhibition called "Ghosts of freedom – territories and experimentations" by French artist Laurent Mulot, which will be on display until 5 January 2014.

Mulot has been a resident artist in Grenoble since 2012, where he spent six months recording encounters between ESRF staff and inhabitants of the surrounding community. The exhibition includes elements of these moments, along with similar happenings recorded at CERN and at neutrino particle physics experiments around the globe.

Further ESRF beamline components, including the micro-diffractometer used by Brian Kobilka of Stanford University in experiments that led to him receiving the Nobel Prize for Chemistry in 2012, are on display at Grenoble's Natural History Museum as part of an exhibition called "The Chamber of Echo", also by Mulot.

Discovery opens new era in materials design



Crystal structure of the new superconductor FeB_4 showing iron atoms (green) embedded in a rigid 3D framework formed by boron atoms (grey). No compound was known that had the FeB_4 composition.

The ESRF has verified the existence of the first superconductor, iron tetraboride, to be predicted from first-principle calculations.

Trial and error has been the traditional way to discover new materials – especially superconductors. Ever since Heike Kamerlingh Onnes found in 1911 that mercury loses its electrical resistance below a temperature of 4.2K, researchers have striven to find materials that would allow this valuable property to be harnessed in everyday applications. In 1987, physicists discovered a new class of ceramic materials that become superconducting at 93K, but the prospect of a room-temperature superconductor seems far away.

Recent experiments at the ESRF provide new impetus towards this goal: the confirmation of a novel superconductor – iron tetraboride (FeB_4) – whose existence was predicted from first-principle calculations (*Phys. Rev. Lett.* **111** 157002). FeB_4, which has a brand-new crystal structure, is almost as hard as diamond and represents the first time that a superconductor was predicted from scratch, says Natalia Dubrovinskaia of the University of Bayreuth in Germany, who co-led the study. "It opens a new class of highly desirable materials that combine advanced mechanical properties and superconductivity."

Conventional surprise

Superconductivity is a subtle phenomenon based on correlations between pairs of electrons. It was first described by the Bardeen-Cooper-Schrieffer (BCS) theory, but this picture only explains conventional low-temperature superconductors in which the electron pairing is mediated by phonons. Calculating the superconducting critical temperature of a material is exceedingly demanding and, because BCS superconductivity requires a high density of states, candidate materials tend to undergo structural or magnetic transformations that make them too unstable to exist.

Iron-based superconductors were discovered in 2006 by Hideo Hosono of the Tokyo Institute of Technology in Japan, and attracted a lot of attention because iron-based materials tend to be magnetic, which makes them unlikely candidates as superconductors. The pairing mechanism in such systems is not fully understood, but it is widely attributed to unconventional electron spin and/or orbital fluctuations. "FeB_4 is a bulk superconductor based on iron, but iron does not play a central role," Hosono told *ESRFnews*. "Aside from its classification, though, this superconductor looks interesting because the magnetism of iron is quenched and conventional superconductivity comes out."

The discovery certainly was unexpected, and could help understand what drives superconductivity in iron-based materials, says Aleksey Kolmogorov at Binghamton University in the US, who predicted the new superconductor. In 2010, he and his co-workers carried out detailed CPU-intensive electron-phonon coupling calculations based on an extension of the BCS theory called the Eliashberg formalism (*Phys. Rev. Lett.* **105** 217003). The principal challenge, explains Kolmogorov, was to find a stable material with the brand-new crystal structure, which involved extensive high-throughput and evolutionary searches. "I am aware of only a handful of predicted materials – not just superconductors – with previously unknown structures that have been confirmed experimentally," Kolmogorov told *ESRFnews*.

"Superconductivity is a subtle phenomenon."

Putting theory to the test

In a series of experiments carried out at the ESRF's ID09A beamline last year, the Bayreuth team used single-crystal X-ray diffraction to investigate the structure of FeB_4 and how it behaved under pressure. Because the material does not occur naturally, it was produced in advance by compressing its chemical ingredients to pressures above 8 GPa and temperatures of 1500 °C.

"ID09A has pioneered the development of high-pressure single-crystal diffraction in the megabar pressure range," says Dubrovinskaia. "Our success in studies of single crystals of various materials and achievements of other groups carrying out such high-pressure experiments also stimulated the development of this technique on other synchrotrons, for example at Petra III at DESY." The team's measurements revealed that FeB_4 is stable under pressure to at least 40 GPa and is very stiff in one crystallographic direction, making it a member of the class of superhard materials.

The transition to a superconducting state took place at a temperature of 3K, which is lower than predicted by the original calculations. By doping FeB_4 with electron-rich metals such as cobalt or nickel, however, it could be possible to raise the transition temperature to the theoretical value of 15–20K, says Kolmogorov.

The low-pressure synthesis of FeB_4 and its stability under ambient conditions raise the possibility of producing thin films that combine superhardness and superconductivity. But potential applications were not the focus of the work, explains Dubrovinskaia. "Rather, it offers a change in thinking regarding the possibility of designing superconductors from scratch."

Matthew Chalmers

Lattice leap reconfigures landscape

Sweden's MAX IV facility will be the world's first storage ring to employ a multi-bend achromat lattice, but many other facilities – including the ESRF – are drawing up similar plans in the race towards the X-ray diffraction limit.

P NORDENG

Fourth-generation light: the MAX IV storage ring near Lund in Sweden, pictured in September 2013, will be the world's brightest source of X-rays.

When Mikael Eriksson, machine director of MAX-lab in Sweden, first presented his plans for a radical new storage ring, people working at other light sources thought he was crazy, he recalls. The new lattice promised a factor of 10–100 increase in brightness over existing facilities at the time, offering users unprecedented spatial resolution and taking storage rings within reach of the diffraction limit – whereby the size and spread of the electron beam approaches the diffraction properties of X-rays.

Fast-forward 10 years to today, and not only is the "MAX IV" storage ring soon to produce its first beams, but most of the world's top facilities are planning to upgrade to the same lattice concept. "I took note of all of the initial criticisms and we adjusted our plans accordingly," Eriksson told *ESRFnews*. "If we knew then what we know today, we could have pushed our lattice even further, but now the other labs can benefit from our knowledge – that's how science works."

Lattice evolution

Most third-generation storage rings employ double-bend achromats, a design that has been around since the mid-1970s, while a few employ triple-bend achromats. But in the early 1990s Dieter Einfeld, now a consultant at MAX-lab, and others recognised that a storage ring's emittance – the product of the electron beam size and its divergence – could be reduced (and its brightness therefore increased) by incorporating a higher number of bends into a lattice. Such a multi-bend achromat (MBA) guides electrons around corners more smoothly, thus decreasing the degradation in horizontal emittance. To avoid prohibitively large storage rings, the MBA demands much smaller magnets than are currently employed at third-generation synchrotrons.

In 1995, Einfeld's calculations showed that a seven-bend achromat could yield an emittance of 0.4 nm rad for a 400 m circumference machine – 10 times lower than the ESRF's current value. At around the same time, a six-bend achromat for the Swiss light source and a five-bend achromat for a Canadian light source were also investigated, but the small number of cells in these lattices meant that it was difficult to make significant progress towards a diffraction-limited source, explains Einfeld. "Mikael Eriksson took the idea and turned it into a real engineering proposal for the design of MAX IV," Einfeld told *ESRFnews*.

In 2002, Eriksson presented the first layout of a potential new source: a 277 m circumference, seven-bend lattice that would reach an emittance of 1 nm rad for a 3 GeV electron beam. By 2008, the MAX-lab team had settled on an improved design: a 520 m circumference, seven-bend lattice with an emittance of 0.31 nm rad. This, says Einfeld, is more or less the design of MAX IV that is now under construction.

The MAX-lab team spent almost a decade finding ways to keep the lattice circumference at a value that was financially realistic, says Eriksson, and even constructed a 36 m-circumference storage ring, MAX III, to develop the necessary compact magnet technology. "There were tens of problems that we had to overcome," he recalls. "We also needed a different vacuum system, plus the electron density was so high that we had to elongate the electron bunches."

Construction of MAX IV began in 2010 and commissioning is expected to start in August next year, with official operations scheduled, appropriately, for midsummer's day in 2016. But MAX IV's reign as the world's brightest storage ring could be short-lived. "A lot of other synchrotrons are planning to upgrade to MBAs, for example, the APS in the US, Spring-8 in Japan, LNSL in Brazil and Diamond in the UK," says Einfeld.

Too good to resist

Phase II of the ESRF Upgrade Programme, currently at the stage of a technical design study, envisions a hybrid MBA lattice that would enter operation in 2019 and achieve a horizontal emittance of 0.15 nm rad. Unlike MAX IV, in which the dipoles all have the same strength, the dipoles in the proposed ESRF lattice will have varying strengths.

"The ESRF lattice is quite different to MAX IV, with many original features," says APS associate division director Michael Borland, who first heard of the ESRF design at a workshop in Beijing last year. The APS has decided to pursue a similar design that, should funding be secured, could enter operation by the end of the decade. Being larger than the ESRF, the APS can strive for an even lower emittance of around 0.07 nm rad. Meanwhile, staff at Spring-8 have settled on a hybrid MBA that will enter operation on a similar timescale.

"Everyone is coming to the same conclusions: these machines are challenging, but there are no showstoppers," says Borland, admitting that people will breathe a big sigh of relief when they see MAX IV operating. "Years ago a lot of people thought that the MAX-lab guys were nuts to go to a seven-bend lattice, so they deserve full credit for being bold enough to take that leap," he says. "But all of their machines are slightly unusual – that's their style."

Matthew Chalmers

Soft X-rays move to new pastures

On 9 October the ESRF's soft X-ray beamline, ID08, was closed. Yet this is not the end, nor indeed the beginning, of the story of soft X-ray provision at the ESRF. ID08's pioneering capabilities evolved from a previous beamline, ID12B, and are now to enter an exciting new phase at the ESRF upgrade beamline ID32.

Being a 6 GeV machine, the ESRF is used mostly for hard X-ray studies. But it was realised from the start that a soft X-ray beamline served by exotic insertion devices that produce circularly- and linearly-polarised beams was well matched. In particular, the photon energy range covers the absorption edges of the 3d and 4f elements, which are important in magnetic systems. This allows soft X-ray dichroism studies that provide unique element-specific information about the magnetic moments of electronic shells, while hard X-rays allow other electronic shells to be probed. These two complementary capabilities were realised at the dual-branch beamline ID12, which was one of the first six beamlines to be built at the ESRF. ID12 was also the ESRF's first canted beamline, allowing parallel operation of both branches.

The soft X-ray branch, ID12B, opened to users in early 1995 and generated a lot of demand because a high flux of highly polarised photons was essentially unique at the time. Users brought their own instruments to take full advantage of the facility, and an in-house ultra-high vacuum 7T superconducting magnet shared between both branches allowed X-ray magnetic circular dichroism studies (XMCD). This end-station also gave rise to new surface science techniques on ID12B, which remains a major activity at the ESRF.

P GINTER/ESRF

The ESRF's ID08 beamline is to be replaced by an upgraded facility at ID32.

Science focus

The scientific focus of ID12 was the electronic and magnetic structure of materials, such as transition metal oxides, thin film magnets, exchange bias materials and dilute magnetic semiconductors, which have potential importance for recording devices. More fundamental research on dimensionality effects in magnetic systems ranging from impurities to dots, wires and thin films was also carried out. The beamline also took its techniques into new scientific fields involving molecular magnets and magnetic/electronic phenomena at interfaces. Due to the success of ID12, in 1998 it was recommended that both beamline branches would benefit from dedicated straight sections and that the soft branch should move to the then vacant ID08 port. ID12B closed at the end of December 1999 and the first user experiments at ID08 started early in 2000.

The new ID08 beamline offered a factor 10 increase in photon flux, with two purpose-built undulators allowing much better control of the polarisation. The move offered users a soft X-ray diffractometer allowing studies of orbital ordering phenomena, better sample characterisation facilities and a scanning tunnelling microscope for prior characterisation of samples. The beamline also opened up soft X-ray angle resolved photoemission studies – a field that was in its infancy at the time – and an inelastic X-ray scattering instrument with its own pre-monochromator was installed thanks to a long-standing collaboration with users from the Politecnico di Milano in Italy. The diversity of instruments combined with ID08's flexible layout made for highly efficient use of beam time.

In 2007 a review of ID08 recognised its scientific excellence and, following the consultation process for the ESRF Upgrade program, a new beamline (UPBL7) was selected. The ESRF's brand new soft X-ray facility, which is nearing completion at port ID32, contains an additional undulator that will increase the flux and offer even greater control of the X-ray polarisation. Its scientific focus is twofold: XMCD studies with high magnetic fields, and very high-energy resolution resonant inelastic X-ray scattering (RIXS) studies.

Open for business

The XMCD branch is already equipped with a new 9T/4T fast sweeping UHV high-field magnet, and extensive improvements to the sample preparation facilities are under way. The state-of-the art facility, which will open to users in the first part of 2014, will be ideally placed to address challenges in nano-magnetism and electronic structure. The RIXS branch will provide unprecedented energy resolution, a much improved sample environment and a 10 m scattering arm that can be moved over a range of 100° under vacuum. This instrument (which will come online in the second scheduling period of 2014) will allow scientists to address problems such as high temperature superconductivity at a level of detail not possible today, and will complement the ESRF's hard X-ray upgrade beamline at ID20. The new home of soft X-ray science at the ESRF will also provide a third experimental area to allow for developments such as microscopy experiments, soft X-ray coherence and, most importantly, experiments that we have not yet thought of.

Nick Brookes, scientist in charge at ID12B, ID08 and ID32

ESRF UP Phase I enters final straight

- New beamlines at ID02, ID30, and ID32 are under construction in the Belledonne and Chartreuse halls, with contracts for ID01 and ID31 construction under preparation.
- Reception of works for the site entrance building on Avenue des Martyrs took place on 4 November, with the new entrance opening in February 2014.
- The laboratory and office building, which flanks the Belledonne hall and will house personnel from the CRG beamlines, experiments division, ISDD and TID, was commissioned at the beginning of October.
- The Science Building, which will host laboratories used by the ILL and ESRF, was handed over to its users during November. A covered walkway between this building and the storage ring is expected to be available in March 2014.
- A new cafeteria in the laboratory and office building is now available with vending machines for cold and hot drinks and sandwiches, while renovations to the kitchen areas in the common building are complete.
- Preparations for placing contracts concerning the Vercors building, which was cancelled in 2011, are under way. The building will be constructed if Phase II of the ESRF upgrade is approved.

Bright times ahead for structural biology

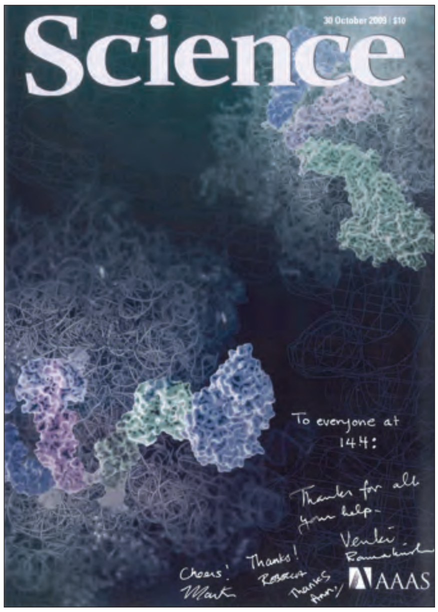
The closure of ID14-4, the last operational beamline in the ESRF's ID14 complex, on 29 November marked the end of an era. Since it first welcomed users in 1998, ID14 had been a major driving force for structural biology, providing Europe's first dedicated undulator-based facilities for macromolecular crystallography (MX) and latterly BioSAXS, which allowed the study of a previously unimagined range of systems. Diffraction data from ID14 has been responsible for a staggering 50% of structures deposited in the Protein Data Bank based on data collected at the ESRF and for almost a quarter of total European deposits.

ID14 has contributed to many breakthroughs in structural biology. Users have gained seminal insights into the activation mechanism of the innate immune complement system, the X-ray induced radiation damage of biological samples and the structures of large membrane protein complexes and viruses, to name a few. Pioneering experiments at the facility also combined UV-visible spectroscopy and X-ray crystallography, and drove the development of instruments including the SC3 sample changer (which ushered in the era of high throughput crystallography at the ESRF) and the MK3 mini-kappa goniometer (which enabled the automatic reorientation of crystals).

Crowning achievement

The crowning achievement of ID14 has undoubtedly been its role in elucidating the atomic structure of the ribosome – the complex molecular machinery that links amino acids in the order specified by messenger RNA molecules. This led to fundamental insights into the mechanisms of biological protein production and, ultimately, to a share of the 2009 Nobel Prize in Chemistry for two long-term ESRF users: Venki Ramakrishnan of the MRC Laboratory of Molecular Biology in the UK and Ada Yonath of the Weizmann Institute of Science in Israel.

While the tunable, high-intensity ID14-4 beamline – which was operated jointly by the ESRF and the European Molecular Biology Laboratory (EMBL) – should take the lion's share of the glory in this regard, the role of the other three ID14 end-stations should not be forgotten. These were often used to evaluate ribosome crystals before proper data collection, demonstrating a highly productive symbiosis that enabled users to tackle projects that require large amounts of synchrotron beam-time.



Researchers thank ID14-4 staff with a signed copy of *Science* magazine, in which the structure of the ribosome was published.

A one-day symposium dedicated to ID14-4 and its contributions to European structural biology will be held on 3 February 2014 in conjunction with the ESRF Users' Meeting.

Of course, the closure of ID14 does not mean the end of structural biology at the ESRF. The highly productive beamlines ID23-1, ID23-2, ID29 and BM29 (BioSAXS) will continue to operate with upgraded functionality. Moreover, a new generation of MX beamlines is currently under construction at ID30 as part of the ESRF upgrade beamline UPBL10 ("MASSIF"), which will use canted undulators pioneered on ID23 to provide four new end-stations for MX at the ESRF.

MASSIF status

The first of the canted undulators will provide the source for the MASSIF suite of beamlines. This will comprise three end-stations: two (MASSIF-1 and MASSIF-2) offering rapidly variable focal spot sizes of 120 μm^2 – 50 μm^2, and the third (MASSIF-3) providing a microfocus beam less than 10 μm^2 at the sample position. These focal spots will contain much higher photon fluxes than were previously available on ID14, and the end-stations will also be equipped with ESRF-developed direct data collection (DDC) robots that act as both sample changer and goniometer. Coupled with the availability of the latest pixel detectors, next-generation automation and an array of complementary techniques (such as crystal dehydration, X-ray fluorescence analysis and on-line UV-visible/Raman spectroscopy), the MASSIF beamlines will allow an extremely high sample throughput. MASSIF-1 and MASSIF-3 will take first users in the summer and autumn of 2014, respectively.

The second of the canted undulators will provide the X-ray source for ID30B. Due to open to users towards the end of 2014, the beamline will be tunable and operated, as was ID14-4, as a joint ESRF–EMBL collaboration. ID30B will be equipped with a Pilatus3 6M detector and will provide a high photon flux in a focal spot of rapidly variable size. Current plans include the installation of an *in situ* sample screening and data collection facility on the beam-line, thus further extending its functionality compared to its predecessor ID14-4.

This new generation of MX beamlines will complement the continued refurbishment of the existing portfolio and ensure the routine performance of the sophisticated experiments that are now required in structural biology. Important research targets include elucidating the structures of eukaryotic macromolecular complexes and membrane proteins such as G-protein coupled receptors.

Andrew McCarthy (EMBL-Grenoble; ID14-4 and ID30B), Gordon Leonard (deputy head of structural biology group) and Christoph Mueller-Dieckmann (ID30, manager UPBL10)

Visitor centre opens

With its large glass wall providing an impressive panorama of the new experimental halls, where four upgrade beamlines are presently under construction, the ESRF's brand new visitor centre takes visitors straight to the action. Inaugurated on 7 November, the 200 m^2 facility conveys the essence of the ESRF and X-ray science through a mixture of animated panels, video presentations, interactive models and exhibits. Its flexible configuration offers a 50-capacity lecture theatre, functions for up to 100 people and a learning space for up to 36 school pupils. "We expect a throughput of around 5000 persons per year in guided group visits, plus substantial use by ESRF users," says Claus Habfast, head of communications.

The centre received its first visitors during the ESRF open days held on 12–13 October, when more than 1000 members of the public were welcomed to the EPN campus, coinciding with a shutdown of accelerators.

C JARNIAS

Claus Habfast welcomes visitors to the new centre, located above the Belledonne hall.

Synchrotrons face up

The vast range of science experiments carried out at the ESRF combined with advances in fast detectors present a unique big-data challenge.

A petabyte is a lot of data. Translated into piles of DVDs, the popular currency of comparison, a petabyte (1 million GB) corresponds to over 200,000 disks. Or, in more modern terms, a petabyte is roughly the volume of data processed by Google every hour.

Experiments at the ESRF are now producing an average of 3 TB every day, equating to more than 1 PB per year. Some 75% of this is taken up by the imaging group, with protein crystallographers the second biggest consumers at around 10%. Furthermore, data volumes are expected to continue to double every 18–24 months for the foreseeable future. The exponential growth, driven mainly by advances in detectors, presents major challenges in data capture, curation and analysis (see p18). For some users, data collected during their stay at the ESRF are already becoming too voluminous to be physically transported back to their home institute on a single disk.

Storage capacity is the lesser problem, however. More challenging is storage latency, which is the speed at which one can write and read data back from disk when dealing with detectors that stream several TB per hour. This puts the ESRF firmly in the realm of "big data", a fashionable term describing datasets that are too large for traditional data management and analysis tools to be effective.

Data-intensive science

The rise of big data is sometimes referred to as the "fourth paradigm" in science after theory, experimentation and computation. Give researchers unprecedented access to data and the necessary computational muscle, many argue, and accelerated breakthroughs will follow thanks to new ways to combine, manipulate, analyse and display data. Companies of all shapes and sizes also view big data as a business opportunity.

Big data is familiar territory for astronomers and also for particle physicists, who have had to invent "grid" technologies in order to cope with the 15 PB per year that stream from the Large Hadron Collider. Grid technology has spread to other communities under the European Grid Infrastructure (EGI) but a 2011 report by members of EIROforum concluded that the Grid offered no benefits for synchrotrons, mainly because synchrotron data is far less uniform than that produced by the particle physics and astronomy communities. The potential of scientific cloud computing – an extension of the grid – is also being explored (see p19).

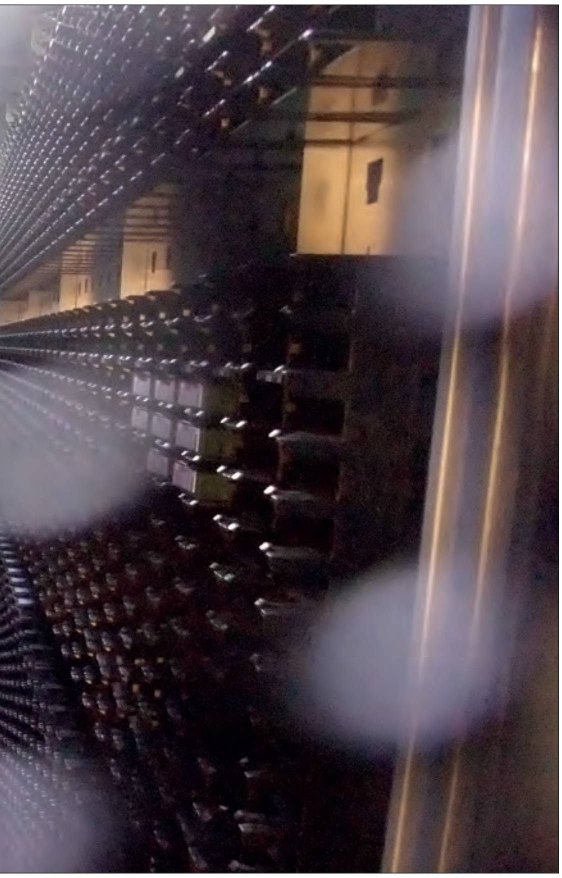
The numbers speak for themselves: the ESRF has more than 30,000 users in Europe coming from more than 10,000 institutes and working in tens of different fields, making the data highly scattered and heterogeneous

ESRF's data centre keeps pace with the data onslaught

In 2011, the ESRF inaugurated a new data centre to help cope with future demand. Although experiments generate some 3 TB each day, the need to visualise and process these data in order to give users the best possible service means that the ESRF's IT infrastructure shifts around 10 TB of data every day. With this volume set to grow, the availability of electrical and cooling power is vital. The new data centre has a capacity of 370 kW, almost five times more than before, plus it allows new equipment to be installed on demand.

All data are backed up via two robotised tape libraries (pictured) that can each host up to 64 magnetic-tape drives and 8500 2 TB tapes. The data copied to tape come mainly from the disk storage systems, which are located in a separate computer room to safeguard against fire. Currently, the two libraries store a total volume of 4 PB and this volume will rapidly increase with the archiving of in-house research data. A full mesh network is now in place, offering up to 25,000 1 Gbit/s copper ports and up to 3500 10 Gbit/s fibre ports for server connections, storage systems and network backbone connections. Data is analysed on computational clusters based on Intel CPUs and Nvidia GPUs, with thousands of processor cores working in concert.

to the petabyte era



B LEBAYLE

"The facility that provides the best computing environment has a competitive edge."

(see p17). "It is crazy that we are generating datasets in formats that few can read," says Rudolf Dimper, who is head of ESRF's Technical Infrastructure Division. "Why can't we define between synchrotron and neutron sources a standard data format? It's a very simple idea, but we've been trying to do this for 20 years."

According to Steve Aplin of DESY's scientific computing group, standardisation of synchrotron data will only follow where there is a genuine need for people to share data – otherwise it could be seen as an unnecessary burden to scientists. " 'Big data' is part of a current buzz that certainly helps to shift to bigger and faster servers and software in industry, but for synchrotrons big data is a real problem that needs a different approach," he told *ESRFnews*. "New 2D pixel array cameras have provided a massive jump in the amount of data, leading to a real shift in how researchers deal with their data, how they share it with other collaborators, and how they access it."

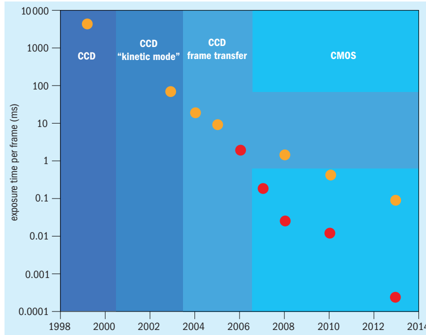
Opening up

Many governments want to see scientific data made as open as possible in the hope that it leads to higher returns. It is a natural extension of the open-access publishing model, yet the challenges of making data open are much tougher (see page 21). "It is undeniable that the ability to store and process large amounts of data helps to produce tangible results, says Aplin. "Assuming that the data produced is not sensitive, the only real risk in storing all the data long term seems to be that of producing huge data silos filled with data dumped in there simply because of well intentioned yet poorly thought-through policy."

Head of the European Commission's e-infrastructures unit, Kostas Glinos, told *ESRFnews* that the challenge is to bridge the gap between the user communities and those that develop and operate e-infrastructures. "Technology allows levels of abstraction and interoperation between heterogeneous systems so that communities can largely share solutions," he says, pointing to the Research Data Alliance and EUDAT. "So far we see that building networks of smaller facilities can be a good solution in many cases, for example networking (GEANT), generic data and computing (EGI and EUDAT), as well as community-driven data infrastructures (PaNdata) where the federated model seems to work well."

Like most synchrotron facilities, and in contrast to neutron sources, the ESRF does not yet have an official policy stating whether data should be stored and/or be made openly available. But the issue of how much time users can spend carrying out online analysis could become one of the main differentiators between labs, thinks Mark Könnecke of the Paul Scherrer Institute in Switzerland. "The complexity of experiments, data handling and analysis is increasing while the knowledge of users about scientific computing is decreasing," Könnecke told *ESRFnews*. "The facility that provides users with the best computing environment and gets them physical meaningful data fast has a competitive edge, and such an environment might result in more science coming out of the data."

Matthew Chalmers



The evolution of sample exposure times at ID15A at spatial resolutions of 0.02 mm (red) and 0.002 mm (orange) thanks to advances in photon detectors, optics and the source.

Detectors race ahead

Advances in detectors have completely transformed the way X-rays are detected at synchrotrons. The Pilatus 6M photon detector, which is currently installed at two beamlines at the ESRF, is capable of churning out 25 six mega-pixel images per second, corresponding to more than 175 MB/s, with high-duty cycles making data rates of 10 TB per day possible.

The Pilatus detector will soon be superseded by the Eiger detector, two of which are expected to be installed on ID13 and ID30 next year. The four mega-pixel version will generate up to 750 images per second, corresponding to data streams of up to 8 GB/s depending on the number of bits per pixel. Even after compression, this represents almost a factor 10 increase over the Pilatus detector. Another example is the PCO.dimax S camera, which generates four mega-pixel images at a rate of 1300/s. Data bursts of 10 Gbit/s are stored in internal memory before being sent to disk via a dedicated camera link.

High-resolution pixel detectors demand increasingly high-speed write access to storage media, which is not possible with large shared storage systems. The ESRF has therefore developed a solution based on off-the-shelf commodity storage, which acts as a large buffer before sending the data into the much bigger central storage system. This stop-over storage is used to empty the detector memory but it also allows quick reading of data to allow users to make early decisions about data quality.

Datasets produced by next-generation detectors may simply be too large to transport back to a user's home institute, and increasingly will demand dedicated central IT infrastructures.

Number of unique users shared between European photon and neutron facilities compared to the total number of users

	ALBA	BER II	DESY	DLS	ELETTRA	ESRF	FRM-II	ILL	ISIS	LLB	SINQ	SLS	SOLEIL	neutron	photon	all
ALBA	773	7	61	58	51	281	2	51	13	5	10	77	105	69	400	773
BER II	7	1563	115	46	27	179	157	383	198	98	191	62	36	580	329	1563
DESY	61	115	4197	137	222	851	116	255	113	62	95	315	188	469	1294	4197
DLS	58	46	137	4407	102	810	30	267	399	33	52	229	192	546	1130	4407
ELETTRA	51	27	222	102	3167	433	11	77	35	20	18	179	367	141	900	3167
ESRF	281	179	851	810	433	10287	139	900	369	190	174	963	1286	1313	3586	10287
FRM-II	2	157	116	30	11	139	1095	347	137	89	161	33	29	509	259	1095
ILL	51	383	255	267	77	900	347	4649	731	301	395	156	222	1518	1347	4649
ISIS	13	198	113	399	35	369	137	731	2880	89	233	94	56	936	745	2880
LLB	5	98	62	33	20	190	89	301	89	1235	74	39	151	391	323	1235
SINQ	10	191	95	52	18	174	161	395	233	74	1219	224	31	590	415	1219
SLS	77	62	315	229	179	963	33	156	94	39	224	3827	399	371	1470	3827
SOLEIL	105	36	188	192	367	1286	29	222	56	151	31	399	4568	394	1817	4568
neutron	69	1563	469	546	141	1313	1095	4649	2880	1235	1219	371	394	10023	2334	10023
photon	773	329	4197	4407	3167	10287	259	1347	745	323	415	3827	4568	2334	25336	25336
all	773	1563	4197	4407	3167	10287	1095	4649	2880	1235	1219	3827	4568	10023	25336	33025

SOURCE: PANDATA CONSORTIUM (PAN-DATA.EU/USERS2012-RESULTS)

Shared tools bring better science

PaNdata co-ordinators envision an open-data infrastructure for European photon and neutron sources.

Research carried out at neutron and synchrotron facilities is rapidly growing in complexity. Experiments are also increasingly being carried out by international research groups and in more than one laboratory. Combined with the increased capability of modern detectors and high-throughput automation, these facilities are producing an avalanche of data that is pushing the limits of IT infrastructures.

In addition, there is a push from policymakers and some scientific communities to make data "open" in order to encourage transparency and sharing between scientists. It therefore makes sense to build a common data infrastructure across different photon and neutron facilities that makes data management and analysis more efficient and sustainable, maximising the science throughput of their user communities.

Established in 2008 by the ESRF, ILL, ISIS and Diamond, the PaNdata consortium brings together 13 large European research infrastructures that each operate hundreds of instruments used by some 33,000 scientists each year. Its aim is to provide tools for scientists to interact with data and to carry out experiments jointly in several laboratories.

Research undertaken by PaNdata partners shows that more than 20% of all European synchrotron and neutron users make use of more than one facility (see table). It is therefore of considerable value to offer scientists a similar user experience at each facility, and to allow them to share and combine their data easily as they move between facilities.

At the heart of PaNdata are shared data catalogues that allow scientists to perform cross-facility and cross-disciplinary research with experimental and derived data almost in real time. The catalogues follow the data from its creation and analysis to the final publication of results, which feed back into new research proposals.

Next stage

The first stage of the project, PaNdata Europe, which ran from 2009–2011, focussed on data policy, user information exchange, scientific data formats and the interoperation of data analysis software. Essential elements of a scientific data policy were agreed, covering aspects such as storage, access and the acknowledgement of sources. A second project, PaNdata Open Data Infrastructure, which will conclude in 2014, includes a common data catalogue and user authentication system across the participating facilities. The project is also tracking the analysis process so that users will be able to trace how data are used once they have been collected. Since a data infrastructure must be sustainable, the consortium is investigating which processes and tools need to be changed in order to allow a facility to move towards long-term data preservation.

These and other topics may well form the basis of further collaboration in the European Commission's Horizon 2020 programme. Furthermore, the recently established Research Data Alliance (RDA) will allow other data managers and scientists throughout the world to collaborate. The RDA is a new organisation backed by the EC, the NSF in the US and the Australian National Data Service and includes several working groups in areas such as metadata, data publishing, digital preservation and policy enactment. PaNdata is establishing a RDA interest group to develop best practise for data management in the photon and neutron community.

The ultimate vision of PaNdata is to allow users to move within and between facilities without having to learn how to use new computing systems. By allowing data to be moved, shared and mixed together simply across the complete life cycle of an experiment, scientists can concentrate on getting the best science from the facility.

Juan Bicarregui, Brian Matthews and Erica Yang of the STFC Rutherford Appleton Laboratory in Harwell Oxford, UK

PaNdata's goals include:

- Allow users registered at one facility to access resources across the consortium using one identity;
- Create standard formats so that data generated by one instrument can be readily combined with data from others;
- Build a federated data cataloguing system with a common metadata standard, allowing users to access data generated from different sources in a uniform way.

Data fit for the future

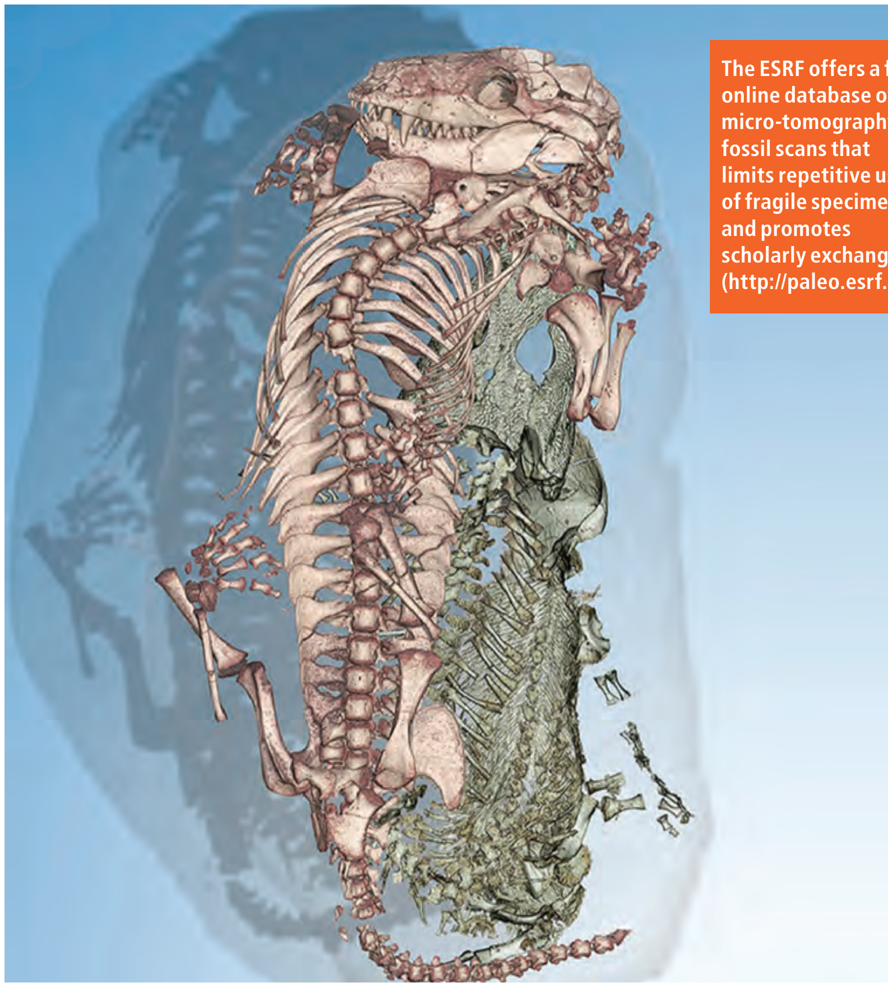
Synchrotrons face technical and sociological hurdles in preserving their soaring datasets for future users.

In a bank of disks and tapes located in the data centre of the Institut Laue–Langevin (ILL) are stored every bit of data generated since the facility's first experiment in 1973. The UK's ISIS neutron source has done similarly since 1985, and the Swiss neutron facility SINQ since 1996. Most synchrotron facilities, by contrast, tend to delete all data after a period ranging from a few days to a few months.

"For a very long time, neutron data rates were so small compared to the increase in disk capacity that it was the lazy option to just copy the stuff over rather than think up deletion policies and haggle with users," says Mark Könnecke of the Paul Scherrer Institut in Switzerland, where he is responsible for SINQ data acquisition. "This is only now changing, but a pattern has been set: the neutron user expects to be able to ask for data collected 10 or 20 years back."

Neutron facilities also have a clear stance on making data openly available, with the ILL providing online access to all data since 1995 and, from last year, linking its datasets to publications via the digital object identifier (DOI) system. "Big data means two very different things for us: a large increase in storage/network capacity and computational resources, and proper curation and open-access to the data," says Jean-François Perrin, head of computing at ILL. "Big data is probably a real scientific opportunity, but there is still a long way to go for it to demonstrate its full value."

The brute fact facing synchrotron data policy, in addition to a highly scattered user base, is the sheer size of today's datasets: soon the ESRF will generate the equivalent of the ILL's entire dataset in a single day.



The ESRF offers a free online database of micro-tomography fossil scans that limits repetitive use of fragile specimens and promotes scholarly exchange (http://paleo.esrf.eu).

EVOLUTIONARY STUDIES INSTITUTE, JOHANNESBURG

Brute volume

Archiving is not just about making data available to anyone and everyone, though. Paleontology research at the ESRF, for instance, is carried out both in-house and by external users so it is useful to store and curate data without necessarily making data accessible beyond Grenoble. The ESRF has established an online open-access archive for certain fossil scans (see image), but still the hutches of the imaging beamlines are overflowing with individual USB disks containing unique datasets.

There are several steps required to make data persist. The first is to provide access to the data for a specified group of people, which requires robust identity management. The second challenge is to capture metadata so that data are sufficiently well described and that they can actually be reused. A third issue is the data format, which needs to be standardised. Finally, a database and a DOI identifier are needed to locate data and link them to a corresponding publication.

Neutron sources have benefitted from the NeXus file format in this regard, and the ESRF's IT group is working on similar forms of metadata capture and identity management. What synchrotrons lack, says head of technical infrastructure Rudolf Dimper, is the push from scientists to make sure that such efforts are "not something IT people do in their ivory towers." One technical bottleneck, he says, is the speed at which data can be written to tape, which at the ESRF is currently around 80 MB per second. Another challenge is to manage the size of the backup catalogues, which contain hundreds of millions of entries per year.

Newer facilities are perhaps better prepared. Diamond in the UK, for instance, has all its data stored on disk in over 100 million files, and according to Könnecke has "set the stage" for synchrotron data management. Meanwhile, the European XFEL, which is expected to initially generate 10 PB per year rising to more than 50 PB as a result of detector upgrades, will store all data in a large disk system similar to those used by companies such as Google.

New way of thinking

Steve Aplin of DESY's scientific computing group says that the data volumes generated by next-generation photon detectors require a new way of thinking – pointing out that previously it was scientists, not the facilities, who were the custodians of the data. "Archiving is pretty much a solved problem, although long-term media migration remains a significant challenge, but curation is certainly one of the hardest things to solve because it requires a large amount of work that cannot easily be scaled up using purely technical solutions," he told *ESRFnews*. "How this will evolve, I expect, will largely be determined by the more general policy regarding access to scientific data as determined by funding agencies."

Matthew Chalmers

Science heads for the clouds

The Helix Nebula initiative is exploring the potential of linking European research facilities with commercial IT providers. Should synchrotrons get on board?

Twenty years ago, computing "farms" were the must-have resource for processing large datasets. As high-speed networks became available, attention shifted to distributed computing or "grid" technologies developed in the late 1990s to cope with data from CERN's Large Hadron Collider. Today, "cloud" computing – the ability to store and analyse data on remote platforms – is the latest big-data buzzword. Whereas the grid is about many user communities sharing their resources, the cloud is about commercial parties selling digital products such as storage, analysis or software to scientists.

In 2011, the European Molecular Biology Laboratory (EMBL), European Space Agency (ESA) and CERN joined forces with over 20 commercial IT providers to launch "Helix Nebula – the Science Cloud" to investigate the potential of scientific cloud computing in Europe. The project takes a "big vision" of what science will look like in the future, says Bob Jones, head of CERN openlab and member of the EIROforum IT working group. "It's all about making data accessible to as many scientists as possible."

Going live

Computing for science has become an operational cost just like offices and other things that have to be paid for, explains Jones, but a scientist or institute is unlikely to invest in computing unless they have very large computing demands. "In the future, governments might pay their scientists to buy their own IT services or even give them credits to buy digital commercial products, allowing researchers to focus on their core business," he says.

There are various ways in which the cloud can be constructed, including a federated model similar to the grid whereby lots of individual clouds are linked. But the basic idea is simple: put computing resources where power and cooling is cheap and deliver it to where it is needed.

The Helix Nebula is due to go live to users in early 2014, following a two-year pilot phase that has seen ESA test the technology for environmental disaster monitoring, EMBL for next-generation sequencing experiments and CERN for LHC data simulations. Neuroimaging, oceanography and European medium-range weather forecasting are other communities keen to make use of the cloud, says Jones. "We are now discussing the potential of the cloud with other members of the EIROforum [which includes the ESRF] because we have to show the providers that it is commercially viable," Jones explains. "Were synchrotrons to link up with cloud e-infrastructures they could move the data to the users, who would benefit from a single data-storage and analysis facility."

"The problem of how to get hundreds of terabytes to the cloud efficiently is not solved yet."

Business case

Cloud computing offers a radically different business model to that currently employed by scientists, who traditionally contact their own laboratory if they have a problem with their IT, and for many communities the benefits of the cloud are yet to be proven. Indeed, a financial study of e-infrastructures recently carried out by the FP7-funded e-FISCAL project (efiscal.eu) found "strong evidence" that cloud prices are not necessarily and unanimously lower than the costs for in-house service provision.

Since synchrotrons and X-ray free electron lasers already have well established IT infrastructures of their own, the big test of cloud models will be how seamlessly these resources can be integrated to make the initial effort needed by the researcher worthwhile. The network topology of synchrotron science is not ideally adapted to the cloud, for instance. Whereas an experiment at CERN typically can run for decades, justifying investment in high-speed links and storage space, the ESRF has a steady turnover of experiments lasting a few days at most. Afterwards, users take the data away to their home institutes, which often have low-speed Internet connections.

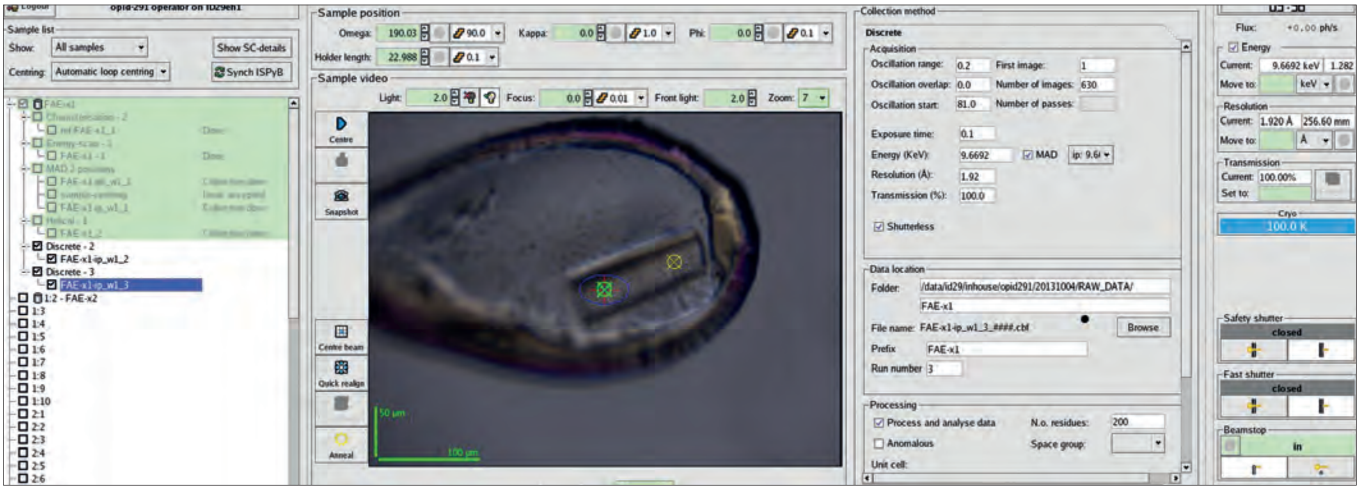
"The problem of how to get hundreds of terabytes to the cloud efficiently is not solved yet," says Andy Götz, head of ESRF's software group, who also questions the benefits of the Helix Nebula over existing open-source solutions like OpenStack. "One solution might be to set up a cloud at the ESRF that provides a data-analysis service, potentially serving the wider EPN campus if we had the necessary hardware, software and staff," he says. "There are lots of cloud solutions being discussed, but things like zenodo (zenodo.org) are also interesting because they offer open-access data storage for everyone."

Data hotspots

Open data and algorithms made available in a pre-configured processing environment are important aspects of the Helix Nebula, says Rupert Lück, who is head of IT services at EMBL Heidelberg. "Helix Nebula could be a useful platform for light-source labs by allowing scientists to store, analyse and share their data," says Lück, adding that other EIROforum facilities plan to join the initiative. "At EMBL, in addition to large-scale genomic analyses it will be interesting to see if we can leverage the technology's potential for the massive quantities of data coming from other life-science domains such as biological imaging or research at synchrotron and XFEL sources."

Cloud services can only be contemplated thanks to the enormous advances in high-speed networks. Back in the 1990s, explains CERN's deputy head of IT, David Foster, LHC planning relied on a few 622 Mbit/s circuits whereas today the lab employs 100 Gbit/s circuits. "The grid continues to be upgraded to cope with the LHC's demands and has spread to other communities," says Foster. "Helix Nebula is about managing what comes next."

Matthew Chalmers



The ESRF's graphical user interface MxCuBE has recently been redesigned to have an even closer symbiosis with the ESRF's ISPyB database.

Biologists enter the fast lane

The ESRF's structural biology group is setting new standards in data management.

The ESRF's macromolecular crystallography (MX) beamlines evaluate more than 100,000 samples every year, contributing half of all European and a fifth of global Protein Data Bank deposits. Each new sample produces roughly 15 GB of raw diffraction images and corresponding processed data, making MX the second biggest consumer of the ESRF's computing resources after the imaging group. This data explosion is set to continue, with the highly automated ESRF upgrade facility MASSIF expected to process up to 500,000 samples each year.

Fortunately, the ESRF has pioneered a highly efficient system with which to manage the MX data deluge, while also providing users with advanced online data analysis. ISPyB (Information System for Protein crystallography Beamlines) allows users to track in real time all stages of the progress of a synchrotron diffraction experiment. From its initial status eight years ago as a database confined to simple sample tracking and experiment reporting, ISPyB has evolved into a sophisticated and interactive system for experiment management and data-analysis retrieval.

Next-generation structures

"Without ISPyB we would not be able to consider next-generation experiments such as those offered by MASSIF, where you are going to have many thousands of diffraction images coming from various small and anisotropic crystals," says Stéphanie Monaco, ESRF industrial liaison engineer for structural biology. "It will allow us to automatically apply previously determined data-collection strategies, provide advice to users on what to do next with their samples and display the results of automatic data processing and structure solution pipelines."

> **"It took 18 months of intense coding."**

Currently, the ESRF is collaborating with Diamond Light Source in the UK where ISPyB is integral to MX experiments. Other European light sources have also expressed an interest in making use of the database, which has recently been extended to provide similar features for small angle X-ray scattering experiments from solutions of biological macromolecules (BioSAXS). These developments pave the way for a pan-European information-management system at the hub of structural biology research.

The relentless rise of big data, however, demands a continual evolution of the system. Future projects in structural biology will require screening of the diffraction properties of several thousand, rather than several hundred, crystals in order to select those most suitable for data collection. Furthermore, users are likely to require combinations of partial datasets collected either from different crystals or from different parts of the same crystal. Indeed, sample pre-screening will increasingly take place at different beamlines – if not different synchrotron facilities – prior to full data collections, explains Monaco. "Future developments of ISPyB will therefore require workflows to be modified to take into account previously collected data, using a revised data model based on projects rather than experimental sessions."

Without the ESRF's beamline control and data acquisition software MxCuBE, however, the ISPyB database would be empty. "MxCuBE is the brother of ISPyB," says Daniele de Sanctis of the ESRF's structural biology group. "One cannot live without the other."

Since its release in 2005, MxCuBE has become the preferred data-acquisition software for structural biologists, providing a platform for modern high-throughput MX and allowing users to carry out experiments remotely. It is also installed at other European synchrotrons including Soleil, EMBL@Petra III, BESSY and MAX-lab. In collaboration with the ESRF, EMBL Grenoble and Global Phasing Ltd, these facilities have helped develop next-generation graphical user interfaces and control software that are capable of interfacing with a variety of low-level control systems.

MxCuBE mark II

After 18 months of intense coding, experiment design and testing, the ESRF has recently launched MxCuBE2, which is available on all the ESRF's MX beamlines. Written in Python, the interface has a radical new appearance and provides an updated environment for performing complicated multi-crystal and multi-position MX experiments in a modular, logical and automatic fashion. The new interface also provides access, via the open-source software DAWN, to a set of recently developed experiment workflows designed to help users collect the best data possible.

"MxCuBE2 provides a new standard for data-collection graphical user interfaces in Europe, and thanks to its collaborative structure it is in constant evolution," says de Sanctis, who led the MxCuBE2 project. "New functionality and features for the graphical user interface are already under development, so stay tuned."

Matthew Chalmers

Open data divides opinion

Whether or not synchrotron data should be made freely accessible depends on who you talk to. *Matthew Chalmers* spends a column in each camp.

"An opportunity for new knowledge"

When the UK government released a database containing unique identifiers for more than 350,000 points of access to the country's transport infrastructure, members of the public began to interact with the data. Some people corrected the precise locations of bus stops, while others developed mobile apps that helped people get to work on time. The move demonstrated the power of "open data" to improve the lives of individuals.

Announcing its open-data strategy in 2011, the European Commission estimated that €40 bn per year could be added to the European economy if big data was made more open. Recently the commission launched an "open data portal" containing datasets covering everything from work accidents to wine production, and similar initiatives are under way in the US and UK.

Science, which is increasingly becoming a more collaborative, interdisciplinary and data-intensive enterprise, is seen as a perfect fit to the open-data movement. According to a 2012 report by the Royal Society, the ability of computers to create and explore massive datasets is a source of knowledge that could unleash a second "open" science revolution as great as that triggered by the creation of the first scientific journals.

Astronomers are already heading in this direction, with "citizen science" projects such as Galaxy Zoo and SETI@home. Although X-ray data are far less homogeneous than data in astronomy or particle physics, important progress is being made in identity management, standardisation, metadata-capture and archiving, which are vital steps to make it easier to build on the findings of public-funded research.

The open-data movement also has implications for the credibility of science. By linking publications to the underlying data, scientists can be held to account by making it easier to detect fraud.

In a short period, the Web has transformed the way that we communicate and interact with information. Scientists have pioneered open movements in software, hardware and publishing. It can be argued that communities or facilities that do not adapt to this changing environment risk falling behind. ●

"A threat to intellectual property"

It is infinitely easier to talk about the merits of open-data than it is to put the idea into practice – at least for synchrotron science. At one level, which is already a major challenge, vastly different types of data need to be standardised and then archived and curated for periods of decades or more – incurring significant financial costs. Another important issue, however, is intellectual property – the lifeblood of science.

Open-access data makes sense for astronomy, where the samples (the stars) are persistent and the property of everybody, and the model also suits the one-off nature of particle physics experiments. The unique samples studied by paleontologists are similar in this respect. At synchrotrons, once a fossil has been scanned at micrometre resolution there is often little further information to extract from it. Combined with the potential damage to samples caused by frequent handling, it makes sense to make fossil scans available in an open-access database – as the ESRF has pioneered.

But consider a biologist who has spent many months at a laboratory bench trying to obtain a protein crystal, or a physicist who has carefully prepared a semiconductor wafer, before it was possible to collect any data in the first place. Were data suddenly to become publicly available, it could effectively disown scientists of everything that they did beforehand and reduce the motivation for carrying out preparatory work. Taking this argument to the extreme, one could imagine scientists trawling open data in search of a quick publication. But anyone who tried to interpret data collected from somebody else's sample might draw wrong conclusions.

The problem with making synchrotron data open is deciding which data to keep. Published data are usually only part of the full data involved, such as alignment scans and datasets generated by online analysis, but it is unfeasible to store every single piece of scientific data. As for open data reducing scientific fraud, the handful of cases already detected must be weighed up against the costs of making data open.

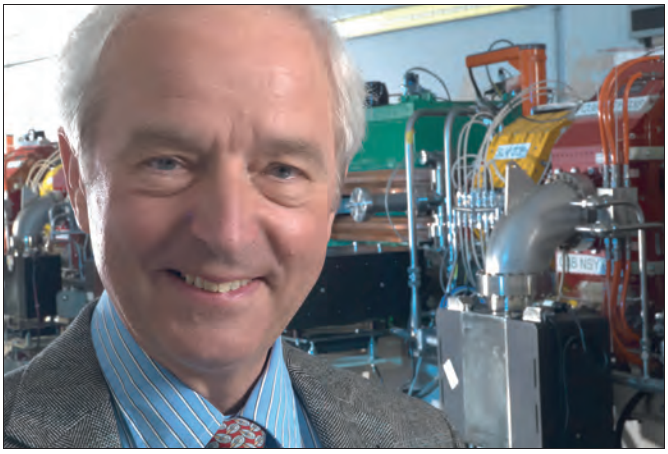
The case for open data in science is not yet proven, and for facilities such as synchrotrons it presents a formidable challenge. ●

X-ray visionary

Gerhard Materlik, who has recently stepped down as CEO of Diamond Light Source, reflects on the origins and future of the light-source landscape

It was 1973, and 28-year-old PhD student Gerhard Materlik was returning home from a visit to the DESY laboratory in Hamburg where he had just seen his first synchrotron X-ray source. His supervisor was driving, and the pair was so excited about the scientific potential of the machine that they lost their directions. "In the car we composed a list of all the things we could do with such a brilliant, continuous source of X-rays," says Materlik. "Everything suddenly became possible – we could see the structure of matter. Right there I decided that synchrotron light was going to be my future."

What he couldn't know is that his list merely scratched the surface of synchrotron X-ray applications, nor that he would wind up building a £260 m X-ray source from scratch in the English countryside.



Gerhard Materlik in brief

Born: 1945, Marl, Germany.

Education: PhD, Dortmund University (1975).

Career: Post-doc, Cornell University (1975–1977); research associate (1978–1985), director HASYLAB (1986–1993), associate director and scientific director (1990–2001), DESY; CEO Diamond Light Source Ltd (2001–2013); professor, University College London (since 2013)

Family: Married, two children.

Interests: Reading, theatre, opera, hiking.

In search of a source

It was aspiration, but also frustration, that drove Europe's first dedicated X-ray facilities, explains Materlik. In the early 1970s, he was one of a growing number of "parasitic" X-ray users who took advantage of particle physics machines, in which bunches of electrons were accelerated and struck a fixed target. The resulting X-ray beams were highly unstable, however, so scientists started to think about an X-ray source for themselves.

Having completed his PhD using the DESY synchrotron, Materlik went to the US where he visited SLAC at Stanford and saw the advantages of storage rings, in which electrons circulate continuously at fixed energy. "There was a clear need for a high-energy storage ring dedicated to synchrotron X-rays because we were fed up with having to argue for things from particle physicists," says Materlik, who is quick to point out that storage rings are not synchrotrons.

Returning to Hamburg in 1978, he got involved with the European Synchrotron Radiation Project with the aim of building a high-energy storage ring for Europe. Permission to build the ESRF was granted, and in 1985 the French government decided to site it in Grenoble. "Some people said we should pick out one experiment and base the whole science case on that, like the particle physicists did, but I and others argued heavily against that," he says.

Afterwards, Materlik was invited to take over the HASYLAB facility at DESY. He was 40 years old and had "just got his science working", but when he realised that there were ideas to build a UV source at HASYLAB, rather than a X-ray storage ring, he decided that it was time he entered science management. Having spent the next few years working on the DORIS III upgrade, Germany's first national synchrotron-radiation source, Materlik was lured towards X-ray free electron lasers (XFELs). In tandem with his research into atomic-resolution X-ray holography and white-beam imaging, he co-ordinated plans for Europe's first XFEL facility now under construction in Hamburg.

> **"You've got to be in it to win it."**

Shortly after the technology had been demonstrated at DESY, however, he received a phone call offering him the "opportunity of a lifetime": the chance to build a new third-generation storage ring, Diamond in the UK, on an empty site. It was 2001 and, reasoning that "it would be at least a further seven years of politics" until construction of the European XFEL could begin, Materlik jumped at the opportunity. "It was a chance to tie all my experience together and we established a great team who all wanted to make it a success," he says. "If you build something like this and you're waking up at three in the morning worrying about it, then you really should find another job."

Change of scene

Today's light-source scene is unrecognisable compared to the situation in the 1970s, mainly thanks to the increased stability of the machines, explains Materlik. The other major shift, he thinks, is the increased focus on users. "At Diamond we coined the term 'next-generation user facility' because the user doesn't need to know how the beamline is working, which is different to the situation in the 1990s."

XFELs are the next big thing , he says, because they are completely new facilities, complementary to storage rings, that will also serve the laser community. "These machines will eventually reach the attosecond regime where we can see electrons doing their stuff," he says, lamenting the UK's decision not to participate in the European XFEL, which is due to switch on in 2016. "You've got to be in it to win it, and politically there is this big embarrassment that the UK is not really a member of something that by now is a 12-country enterprise."

Getting more involved with the European XFEL is one of Materlik's post-Diamond plans, in addition to taking more time for his private life. He is also establishing the Centre for Facility Sciences at University College London. "I find it very interesting to see things from the point of view of the users and the universities," he says. "Definitely being CEO of Diamond from start to full operation will remain a once in a lifetime job!"

Matthew Chalmers

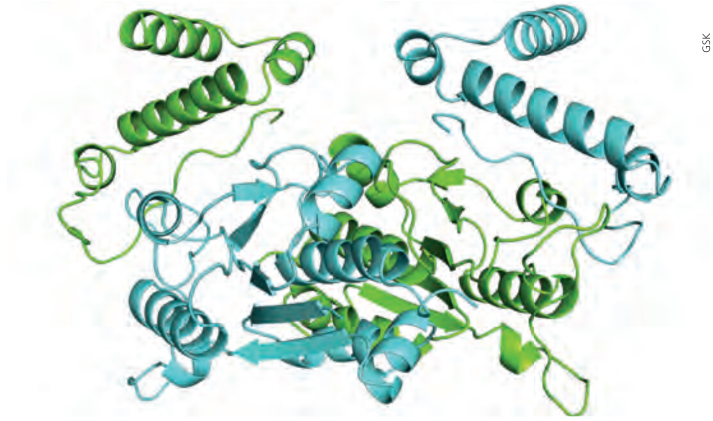
Novel partnership tackles bacterial growth

Researchers at GlaxoSmithKline have taken part in a collaboration with the ESRF and the ILL to train a young scientist and target drug-resistant bacteria.

Antibiotics save countless lives by stemming infections and making invasive surgery possible. But the ability of bacteria to evolve and mutate is making them increasingly resistant against antibacterial agents, despite antibiotics being in widespread use for just 70 years. Recent studies estimate that about 25,000 people in the EU, and a similar number in the US, die each year as a result of infections caused by multi-drug-resistant bacteria.

Last year GlaxoSmithKline (GSK) scientists, who have been regular ESRF users for 10 years, saw a new opportunity to address this considerable scientific challenge by teaming up with the ESRF and ILL to fund a placement student for a year. The ESRF and ILL can provide key information about the structural processes through which bacteria multiply, helping researchers to find compounds that inhibit this growth.

Describing the experience as the best year of her life, undergraduate student Jennifer Channell of the University of Bath in the UK has recently returned home from the EPN science campus in Grenoble where she carried out research to understand the interaction between the bacterial protein TrmD and tRNA – a highly efficient but poorly understood process that is essential for bacterial growth.



GSK

Two molecules of the bacterial protein TrmD associate intimately to form a dimer, which is critical for association with tRNA.

"The project with Jennifer came about through us wanting to explore and immerse ourselves in BioSAXS and its associated technologies and techniques over a longer period than a simple experimental beam-time allocation would have allowed," says Paul Rowland of GSK Stevenage in the UK.

Detailed focus

Despite previous successes in finding the 3D structure of the TrmD protein in complex with a small molecule, there are no structures that catch TrmD in the act of transferring a methyl group to tRNA, explains Trissa Elkins of GSK Pennsylvania in the US, who co-supervised the project with Onkar Singh of GSK Stevenage and ILL and ESRF staff, including Elspeth Gordon and Michael Haertlein. "We have carried out simple experiments but didn't have time to do the detailed work needed, so I was thrilled to get Jennifer's help," says Elkins.

Channell carried out several SAXS experiments on the ESRF's BM29 beamline to characterise the size and shape of the complex in solution, and similar experiments using neutrons at the ILL. The high point, she says, was solving the crystal structure of a TrmD crystal bound to the natural inhibitor sinefungin – a small molecule that also inhibits human protein synthesis – at very high resolution. "It's the first sinefungin structure to be obtained and we hope to publish the work," says Channell. On the other hand, the team was unable to obtain the structure of the crucial TrmD-tRNA complex.

"The interaction between the protein and tRNA is weak and the tRNA wasn't pure enough, but it took some time to realise this," explains Channell, who is now in her final year of a degree in biochemistry. "It was inspiring and motivating to be doing real science rather than pottering around in a teaching lab, and has confirmed my decision to pursue a PhD in molecular biology."

"The collaboration with ESRF and ILL provided a great opportunity to be involved with world class experts in SAXS and SANS," concludes Singh. "It paves the way for future shared collaborations between ESRF, ILL and GSK to help train future scientists and conduct some basic but fundamental research."

Ed Mitchell of the ESRF's Business Development Office welcomes such collaborations. "The project was an excellent opportunity to train an exceptional young researcher in a joint industry–academic environment," he says.

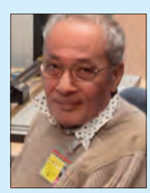
Matthew Chalmers

Movers and shakers

John Helliwell of the University of Manchester in the UK, has been awarded the 2014 American Crystallographic Association's Patterson Award for his pioneering contributions to synchrotron instrumentation, methods and applications for macromolecular crystallography (MX). Helliwell, who is currently chair of the Science Advisory Committee of ALBA, helped define the ESRF's MX activities in the early 1980s and has been a regular user ever since.

Giacomo Ghiringhelli of the Politecnico di Milano in Italy has won the Vacuum Ultraviolet and X-ray Physics (VUVX) Conference award in condensed matter physics for his contributions to the development of new soft X-ray techniques and their application to correlated electron materials. Together with the ESRF's Nick Brookes, Ghiringhelli is pioneering the European resonant inelastic X-ray scattering (RIXS) facility at the ESRF.

Catherine McCammon of the University of Bayreuth in Germany has won the European Geosciences Union's 2013 Robert Wilhelm Bunsen Medal "for her outstanding contributions to understanding the redox and spin state of iron in the Earth's interior", based on her innovative advances in Mössbauer and X-ray spectroscopies. McCammon is a longstanding user of the ESRF's ID18 beamline.



Ahmed El Goresy of the University of Bayreuth in Germany has been awarded the Meteoritical Society's 2013 Leonard Medal "for his many contributions to improving our understanding of the mineralogy and petrology of chondrites and shocked meteorites." Goresy, who turned 80 this year, is a user at the ID21 beamline where he has carried out micro X-ray fluorescence studies into the origin of EL-3 chondrites in the ureilite Almahata Sitta.

Multi-element X-ray detectors for beam-line applications

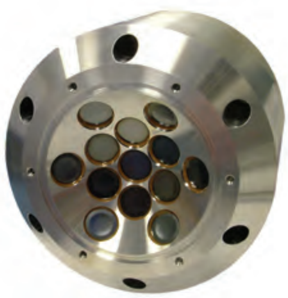
SGX Sensortech have a distinguished heritage in the manufacture of Silicon Drift (SDD) and Si(Li) detectors. Previously known as e2v scientific and Gresham Scientific, SGX specialises in producing detectors from standard designs through customised assemblies to complex multi-element detectors. All detectors are designed to meet the highest specifications, including the largest single active area (100mm^2) available of its kind.

Applications include:

- Extended X-ray Absorption Fine Structure (EXAFS)
- X-ray Absorption Near Edge Structure (XANES)
- Total Reflection X-Ray Fluorescence (TXRF)
- Particle Induced X-Ray Emission (PIXE)
- Micro X-ray fluorescence (µXRF)
- X-Ray Fluorescence (XRF)

Features

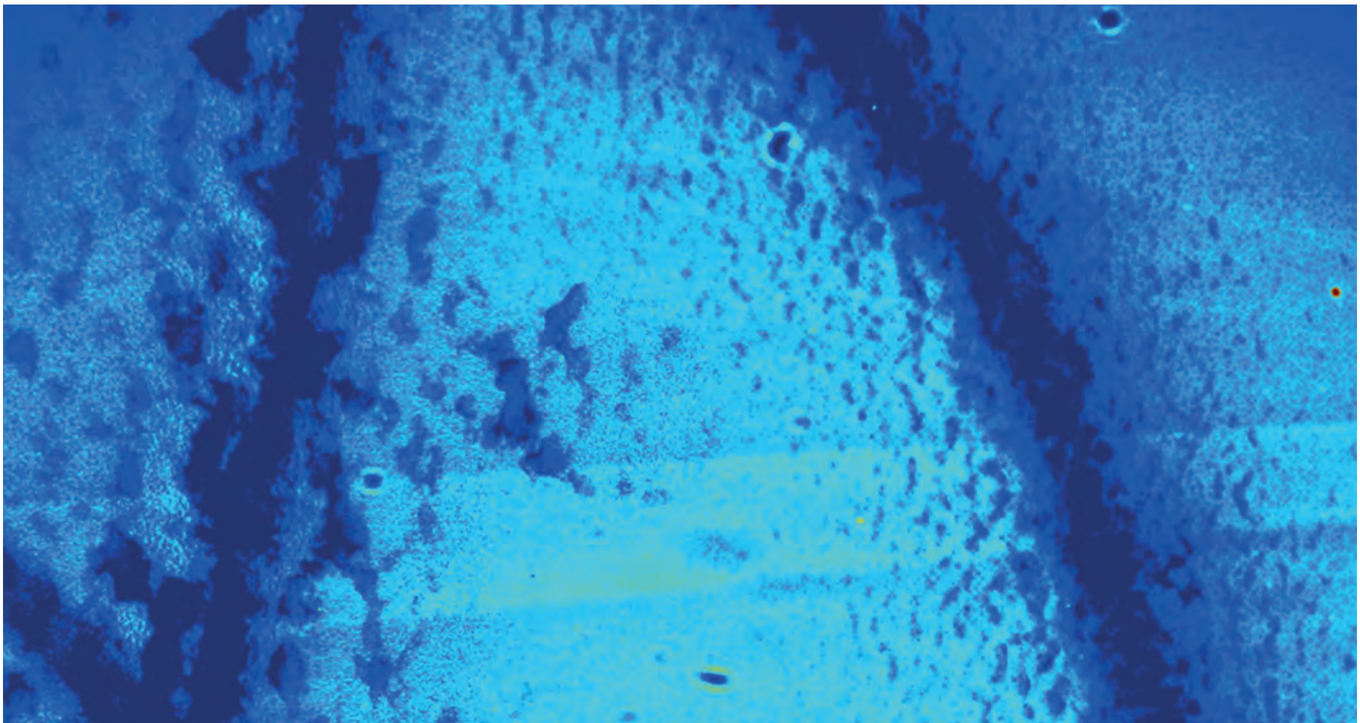
- Available in 1 to 13+ channels
- Individual sensor areas 10mm^2 - 100mm^2
- Resolutions from 127 eV
- Count rate > 1 Mcps
- Windows: Thin Polymer / Beryllium
- Focused or planar sensor arrangements
- Custom collimation
- Application specific designs
- High solid angle
- Vacuum compatible (including UHV)
- Slide options: manual / adapted for translation tables



Also available is a choice of close packed 3 x 100mm^2 or 6 x 100mm^2 sensor arrays

SGX Sensortech (MA) Ltd, Watery Lane, High Wycombe, Bucks, HP10 0AP, UK
www.sgxsensortech.com

Tel: +44 (0) 1628 533 060
Email: sales.ma@sgxsensortech.com

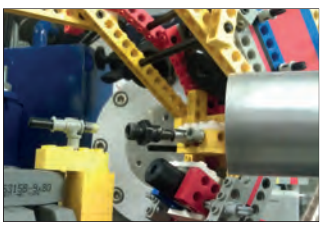


T THI *ET AL.*/ESRF; S DUBOIS AND N ENJALBERT, CEA-INES

Solar cell distortion: Researchers at the ESRF's BM05 beamline have used X-ray diffraction to characterise the structure of a solar cell comprising a layer of silicon in contact with an aluminium backplane. The complete structure is present on the left, while on the right the aluminium backplane has been removed by etching. The integrated diffracted intensity (image represents a region 1.3 mm × 0.6 mm) shows two completely different lines where the silicon diffracts. Where the aluminium backplane is present (left) there is a lot of distortion induced in the silicon, whereas there is much less distortion where the aluminium back-plane has been removed. The results reveal a correlation between the photovoltaic efficiency and the lattice distortion of the silicon in contact with the eutectic and aluminium layers. (*Progress in Photovoltaics: Research and applications* submitted for publication).

In the corridors

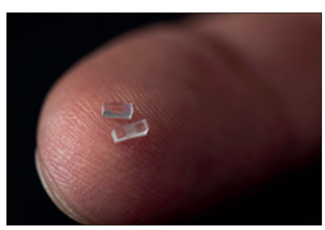
LEGO to the rescue



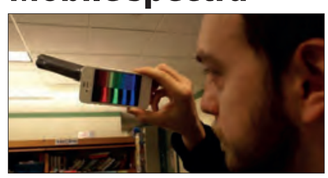
A rather unusual sample holder has made its way to the ESRF's DUBBLE beamline thanks to users from the LumiLab group at Ghent University in Belgium. The team was investigating the energy-storage mechanism in persistent luminescent materials, which are widely used in emergency exit signs and more playful applications, using a combination of XANES and optical pumping. The experiment demanded a highly accurate and stable sample environment in order to measure the emitted light – a task made more challenging due to the small space between the X-ray beam and the fluorescence detector. So instead of turning to a university workshop, LumiLab's Philippe Smet decided to raid his children's toy box. The result was an optically pre-aligned sample stage made from LEGO, which contained lasers, filters, lenses and an optical fibre. "Both the optics and X-rays are then perfectly aligned on the sample, leaving more time for the actual measurements, and if we need to change something it is much easier, cheaper, faster and more environmentally friendly than with a single-use, CAD-made, custom-milled metal sample stage," says Smet. "Back at home, we take it apart and build something else."

Chip-scale accelerators

Electrons can be accelerated through enormous field gradients in a nanostructured glass chip smaller than a grain of rice, according to a recent experiment by physicists at SLAC in California. Joel England and co-workers first accelerated electrons using conventional technology and then focused them into a half-micron-deep channel patterned with regular nanoscale ridges. Electric fields generated by an infrared laser produced an acceleration gradient of 300 MeV per metre – roughly 10 times that of SLAC's 3 km-long linac. At its full potential, says the team, the technology could match the accelerating power of the linac in a distance of 30 m while delivering a million more electron pulses per second, though the team admits that a number of challenges must be overcome before the technology could enter real-world use (doi:10.1038/nature12664).



Mobile spectra



A new app developed by the American Physical Society's (APS) public outreach department turns your iPhone or iPad into a personal spectroscope. SpectraSnapp requires users to attach a paper tube with a thin diffraction grating at one end (which is available for "less than a dollar", according to an APS press release) to their phone's camera, which breaks the light into its component wavelengths. The app comes with a spectral library of 20 of the most common light sources, allowing users to determine what the light source is made from. Pictured above, co-developer James Roche of the APS compares the spectrum from a fluorescent light (bottom image) to the spectrum of neon, showing conclusively that neon was not the source.