CWI

Centrum Wiskunde & Informatica

Annual Report **2013**

The world around us shows that the importance of mathematics and computer science is growing. We are surrounded by interactive objects and systems with ever-expanding functionalities. Everything becomes smart, contains digital components or is completely digitized. Examples are large software systems, logistic planning systems, risk analysis and the World Wide Web. Mathematics and computer science are the essential tools to understand, control and improve this complexity. By persisting in long-term research projects, we have been able to realize many ground-breaking new ideas and technologies over the years. Many of these ideas are now incorporated in the products and services that are used by millions in their everyday life. In the current digital age, this continuous stream of fundamental new concepts and ideas towards progress in important societal and technical challenges is of the utmost importance. We are devoting all of our resources and efforts to keep this stream flowing. We educate new talented researchers to become tomorrow's professors, corporate researchers, CTOs and entrepreneurs, we mine promising, unexplored research fields and we cooperate with partners from society and industry to realize new technologies. In 2013, we have been able to add new achievements to our growing list of contributions to fundamental research and society. This Annual Report gives a summary of our efforts, and accounts for what we have done with the resources entrusted to us by society and our public and private partners. This document is intended to be browsed in no particular order. In it, you will find four articles where we highlight our research on algorithms for quantum computing, simulations for energy transport, online disclosure of cultural heritage and digital forensics. In five overviews, we further present a bird's eye view of our work on the research themes Software, Information, Life sciences, Logistics and Energy. These texts are interwoven with snippets of information concerning other events and activities at CWI in 2013 and facts and figures about our institute. I hope you will enjoy reading and browsing through our Annual Report 2013.

Jos Baeten
General director

About Centrum Wiskunde & Informatica

Mission

Centrum Wiskunde & Informatica (CWI) is the national research institute in the Netherlands for mathematics and computer science and is part of NWO, the Dutch Science Council. The mission of CWI is to conduct pioneering research in mathematics and computer science, generating new knowledge in these fields and conveying it to trade and industry and society at large.

Vision

Results of mathematics and computer science are the invisible driving forces behind our economic growth and welfare, and are instrumental to developments in other scientific disciplines. They provide new insights and powerful tools for societal problems in energy, health care, climate, communication, mobility, security and many other domains. As national

Research theme

Software

Developing RDFa

It is increasingly useful to formulate web content in machine-readable form so it can be automatically processed by applications. W3C web standard RDFa adds machine-readable hints to visual web information to identify information like people, places and events. Search engines and Web services can use RDFa to improve their search results. CWI researchers contribute to the development of RDFa and the related standards HTML+RDFa and XForms. In cooperation with ISOC.nl, Waag Society and CMPro, the W3C Benelux Office also organized events on XForms and CSS at CWI.

Detecting forged signatures

CWI released open source software that can detect and block forged digital signatures: CWI's hash collision detection library. The use of digital signatures is very common: secure websites, documents, e-mail and other software can contain them. It is therefore very important to identify forged signatures and to ward off cyber-attacks that use them – like the recent super malware Flame. To enable this, CWI developed platform independent software that can easily be incorporated in existing software for processing digital signatures.

continued on p. 3

platform for mathematics and computer science CWI wants to expand its position in safeguarding the interests of these research fields and play a leadership role in science policy. To achieve this, CWI is in the forefront of developing new lines of long-term research in high risk areas, inspired by problems in society and industry. We also serve as a breeding ground for academic staff and young talented researchers, and give high priority to knowledge transfer. This is not only achieved by scientific publications and public lectures, but also through training PhD students to become high-potential researchers in science and industry, founding spin-off companies, collaboration with private and public partners and making innovative software tools available for researchers, companies and the general public.

Milestones

CWI has a unique talent pool of researchers. Since its foundation in 1946, more than 190 of its researchers have become **full professor**. In 2013, our researchers included a **Spinoza Prize** winner, 17 researchers with one or more **NWO Innovational Research Grants**, two **KNAW** members, three members of the **Academia Europaea**, two members of the **Koninklijke Hollandsche Maatschappij der Wetenschappen**, a **SIAM Fellow**, an **AMS Fellow** and two **honorary doctorates**.

CWI has a long-standing tradition of excellence in research that is both fundamental and societally relevant. CWI's track record includes building the **first computer in the Netherlands**, computing the **dike heights for the Dutch Delta Works**, connecting Europe to the **internet**, developing the **Python** programming language, computing the **train timetables for the Dutch Railways**, breaking factorization records of **RSA encryption** for internet security and developing the open source database system **MonetDB**. Recent highlights include analyzing the **Flame** malware, using domain specific programming languages in **forensic research**, investigating **smart energy networks** and modelling and simulating phenomena such as **lightning, ocean currents, financial products, wind parks, proteins** and **quantum computing**.

CWI plays a central role in various programs and organizations, including the **W3C Benelux Office, Platform Wiskunde Nederland**, research programme **COMMIT/, EIT ICT Labs, ERCIM** and **Informatics Europe**. Since its foundation CWI has commercialized its research in the foundation of **22 spin-off companies** that have generated millions of turnover to date.

Research theme Software

Audio augmented reality

The Talescape app, developed at CWI, is able to bind virtual sounds to geographical locations and play them back to users who approach them. An open application programming interface (API) allows users to create their own scenarios. In 2013, Talescape won first place at Startup Weekend Enschede and was demonstrated during the CWI in Bedrijf event. A scenario has since been created for the Dam square in Amsterdam, where users can experience sounds (recreated by Maastricht University) as they could have been heard around 1890.

Standards for open data

Organizations like the government are increasingly building upon an open data environment in which data is shared, connected and re-used. Semantic data standards, such as RDF and RDFa, help to share and re-use data between different organizations. These standards have recently been adopted by leading search engines such as Google, Bing and Yahoo!. CWI cooperated with OpenLink Software to apply advanced database methods developed at CWI in the Virtuoso database product OpenLink, strengthening the tools and technology for semantic data management.

Open source quality

Assessing the quality of open source software projects is hard since many factors are important: internal quality, the community's behaviour and its impact. The OSSMETER project combines fundamental questions in software engineering with the pragmatic goal of developing a meta-platform to assess the quality of open source projects. It integrates both formal sources of information (code, metadata) as well as informal sources (e-mails, bug reports). CWI contributes source code and programmer activity measurements.

continued on p.2

Highlight

Solving crimes with domain-specific languages

Digital evidence plays an increasing role in forensic investigations. Data files found on confiscated devices can contain valuable evidence of crimes, from fraud and theft to child abuse, terrorism and murder. These data files need to be recovered, even when the device is damaged, or the files have been deleted by owners fearing prosecution. The Netherlands Forensic Institute (NFI) has to deal with terabytes of such digital files in a wide range of file formats, and called in CWI's help to automate the process.

NFI develops forensic software to aid police and the judiciary in acquiring, recovering and analysing data relevant for solving crimes. "The main problem

Research theme

Information

3D video communication

Live communication rarely makes use of 3D representations. Although it is technically feasible to produce full 3D geometries using multiple Microsoft Kinects, it requires too much bandwidth. Researchers from CWI and QML have developed a prototype that can efficiently stream live-reconstructed 3D geometry over the internet in real-time. Using a novel source coding method and a channel coder based on rateless codes, real-time transmission was achieved. This advance opens up the possibility of 3D interaction and real-time activities over the internet for remote users.

Big Data in a ring

Researchers from CWI challenged the common belief that network is your bottleneck in Big Data. In modern hardware, sending data through a network is much quicker than reading it from a disk. The researchers exploit this to build a fast and efficient database architecture consisting of computers connected in a ring. The most relevant data is not stored anymore in one central location, but continuously flows through the network ring. This experimental Data Cyclotron architecture is faster and gives better throughput than systems that read and write data from a traditional disk subsystem.

continued on p.5

Meta-programming with Rascal

The domain-specific programming language Metal is written in Rascal. Rascal is a meta-programming language that is developed by the Software Analysis and Transformation group at CWI since 2009. It is the result of a long-running research line on meta-programming. Meta-programming is essentially writing software for writing software. Applications include analysing programs to better understand them, or generating programs from high-level descriptions. Meta-programs can detect where and how the essential knowledge of software systems is implemented. This is very useful when improving and restructuring large software systems. These systems, which can for instance be found at banks, insurance companies or the government, are usually written and added on to over a long period of time to a point where nobody can fully understand the system. Rascal uses concepts like pattern matching and syntax analysis to determine the relations between the parts of a software system, and rewrites it using an easy to use Java-like language. This can result in a huge improvement in software reliability and quality, and a reduction in size and maintenance.

The successful company Software Improvement Group (now over a hundred employees) was founded by CWI in 2000 based on software engineering research in the SWAT group. SIG uses the principles of meta-programming to improve the software of large companies. Customers include ING, ABN Amro, Nationale Nederlanden, CZ, Eneco, KLM, PostNL, IKEA, KPMG, Capgemini and the Dutch government. There is still close cooperation between SIG and CWI.

at NFI is that we have to deal with an enormous number of file formats from all kinds of devices," says Jeroen van den Bos, researcher at both NFI and CWI. "For images alone there are hundreds of different file types. And to make matters even more complicated, each individual file type exists in different variations and can be damaged. This makes it very hard to automate data recovery or perform an automated search for a specific file." In 2009, Jeroen van den Bos joined CWI to work on a solution for this problem. The Software Analysis and Transformation (SWAT) group at CWI has great expertise in domain-specific languages (DSLs). Together with CWI

continued on p. 6

Research theme Information

Firefighting with Big Data

By combining satellite images with geographical data such as vegetation, land-coverage, cadastral information and environmental data, CWI and its European partners were able to develop an online forest fire monitoring tool. The service calculates where fires can occur and how they can develop. CWI developed the technical infrastructure, including advanced database and semantic web technologies for storage and quick access of the data. The application is now in use by the Natural Observatory in Athens and assists the Greek army, fire brigade and Civil Protection Agency.

Recommending news

Most news websites or web shops feature recommendations: if users like a certain item or article, a list of other articles or items is generated that might also interest them. This year, a team of CWI researchers won the News Recommendation Challenge, an international contest for developing the best recommendation algorithms, a contest for developing the best recommendation algorithms. When featured on several German news websites, the algorithms submitted by CWI achieved the most clicks, showing that they did the best job in reflecting the actual interests of the users.

Monitoring emotions

Computers are getting better in recognizing human emotions. By collecting the emotions of a large group, for instance by analysing photo's published by the visitors of a festival, a crowd emotion monitor can be developed. This data can be used in applications like mood maps, allowing festival visitors to find a place to go based on their current mood. In the COMMIT/ project SWEET on this topic, CWI provided the database technology necessary to collect, store and exchange the sensor data. For this, the MonetDB system was expanded to support the JSON format for data exchange.

continued on p.4

researchers Paul Klint and Tijs van der Storm, Van den Bos wanted to design a DSL for digital forensic investigations that could automatically add new file formats to NFI's forensic software. At the end of 2013, this work was completed.

Metal code

In Van den Bos's approach, the structure of file formats are modelled in a DSL called Metal. These descriptions are designed in such a way that they are easy to understand by non-programmers. When necessary, it is also easy to modify the descriptions with only a few lines of code when new versions or variants of file formats are encountered. The code generator then automatically transforms the Metal code into full-fledged data recovery software. In comparison: using traditional techniques this would involve the modification of thousands of lines of low-level program code. Nevertheless, the generated tools can compete with industrial forensics tools in terms of accuracy and run-time performance. Metal was designed and implemented using the meta programming language Rascal (see text box), which is specifically designed for building DSLs. Metal is a declarative language that defines all relevant information in a single location. Van den Bos: "Compare it for instance to a program for recognizing images of cars. There are many different cars in all kinds of sizes, shapes and colours. With traditional programming techniques, you have to describe every new type of car completely in order for it to be recognized. But with meta-programming, you only change the general description of a car in order to include the new type you encountered. The rest of the code is adapted automatically everywhere the description is used. This is the reason that we can suffice with a few lines of code, removing everything from the description that is not relevant to the forensic investigation." Starting in 2014, Van den Bos will implement the results of this research at NFI, after defending his PhD thesis on this topic on 9 January 2014 at the University of Amsterdam. He will remain affiliated to CWI to continue improving forensic software. The ultimate goal is a software tool that can readily be used by the police without help from experts at NFI. This could significantly improve the speed with which cases involving digital evidence can be solved.

cwi in 2013

MonetDB Solutions

CWI's 22nd spin-off company was launched this year: MonetDB Solutions. This consulting company is specialized in the development of database technologies for business analytics. It is is established by the founders of MonetDB, the world's first and most widespread open-source column-store database management system, and led by Ying Zhang. Bundling the expertises of world-leading researchers and engineers, its core services involve technical consultancy, architectural review, performance assessment, migration & upgrade services and training.

cwi in 2013

Rota's Conjecture

CWI researcher Bert Gerards and colleagues Jim Geelen (University of Waterloo) and Geoff Whittle (Victoria University of Wellington) proved the famous Rota's Conjecture. They have been working on solving this mathematical problem posed by Gian-Carlo Rota in 1970 for more than 15 years. This year the trio realised that they had collected all the essential ingredients for the proof. Rota's Conjecture relates to a specialised area of mathematics known as matroid theory, a modern form of geometry. The team expects that writing down the results will take several more years.

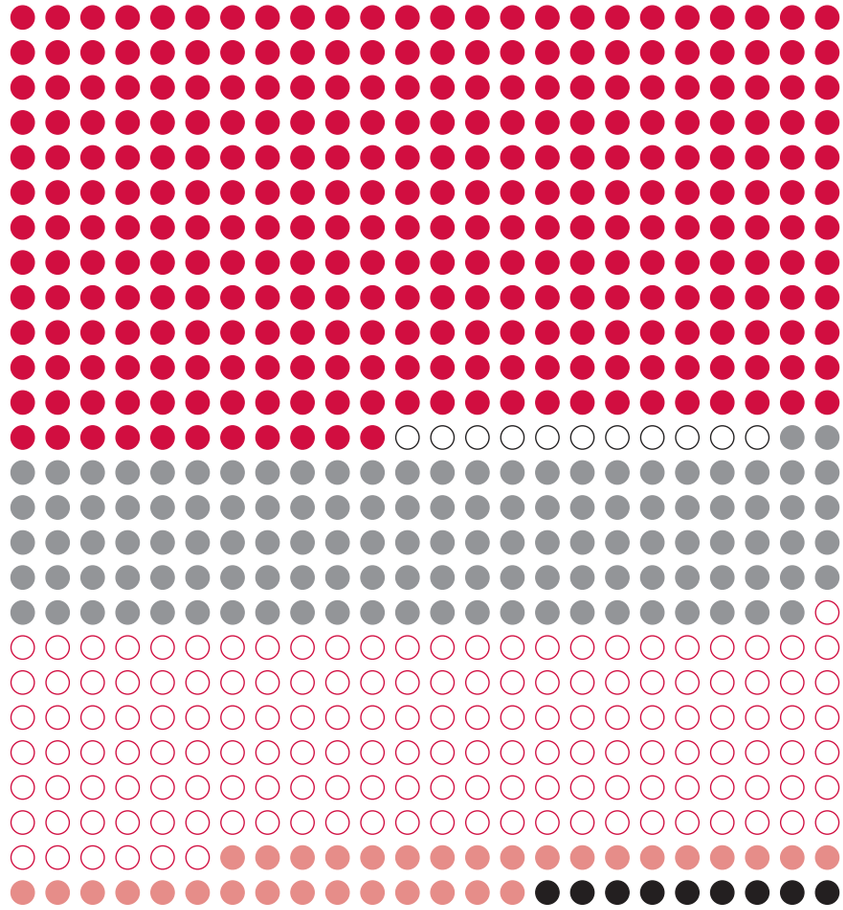
cwi in 2013

Media coverage

Several research projects were covered by national media this year. In January, PhD student Willem Haverkort presented his research on national television in *De Wereld Leert Door*, a spin-off of the popular show *De Wereld Draait Door*. National newspapers published about our research on cloud simulation (*NRC*), windpark simulation (*NRC*), police service planning (*Spits*) and lightning (*Volkskrant*), and numerous more topics were published in professional magazines and websites. There was also a wide coverage of 25-year anniversary of the open European Internet on 17 November 2013.

Facts & figures
Output 2013

Papers	299
Book contributions	11
Media appearances	121
Lectures	151
Software products	33
PhD theses	9



cwi in 2013

Royal decoration

Computer scientist and CWI Fellow Paul Klint was appointed Officer of the Order of Oranje-Nassau on 9 September. Gerard Boekhoff, deputy Mayor of Bussum, presented the Royal Decoration to Paul Klint for his outstanding contribution to science and society at the reception that was held in honour of his 65th birthday. Boekhoff praised in his speech Klint's role as top researcher in software engineering, his educational qualities and influence on the profilation of the field of computer science in the Netherlands and beyond.

cwi in 2013

Prizes and awards

A selection of prizes and awards won by CWI researchers in 2013:

- **Andreas Bonn Medal**
  Jop Briët
- **Nomination Huijbregtsen Prize**
  Scientific Computing group
- **Humboldt Research Award**
  Peter Boncz
- **SIGMM Award**
  Dick Bulterman
- **Van Marum Prize**
  Marc Stevens

cwi in 2013

Sino-European cooperation

CWI officially joined research network LIAMA. The Sino-European Laboratory in Computer Science, Automation and Applied Mathematics (LIAMA) is a research lab consisting of European and Chinese research institutes in the field of mathematics and computer science. LIAMA conducts research, training and transfer projects in these fields. CWI joins two projects, on cryptanalysis and embedded systems. CWI was introduced to LIAMA by founding member Inria, the French national research institute for mathematics and computer science, and strategic partner of CWI.

Highlight

The possiblities of a quantum computer

There seems to be no end to the growth in computing power. Your smartphone is more powerful than a supercomputer was 25 years ago, and your next laptop could be faster than all your previous machines combined. According to Moore's Law, which predicts that the number of transistors on a microchip doubles every 2 years, we will reach atomic-size transistors around the year 2030. And yet even this might not be the limit. In the past decades, researchers have started investigating the quantum computer, a machine that uses sub-atomic quantum mechanical effects for computation.

Research theme

Life sciences

Crystalline tissues

How biological cells cooperate to build biological tissues and organs is a fundamental problem in developmental biology. Recent studies suggest that liquid crystals are good analogies for understanding cell ordering in biological tissues. Computational models developed at CWI show how crystalline ordering of cells contribute to the formation of blood vessel networks, causing a system of elongated, cohering biological cells to get 'stuck' into a network pattern. The model provides a new explanation for the formation of cellular networks in systems without exchanging chemical signals.

Tropical outbreak

Over ten years ago, the tropical pathogen Cryptococcus gatti appeared out of nowhere in the temperate climate of the Canadian West Coast. Hundreds of otherwise healthy people were infected. CWI participated in a large research team consisting of microbiologists, geneticists and mathematicians that was able to prove that the lineage that caused the outbreak originated from the Amazon rainforest. This lead to the conclusion that not the pathogen itself, but the environment has changed, causing the exotic species to thrive outside tropical South America.

continued on p.9

A quantum computer is fundamentally different from a classical computer, and researchers are only starting to grasp the possibilities of such a machine. CWI researcher Ronald de Wolf investigates possible algorithms for quantum computers. Quantum algorithms use a wholly different, very counterintuitive logic. While classical bits are either 0 or 1, quantum bits (qubits) can be in any superposition in between 0 and 1, which means that they are 0 and 1 'at the same time.' A collection of qubits can therefore be in a superposition of all possible states at the same time. This allows for unrivaled speed-ups for some computations, but it comes at a cost. "Unlike classical computers, quantum computers use massive interference effects that depend on the entire computation", says Ronald de Wolf. "The algorithms therefore cannot be broken into smaller pieces like classical algorithms can, but have to be developed as a whole. This makes them much harder to design." Up till now a only small number of quantum algorithms have been found. De Wolf: "Most famous is Peter Shor's algorithm from 1994 that factorizes large prime numbers. This problem is believed (though not proved!) to be hard for classical computers: the widely used RSA encryption method is based on the fact that it is easy to generate a large number by multiplying

continued on p. 10

Quantum techniques for classical problems

The techniques developed for analyzing quantum computers turn out to be surprisingly helpful in other areas too. One of De Wolf's specialties is to apply such techniques to solve problems in classical computer science and mathematics: reformulating classical problems in quantum terms sometimes gives surprising new insights. In 2012, with colleagues from Brussels and Erlangen, he managed to prove that linear programs are unable to efficiently solve the Travelling Salesman Problem (TSP), a famous optimization problem, refuting a decades-old claim. The proof was based on a surprising connection to quantum communicating protocols. The resulting research paper shared the Best Paper Award at the annual ACM Symposium on Theory of Computing (STOC'12).

Research theme Life sciences

Personalized medicine

DNA differences between individuals influence the effectiveness of drugs. Researchers from CWI are working on an 'encyclopaedia' of Dutch DNA that can serve as a guide for more effective, personalized medicine. Using the genomic data gathered in the Genome of the Netherlands projects, the researchers work on revealing and categorize the DNA-related differences in the Dutch population. This research is funded by the NWO Vidi grant that Alexander Schönhuth received in 2013 for his research project 'A Dutch encyclopaedia of genetic variation'.

Quality of tomography

When performing a CT scan of a patient, the judgment of a doctor relies on the quality of the 3D image mathematically reconstructed from the 2D scans. Any measure of the quality of the reconstructed image is therefore most welcome. CWI developed methods that give a theoretical guarantee that a reconstruction resembles the actual scanned object. The method makes clever use of very basic prior knowledge about the nature of the scanned object, and can be used if the scanned object consists of a single material. It is the first time that such a guarantee can be formulated.

Human Brain Project

The prestigious billion-dollar Human Brain Project aims to connect all European forces to understand one of the greatest challenges in modern science: the human brain. It will bring together all existing knowledge about the human brain and reconstruct it using models and simulations on a supercomputer. This will provide new insights into the human brain and help to develop completely new computer and robotics technologies. CWI will contribute to the project by examining new forms of data storage and processing for extensive data streams released by the project's simulations.

continued on p.8

ERC Consolidator Grant

In 2013, Ronald de Wolf was awarded a prestigious ERC Consolidator Grant for his quantum computing research. This individual grant of approximately 1.5 million euro is awarded by the European Research Council (ERC) to support talented researchers in expanding their independent research program over a five-year period. De Wolf was previously awarded a Veni (2005) and Vidi (2008) grant from the Netherlands Organisation for Scientific Research (NWO), as well as the Cor Baayen Award (2003) for Europe's 'most promising young researcher in computer science and applied mathematics'.

two smaller prime numbers, but probably very hard to do the reverse: to deduce these two smaller prime numbers from just the larger number. Shor's algorithm however can compute this very fast."
A second famous quantum algorithm is Grover's search algorithm from 1996, which can quickly find a specific item in an unsorted list. A few other quantum algorithms have been found since then, and De Wolf is working on a couple more: "I am now working on quantum algorithms to test if a given function depends on only few coordinates. This is interesting for instance if you know that only a few genes in a large DNA-string are active, but you don't know exactly how many or which ones."
For most problems, quantum computers will be no faster than classical computers. However, for many of the problems for which they are suited, including factorization, and simulations for drug discovery and nanotechnology, they will probably be exponentially faster. Provided that large-scale quantum computers are eventually built (the QuTech Institute in Delft aims at doing so within the next ten to fifteen years) they will be vastly more powerful for solving such problems. And computer scientists like Ronald de Wolf might just come up with new quantum algorithms that open up a whole new range of applications.

cwi in 2013

New KNAW member

Cryptographer Ronald Cramer was elected as Member of the Royal Netherlands Academy of Arts and Sciences (KNAW). Cramer is best known for his work with Victor Shoup on the Cramer-Shoup encryption (now an international standard), hash proof systems, cryptographic protocol theory and mathematical aspects of secure computation. In 2013 he was also appointed Fellow of IACR, the International Association for Cryptologic Research, 'for fundamental contributions to cryptography, for sustained educational leadership in cryptography, and for service to the IACR'.

cwi in 2013

CWI Lexures

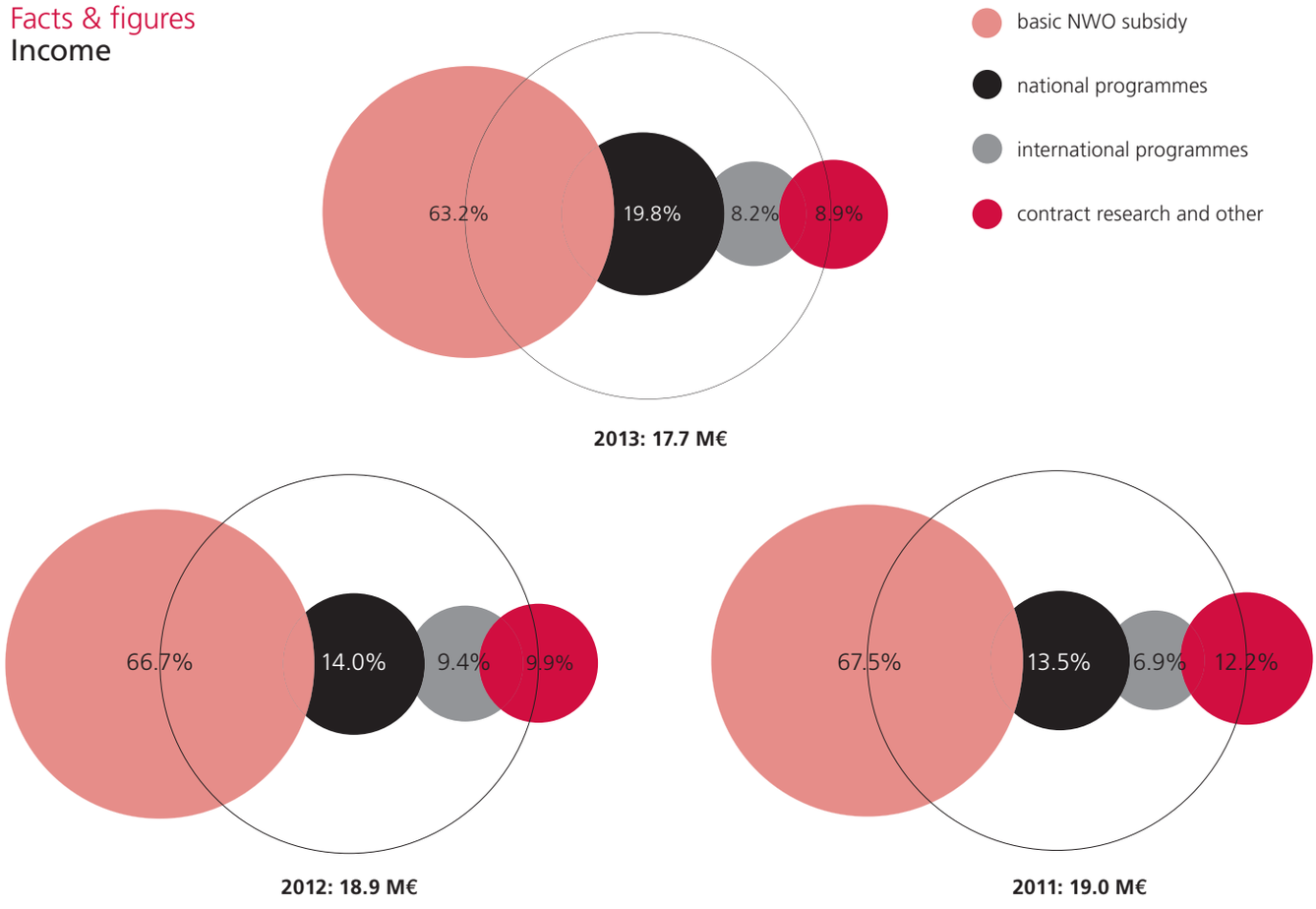
To honour mathematician and Spinoza Award winner Lex Schrijver, CWI organized the CWI Lectures 2013 – the 'CWI Lexures' – on the occasion of his 65th birthday. Schrijver is a world-renowned authority in the field of discrete mathematics and optimization and is best known to the general public for his work on the optimization of the Dutch railway timetable. Theme of the CWI Lexures was Discrete Mathematics and Optimization. Keynote speakers included Martin Grötschel (Zuse Institute Berlin), László Lovász (ELTE) and Leo Kroon (EUR and Dutch Railways).

cwi in 2013

Demonstrating research

Business networking day CWI in Bedrijf on 22 november was themed 'Innovation in ICT' and featured speakers Rik Bleeker (Amsterdam Economic Board), Taco Dibbits (Rijksmuseum), John Koster (ASML) and Michel Cosnard (Inria). The introduced research vouchers resulted in the initiation of several new mini-projects on the match-making market. Researchers of CWI introduced the general public to their research on this year's Open Doors Day on 5 October. In a new set-up, all groups organized a demo, game or other activity based on their current research. Around 1000 visitors came to visit CWI.

Facts & figures
Income



cwi in 2013

25 year open European Internet

On 17 November 2013, the open Internet in Europe celebrated its 25th birthday. On that day in 1988, CWI system administrator Piet Beertema received an e-mail stating that CWI, as first organization outside the USA, had gained access to NSFnet, an academic computer network that later evolved into the Internet. CWI celebrated this anniversary during CWI in Bedrijf, where a plaque was unveiled in the presence of a group of Internet pioneers. The anniversary was covered by most national newspapers and news shows, including *NOS Journaal*, *Jeugdjournaal*, *EditieNL* and *Hart van Nederland*.

cwi in 2013

Networks grant

Six research consortia in which prominent scientists from various Dutch universities work together received a combined sum of 153 million euros for long-term and large-scale research in the Gravitation programme of NWO. CWI is involved in the Networks proposal through researchers Harry Buhrman and Lex Schrijver. Buhrman will focus on the behaviour, performance and security of quantum networks, Schrijver on the dynamics of large networks, including transport and logistic networks. Gravitation provides financial support for a maximum period of ten years.

cwi in 2013

ICT Awards for Python and SIG

Two successful long-term results from CWI research were awarded this year at the Nationale ICT Awards. Programming language Python was awarded the COMMIT/ Award for the most valuable product resulting from ICT research that is over ten years old. Software consultancy company SIG received the ICT Milieu Award for their energy profile register, an application that determines the energy profile of software-hardware combinations.
Python was developed in the early '90s by Guido van Rossum, who was at CWI at that time. SIG was founded in 2000 as a CWI spin-off company.

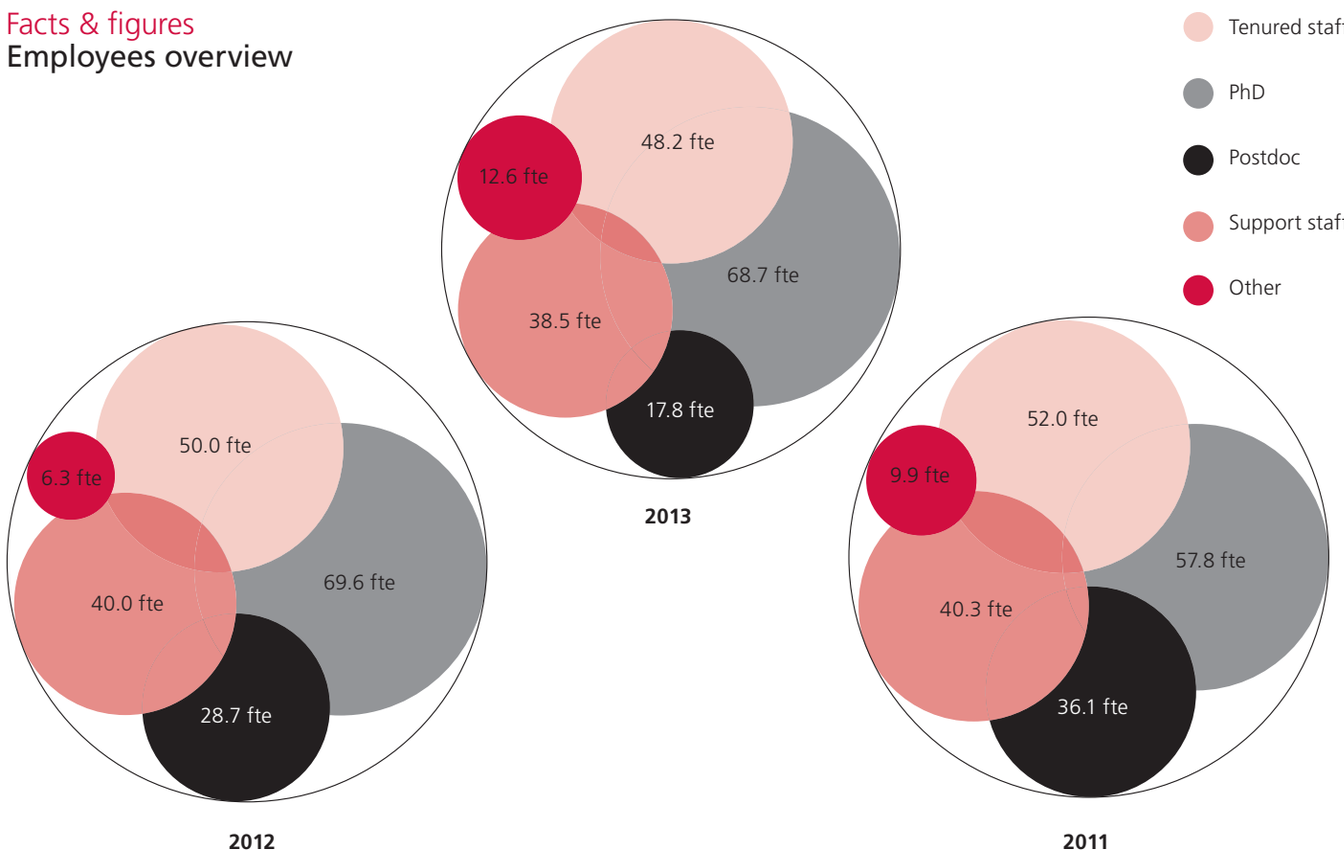
Highlight

The collection is yours

Our information society has a major impact on the way museums record, conserve and manage their collections. In this trend, interaction with the audience is no longer restricted to visits on the physical location. Worldwide, museums and cultural heritage organizations innovate and share collections in multiple ways with their users. To respond to changes in this dynamic environment, the Rijksmuseum in Amsterdam is playing a leading role in developing new technologies. For more than 10 years, the museum has joined forces with CWI researchers from the Information Access group.

Big data
The Information Access group of CWI focuses its research on complex information systems and investigates how access to information can be improved with humans 'in the loop'. "This concerns

Facts & figures
Employees overview



tasks that have too much volume to be considered manually by people," explains researcher Jacco van Ossenbruggen, "but are, on the other hand, too complex to fully automate, and leave to computer processing only. In our research, humans are the vital link in the automation processes." Also the heterogeneity of the data is a challenging factor. Big Data is often defined 'BIG' because of its volume, velocity and variety. "In our research we specifically focus on challenges in the variety in data sources," says Van Ossenbruggen. "The cultural heritage domain is a good example, where data not only varies across different organizations, but even within a single collection the conventions on how to represent information vary enormously."

The Rijksmuseum online

In the Netherlands, the Rijksmuseum is a real pioneer and at the forefront of opening up our cultural heritage. Traditionally, a historical object in a database is linked to a unique number and described in splendid isolation. In the current development, the focus is much more on the networked context of the object, with multiple dimensions. "Think of the story the tour guide tells you during a museum visit," says Van Ossenbruggen. "The guide adds value by explaining the historical role of an artwork, but also how it is related to the present context and to other works of art."

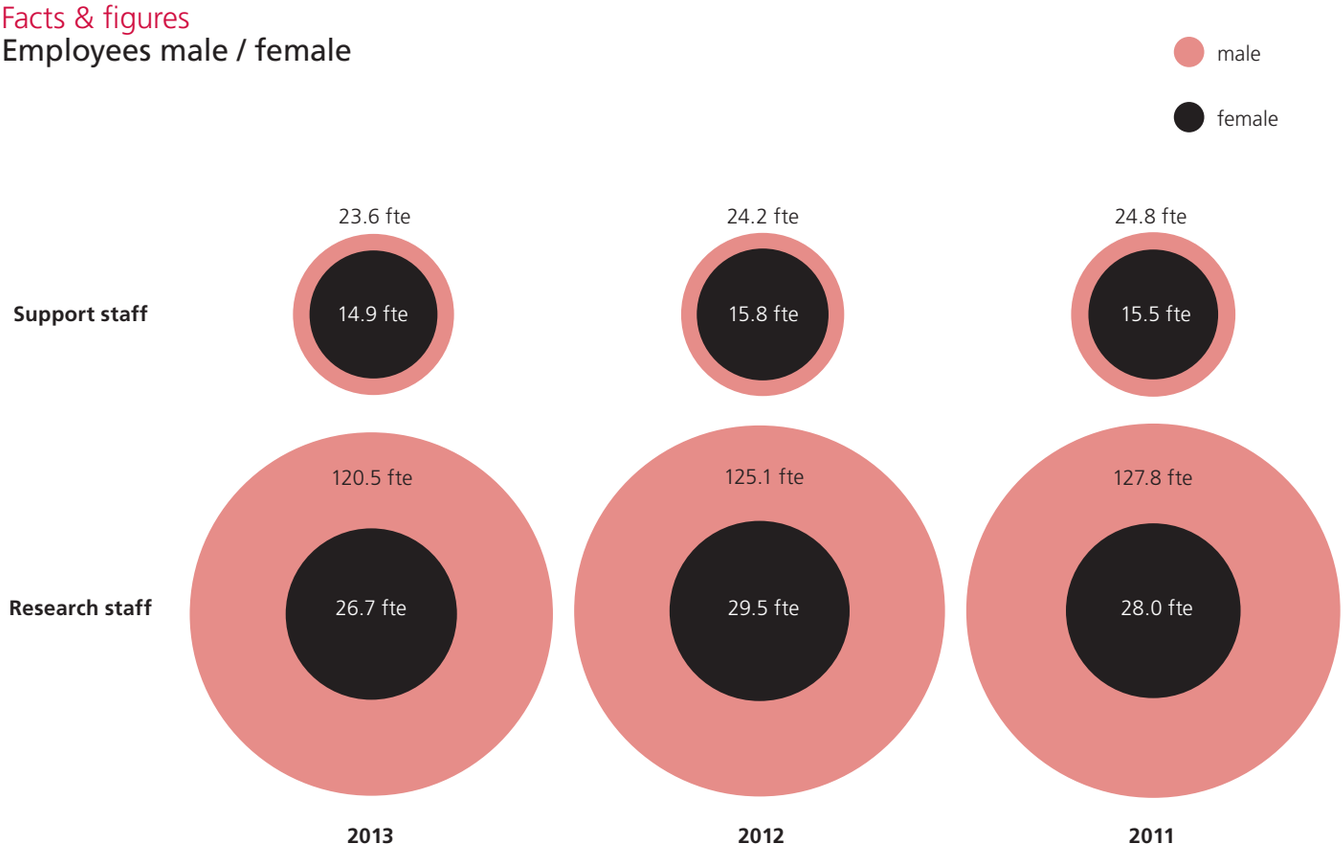
In this concept of networked heritage, the relationship with the audience has become very important. The importance of public participation is also shown by the new Rijksmuseum building. Its open architecture is reflected by the transparency in the online presentation of the collection. Out of the total collection of over 1.000.000 objects that are stored in the depots of the Rijksmuseum, only 8.000 objects are displayed in exhibitions. Since the Rijksmuseum started to digitize the artwork, more than 400.000 artworks are now available online.

"To make this quantity easily accessible online is a huge job," says Van Ossenbruggen. "There is little time per image for data entry, describing and annotating the objects with art-historical information. One problem is that cataloguers are familiar with the jargon of the internal thesaurus. But that thesaurus is often incomplete or uses terms not matching the keywords used by the general public. Additionally, for many prints, specialist knowledge is required to describe accurately what is depicted, for example in

continued on p. 14

Facts & figures

Employees male / female



prints depicting exotic birds or flowers. We investigate to what extent we can achieve this by involving experts in the crowd in the annotation process. This requires new technologies to find the best match between annotation task and expert, to develop user interfaces for these experts and to measure the usefulness of the provided annotations afterwards."

Niche sourcing

In the current research project with the Rijksmuseum, SealincMedia, the relationship between the collection and the public is explored. The focus is on acquiring crowdsourced descriptions. "Traditionally, crowdsourcing projects target tasks that are difficult for computers but easy for humans, requiring no specific expertise from the crowd workers. Our research focus is on tasks that are also difficult for humans because they require specific domain expertise. We often use the term niche sourcing to stress why our approach is different," explains Van Ossenbruggen.

"We studied, for example, the use of machine learning algorithms to do classification of paintings on subject type. These algorithms need to be trained on data provided by (expensive) domain experts. However, in many realistic cases there is too little data for training and results are insufficient to find the single correct answer fully automatically. In response, we developed a game in which we use the same algorithms, but now to preselect a few, likely to be correct, answers from a large set of potential answers, and then ask humans to pick the single correct answer. Interestingly, our research shows that in this setup, non-experts can learn to select the correct answers very quickly just by playing the game. The human player needs the software to help him play the game, but the software also needs the human to achieve the required level of accuracy that was missing before."

www.commit-nl.nl/projects/socially-enriched-access-to-linked-cultural-media

Research theme

Logistics

Dynamic pricing

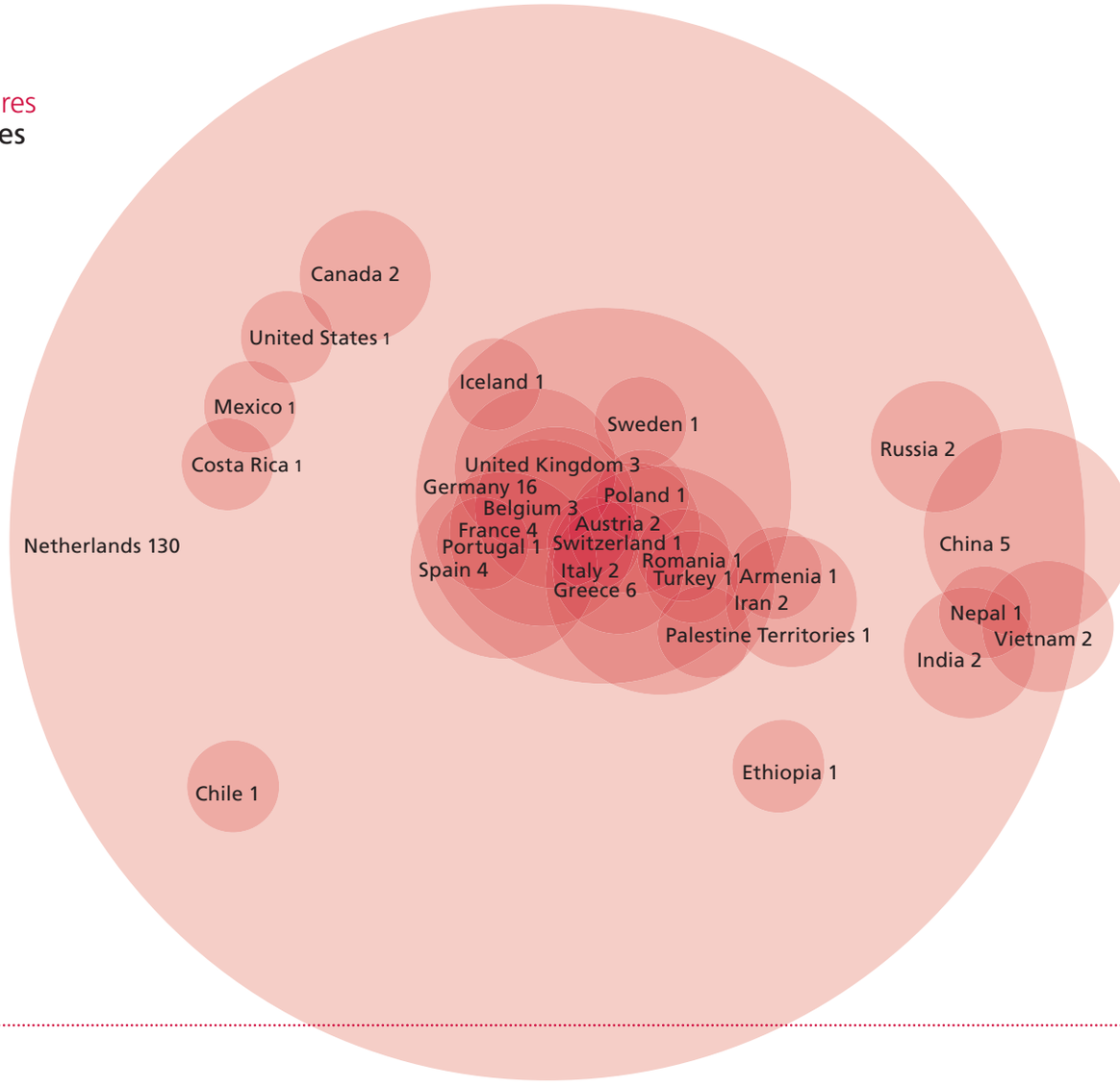
With dynamic pricing, vendors can adapt retail prices on the fly in order to raise their revenue. This strategy is for instance used in tourism, entertainment and the airline industry. Vendors often do not know how to exploit the enormous benefits of dynamic pricing for optimizing revenue. CWI studied optimal pricing strategies to realize maximum revenue. By sacrificing profits on a short term in order to learn more about customer behaviour, better pricing decisions can be made in the long term. The developed algorithms show the optimal balance between learning and profit maximization.

Planning police service

Researchers of CWI and VU University Amsterdam joined forces with the Amsterdam-Amstelland police force to optimize the planning of police services. New techniques allow a more efficient capacity planning and quicker response times by matching available resources with the probable occurrence of incidents. The researchers will develop advanced planning and prediction methods. Recent research has shown that historical data can predict the number and location of accidents, including times of occurrence. This research project is initiated as part of the Police and Science Programme.

continued on p.15

Facts & figures
Nationalities



Research theme Logistics

Minimizing airport noise

With increasing air traffic, more people living around airports experience noise pollution. CWI and NLR investigate ways to compute flight schemes that maximize the number of flights while keeping the noise pollution at a minimum. Simulations on data reveal that the developed algorithms compute optimal solutions for all tested airports. First results reveal that the number of flights could be increased while at the same time fewer households suffer from noise excess, thus increasing the quality of life around the airport.

Traffic management

In the event of traffic accidents or road works, traffic is often directed along alternative routes by digital signs. This can lead to large congestions along these routes. Using real-time coordination of traffic lights and digital road signs, traffic streams can be monitored and directed along different routes. By changing the information on the digital signs and the frequency of the traffic lights, such congestion might be prevented. CWI cooperates with Trinité Automation, a company that develops dynamic traffic management systems, to realize this system in the Amsterdam region.

Efficient networks

Even with the fastest computers, key networking problems remain very challenging. This includes problems such as finding the cheapest way to connect specific terminals in a given network, or designing networks that can handle traffic that varies substantially over time. The latter is often an issue when designing cost-effective communications networks. CWI investigates the design of efficient networks from a computational approach. This research is funded by the NWO Veni grant that Neil Over received in 2013 for his research project 'Designing better, faster networks'.

continued on p.14

Highlight

Transcontinental electricity nets

Our future energy supply is likely to rely heavily on sustainable sources such as wind, sun and water. Efficient transcontinental electricity nets could average out the fluctuations of natural energy sources and transport the energy to where it is needed. This requires a new high-voltage net structure, with an alternation frequency reduced from 50 to 16.7 Hz or even to direct current (DC). The Multiscale Dynamics group at CWI develops mathematical discharge models for developing the components of this new high voltage technology.

Sources like wind and sun are highly variable by nature and are therefore not constantly available. Also, optimal circumstances for producing this power are spread unevenly across the globe, like hydropower in wet and mountainous areas, wind power at sea and solar power in dry regions near the equator. One way

Research theme

Energy

Stabilising nuclear fusion

In order to make nuclear fusion possible, the plasma in the fusion reactor must be heated to over 100 million degrees. No material can withstand those temperatures, so the plasma is confined by strong magnetic fields. This causes the plasma to rotate in the reactor. Researchers from CWI and FOM institute Differ used advanced mathematical analyses and numerical simulations to determine the effect of different forms of rotation on the plasma stability. One of the results is that the so-called Coriolis Effect resulting from a certain type of rotation can stabilise the plasma, preventing energy loss due through instabilities.

Efficient wind farms

Wind turbines in off-shore wind farms affect each other through their wakes, highly variable turbulent wind flows which can be hundreds of meters long. Wakes cause a lower energy production and higher loading of downstream turbines, but on the other hand cause an extra inflow of air, amplifying the amount of wind in the farm. CWI and ECN developed mathematical models capable of predicting the turbulent air flow in the wakes, and simulating different designs and configurations in a 'virtual wind farm'. These simulation tools allow for the efficient design of large-scale wind farms.

continued on p.17

to provide a stable energy system is to set up a large, transcontinental power grid where power can be distributed over large distances. In this way, natural variations in availability can be averaged out, and electric power from distant areas like seas, mountains and deserts can be transported efficiently to more populated and industrialized regions.
The present-day alternating current (AC) nets are however unsuited for this, as they would lose too much power over such long distances. Using direct current (DC) nets is currently in development, though not all required parts of the grid are fully developed and tested. China has recently started construction of an 800 kV DC transportation line, connecting the energy-rich west of the country to production sites in the east. Germany is presently planning a DC electricity line to transport wind power from the north to industrial sites in the south.
One of the unsolved problems in building a DC net is circuit-breakers. These electrical switches turn off the network in cases of overcapacity or short-circuiting to prevent damage. AC nets can be safely switched off when the current passes through zero. In DC nets however, the current flows continuously, making it much more difficult to interrupt.
In a new NWO-Shell project (see text box), the Multiscale Dynamics group at CWI is working on this problem. The project continues work with multinational ABB, a world leading producer of high voltage equipment and extends two ongoing STW-projects on electric flashover, a poorly understood phenomenon that can ruin the insulator layers of electric equipment, and on finding an environmentally-friendly replacement for the extremely strong greenhouse gas sulfur hexafluoride

continued on p. 18

NWO and Shell for energy research

This project for modeling DC circuit-breakers is funded through the NWO-Shell programme 'Computational Sciences for Energy Research'. The program is a joint initiative by NWO and Shell, executed by FOM within the framework of the Dutch top sectors Energy and Chemistry. An important driver for Shell to invest in this initiative is to use the program as a recruitment pool for computational scientists for its technology centre in Bangalore, India. The PhD positions in the program are filled with young scientists hired by Shell Bangalore and interested in building a career there after finishing their PhD in the Netherlands.
A second NWO-Shell project at CWI focuses on rare event simulation for energy grids, and is carried out jointly by the groups Stochastics and Scientific Computing. It studies the occurrence of rare events like black-outs or overloading in energy grids due to a fluctuating supply of power from sustainable sources.

Research theme Energy

Efficient solar cells

Polymer solar cells are composed of a mixture of polymers that generates energy by exchanging electrons. The exact composition of the mixture is essential to the cell's efficiency, but couldn't be imaged accurately because of the destructive nature of the imaging process. A new algorithm developed by CWI is able to produce a detailed 3D image of the cell based on only a small number of electron microscope images, so the cell remains intact. The new technique now allows researchers to investigate and improve the process of mixing with the aim of producing more efficient solar cells.

Uncertainty quantification

Better mathematical methods are needed to reduce the impact of the increasingly uncertain availability of sustainable energy sources. The resulting computer algorithms are necessary to design the robust energy technologies of the future, which prevent or compensate these uncertainties. CWI develops computational models for quantifying and reducing uncertainties in wind power generation, nuclear power stations, and power grid applications. Partners in this research line are ECN, KEMA, NRG and TNO.

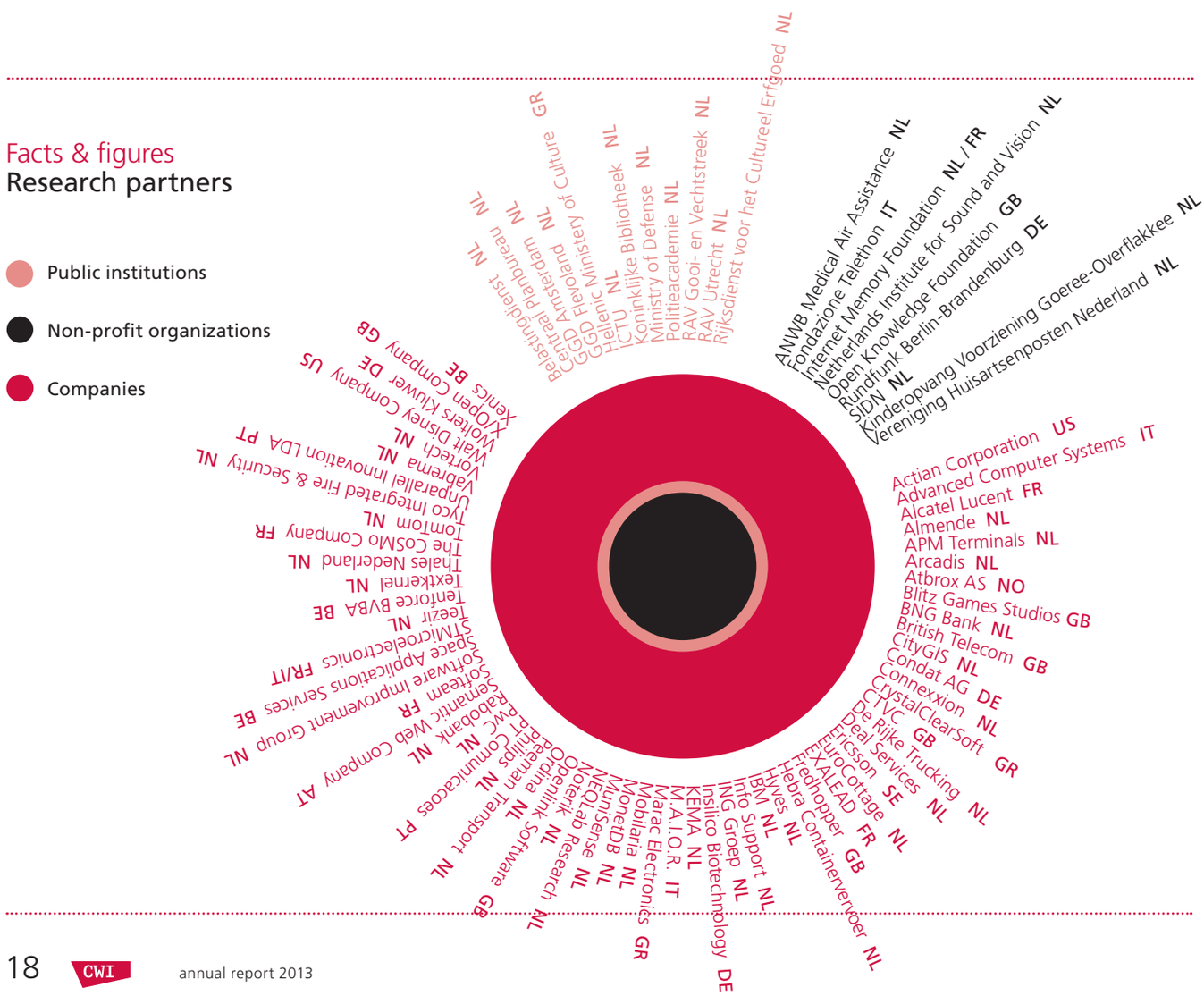
Energy market

The switch to a sustainable society will cause major changes to our energy system. In an EIT ICT Labs project with TU Delft, Imperial College London, TU München and TU Berlin, CWI developed software platform Market Garden for the simulation and demonstration of future smart grid market scenarios. It allows researchers to study and evaluate the effect of market mechanisms under various circumstances and additionally gives user groups, including network operators, consumers and policy makers, a better understanding of future energy distribution and the impact of market mechanisms on energy prices and consumption.

continued on p.16

which is currently used in high-voltage switches. As a result, the group is well equipped to now also access the problem of DC switches. “We have great expertise in the mathematical modelling of electric discharges in high-voltage systems and plasma technology as well as in nature,” says Ute Ebert, leader of the Multiscale Dynamics group. “Our models are able to simulate the discharges down to the molecular level. It is very hard and costly to do lab experiments to figure out the full dynamics, so simulation experiments are a valuable tool to help understand how these discharge phenomena evolve.”
The most reliable simulation results are achieved when all single electrons and their interactions with neutral molecules are simulated. But as it requires a large computational effort, this is usually not feasible. Ebert: “The discharges typically include huge and ever increasing numbers of electrons. We therefore use several mathematical techniques to simplify the problem. For instance, we can redefine the model by using electron densities instead of the single electron motions for all but certain areas. We use models that are even more macroscopic when examining the tree structure of discharge channels, and we study the coupling of discharge and fluid dynamics of the medium.”
The resulting computational models reveal how different media would behave within a high-voltage discharge, and more specifically whether it is suitable to use as a DC switch. To function as such, it has to be possible to efficiently create and extinguish a discharge within the medium. Using this knowledge, electrical engineers like those at ABB should eventually be able to design reliable circuit-breakers for DC nets.
Simulating discharge and plasma phenomena over different scales in space and time is the core expertise of the Multiscale Dynamics group at CWI. Over the past years, the group has applied these techniques to various problems in plasma technology and to lightning-related phenomena in geoscience.

Facts & figures
Research partners



Facts & figures
Organogram

per 31 December 2013

Governing Board

Peter van Laarhoven, Schiphol Group, chairman
Anton Franken, Hogeschool Utrecht
Lorike Hagdorn, TNO
Frank den Hollander, Leiden University
Joost Kok, Leiden University
Rineke Verbrugge, University of Groningen

NWO
Netherlands Organisation for Scientific Research

General Director

Jos Baeten

Scientific Advisory Committee

Management Team

Dick Broekhuis
Lynda Hardman
Han La Poutré
Rob van der Mei
Kees Oosterlee

Research Groups

Algorithms & Complexity	**A&C**	Harry Buhrman
Cryptology	**CR**	Ronald Cramer
Database Architectures	**DA**	Stefan Manegold
Distributed & Interactive Systems	**DIS**	Dick Bulterman
Formal Methods	**FM**	Frank de Boer
Information Access	**IA**	Arjen de Vries
Intelligent Systems	**IS**	Han La Poutré
Life Sciences	**LS**	Gunnar Klau
Multiscale Dynamics	**MD**	Ute Ebert
Networks & Optimization	**N&O**	Monique Laurent
Scientific Computing	**SC**	Daan Crommelin
Software Analysis & Transformation	**SWAT**	Jurgen Vinju
Stochastics	**ST**	Bert Zwart

Service Departments

Communication
Peter Hildering

Financial Administration
Edwin de Boer

Information & Documentation
Lieke Schultze

IT & Facilities
Niels Nes

Personnel & Organization
Angelique Schilder

Secretariat
Hans Hidskes

Centrum Wiskunde & Informatica (CWI) is the national research institute for mathematics and computer science in the Netherlands. The institute's strategy is to concentrate research on five broad, societally relevant themes: Software, Information, Life Sciences, Logistics and Energy.

Visitors address
Science Park 123
NL-1098 XG Amsterdam
The Netherlands

Postal address
P.O. Box 94079
NL-1090 GB Amsterdam
The Netherlands

Phone +31 (0)20 592 93 33
info@cwi.nl

www.cwi.nl

Design
Kitty Molenaar
Printing
Drukkerij Palteam
Publication date
April 2014

Credits

Benschop, Guido	3, 8, 12, 15, 18, 19
CWI	6
Deelman, Paul	10 - 11
Google Earth / Michiel Helvensteijn	2
Lowie, Peter	10, 11
NASA	4
NFI	5
Novum Foto	9
Omroepvereniging VARA	7
Rijksmuseum	13
Shutterstock	14, 16, 17
Z2 / Jolanda Muilenburg	cover

CWI
Centrum Wiskunde & Informatica

Contents

Preface 1
About Centrum Wiskunde & Informatica 2
CWI in 2013 ..... 6, 7, 10, 11

Highlights
Solving crimes with domain-specific languages 4
The possiblities of a quantum computer 8
The collection is yours 12
Transcontinental electricity nets 16

Research themes
Software 2, 3
Information 4, 5
Life Sciences 8, 9
Logistics 14, 15
Energy 16, 17

Facts & figures
Output in 2013 7
Total income 11
Employees 12, 13
Nationalities 15
Research partners 18
Organogram 19