I/O magazine
ICT RESEARCH PLATFORM NEDERLAND
VOLUME 17 DECEMBER 2020 NO 3

GREEN ICT

RESEARCH

16

Understanding the social side of language

In conversation with Dong Nguyen.

24

SURF as innovation broker

Vision of Ron Augustus, Chief Innovation Officer at SURF.

COLOFON

I/O Magazine is a publication of the ICT Research Platform Nederland (IPN) and is sent free of charge to ICT researchers and relations of IPN four times a year. IPN consists of the ICT research schools ASCI, IPA, and SIKS; the ICT-related themes of NWO domains Science (ENW) and Applied and Engineering Sciences (AES); the institutes of the technical universities, united in NIRICT; the institutes of the general universities; SURF; e-Science Center; CWI; Dutch Platform for Mathematics; Data Science Platform Netherlands; Dutch Tech Center for Life Sciences; VERSEN; TNO and COMMIT.

IPN (ICT Research Platform Nederland) unites all Dutch academic research groups that have ICT science as their core, and as such acts as a single point of contact for all matters relating to ICT innovation and its importance for our current and future society. IPN is supported by the NWO Domain Science.

EDITORS Eveline van den Bos, Sonja Knols, Avital Lievendag, Femke Stephan, Yvette Tuin **MANAGEMENT** Avital Lievendag **FINAL EDITING** Sonja Knols **TEXT CORRECTION** Dave Thomas, NST Science **CONTRIBUTORS** Leendert van der Ent, Sonja Knols, Rianne Lindhout, Reineke Maschhaupt, Bennie Mols, Amanda Verdonk, Jurgen Vinju **DESIGN AND LAYOUT** WAT ontwerpers **PHOTOGRAPHY** Collectie Naturalis (cover, p. 18), Shutterstock (p. 4, p. 7, p. 9), WAT ontwerpers (p. 8, p. 19), Bram Saeys (p. 6, p. 7), uCrowds (p. 10), Ivar Pel (p. 12, p. 24), Elodie Burrillon (p. 16)

EDITORIAL ADDRESS Secretariat IPN, c/o NWO Science, PO Box 93460, 2509 AL The Hague, the Netherlands, +31 70 344 0772, ipn@nwo.nl, www.ict-research.nl

JURGEN VINJU - CWI & TU/E

Noisy numbers

We use computers to do all of our computations. The faster and more ubiquitous computer hardware becomes, the more we use it to compute all kinds of numerical output. But what if those computations are wrong?

Unfortunately, many computations produce inaccurate results. In two cases, the computer is the source of the increasing inaccuracy of the output numbers:

(1) The input data is sometimes inaccurate. For example, let us take daily maximum temperatures measured over the last century. These data undoubtedly contain both random errors and structural errors that we do not know about. When we use a statistical procedure to confirm a correlation, then the output of this computation could be wildly different given accidentally different numbers. Many errors are heterogenous, non-constant and non-independent. So even averaging the result of multiple measurements does not help.

(2) Computers do not compute all that accurately. Most of the computations done by computer chips are based on floating points. These floating points have a fixed set of binary digits, leading to rounding errors, and they might differ in significance. As a result, floating-point arithmetic behaves nothing like high school mathematics. Calculations are not always associative or distributive, as is the case with exact numbers. Calculations based on floating points might therefore lead to mistakes that are not immediately obvious. Practically all computers implement the same kind of standardised floating-point arithmetic, and so they all make the same mistakes.

Experts in numerical methods know exactly how to minimise cumulative floating-point errors and the effects of inaccurate inputs. However, we now have a huge programmer population who neither have the time nor the necessary training to do the same. Computations on inaccurate data with inaccurate floating points fail us silently, since they produce numbers that look plausible. What if programming languages would make all that inaccuracy explicit? That at least would make the issue visible. Yet that would come at a price. Perhaps the computation performed would be ten times slower. However, it would be very good to know exactly how noisy the numbers we build our lives around are.

NWO
IPN ICT-ONDERZOEK PLATFORM NEDERLAND

IN-DEPTH INVESTMENT FOR DUTCH COMPUTER SCIENCE

For a period of five years, the Dutch government is investing an extra 8.4 million euros per year in academic computer science. Main objectives: strengthening fundamental research, increasing the number of permanent appointments and increasing diversity.

By Bennie Mols Images Shutterstock, Bram Saeys

Over the past ten years, the number of STEM students in the Netherlands has grown almost eight times faster than the permanent academic staff. This has led to a high workload and, in particular, a high educational burden: too few lecturers for many students. This has been at the expense of academic research. At the same time, the demand for computer scientists from the labour market increased sharply, due to the ever further digitisation of society.

These two reasons have urged the Dutch government to invest an extra 8.4 million euros per year in academic computer science from 2020 till 2024. This extra funding is part of the so-called *Sectorgelden voor Bèta en Techniek* (see box). Professor Hans van Duijn (Eindhoven University of Technology) chairs the committee that will monitor the *Sectorplan Bèta en Techniek*. Professor Ineke Braakman (Utrecht University) is the chairman of the *Kamer Bèta*, which also includes computer science. In a double interview, Van Duijn and Braakman explain the choices made for the computer science part of the Sectorplan.

'In addition to reducing the workload and responding better to the demands of industry, the Sectorplan has two other objectives', says Van Duijn. 'One is to renew scientific research. The positions need to be filled with new personnel that will bring in their own new research line. Simultaneously, we are striving for greater staff diversity in terms of gender and ethnicity. In summary, it comes down to strengthening the foundations of computer science in the Netherlands.'

Braakman adds: 'In five years' time there will be around 60 to 80 new permanent appointments for computer science. The first funds have already been paid out in 2019, and that's also when the first recruitment began.' Braakman and Van Duijn do not think that the imbalance between the number of lecturers and the number of students will be completely resolved. However, according to Van Duijn: 'This is much more than a drop in the ocean.'

Bottom-up

On the basis of which criteria have the funds been distributed to the computer science departments at the various universities? Van Duijn explains: 'The deans have made an initial budget distribution, taking into account the size of the departments and the ratio between the number of students and the number of lecturers. Based on plans that the individual universities presented to the committee and, of course, the quality of the research, the Kamer Bèta has determined the final distribution and awarded positions to each university.'

No additional applications were required. The Sectorplan budget is earmarked money and not a lump sum. The number of junior and senior positions per university in each discipline is determined, as well as how much money is available per position. The ratio between the number of junior and senior positions is tailored to local needs and was part of the plans presented by the deans. Besides funding for new junior and senior positions, there is also budget for temporary positions and (small scale) infrastructure, as a starting package for the new Sectorplan positions. 'The universities have decided which focus areas they want to invest in using the Sectorplan budget', says Braakman. 'It is great to see that some universities complement the new Sectorplan positions themselves with new appointments.' In this way, these focus areas really get a whole new impulse. This should also apply to the diversity of the workforce. Braakman: 'Attracting women is already going much better in computer science. The deans told us they were aiming for 50% women, but that we could keep them to the promise of at least 30%. Many new people are coming in from Eastern Europe, Southern Europe and Asia, so that automatically gives you a lot of diversity, but what is still missing are second or third-generation migrants from our own country. Monitoring ethnicity, however, is not allowed, so it is much more difficult to oversee and redirect this.'

Ineke Braakman:

'In five years' time there will be around 60 to 80 new permanent appointments for computer science'

Compared to the other science subjects, there is a specific problem with computer science: the demand for computer scientists from industry is so high that it is very difficult to keep good people at the university. 'We need to think more creatively to do something about this,' says Braakman. 'You can think of hiring people with double appointments both at a university and at a company. That happens much more often in the US, and that already makes it much more attractive to some people.'

Critical friend

To stay on top of things, the Kamer Bèta will continuously monitor progress and evaluate whether there are any bottlenecks in the roll-out of the Sectorplan. Van Duijn: 'If we can make this Sectorplan a success, the funding will become structural. In the meantime, our role is that of a critical friend.'

'We look over the shoulder of the department deans', adds Braakman, 'we ask critical questions, but all in good faith and with the intention of making sure that the investments really become a success. We have also been able to speak to a number of newly appointed computer scientists, and we heard a great deal of enthusiasm. But we also became aware of concerns from people who have come here from abroad. We take those concerns seriously and see what we can

do to help and improve. For example, we give advice on how to create a closer community between all those new people.'

Drawing lines

Another point of attention is the extent to which boundaries between disciplines and sub-disciplines must be maintained. These boundaries are often arbitrary, and precisely at these boundaries between disciplines is where the most ground-breaking research takes place. Braakman: 'There is hardly anyone who works purely mono-disciplinary anymore. That's why you have to think carefully about where to draw the lines between different Sectorplans and realise that they are always artificial. Is everyone who works in computer science in the Netherlands now served with this Sectorplan? No, probably not. But that also means that people have to start organising themselves. And if you're thinking about a follow-up to the present Sectorplan, I would like to be pragmatic and draw the lines generously between disciplines. You want everyone to be able to get funding from the Sectorplan once. Not twice, but not zero times either.'

When will Van Duijn and Braakman be satisfied with the allocation of the Sectorplan funds for computer science? Braakman: 'When we see that the computer science base at the universities has been strengthened so that there is a stable structure.' Van Duijn adds: 'And when the workload has been structurally reduced, and diversity has been structurally increased. I also hope that the Sectorplan initiative as a whole will have such an impact that it will be continued. There are other disciplines that do not currently fall under the Sectorplan funds, but which would also deserve it.'

Hans van Duijn:

'We aim to renew scientific research and strive for greater staff diversity in terms of gender and ethnicity'

SECTORPLAN FUNDING

A few years ago, the Dutch government decided to structurally invest 70 million euros per year for a five-year period to strengthen the basis of scientific research. Bert Meijer (Eindhoven University of Technology) together with the deans drew up sector outlooks of the natural sciences for four science disciplines (computer science, physics, chemistry and mathematics) and for three technology disciplines (civil engineering, electrical engineering and mechanical engineering). In these disciplines, the need to provide additional financial support was considered the greatest.

These sector outlooks were then translated by an advisory committee headed by Hans van Duijn into a *Sectorplan Bèta en Techniek*, describing which university should receive which part of the available budget and for which particular positions. For computer science, 8.4 million euros was made available for new positions in seven focus areas, namely Data modelling and analysis; Machine learning; Machine reasoning and human-machine interaction; Algorithms; Software; Security and privacy; and Networked computer systems and embedded systems.

National Growth Fund

By Sonja Knols
Image WAT ontwerpers

Last September, the Dutch Cabinet launched the National Growth Fund. Over the next five years, twenty billion euros will be available for investments that contribute to economic growth. The National Growth Fund is aimed at knowledge development, infrastructure and research, development and innovation.

The National Growth Fund is intended for one-off public investments that structurally increase the growth capacity of the Dutch economy, but are not commercially attractive yet. Investments from the National Growth Fund are gifts, not loans. There are no return criteria. The money for this fund does not come from existing funds, but is borrowed outside of the regular budget. The idea is that the loan will be recovered in the long term in terms of higher economic growth.

The Cabinet expressed the intention to fund a first round of projects within its current term of office. For the sake of time, proposals for the first round have arisen from sectors that already had well-developed strategies in place. For the ICT field, topics like artificial intelligence and quantum computing could perhaps be promising candidates for this fund.

Proposals for second round

After evaluation of the first round, a second submission round will start in the spring of 2021. In principle, for the National Growth Fund, ministries are in the lead. Knowledge institutes, companies, entrepreneurs and researchers can submit proposals for the next round of investments to the ministry responsible for the subject of their application. However, subsidy schemes will also be designed that enable field parties to submit proposals themselves, without the intervention of a ministry. The exact design of this process will be announced this autumn. Information about this process can be found on the website of the Dutch government. In addition, the National Growth Fund letter to the Dutch House of Representatives ('kamerbrief Nationaal Groeifonds') provides more information about the conditions that proposals should meet.

To ensure an objective and politically independent assessment of the investment proposals submitted, an independent committee has been set up. Money will only be made available after a positive advice from this committee and a subsequent decision by the Cabinet. During an annual "State of the Economy", the Minister of Economic Affairs and Climate Policy and the Minister of Finance will jointly announce which plans will receive funding.

Process

After submission, each proposal is screened to see if it fulfills all entry requirements, such as that it focuses on knowledge development, innovation or infrastructure and concerns a one-time investment that is not covered by other (subsidy) regulations or regular expenditure. An independent committee will assess eligible proposals on their social benefits and costs and their effect on the growth of the Dutch economy. Based on the committee's advice, the Cabinet will decide upon the final investments.

More information (in Dutch):
www.rijksoverheid.nl/onderwerpen/nationaal-groeifonds

ICT.OPEN2021

The field of ICT is rapidly developing. There are many open research questions and new opportunities to apply research results to further transform and improve our society. To bring together fundamental ICT research, as well as novel and innovative applications of ICT research, the Dutch Research Council (NWO) and the ICT Research Platform Nederland (IPN) are organising the 9th edition of ICT.OPEN. More than 500 scientists from all ICT research disciplines and the ICT industry come together to learn from each other, to share ideas and to expand their network. The 9th edition of ICT.OPEN will take place online on Wednesday 10 and Thursday 11 February 2021.

RESPONSIBLE USE OF ARTIFICIAL INTELLIGENCE

Last summer, six applications were awarded in the multidisciplinary call Artificial Intelligence Responsible Use. All six projects are public-private partnerships and address major AI issues in Dutch society and the Dutch economy. The funding of these projects is made possible by the Ministry of Economic Affairs and Climate Policy, the Top Sector ICT and the NWO Domains Science and Social Sciences and Humanities. The projects cover areas as diverse as culturally aware AI for automatically generated classifications and descriptions of collection data, causality in multi-agent systems, trustworthy implementation of AI, using AI to help medical teams in finding clinical trials for patients, AI to improve customer service, and explainable AI to assist auditors who have to check financial self-reporting on credibility.

VENI AND VIDI GRANTS AWARDED

A total of 161 highly promising young scientists have been awarded a Veni grant worth up to 250,000 euros. The grant provides the laureates with the opportunity to further elaborate their own ideas during a period of three years. A total of 1,127 researchers submitted an admissible research proposal for funding. 161 of these have been granted. That comes down to an award rate of 14%. 81 experienced researchers received a Vidi grant worth 800,000 euros. The grant enables them to develop their own innovative line of research and set up their own research group in the coming five years. A total of 503 researchers submitted an admissible research project for funding during this Vidi funding round. Eighty-one of these have received grants. That amounts to an award rate of 16%.

DATA AND INTELLIGENCE CALL OPEN

NWO has launched two KIC calls within the Knowledge and Innovation Agenda (KIA) Security: Data and Intelligence and Maritime high-tech. The Data and Intelligence call focuses on a safe society using data and intelligence, with an emphasis on security. Three million euros will be made available for this call. The deadline for submitting a pre-proposal is 22 April 2021. In January 2021, NWO will organise a virtual matchmaking event with the aim of bringing together researchers from different scientific disciplines and organisations. For the KIA Key Technologies, a previously launched call is currently in progress.

WISDOM ABOUT THE CROWD

By Leendert van der Ent
Images uCrowds

Many companies rely heavily on the events market. But how many of those managed to compensate their COVID-19 invoked loss of turnover by working to mitigate the corona-virus instead? Utrecht University spin-off uCrowds did. Its crowd simulation software SimCrowds® now comes in handy for optimising social distancing scenarios.

UCROWDS TIMELINE

2006
Geraerts receives his PhD in sampling-based motion planning for robots, which inspired him toward real-time and interactive simulation of 100,000s of agents.

2009
Three industrial parties ask Geraerts to make a demo of crowd simulation in a virtual environment.

2014
The underlying software engine is used to create a simulation product, SportsEvac, which was used by NCS4, an organisation linked to Homeland Security (USA).

SWIFT SWITCH

Imagine you are a start-up simulating routing and crowd behaviour during events such as rock festivals, the Vuelta or the Eurovision song contest. Then COVID-19 comes along and this market collapses. What to do? 'We quickly realised that simulations might help to study the effects of social distancing on buildings or shopping areas', says founder Roland Geraerts. As an example, uCrowds now helps the city of Utrecht to optimise the capacity of its 4,000 FTE city office under the 1.5-metre distancing guidelines. Geraerts: 'In another project, we are working on adding predictions to real-time crowd flow data. We need three minutes to calculate fifteen minutes ahead, for instance in a shopping district. This buys enough time to allow for more pleasant control measures than locking down a whole area. The result: law enforcement officers are happy, shop owners are happy and the public is happy.'

HIGHLY SPECIALISED

uCrowd's domain is highly specialised – its competition comprises no more than ten parties worldwide. Geraerts argues that its product, SimCrowds®, has a disruptive licensing system. 'A number of characteristics make us stand out. One of these is that there is always an external expert user involved to create the impact for the end client by using the software. Typical experts are consultants in crowd management or logistics whom we empower to reach the proper interpretation of the simulation by training them in using the software', Geraerts comments. uCrowds therefore licenses end clients such as municipalities, the central government, and festival organisers. 'That is a disruptive approach in our market.'

The software distinguishes itself by its ease of use and its efficient learning curve, says Geraerts. 'Typically, other crowd simulation software directly connects its Graphical User Interface (GUI) to the underlying Application Programme Interface (API). Our user-centric approach of GUI-design focuses on the ease of use that minimises the number of actions – such as mouse clicks – and supports interaction by touch to get the job done. Users can get a full grip on the software within one or two days.'

INTERACTIVE

Geraerts points out that SimCrowds allows for the creation of highly realistic models while keeping the calculation time low. 'For instance, we include groups with a special social connection in our crowds, such as families or groups of friends. Their behaviour differs from that of individuals. Despite this level of detail, we run our simulations ten times faster than other simulation software, thanks to our computer science research conducted at Utrecht University.'

The use of game technology in SimCrowds makes the programme interactive. 'This allows us to introduce on-the-spot scenario changes in an interactive simulation while we sit together with our end client.'

Roland Geraerts

ELECTIONS

Clients from all over the world come to uCrowds. For London's St. Pancras station, the predictive software helps to manage the flow of people during corona restrictions through instructing train arrivals and other means of spreading travellers. Currently, uCrowds is working for the Dutch Ministry of the Interior and Kingdom Relations to support next spring's election process. What does social distancing mean for voting throughput and waiting times? In February 2020, Geraerts could not have imagined that he would be asked to apply SimCrowds in support of the democratic process.

2015
The software engine is used to predict crowd behaviour during the Grand Départ of the Tour de France in Utrecht and during the evacuation of the Amsterdam North/South subway line. An NWO valorisation grant enables completion of the Application Programme Interface (API).

2016
Geraerts participates in the Science Venture programme of Utrecht University's incubator, UtrechtInc, to prepare for spin-out.

2017
A NWO Take-off 1 grant allows him to write a business proposal and do market research.

2018
uCrowds starts after signing a contract with Utrecht University. An NWO Take-off 2 loan allows the creation of SimCrowds.

Sustainability in software engineering

By Bennie Mols Images Ivar Pel

GROUP PASSPORT

RESEARCH FIELD

Software and Sustainability: creating software engineering knowledge that makes software better, smarter, and more sustainable; Software architecture; Software design and modelling; Software quality assessment.

INSTITUTION

Vrije Universiteit Amsterdam

WEBSITES

http://patricialago.nl
http://s2group.cs.vu.nl

EMPLOYEES

1 Professor
9 assistant and associate professors
6 postdocs and PhD students
1 junior lecturer
2 research assistants.

FACILITIES

Green Lab: a lab for experiments with the optimisation of the energy efficiency and other quality properties of software deployed on servers and mobile devices like smartphones.

Software engineering is about more than getting the software to function properly. It is also about getting software to consume as little energy as possible and to take care of the diversity of developers and users. The research group Software and Sustainability at VU Amsterdam studies these and other dimensions of sustainability in software engineering.

Ivano Malavolta

Darkness falls in the early afternoon in the Swedish town of Uppsala during the winter. Not particularly attractive for children to play outside. However, in 2015 the municipality of Uppsala, together with Philips, showed that smart LED-lighting of a local playground, by installing a system which allows remote control of light levels and tones, increased the time kids played outside and improved the children's mood. This is an example of the transformational power of software and IT to improve people's lives. Professor Patricia Lago gave this example during her inaugural speech in January 2016 at VU Amsterdam, one year after she started her position in the new field of Software and Sustainability, which at that time was mainly focused on the energy efficiency of software.

Better, smarter, more sustainable

Four years later, Lago has expanded her VU Amsterdam research group both in size and research scope. 'The goal of our group is to create software engineering knowledge that helps software to become better, smarter and more sustainable', says Lago. In essence, sustainability is the ability to continue a defined behaviour for a long time. Nowadays, sustainability is considered to have four dimensions: not just the environmental dimension that is traditionally associated with sustainability, but also the social, technological and economic dimensions. Lago: 'In all these dimensions we should treat sustainability as a quality property of software. They should already be taken into account during the design phase.'

In the environmental dimension, engineers try to reduce the energy consumption of software. 'Surprisingly there is a lot to be gained in efficient software', says Lago. 'We once showed how optimising the software of a website saved up to 45% of the energy consumed by the webserver.' Lago and her colleagues set up their own Green Lab where they experiment with the energy consumption of software in servers and more recently also the energy consumption of software in smartphones and in robots.

Lago's research group was the first in the world to develop a Sustainability Quality Assessment Model for all four dimensions. 'Basically, it is a set of instruments to help software architects and other professionals to identify which of the software quality properties are relevant for them. At the same time, it is an instrument to measure the impact on the target sustainability goals.'

Patricia Lago

Emitzá Guzmán Ortega

Lago believes in the power that software has in shaping our society. 'That's also the reason why we cooperate with many industrial partners. Connecting software with sustainability is quite a new IT field, but my dream is that the societal relevance of software engineering will be increasingly recognised by the international community. I already see how the new generation of students wants to have a meaningful impact on society.'

Green Lab

Ivano Malavolta is assistant professor in data-driven software engineering in Lago's group. His work is structured around the Green Lab. 'If you entered the lab, you would see a rack of servers for collecting energy measurements. And you would also see a large table with twenty smartphones for doing experiments. All these devices can be controlled remotely via a dedicated website.'

The Green Lab is also a mental gym for students, as Malavolta calls it. 'We teach a master course where around forty students per year learn to design and conduct scientific experiments in the energy efficiency of software. Finally, the Green Lab is also a platform to collaborate with companies. They come to us with their problems, we talk with them and design an experiment to investigate the energy efficiency of their software. It's a form of scientific consultancy. Since software engineering is such an applied field, we need this type of cooperation with companies. It keeps us fully updated with professional practice.'

What Malavolta especially likes in the group is the fact that not just the publication of academic papers counts, but most of all the impact they have on society. 'For us, academic papers are a means to achieve something else. That is Patricia's school of thought that all our group members share. So, societal relevance is really the backbone of our group. And then, in the informal sense, there is an Italian backbone: eating pizza is another unifier, as both Patricia and I are Italian.'

Social dimension

Mexican-raised Emitzá Guzmán Ortega is another assistant professor in Lago's group. Her research focuses on the social dimension of software sustainability, studying both the side of the developers and the side of the users. 'I investigate, for example, which social barriers software developers encounter when working in the field. I did a survey about the differences that diverse genders encounter. How often did they get interrupted? How often were they excluded from social or networking events? How often did they have the feeling that they were not met with eye contact during a discussion? Micro-inequities might not make you cry, but they affect your confidence and ultimately, your work performance. We expected a gender gap, but it turned out to be much larger than expected.'

At the side of the user, Guzmán Ortega investigated, for example, whether user feedback on certain software applications is culturally influenced and how companies deal with this. Finally, she also studies how ethical concerns can best be integrated in the early development of software.

'In previous groups where I was working', tells Guzmán Ortega, 'I was usually the only one working on the social impact of software. Now I am surrounded by many more people interested in this. And although I do not yet work in the environmental dimension of sustainability myself, in my private life, I care a lot about the environment, so I also really like the fact that other group members concretely work on the environmental impact.'

Dong Nguyen is an assistant professor at the Department of Information and Computing Science at Utrecht University, where she is leading the NLP and Society Lab. She graduated cum laude as a PhD at the University of Twente and was also a research fellow at the Alan Turing Institute in the UK. Last year she won a KNAW Early Career Award.

UNDERSTANDING THE SOCIAL SIDE OF LANGUAGE

Dong Nguyen, assistant professor at Utrecht University, is working on Natural Language Processing (NLP). She wants to improve NLP models, so they become better at understanding the social side of language. This could be used to prevent escalated discussions online.

By Amanda Verdonk
Image Elodie Burrillon

What is Natural Language Processing exactly and what makes it so interesting?
'NLP is about automatically processing and analysing human language. It is used, for instance, in voice assistants, spam filters and translation software. However, NLP models are not yet very good at analysing the social side of language, such as informal language, dialects, the social background of people and developments in society. I'm trying to change that. If we can make models that adapt to the social context of the interaction, they become better at processing, analysing, and generating natural language."

You are especially interested in social media. Why is that?
'Because this is where people with different interests and different backgrounds interact. When do discussions escalate and how can we bring people together? In a new research project funded by NWO, we want to discover subtle patterns in the conversations that you cannot see manually. Apart from online data research, we will also perform lab and field experiments together with sociologists. Perhaps the outcomes of our research could help moderators or lead to an improved design of social media platforms.'

What are the main challenges in your research?
'We are using models to measure patterns in language, but how can we be sure that a model does what it is supposed to do? When you analyse hate speech, does the model really seek for hate speech, or does it wrongly correlate with, for instance, the topic of the message or the people mentioned in the text? It is like Clever Hans, the horse that was believed to do maths by lifting its leg. But later it turned out that he was responding to a signal given by the owner. We should pay more attention to the input data and be more critical on how we validate our models.'

You have worked as an intern at Facebook, Google and Microsoft. What was that like?
'It was very inspiring. The setting is very informal and you immediately get a lot of responsibility. But this was several years ago when people were just very excited by the positive impact of social media. Things might be different now. In my current work, I am also trying to give a lot of freedom and responsibility to the PhDs that I supervise. I have deliberately chosen to work in science because it gives me the most satisfaction. You have more influence and are allowed to be critical. Besides that, I really enjoy being in contact with students.'

MULTIMODAL PUZZLES IN **A NATURAL HISTORY ARCHIVE**

Historian Andreas Weber uses AI techniques to understand the secrets of the natural history archive of the Natuurkundige Commissie (Natural History Committee). 'For computer scientists, this is highly challenging material.'

By Reineke Maschhaupt
Images Collectie Naturalis

In 1820, King William I established the *Natuurkundige Commissie.* Eighteen scientists were sent to the Dutch East Indies to study the local flora and fauna there. That yielded an incredible amount of archive material, which is now stored at Naturalis Biodiversity Center in Leiden. Andreas Weber is assistant professor of Science, Technology and Culture at the University of Twente, where, amongst other things, he is working on the project "Making Sense of Illustrated Handwritten Archives". He has yet to make a spectacular find, such as an extinct animal, says Weber laughing. However, the team did manage to enrich the archive and make it digitally accessible. In collaboration with academic publisher Brill, the entire archive has been available online since the end of November and can be searched by everyone free of charge.

Linking animal to diary

Weber: 'Worldwide, this is the most complete nineteenth-century archive of biodiversity in South-East Asia. It contains 20,000 scanned pages with handwritten texts, illustrations, publications and thousands of objects including stuffed animals, specimens on alcohol, skins, skeletons and eggs.'
'Archives like this are only meaningful if you view them in their multimodality. For example, the nameplate of a stuffed crocodile just states Borneo as its origin. Only after linking the object to the diaries and the field notes, can you discover exactly which river that animal was found in. There once did exist links between the images, text and objects in the archive, but many of those have been lost. We have tried to reconstruct those links with the help of learning and semantic computer techniques.'

Interaction

The project made use of MONK, a handwriting recognition system that actively learns. 'The system divides a text into lines. Then it starts to label the words, interpret these and to give meaning with levels of probability. Twelve volunteers helped to train the system. At a certain point, a self-learning process is initiated. However, people need to continuously point the system in the right direction in that visual labyrinth.'
The team added metadata previously recorded by Naturalis to this handwriting recognition technique. For each scan, somebody has once documented what was on a drawing and who had produced it. Finally, a semantic system was linked to this procedure. Now, if you come across a scientific animal name, the system can link this to the overarching group.

Collaboration

Collaboration between the three disciplines of computer science, biology and history is vital to smartly digitalise such a diverse archive. 'The computer scientists, for example, found it difficult to deal with the fact that scientific names for animals changed over the centuries. Then you need biologists who can look very carefully at these images to identify which animal it is.'
Archives like these still contain a vast amount of hidden information. Together with colleagues, Weber is now working on a large new grant proposal for AI and digital heritage. 'These techniques can also play an important role for other collections. For example, you can link objects in museums with handwritten documents or with ships' passenger lists. We want to make further use of this knowledge.'

The archive is available via: **dh.brill.com/nco/**

Red-throated barbet, drawn by **Pieter van Oort** in Buitenzorg (Java) in 1827.

BETTER BREAST PUMPS AND **SMARTER** BIKES

Several universities and companies, new technology and comprehensive insights too. Many different things come together in projects from the research programme Smart Industry. Two project leaders tell us more.

By Rianne Lindhout
Image WAT ontwerpers

More than half of all Dutch bikes and, as a matter of fact, also more than half of those in Belgium, France and Germany, come from the Accell Group, known from brands such as Batavus and Babboe. The company approached Paul Havinga, Professor of Pervasive Systems at the University of Twente and scientific director of TNO-ICT. They wanted to know whether Havinga could help them to improve the experience of cycling. 'There is a limit to how much you can stay ahead of the competition by making an increasingly better bike.'

That is how the project Smart Connected Bikes started. 'We want a bike that not only monitors the speed and battery level but can also see the surroundings. For example, you can use radar to estimate what other road users will do, or whether something is approaching from behind. It is about improving the comfort and safety of cycling, making it even more healthy for you, and it is about improving the reliability of bikes as well.'

However, the innovations not only benefit cyclists, says Havinga: 'If bikes can communicate, then sensors can be used, for example to monitor the traffic density for the municipality. Furthermore, via the bike, the cyclist can gain information about the air quality, measured by sensors along the route.' The cycling infrastructure can play an important role in this new generation of bikes. 'Examples are adaptive traffic lights and luminous strips on cycle paths.'

No complex display

The project Smart Connected Bikes received a grant from the NWO programme Smart Industry in 2019. Various disciplines are involved in the project. 'Industrial designers from TU Delft are helping to make the embedded intelligence that my department devises easily accessible for the users. We don't want a complex display, but interaction, for example in the form of a handlebar that pivots a bit to the left to indicate the route. We also sought a connection with the Saxion University of Applied Sciences where they have experience with teaching elderly people to cycle again. We will soon be able to test our innovations on these possible users.'

Valuable and scary

However, the research goes further than simply designing smart bikes. 'A lot of data emerges from our innovations. Information about the physical condition of the cyclist, for example, if we can measure the heart rate and oxygen uptake, or how stably a cyclist cycles. This data is relevant for cyclists, but also for a doctor or insurance company. Moreover, data about road and air quality is interesting for the government.'

Such data streams are valuable, but also sensitive, realises Havinga. 'The police, other government parties and the agricultural sector..., everybody wants to use this type of data, but nobody knows how that can be done within the ethical and legal frameworks. We need to find solutions for that.' The professor from the University of Twente already has in mind which technology he wants to use for that: secure multi-party computation (MPC). That enables multiple parties to make use of the data without accessing the data itself. 'Such techniques are being investigated, for example, within healthcare, where patient data can provide deeper insights into general characteristics and disease patterns, without the need for personal data to be shared. For such cryptographic techniques, the models go to the data and not the other way round, which is a safe solution. TNO has a lot of expertise in this area and has been given an important role in the project.'

The perfect size safety shoe

The data issue emerges in a different way at the Smart Industry project of Daan van Eijk, Professor of Applied Ergonomics and

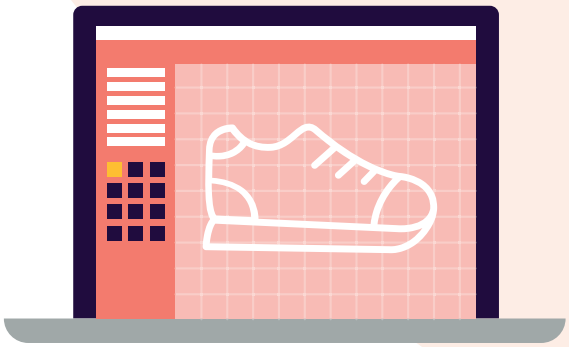
Design at TU Delft. Since 2016, he has been working on ultra-personalised products and services, also referred to as Next UPPS. The initial question was: how can we use newly-accepted technologies, such as 3D printers and cloud services, to provide more customised products and services?

'For instance, we are working with BATA on safety shoes. Currently, those shoes are still made on standard lasts with standard length and breadth measurements. But if you can produce a 3D scan of someone's foot, then you can provide a far more tailor-made product.' Van Eijk is also seeking such a solution for work clothing, such as uniforms for ambulance personnel. 'The mass production method is that a knife cuts 50 jackets in one go from 50 pieces of material. Somebody who is tall and thin often ends up with clothing that does not fit well. A more customised production means that everybody can choose extra compartments or pockets of the size they want and at the place they want.' The workwear company HAVEP from Goirle is the project partner and wants to innovate in this area.

Taking your size

Just as for Smart Connected Bikes, practical questions from industry were the starting point for the research. However, the researchers soon came up against broader issues. For a tailored uniform all of your body measurements, or even a complete 3D image of you, is placed in the cloud. Is that secure? And something or somebody needs to 'take your size'. That is especially sensitive in the case of the tailor-made breast pump that industrial partner Phillips wants to develop. Van Eijk: 'We can develop a method that makes it possible to ensure that such a pump fits perfectly, but perhaps women do not want their breasts to be measured.' Another issue that Van Eijk wants to focus on is the question: is it sustainable? 'A built-to-measure office chair sounds fantastic until the employee for whom it is made leaves and somebody else needs to sit on it.'

These are complex issues that Van Eijk and his partners have not yet fully solved. Nevertheless, the trend towards personalised design is irreversible: 'Take, for example, Blendle and Netflix: in the digital world, we are already used to customised products. We will ensure that this will become a reality in the physical world too.'

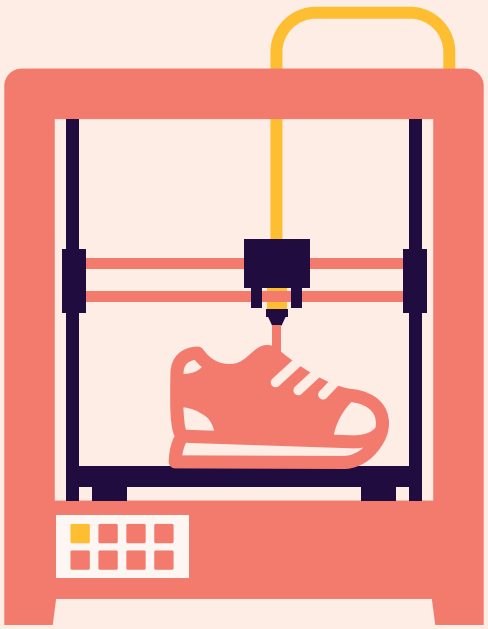


NWO PROGRAMME SMART INDUSTRY

In the programme Smart Industry, researchers and companies are working together on themes in which big data, smart industry and creative industry intersect. The industry concerned co-funds the project. The programme's multidisciplinary research subjects lie at the interface of personalisation, big data, autonomous systems, smart systems and smart equipment.

https://www.nwo.nl/en/researchprogrammes/creatieve-industrie/smart-industry

'If you can produce a 3D scan of someone's foot, then you can provide a far more tailor-made product'



CRYPTOLOGY

OUTSMARTING QUANTUM ATTACKS

By Amanda Verdonk

When large, universal quantum computers are built, our online communication will no longer be safe. Post-quantum cryptographers are putting a lot of effort into building a new kind of quantum-safe encryption. Peter Schwabe, Professor of Cryptographic Engineering at the Radboud University, explains the ins and outs of this young research field.

'Encryption existed before the computer age. Military and lovers used it to secretly communicate with each other. The use of encryption changed dramatically when the internet came. People could agree on a keyword when they met in person, but how can we agree on a key when we use an untrusted channel like the internet? This issue was solved in 1976 in a paper by Diffie and Hellman, who introduced public-key cryptography. It has two parts: public encryption and private decryption.

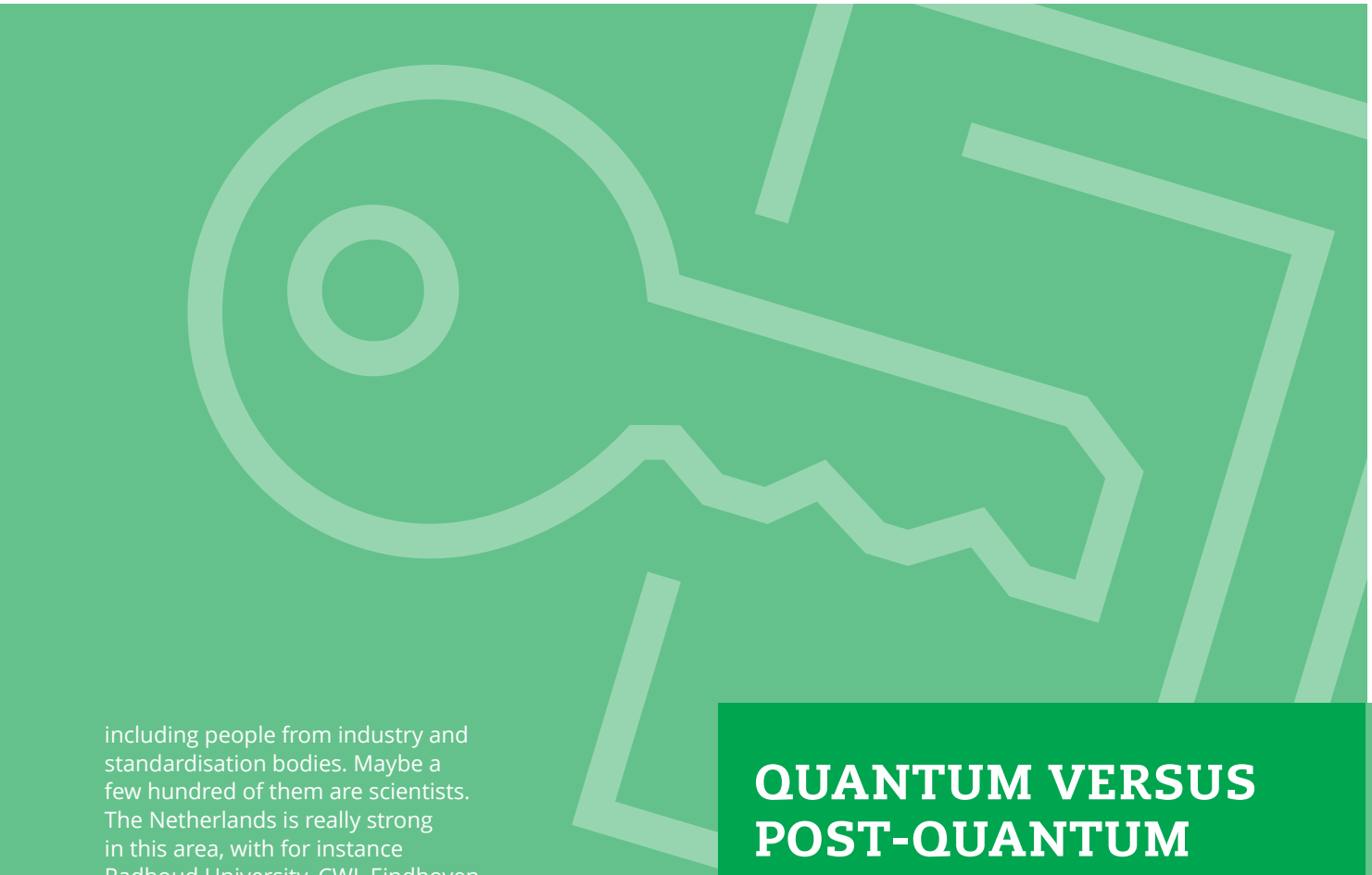
Human rights

'Public-key cryptography is still in use today and is based on two mathematical problems: factoring large numbers and the discrete logarithmic problem. Computers today cannot solve these problems. And though a few small and specialised quantum computers exist already, they are not able to crack current cryptography either. However, if physicists and engineers manage to build a large and universal quantum computer, these two mathematical problems could be solved, and our communication would no longer be secure. For a long time, many people believed such a computer would not be built in the foreseeable future, but this has changed. Big companies such as Google, IBM and Microsoft, and also the EU, are now heavily investing in quantum computers. And the secret service would be one of the first users for sure. So now the concern is growing that, in twenty years' time perhaps, all trusted digital infrastructure will be broken. In fact, this could mean that a conversation that we have today may already be insecure. Imagine a Chinese dissident communicating with a human rights activist. If this encrypted communication is intercepted by a secret service today, it could be decrypted in the future.'

Strong position

'This is where post-quantum cryptography comes in. The idea is that we replace public-key cryptography with other mathematical problems that we believe remain hard to solve, even when faced with the power of a large universal quantum computer. Although they are often presented as such, quantum computers are not magical devices that make every computation faster. They are developed for specific problems. We are now working on five different directions for encryption replacement with their own underlying mathematical problems. As a generalist within this field, I am working on all five directions.'

'Academic cryptography is not a big field. The international conference on Post-Quantum Cryptography in Utrecht in 2018 had about three hundred and fifty participants,



including people from industry and standardisation bodies. Maybe a few hundred of them are scientists. The Netherlands is really strong in this area, with for instance Radboud University, CWI, Eindhoven University of Technology, and also Philips and NXP. For me, the biggest challenge today is how to safeguard the secure implementation of Post-Quantum Cryptography on a device, for instance, a laptop or banking card. An attacker could measure the power consumption or electromagnetic radiation of the device while it's doing encryption. From this data, you might be able to extract

'If physicists and engineers manage to build a universal quantum computer, our communication would no longer be secure'

the key. Our research field is relatively young, so understanding these kinds of attacks will definitely keep us busy over the next years. Moreover, you might be using a bit of Post-Quantum Cryptography already today; Google has implemented it in its Chrome browser. This is partly for marketing, but also to be ready for future developments.'

QUANTUM VERSUS POST-QUANTUM

'Post-quantum cryptography has nothing to do with quantum cryptography', stresses Schwabe, 'except for occasionally competing for funding.' Although the name might suggest otherwise, Post-Quantum Computing (PQC) is also not the technique that follows after Quantum Computing (QC). Both techniques are currently being developed. Quantum cryptography has nothing to do with a quantum computer either; it just uses the laws of physics and the principles of quantum mechanics to enhance security and detect whether a third party is trying to intercept communication. Quantum encryption is based on the observer effect, which states that it is impossible to identify the location of a particle without changing that particle. One of the applications of Quantum Cryptography is Quantum Key Distribution, which uses light particles (photons) to transfer data instead of bits. These photons cannot be altered or cloned. PQC is specifically aimed at quantum computers, while QC could be used independent of the underlying computer infrastructure.

SURF AS INNOVATION BROKER

Text Reineke Maschhaupt Photo Ivar Pel

Ron Augustus wanted to contribute more to society. He recently made the switch from Microsoft Netherlands to SURF, where he has been Chief Innovation Officer (CINO) since September 2020. 'My question to ICT researchers is: how can we together do more for innovation?'

'Last October, all units of SURF were merged into a single organisation. Now we can more precisely determine where our focus should be. How can we ensure that the Netherlands once again progresses faster in the international economy? And how can we deploy knowledge from ICT research for this? In the 1990s, SURF played an important role in the emergence of the Internet. And the new supercomputer that we will install in 2021 is once again state-of-the-art. Money is increasingly important in maintaining that innovative position. If we want to achieve exascale computing capacity, then we will need to invest in international public networks and computer facilities.

With my background, I can fulfil a bridging role between the public and the world of commerce. There is much discussion in society about public values, such as privacy and the emerging power of large players like Microsoft, Google and Amazon. That sometimes leads to polarisation. There is a lot of uncertainty among our members about how that data can be stored in the public cloud. But you can also make agreements about that with public parties. I think that many benefits will be gained by bringing those two worlds together. We can make use of the innovations from companies, under certain conditions, of course.

DIGITAL DATA PLATFORM

SURF can act as a broker with respect to data management from research. Researchers can then use the national facilities of SURF, but they can also switch to the public cloud if that is necessary. I want to make this as easy as possible for them. I also see a role for SURF as an innovation broker in the storage and sharing of research via a digital platform. I would like to call upon ICT researchers to help us think about the form such a digital platform could take.

During the COVID-19 crisis, it has transpired that communal research capacities could be very important. At present, it is still difficult to offer a researcher a quick research pop-up environment. Our goal is to be able to facilitate that quickly using modern resources. Currently, France, the UK, Germany and even Belgium and Finland are progressing faster in that development. With my contribution, I want to ensure that the Netherlands will be a frontrunner once again.'

Ron Augustus has been Chief Innovation Officer (CINO) at SURF since September 2020. Before that, he was Director Customer Success Unit at Microsoft Netherlands. He studied Management Business Administration with a minor in Information Technology at Erasmus University Rotterdam. In his spare time, he teaches courses in art history and the history of science.