CWI

Centrum Wiskunde & Informatica

Annual Report 2012

I am proud to present our Annual Report 2012. It is a summary of our efforts in 2012 and accounts for what we have done with the resources entrusted to us by society and our public and private partners. Societal relevance becomes more and more a criterion in the allocations of research funds. I believe that the duality of our mission, to conduct pioneering research in mathematics and computer science of the highest quality and put our results at the disposal of society, align very well with the changing research climate. If we emphasize the second part of our mission in the years to come, we will both be able to increase our contribution to society, and safeguard and build on the first part of our mission, as one does not go without the other. In 2012, we developed a new Strategic Plan for 2013–2018, in which we detailed our plans to step up our efforts to cooperate with external partners and anticipate on changes in the research landscape. We realigned our strategic research themes to guide our research in the years to come. They will be Software, Information, Life sciences, Logistics and Energy. This means that this is the last year that we present our research activities through the 2007–2012 research themes.
This document is intended to be browsed in no particular order. In it, you will find four articles where we highlight our research on mobile communication, simulations for drug research, social media and tomography. In five overviews, we further present a bird's eye view of our work on the research themes Earth sciences & energy, Life sciences, The data explosion, Societal logistics and Software as Service. These texts are interwoven with snippets of information concerning other events and activities at CWI in 2012 and facts & figures about our institute. I hope you will enjoy reading and browsing through our Annual Report 2012.

Jos Baeten
General director

About Centrum Wiskunde & Informatica

Mission

Centrum Wiskunde & Informatica (CWI) is the national research institute for mathematics and computer science and part of NWO, the Netherlands Organisation for Scientific Research. The mission of CWI is to conduct pioneering research in mathematics and computer science, generating new knowledge in these fields and conveying it to society, especially industry.

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 platform for mathematics and computer science CWI wants to expand its position in safeguarding the interests of these research fields and playing 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

Research theme

Earth sciences & energy

Smart buildings

Increasingly, smart buildings come equipped with sensors and actuators that control the indoor climate, automate lighting and operate electrical systems. Within the top sector Energy, CWI investigates how this type of infrastructure can be optimized to reduce energy consumption but safeguard – or even increase – the level of comfort for the occupants. The project is carried out with the Eindhoven University of Technology and companies Kropman Installatietechniek BV and Almende BV.

Time-varying prices for energy

The generation of renewable energy is not constant and has a major impact on the reliability of our electricity supply. When switching to green energy the demand should follow the supply of energy instead of the other way around, as is currently the case. By applying mechanism like 'peak shaving', part of the energy consumption can be influenced by using time-varying prices. CWI is investigating different market mechanisms that affect the demand of energy and is developing software that demonstrates these mechanisms in different scenarios.

more on p. 3

Milestones

CWI has a unique talent pool of researchers. Since its foundation in 1946, more than 180 of its researchers have become **full professor**. Our current researchers include a **Spinoza Prize** winner, 15 researchers with one or more **NWO Innovational Research Grants**, one **KNAW**-member, 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 breaking the MD5-hash function, using domain specific programming languages in forensic research, investigating **smart energy networks** and modelling and simulating phenomena such as **lightning**, **plankton growth**, **ocean currents**, **financial products**, **proteins** and **tumour growth**.

CWI plays a central role in various programs and organizations, including managing the W3C Benelux Office and housing public-private research community COMMIT. Since its foundation CWI has commercialized its research in the foundation of **21 spin-off companies** that have generated millions of turnover to date.

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.

Research theme Earth sciences & energy

Ocean currents

Ocean currents play a crucial role in the global climate system by transporting heat. At present, climate models can only simulate large-scale ocean currents, including small-scale features of ocean circulation (eddies) in a very simplified way. They are however crucial in, for instance, mixing of cold and warm water, transport of plankton and salinity of the ocean. CWI develops new mathematical techniques based on random processes to include these small microscale effects in a macroscale model. This research is funded by the Vidi grant of Daan Crommelin.

Predicting desertification

Around one billion people live in areas threatened by desertification. To predict the formation of new deserts, mathematicians at CWI, University of Amsterdam and Leiden University modelled vegetation growth in dry areas near deserts. They modelled patterns of empty spaces that appear at regular intervals between the shrubs and identified the type of instability that causes desertification. They also discovered that the distance between the gaps in the vegetation is constant, so that re-planting gaps during low rainfall is no solution.

Cosmic lightning

Cosmic particles with ultra-high energies penetrate the Earth atmosphere, creating gigantic showers of secondary particles. Researchers from CWI and Kernfysisch Versneller Instituut (KVI) in Groningen investigate the role of these particles in the creation of lightning bolts in thunderclouds. This 'cosmic lightning' project also studies the influence of approaching thunderstorms on cosmic particle measurements. Observations are made with the LOFAR telescope in the Netherlands and the Auger Engineering Radio Array (AERA) in Argentina.

more on p.2

Highlight

The future will be in 3D

Advanced scanning techniques such as electron microscopy, MRI and radiology can be used to make detailed images of objects that are otherwise hidden from view: the inside of the human body, individual molecules in microscopic samples or the contents of a locked suitcase. By combining scans made from different angles of the same object even three-dimensional images can be constructed. The more scans, the more reliable and detailed the resulting image is. But sometimes the number of scans is limited: not only because of time and money, but also because they might damage the object. Constructing reliable images from a small number of scans is one of the challenges of tomography, the field of CWI researcher Joost Batenburg.

Research theme

Life sciences

Flower growth

Researchers of CWI and Wageningen UR developed computer simulations to understand the growth of flowers. The results suggest that flower growth is a self-regulating process initiated by the transport of auxine, a growth hormone. Modelling this process helps researchers to get a better understanding of growth mechanisms and the role of the position of various parts of the flower during development.

Neural networks

Most of the workings of the brain are shrouded in mystery. It has the computational power of a hundred modern supercomputers, but the volume of a soda bottle and the energy consumption of a small light bulb. Researchers at CWI develop models for information processing in the brain. They introduced a dynamic component: the spiked pulses fired by the neurons change in meaning depending on the input of information. This explains why activity in the brain can remain stable, and also suggests novel and efficient models for computation in artificial neural network models.

more on p.5

Tomography based on a few projections is like solving a complex puzzle with what appears to be insufficient information. “I usually start presentations about my work with a cartoon of a rabbit standing in front of a projector, stretching himself as to make his shadow hand-shaped,” says Batenburg. “Nobody in the audience ever guesses that they see the projection of a rabbit. One more projection might show that there is no hand, but how many shadows would you need to reconstruct the actual rabbit? This is the challenge of tomography: not only discovering that a rabbit is posing as a hand but also producing a detailed description of that particular rabbit, and that from only two or three projections.”

Powerful algorithms

Correct interpretation is one of the fundamental problems in tomography, but it is also the key to designing the most powerful algorithms. Batenburg: “A lot of possible objects and shapes can be discarded if you know the context of the scan. If you are scanning a human being, that five-fingered shape will most definitely be a hand and not a rabbit. And if you are scanning a crystal, you expect some orderly arrangement of atoms, and not a randomly folded molecule. The essence of my approach to tomography is to incorporate as much of this human prior knowledge into the algorithm as possible, thus limiting the number of options that need to be considered.”
“When I develop an algorithm, I usually start with making a very detailed mathematical description of the objects that can be expected to be in the scan. The mathematical challenge is then to take away as much information as possible from that description, while preserving its uniqueness. The fewer details needed to describe an object, the more likely is it that it can be reconstructed from a very limited amount of information.”

Nanotechnology

Being able to produce 3D images from limited information is especially important in nanotechnology. In the case of nanoparticles, making only two or three scans is not a matter of efficiency, but an absolute requirement for making any 3D image at all. Atomic-resolution electron microscopy has been feasible for nearly forty decades, but its destructive effect on nanoparticles (mainly electron bombardments and heating) made it impossible to perform multiple scans on the same particle for 3D reconstructions. But in 2011, a group of

continued on p. 6

Research theme Life sciences

Protein comparison

Proteins are very complex molecules, consisting of tens to thousands of amino acid, folded into complex, 3-dimensional shapes. Biologists compare the structure of proteins with an unknown function to those of known proteins to discover their function or their evolutionary relationship. To aid this, CWI developed an online application containing an advanced algorithm that compares any given proteins and gives an optimal solution based on various criteria.

Sleeping sickness

The parasite causing sleeping sickness has very unusual DNA. Information is stored in fragments instead of sequential. Models developed at CWI confirmed the hypothesis that this energy-inefficient system has an important evolutionary advantage. The parasite does not use its entire DNA inside a human. With sequential DNA, individuals losing the unused part of their DNA in the human body would gain the upper hand, only to die in their carrier, the tsetse fly, because of this loss. With fragmented DNA however, undamaged parasites survive. In the malaria parasite a similar mechanism could operate.

Biofuel

Micro-organisms like cyanobacteria can produce energy from sunlight by means of photosynthesis. The energy can be harvested in the form of biofuel. It is a highly sustainable way to produce biofuels, as it does not compete for resources with food production. CWI started a new research project on biosolar cells and is developing mathematical models for optimization of bio-energy production from cyanobacteria. The research is carried out in cooperation with the University of Amsterdam, VU University Amsterdam and Wageningen UR.

more on p.4

Joost Batenburg (32) is one of the leading researchers in discrete tomography worldwide. Besides his position at CWI, he is also part-time Professor of Physics at the Vision Lab of the University of Antwerp (since 2010) and Professor of Discrete Mathematics and Tomography at the University Leiden (since 2012).

researchers including Joost Batenburg succeeded in imaging a gold nanoparticle in 3D from only two Transmission Electron Microscope (TEM) images. Their result, published in Nature, was the first 3D image ever of a nanoparticle with atomic resolution. Using Batenburg's discrete tomography algorithms, the team reconstructed the structure of the particle by cleverly integrating existing knowledge on nanoparticles in the reconstruction. Imaging nanomaterials is essential for understanding and improving their properties. "Material scientists have never seen the nanomaterials they produce in 3D. They know how to make them and know their effect, but up until now, there was no way to see what they looked like at the atomic scale," says Batenburg. "Last year we were able to image the 3D structure of a certain type of solar cell. These cells are composed of a mixture of polymers of only a few nanometres thick, and produce electricity by exchanging electrons excited by sunlight. By imaging the cells in 3D, we saw that the shape and structure of the mixture varied greatly between cells. Consequently, the efficiency of individual cells also varied greatly. Identifying what shape and structure produce the most efficient cells can provide valuable information on how to improve the production process."

Tricorder

The future of tomography looks bright. "The potential applications of 3D scans are ever growing. They are now used for medical imaging, materials science, astronomy, security, even for finding oil far beneath the surface. With efficient algorithms, new areas such as nanotechnology that were beyond the scope of imaging come within reach. And who knows, combining fast and efficient algorithms with powerful microcomputers and high-tech scanning technology might even result in portable scanners. In the Star Trek universe, the tricorder, a portable medical scanner that is able to determine a patient's condition within seconds, is invented in the 23rd century. Wouldn't it be great if we already had one in the 21st?"

Research theme

Societal logistics

Quality of composite services

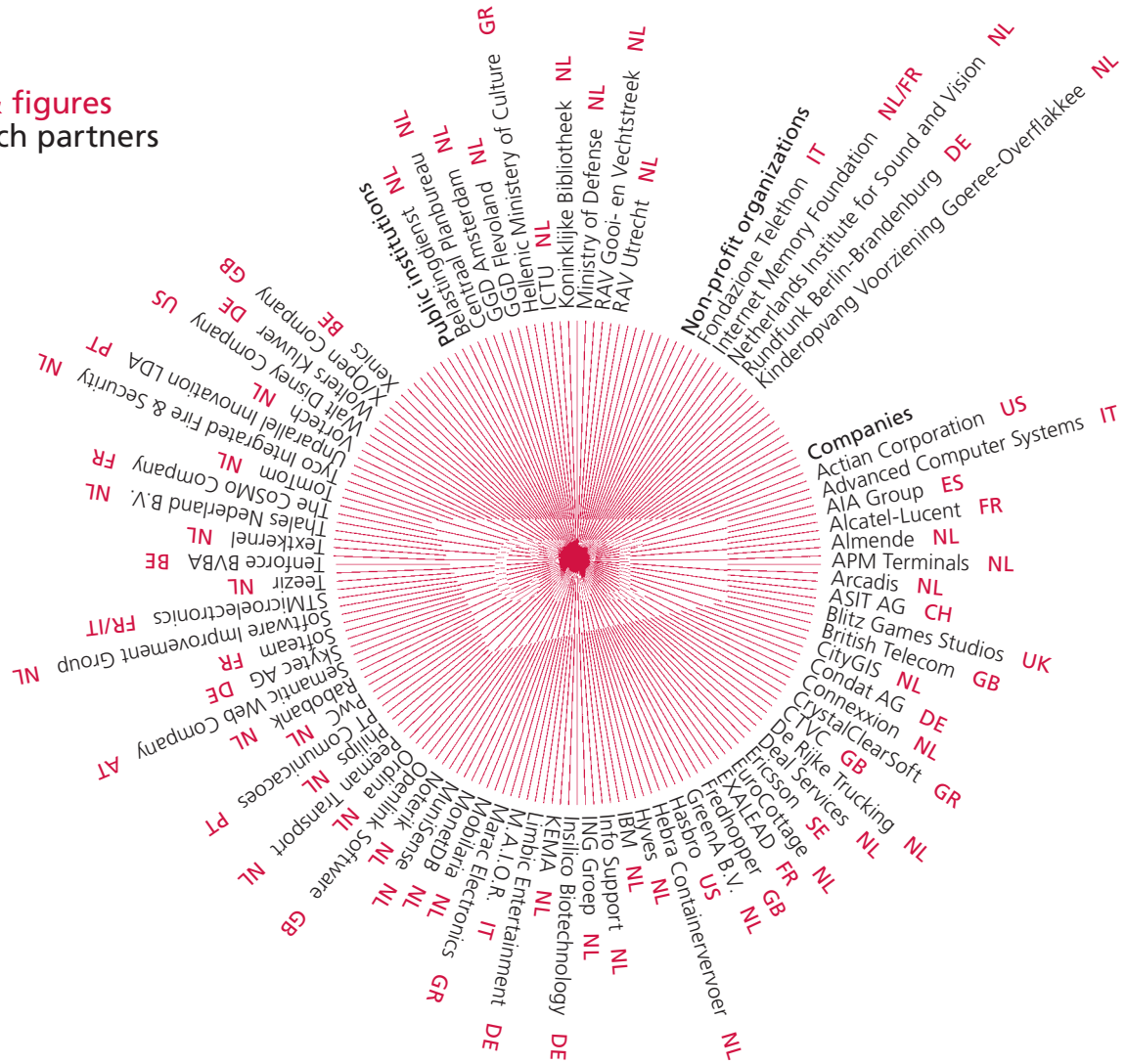
Finding a nearby restaurant with a handy app is a composite service: it combines and integrates the functionality of different services like location ('Where am I?'), database services ('Where are the restaurants?') and navigation ('How to get there?'). However powerful, composite services may lead to long experienced response times if certain services are slower than the rest. Researchers at CWI developed an algorithm for a dynamic composition of composite services, guaranteeing an end-to-end quality for the user.

Traffic flow

Traffic control measures at motorway ramps can lead to a better traffic flow, but need to be coordinated carefully. When too much traffic tries to enter the motorway, ramp metering can for instance lead to highly variable waiting queues between different highway ramps – an unwanted effect. Mathematical models developed at CWI coordinate the traffic control measures at motorway ramps. The control performance is improved and the communication limited by hierarchical control of decentralized systems. This coordination can create a fairer distribution over adjacent ramps.

more on p. 7

Facts & figures
Research partners



Research theme Societal logistics

Google AdWords Optimizer

Companies can place Google AdWords advertisements next to relevant search queries by bidding on certain keywords. The more they bid, the higher the visibility of their ad, leading to more customers clicking on it. Since bidding by trial-and-error gives suboptimal results, CWI researchers analysed the underlying auction mechanism. Based on advanced mathematical tools, they developed a CWI Bid Optimizer, which computes the optimal distribution of the daily budget over the keywords. This can result in 20–30% more clicks – and thus more clients – within the same budget.

Frozen images

A frozen screen is a major source of annoyance when watching web videos. These are often caused by an empty play-out buffer. Streaming services like YouTube use play-out buffers on the receiving side to line up packets ready to be played. This buffer prevents disturbance from other traffic streams crossing the Internet. Researchers at CWI have developed an algorithm that can determine in advance the minimum required buffer level. This prevents empty buffers and therefore frozen images.

Network randomness

Telecommunication, bioinformatics, statistical physics and social networks all deal with networks with a certain amount of randomness. CWI studies the properties of such networks by making random alterations to those networks and investigating their effect. This reveals aspects of the underlying structure and reliability and can be used to understand and prevent network problems that are critical for society, such as the robustness of grid-based power systems and Internet systems. This research is funded by the Veni grant of Ross Kang.

more on p.6

Highlight

Faster downloads with shared networks

The spectacular growth of mobile internet on smartphones and tablet computers has boosted the demand for fast wireless networks. Mobile data exchange is expected to be 26 times as large in 2015 as it was in 2010*. To meet this enormous increase in demand, CWI and technology company Thales Nederland B.V. are working on new methods to make more efficient use of existing communication networks.

* Source: TNO Monitor draadloze technologieën 2011, TNO, August 2011

Research theme

Software as service

Espionage virus analysis

The Flame espionage virus discovered in 2012 infects systems by impersonating a Windows security update, forged using a new, yet unknown variant cryptanalytic attack. Cryptanalysts at CWI discovered this using their novel forensic software. This software can detect whether a digital signature that secures communications is a forgery, so connections can be closed before infection occurs or sensitive information is leaked. Further research also indicates that the cryptographic hash function standard SHA-1, as MD5 before, might be no longer safe to use for digital signatures.

Multi-core processing

Nowadays, desktop computers, tablets and even smartphones have multi-core processors that can execute many tasks at the same time. To make optimal use of the advantages, programmers need new techniques and protocols to make software running on modern multi-core systems satisfy performance and predictability requirements. Software engineers at CWI develop theory and tools to accommodate those needs, in the form of Reo: a graphical, mathematically founded, programming language for specifying interaction among parallel processes.

more on p.9

Most places in the Netherlands are covered by a multitude of overlapping wireless networks from different providers such as KPN, Vodafone and T-Mobile, as well as hotspots and local Wi-Fi-networks. Usually, users can only access one of these networks at a time. But what if several of the available networks could be used simultaneously? That is the question researchers at CWI and Thales asked themselves. Would splitting data over two wireless networks also make mobile communication two times faster? And would it increase the overall performance of the communication? The results of their work were beyond expectation: simultaneous use of two networks did not just double the speed of wireless communication, but could even increase it up to a factor of ten.

High speed gain

The key factor in this speed gain is the strong fluctuation in available network capacity. "We use temporal under-utilization in one network to compensate for congestions in the other." says Gerard Hoekstra, researcher at both Thales and CWI. "Even if both networks are fully in use, there are very high fluctuations in supply and demand of network capacity on a microscale. We exploit this by dividing files in small packages and deciding for every package microseconds before sending which network will be used, namely the one with the highest available capacity at that moment. And of course, sharing networks will drastically decrease downtime, as one network can take over if the other is down." Removing peaks in capacity demand is very effective in reducing congestion. Hoekstra: "As in several logistic problems, speed is not linearly related to available capacity, but exponentially. If the number of cars on a road could be decreased

continued on p. 10

Casimir programme

This research is a joint effort of CWI and Thales Nederland B.V. within the context of the Casimir programme of the Netherlands Organisation for Scientific Research (NWO). This program stimulated the exchange of researchers between companies and public research institutions. In March 2012 Gerard Hoekstra successfully defended his PhD thesis on his work in this project at the VU University Amsterdam, under supervision of prof. dr. Rob van der Mei (CWI).

Research themes Software as service

Developing XForms

CWI is involved in open Web standards that help develop the Web to its full potential. Together with IBM, Inventive Designers, Xerox and others, CWI contributes to the XForms standard of the World Wide Web Consortium (W3C). XForms is used for generating web forms and the design of device-independent applications. It speeds up application production, its code is much smaller than the equivalent JavaScript and it is very flexible and accessible. This year the XForms 2.0 Working Draft was released.

Understanding PHP

PHP is the 6th most popular web programming language and is used on nearly 80% of web servers. It is a very dynamic language, making it hard to perform a static analysis to check the security and vulnerability of websites using PHP. Researchers at CWI have created a corpus of the 19 most popular PHP systems including MediaWiki, Drupal and WordPress, comprising 20.000 files and 3.5 million lines of code. Using Rascal, they analysed the corpus, revealing that circa 75% of the dynamic features can be resolved statically. This creates opportunities for better static analysis of PHP code.

Analysing game software

Lua is a scripting language that is widely used in the game industry to glue game engines and component libraries together. Since Lua is a dynamically typed language, incorrect use of interface functions will only show up during execution. Using Rascal, researchers at CWI have developed a type inferencer for Lua that is able to spot many of those errors. Together with the Hogeschool van Amsterdam and IC3D Media, CWI is currently assessing this approach in practice.

more on p.8

Developing the model

The mathematical challenge of modelling the use of multiple networks at the same time lies in devising accurate models that determine how the file should be split. This requires very complex mathematics, as the networks involved are coupled: the download starts at the same time in all networks, while the transported data through each network depends on the original file size and the statistical (stochastic) variation in available capacity in each network.
The researchers at CWI and Thales first developed an accurate, experimentally verified model of file downloads in a single network. This model is based on queuing: abstractions used in computer science to model waiting lines. The behaviour of multiple networks was modelled with a network consisting of several of these queuing abstractions. By assuming perfect splitting of files, the researchers found a theoretical solution for this model. Experimental results subsequently showed that this theoretical solution indeed provides nearly optimal performance in practice, with a difference between modelled and practical outcomes of no more than a few per cent.

by 10% during rush hours, traffic jams would not decrease with 10% in length, but be halved or even disappear. The same happens in wireless networks: the download times of individual users can be up to ten times shorter when capacity peaks are flattened by using two networks at the same time."

Sharing networks

According to Hoekstra, the developed techniques can be realized on short term. "I could in principle devise a smartphone that can combine any mobile networks, hotspots or Wi-Fi-networks, and a piece of software based on our models, and I am ready to go. But for a more effective implementation, network providers need to adapt their software and hardware to this new mode of operation and more importantly, reach an agreement with competing providers to share networks. Even if they see the potential, I do not see this happen on a large scale within a few years. However, I do think that there are several niches that could highly profit from fast and reliable networking, like emergency and security services, or the military. They could use their own networks to profit from these new techniques."

cwi in 2012

Alan Turing Exhibition

As part of Alan Turing Year 2012, CWI initiated the exhibition 'Turing's Legacy', an innovative format to make computer science and mathematics accessible to the general public. The exhibition told the story of Turing (1912-1954) as World War II hero and founding father of computer science and artificial intelligence. Among the objects displayed were an original WWII Enigma machine, a demonstration of Turing patterns and an X1 computer from 1958.

Highlight was the working LEGO Turing machine, built by researchers Davy Landman and Jeroen van den Bos, that demonstrated Turing's famous theoretical model for a computing machine from 1936. The machine caught the attention of researchers and media worldwide and was subject of a viral web video watched more than half a million times surrounding Turing's centenary in June.

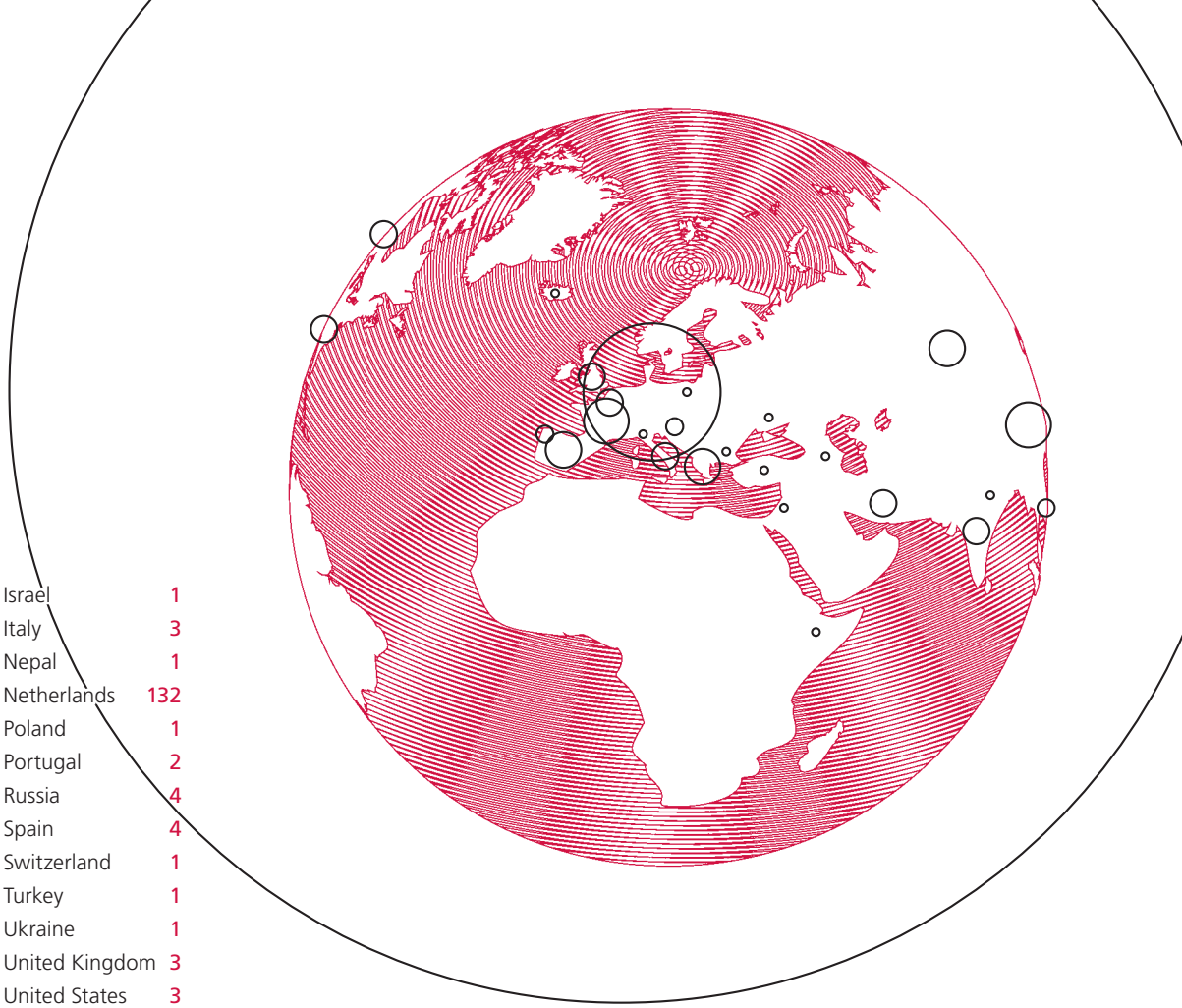
Turing's Legacy was open from 14 June to 30 November 2012. High school groups could combine a visit with a cryptology workshop organized by Stichting Vierkant voor Wiskunde. A total of 40 tours were given to more than 1000 visitors, including Mayor of Amsterdam Van der Laan and President of the ERC Nowotny. Turing's Legacy was realized with support of the Municipality of Amsterdam, Nikhef, Instituut voor Informatica (IvI, UvA), the Institute for Logic, Language and Computation (ILLC, UvA) and the Korteweg-de Vries Instituut (KdV, UvA) and with contributions of IOS Press, King's College Cambridge, LEGO Netherlands, TNO, TU Delft Library and UvA Computer Museum.

cwi in 2012

Media coverage

In 2012 CWI was covered by media numerous times. The LEGO Turing machine developed for the Turing exhibition and Marc Stevens' analysis of the Flame virus was covered by media worldwide and all major national newspapers wrote about 20-year old PhD Floor Sietsma. *Elsevier Magazine* published a 4-page background article on CWI, director Jos Baeten had a long interview on *BNR Nieuwsradio* and various research topics, including quantum computing, neural networks and the genetic code were covered by national radio, TV or newspapers.

Facts & figures
Nationalities



Armenia	1	Israel	1
Austria	2	Italy	3
Belgium	3	Nepal	1
Bulgaria	1	Netherlands	132
Canada	3	Poland	1
China	5	Portugal	2
Ethiopia	1	Russia	4
France	5	Spain	4
Germany	15	Switzerland	1
Greece	4	Turkey	1
Hungary	2	Ukraine	1
Iceland	1	United Kingdom	3
India	3	United States	3
Iran	3	Vietnam	2

cwi in 2012

W3C-ISOC masterclasses

To spread the use of Web standards, the W3C Benelux Office at CWI and ISOC.nl invited two well-known developers from the World Wide Web Consortium to give two masterclasses in September on the newest CSS and HTML5 web standards. Bert Bos from France, co-founder of Cascading Style Sheets, spoke with an enthusiastic audience on the newest techniques of this powerful tool for web designers. Mike Smith from Japan enticed the public with the current work on HTML5, with Anne van Kesteren adding various interesting real-life examples.

cwi in 2012

Veni and Vidi

Three researchers were awarded an Innovational Research Grant by NWO in 2012. Both Stratos Idreos and Ross Kang were awarded a Veni grant (250,000 euro), Daan Crommelin received a Vidi grant (800,000 euro). The Innovational Research Grants are a personal scheme for talented, creative researchers who engage in innovative research, awarded by NWO. The CWI researchers will spend their research money on respectively the topics of Big Data, network randomness and ocean currents modelling.

cwi in 2012

Public events

CWI participated in various events for the general public. Lex Schrijver lectured on discrete mathematics for hundreds of festival-goers during Lowlands University. On the first Science Park Filmfestival in September, CWI hosted the movies Codebreaker and Lightning. And on our Turing-themed Open Doors Day, over 300 guests visited the Turing exhibition, attended lectures on Turing and quantum computing, programmed their own computer game or discovered the secret Enigma code.

Make it personal with a visual story

Reconnecting with high school friends on social networks, playing an online game with your brother, video calling grandma for her birthday… Sharing moments together is no longer limited by distance, as social media provides a range of solutions for maintaining relationships.

Most forms of social media aim at synchronous experiences, like chatting or playing online games. But what if you want to share memories of special events? Even though most people carry around a smartphone with a built-in video camera, the hours of recorded raw material never leave the

cwi in 2012

CWI in Bedrijf

On 5 October 2012, CWI organized 'CWI in Bedrijf', a day for relations in industry. Renowned speakers such as Alexander Rinnooy Kan, Carlo van de Weijer (TomTom), Jacob de Vlieg (eScience Center), Peter Molengraaf (Alliander) and Michiel Buitelaar (Sanoma) spoke about collaboration between science and industry, and innovation in logistics, life sciences, energy and creative industry. On the matchmaking market CWI researchers demonstrated their work. With more than 200 visitors, the day was a huge success.

cwi in 2012

COMMIT/ kick-off

On 20 and 21 March 2012, the COMMIT/-programme had its official kick-off with a two-day meeting in Noordwijkerhout. COMMIT/ is a public-private research community for ICT research and consists of more than 75 partner institutions carrying out 15 large research projects. CWI is involved in COMMIT/ through the projects Information Retrieval for Information Services (INFINITI), Socially-enriched Access to Linked Cultural Media (SEALINC) and Spatio-temporal Data Warehouses for Trajectory Exploitation. CWI also houses the COMMIT/ office.

cwi in 2012

INRIA cooperation

ATEAMS, the joint research team of CWI and INRIA, the French national research institute in computer science and mathematics, was this year formally extended for four years. The cooperation started in 2009 and focuses on identifying and improving software quality. Starting from an initial focus on meta-programming, resulting in the meta-programming language Rascal, ATEAMS is now making a gradual shift from meta-programming to comprehensive, data-rich, programming. In Europe, CWI is also cooperating with other research institutes through the ERCIM consortium, W3C and EIT ICT Labs.

phone memory or the backup disk. The majority of amateur videos are rarely shared or even watched. They are overly long and tedious, when compared to professional material. They do not follow any narrative and show a single, fragmented point of view. In order to make videos interesting to others, there is a need for editing them, and amateur filmmakers do not have the time or the skills. Researchers at CWI have developed a system that facilitates social sharing of videos. The system is capable of automatically creating personalized video mash-ups from raw material captured by multiple amateur filmmakers at an event. The project, called MyVideos, enables users to share stories from an event with people that might not have been there.

Sharing experiences

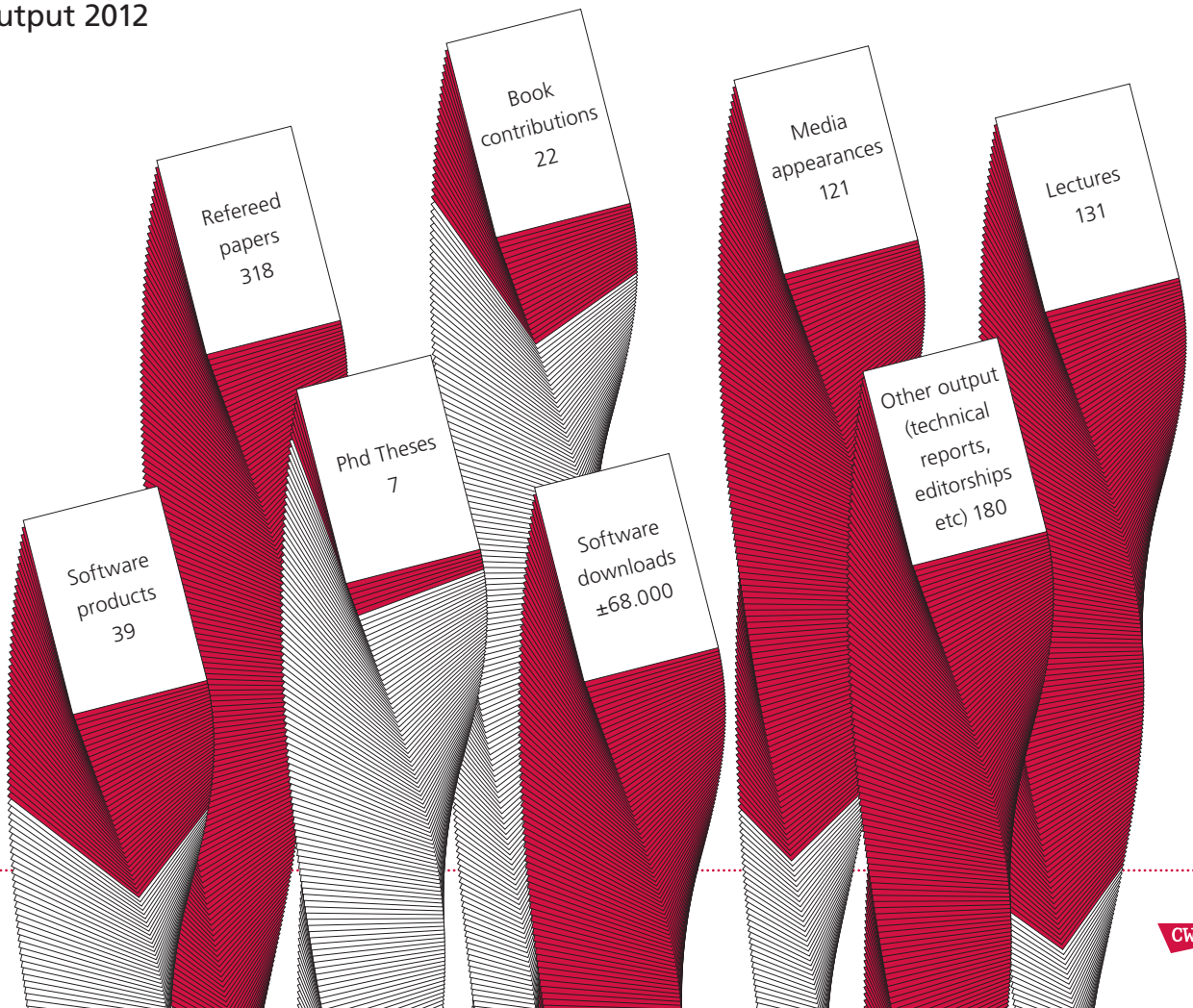
Stories are told differently depending on the audience. Grandma does not have the time to watch an entire school play, but wants to see the scenes where her grandchild shines; whereas the music teacher might prefer to see that song in the play she practiced so hard to get right. "We took a human-centered approach to this project," says CWI researcher Dick Bulterman. "Creating a personalized video memory is all about relationships. Who are those people in the film, why is it shared and who is this person who is going to watch it? A good director does not only know how to put together a good film, he also knows the people in it and the audience that it targets."

School concert

As a pilot event for MyVideos, the researchers at CWI chose an amateur concert at an Amsterdam high school. "The set-up is ideal for our purposes," says researcher Pablo Cesar. "There are multiple parents filming, and the audio track of the concert gives the event a clear timeline." More than 300 raw video clips were collected from twelve different cameras, some fixed and some belonging to parents in the audience. After the concert, the parents assisted in labelling the clips with personal details: who was in the shots and what instrument they did play. The researchers then developed their 'narrative engine' software, which can dynamically compose a fluid and continuous video stream of the concert, targeted to the personal preferences of the viewer. Creating the narrative engine posed severe scientific challenges. Cesar: "We first needed to sort the video clips. A master audio track we recorded allowed us to match the videos to a timeline using their audio stream. Another typical problem with

continued on p. 14

Facts & figures
Output 2012



amateur material is that you are working with bad content. To create the viewing experience, our software needed recognize faces and instruments in low-quality material filmed from different perspectives, with regular blockings of view."

Promising development

Participants were satisfied with the resulting personalized videos. They particularly liked the possibility to focus on a specific person, and agreed that MyVideos would allow them to create more videos and share them with others. "The results show that there is potential in this idea, both from a social and a technological view," says Bulterman. "There are many possibilities to build on the concept of personalized videos. For instance, information about personal relations was now provided by the participants, but in the future users might link their social media accounts to the application." CWI will continue to work on storytelling through social media. Bulterman: "If you look at the landscape of research and development in social media, you see most effort put in synchronized media experiences: sharing an event together at the same time. But I believe that we should focus on asynchronous experiences: telling stories about events that happened. If we continue to investigate these cases, I believe we will not only discover how to develop a good story using technology, but we will discover the technology to share these stories as well."

Together anywhere, together anytime

The MyVideos experiment was carried out by CWI's Distributed and Interactive Systems group as part of the Together Anywhere, Together Anytime (TA2) project. This project, which ran from 2008 to 2012, was funded through the EU's 7th Framework Research Program (FP7). TA2 focused on the concept of togetherness: providing the experience of being together in the same room while being in different locations. A futuristic living room equipped with sensors and monitors was built at CWI to experiment with and to demonstrate new technologies for enabling group-to-group interactions. Sharing personalized videos was one case of such an interaction. TA2's project partners include Alcatel-Lucent, British Telecom, Brno University of Technology, Fraunhofer IIS, Eurescom, Goldsmiths College, IDIAP, Interactive Institute, Joanneum Research, Limbic Entertainment, Philips, Ravensburger and TNO.

Research theme

The data explosion

Big Data at sea

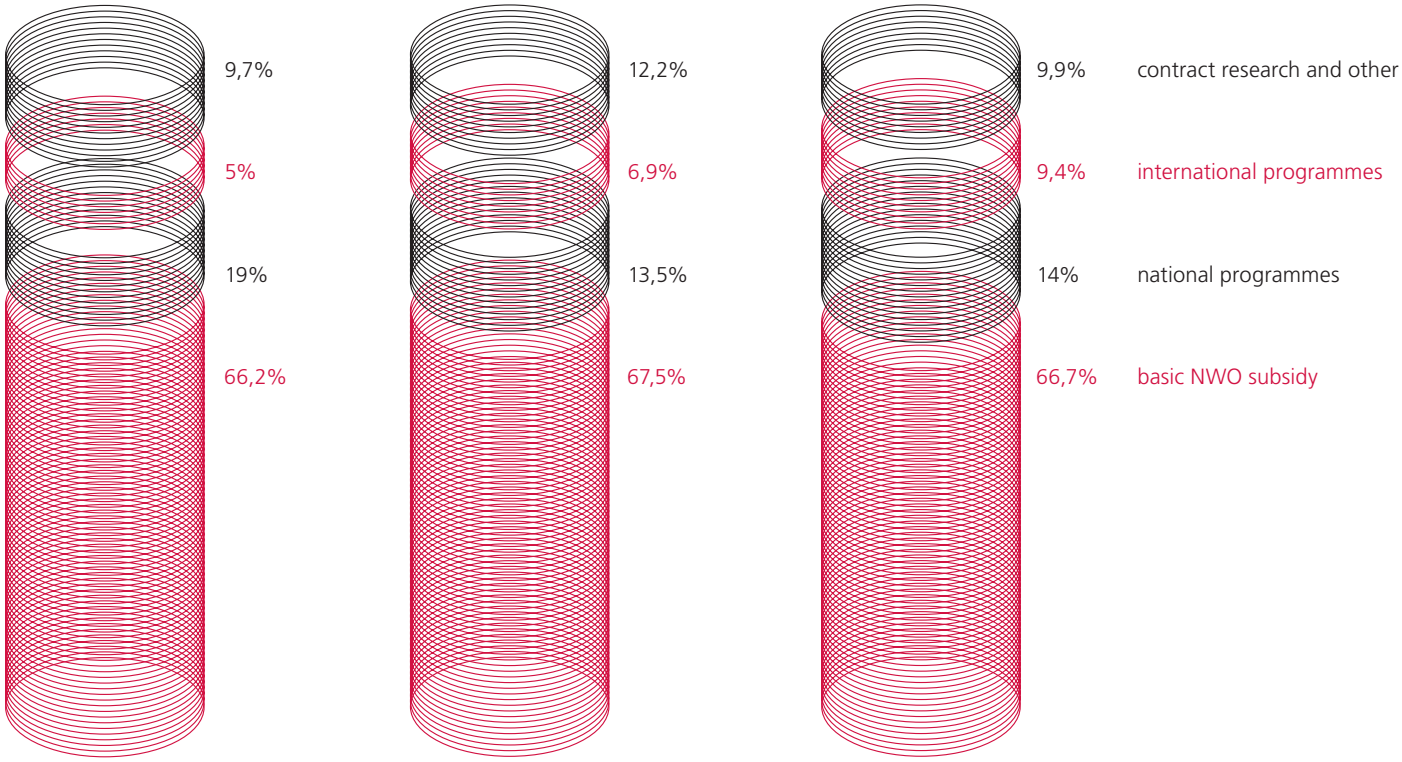
In order to prevent collisions at sea, large ships send their position, speed and heading to other vessels through the Automatic Identification System (AIS). From this system, more than 10 gigabytes of data is collected per hour worldwide. With MonetDB, the open source high performance database management system developed at CWI, a better and faster analysis of these data is possible. Results are of interest for regulatory bodies such as the Coast Guard.

Organizing web archives

Web access to information has become a natural part of our daily lives. Web content is however dynamic, immediately impacting our future cultural heritage. NWO Catch project WebArt studies how web archives should be organised from the future humanities researcher´s perspective. CWI´s Information Access research group develops new approaches and tools to create and access web archives, to maximize the archives´ utility for humanities research. The project is carried out in collaboration with the Koninklijke Bibliotheek and UvA.

more on p.15

Facts & figures
Income



2010: 17.6 M€ **2011: 19 M€** **2012: 18.9 M€**

Research themes The data explosion

Linking TV to the Web

Additional information from the Web can be linked to items in TV broadcasts and presented on the fly. The goal of the Linked TV-project (TV linked to the Web) is to create a rich user experience that seamlessly combines watching television and surfing the web. In this European research project, which is carried out with several academic, public and private partners, CWI is responsible for developing user interfaces for the integration of the additional information.

Attention levels

Distance education opens new opportunities, but comes at a cost. Because of the remoteness, it is difficult for teachers to assess the level of interest of their pupils. Similarly, artists in the future will find it difficult to interact with a remote audience. CWI investigates whether it is possible to make performers aware of the attention level of a remote audience without requiring active input them. Output from bio-sensors, measuring brain waves and skin conduction, were shown to correlate with attention levels of users passively consuming media content or actively playing a computer game.

Linked Open Data

Companies and organizations like the government are increasingly sharing, connecting and re-using data. The RDF web standard is developed to describe and model this open data on the web. Together with company Openlink Software, CWI studied how some of its MonetDB database inventions, such as vectorized query execution, can be applied on RDF data. This research was carried out in LOD2, a European research project in which CWI cooperates with various knowledge partners and companies.

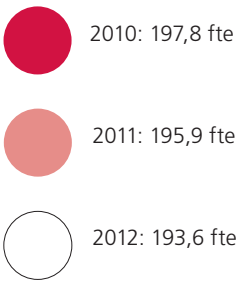
more on p.14

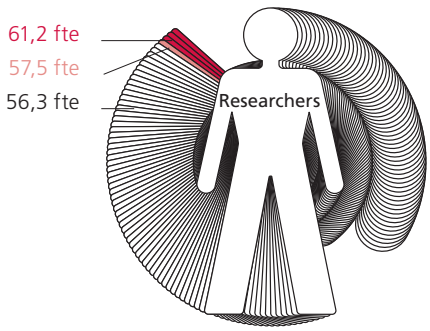
Splitting molecules to find new drugs

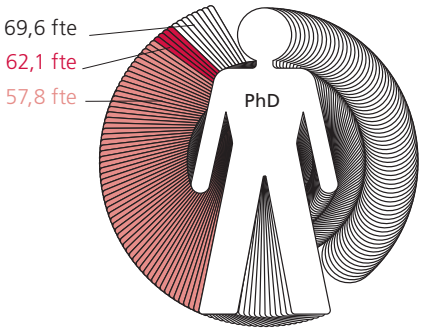
In search of new drugs, researchers increasingly make use of computer simulations. With these simulations, resesarchers can test a large number of molecules on their applicability as a medicine. Current simulation techniques are not yet fast enough to compete with lab tests in both time and costs. CWI develops algorithms that can drastically speed up the simulation process.

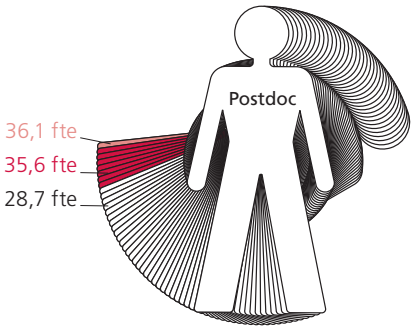
The working of most drugs is based on the medicines's molecules binding to proteins in the body. This causes the protein to change, creating a desired effect such as pain relief, destruction of infectious bacteria or acceleration of wound healing. Finding a molecule that interacts well with a certain protein, without creating side effects by interacting with others, is a long and costly process of informed guessing, research and experiments. Computer simulations that predict the effect of interaction between a candidate medicine and specific proteins offer a promising alternative. With powerful simulation algorithms, millions of potential drugs can be pre-tested on their effect without a single lab

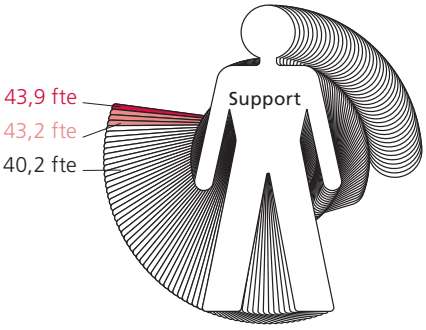
Facts & figures
Employees overview











experiment. This will not only increase the chance and speed of finding new and better drugs, but also prevents wasting expensive lab tests on molecules that in a later stage prove to be ineffective or even toxic.

Chemical intuition

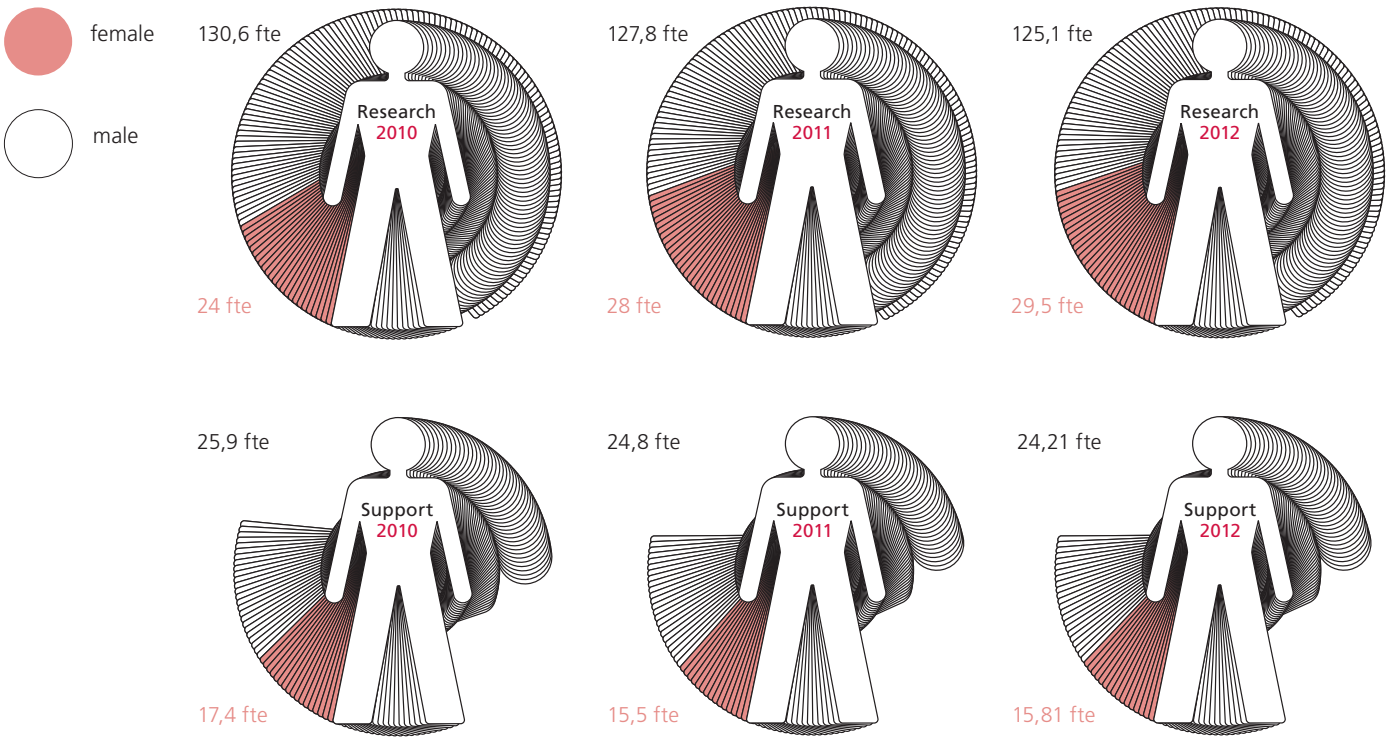
Modelling the interaction between proteins and drug molecules is a very challenging task. Proteins are complex three-dimensional molecules consisting of thousands of atoms. Interactions take place on the atomic scale, so proteins and medicines need to be modelled atom-by-atom during simulation. The molecular shapes even change dynamically over time due to atomic forces and interactions with their environment. This makes exact simulations impossible, even with modern supercomputers. To simplify this situation, researchers model interactions between groups of atoms instead of individual atoms. For this to work, the groups should not be too big and they need to be neutrally charged to prevent inaccuracies in the modelled interactions. Until now, this splitting into groups has been specialist work performed manually by chemists, based on their intuition. It is this chemical intuition that researchers at CWI want to capture in their models. "At CWI, we aim to make the life sciences more exact and quantitative," says Gunnar Klau, leader of CWI's Life Sciences group. "Molecular structure and electronic charges are quantitative properties, so it should be possible to develop an algorithm with the same chemical intuition as chemists, but much quicker. By exploiting several structural properties of molecules (see text box) we managed to devise such an algorithm. It can split molecules into neutrally-charged groups of atoms in a split second with a quality comparable to the work of an experienced chemist."

Drug library

This speed gain opens up the possibility to screen large numbers of possible drug candidates individually for their effectiveness on certain proteins. Drug researchers have large libraries available with millions of molecules that might have an effect on proteins in the human body. Currently, chemists mostly determine experimentally which drugs are best suited for a certain disease. Because of the vast number of molecules in the library, they cannot possibly test all candidates, so the molecules are usually tested in large groups. Computer simulations however can potentially work through the entire library and find the best

continued on p. 18

Facts & figures
Employees male / female



Splitting molecules with graphs

The algorithm for splitting molecules developed at CWI is based on the properties of graphs, mathematical structures that represent the relation between objects by connecting them with links. The researchers modelled a molecule as a graph by defining its atoms as objects with a weight related to their charge, linked together by chemical bonds. The problem of splitting molecules translates to partitioning this graph into groups with charges as close to zero as possible. As the number of possible groups is enormous, the researchers exploited structural properties of the graph to find the optimal charge group partition. "Atoms are not randomly connected, but can only bind to a limited number of other atoms", says Mohammed El-Kebir, researcher in the Life Sciences group. "The graph representing the molecule therefore has a relatively low degree. In addition, the graph structure resembles a tree; it has a relatively large amount of branching and few circular paths. Exploiting these properties makes the problem much easier and quicker to solve."

potential drug. Even the drug's effect on other proteins can be tested to predict possible side effects. Side effects become increasingly important in the pharmaceutical industry, but are now addressed much later during drug development. It might be no surprise that pharmaceutical companies follow the developments around computational drug design very closely. Fast and reliable simulations could decrease the time and cost needed to develop new drugs, and might show better alternatives for existing drugs. Together with chemists from the VU University Amsterdam and the University of Queensland in Australia, the Life Sciences group at CWI continues to improve the algorithms. They work for instance on further error reduction and on determining how to split a molecule based on optimal splits for similar molecules, thereby reducing the need for highly expensive calculations.

cwi in 2012

20-year old PhD

On 13 December 2012, 20-year old Floor Sietsma successfully defended her PhD thesis 'Logics of Communication and Knowledge', making her the youngest PhD in recent academic history of the Netherlands. This fact was widely covered by national media. Sietsma started her PhD research at CWI in early 2010 at age seventeen with a special, personal grant awarded by NWO because of her unique talents.

cwi in 2012

CWI Lectures

On 14 June 2012, over 200 software engineers, students, people from business, researchers and relations met at CWI for the CWI Lectures 2012: Understanding Software. The topic of the day was understanding how to build better software and maintain and improve existing software. The Lectures were organized in honour of Paul Klint's appointment as CWI Fellow, as well as his 40th anniversary at CWI. In the afternoon, NWO's general director Hans de Groene officially opened the exhibition Turing's Legacy at CWI.

cwi in 2012

Cryptanalysis in the media

CWI's cryptanalysist Marc Stevens analysed the Flame virus only briefly after the espionage malware was discovered. His findings were covered by media worldwide, including *BBC News online*, *Le Monde Informatique*, *PCWorld*, *GeekWire*, *redOrbit*, *Automatisering Gids*, *Computable*, *Tweakers.net* and media from China, Russia and India. The news item on CWI's website was viewed more than 20,000 times. Security guru Bruce Schneier wrote on his blog that the SHA-1 standard should be replaced due to CWI's cryptanalysis.

Organogram

per 31 December 2012

Governing Board

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

NWO

General Director

Jos Baeten

Scientific Advisory Committee

Research Clusters

Probability, Networks and Algorithms	Rob van der Mei
• **Algorithms, Combinatorics and Optimization**	Monique Laurent
• **Probability and Stochastic Networks**	Bert Zwart
• **Cryptology**	Ronald Cramer
• **Algorithms and Complexity**	Harry Buhrman
Software Engineering	Han La Poutré
• **Software Analysis and Transformation**	Jurgen Vinju
• **Foundations of Software Engineering**	Frank de Boer
• **Multi-agent and Adaptive Computation**	Han La Poutré
• **Distributed and Interactive Systems**	Dick Bulterman
Modelling, Analysis and Computing	Kees Oosterlee
• **Computational and Stochastic Dynamics**	Jason Frank
• **Scientific Computing and Control Theory**	Kees Oosterlee
• **Multiscale Modelling and Nonlinear Dynamics**	Ute Ebert
• **Life Sciences**	Gunnar Klau
Information Systems	Lynda Hardman
• **Database Architectures**	Stefan Manegold
• **Interactive Information Access**	Arjen de Vries

Service Departments

Communication
Peter Hildering

Information & Documentation
Lieke Schultze

Personnel & Organization
Angelique Schilder

Projects, Finances & Control
Dick Broekhuis

Information Technology & Facilities
Niels Nes

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: earth sciences & energy, life sciences, the data explosion, societal logistics and software as service.

Visitors address
Science Park 123
1098 XG Amsterdam

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

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

www.cwi.nl

Design
Kitty Molenaar
Printing
Drukkerij Palteam
Publication date
April 2013

Credits

CWI 3, 6, 11, 12, 13, 14, 17, 18, 19
Kaspersky Lab 9
Minderhoud, M. 7
Nijhuis, Hanne 10
Shutterstock 2, 5, 8, 15, 16
University of Antwerp 4

CWI
Centrum Wiskunde & Informatica

Contents