I/O magazine
ICT RESEARCH PLATFORM NEDERLAND
VOLUME 20 JULY 2023 NO 2

ICT.OPEN 2023

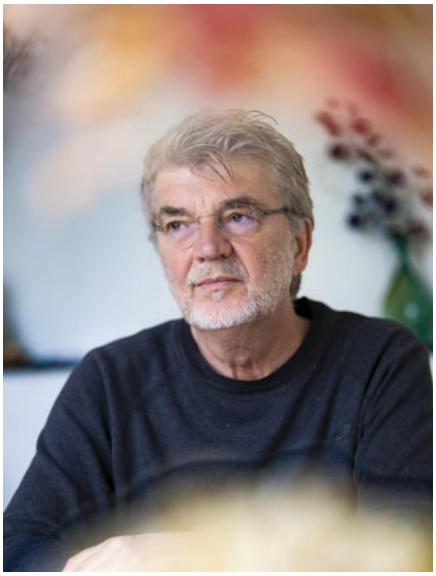
INSIDE THE BUBBLE

12

Looking for smarter ways to test software

Portrait of the Software Testing Group at the Open Universiteit.

16



A remarkable career in a tough field

In conversation with Peter Sloot on the occasion of his retirement.

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; Netherlands eScience 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 with ICT science as their core and 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, Juul Brouwers, Sonja Knols, Femke Stephan **FINAL EDITING** Sonja Knols **TEXT CORRECTION** Dave Thomas, NST Science **CONTRIBUTORS** Marysa van den Berg, Leendert van der Ent, Sonja Knols, Bennie Mols **DESIGN** WAT ontwerpers, Utrecht - Leon Hulst **PHOTOGRAPHY** Thijs ter Hart (cover, p. 4), iStock (p. 9, p. 10, p. 18), Ivar Pel (p. 12, p. 16), WAT ontwerpers (p. 19), Bram Saeys (p. 24)

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

PAUL KLINT

Feed the elephant

Buildings, bikes and Bentleys need maintenance. What about software?

Building new software (or indeed ICT systems in general) is sexy, but maintaining it is not. Maintenance is the proverbial elephant in the room that nobody dares to talk about. During budget discussions, lips remain tightly sealed when it comes to the elephant. Then, “unexpected” maintenance costs are discovered years later, and the budget needs to be revised.

The Dutch government is no exception to this practice. Of the 396 billion euro budget for 2023, roughly 2 to 3 billion will be spent on IT. Half of this budget is labelled “management and maintenance”. Nobody knows the actual annual maintenance costs of government ICT. Making such maintenance costs explicit will require an amendment to the Dutch Government Accounts Act so that a maintenance entry can be given in the government’s annual budget.

A not-so-fun fact: it is common knowledge that the annual maintenance costs of a system are roughly 15 percent of the original costs for creating it.

Meanwhile, the elephant is shaking its head and wonders what all of this has to do with research. Well, dear reader, research needs more and more software that must also be maintained. When you put together the budget for your latest research proposal, could you enter software maintenance as a cost? No! And were there any general provisions for software maintenance? No!

Our elephant is now trumpeting loud and clear: research software must be maintained, and that requires a budget. A clear case in point is the future of “Schoonschip”, the widely used computer algebra system developed by Nobel Prize winner Martin Veltman, which is endangered by a lack of maintenance.

So feed the elephant or research software will no longer run.

NWO

IPN ICT-ONDERZOEK PLATFORM NEDERLAND


ICT!

MASTERING SOFTWARE COMPLEXITY AND QUANTUM COMPUTATIONS

Professors Gail Murphy and Harry Buhrman gave two engaging keynote lectures at ICT.OPEN 2023, on the essential complexities of software engineering and the promise of quantum computing, respectively.

By Bennie Mols
Images Thijs ter Hart

It has taken mankind 3,500 years of ingenuity to go from the first known short wooden footbridge from 1,500 BC to more than a hundred kilometres of steel railroad bridges today. For comparison, it only took mankind sixty years to get from the first simple piece of software for the Manchester Baby computer in 1948 to software with roughly a million lines of code in the first iPhone in 2007.

'Software is often called the most complex thing that humans have ever built', said Gail Murphy, elaborating her comparison of software engineering with the engineering of bridges. 'It is really due to the ingenuity of a lot of the people in the room here, computer scientists, software engineers and hardware engineers, who have figured out how to bring all of the necessary pieces together.'

Fundamental complexities

Gail Murphy is Professor of Computer Science at the University of British Columbia in Vancouver (Canada). She gave a keynote lecture at ICT.OPEN 2023, entitled "Tackling the ever-changing essential complexities of engineering software". Murphy made a warm plea for studying the fundamental complexities of software: 'We focus too much on the building blocks rather than the whole. We should study software more in a holistic, longitudinal and interdisciplinary, both in situ and in its deployment context.'

Murphy built her arguments on a famous 1986 essay by software engineer and Turing Award winner Fred Brooks: "No Silver Bullet – Essence and Accident in Software Engineering". Brooks identified four fundamental complexities in software: complexity itself, changeability, conformity and invisibility. In his essay, Brooks wrote that 'there is no single development, in either technology or management technique, which by itself promises even one order of magnitude improvement within a decade in productivity, in reliability, in simplicity.' In other words: there is no Moore's Law for software as there has been for many decades in hardware.

Gail Murphy:

'We need to consider the build context, the deployment context and the societal context'

Murphy added a fifth fundamental software complexity, namely context, and posed the question of where we stand anno 2023: 'Have all the new developments in the decades since Brooks' essay reduced the essential complexities of building software or not?' Murphy claimed not. She said that developments such as object-oriented programming and verification tools have only addressed accidental complexities. And that also holds for key developments of the past 25 years, such as open source, cloud and automation.

Murphy: 'Open source reduces some development costs but incurs evolution costs and, as a result, doesn't immediately provide an order of magnitude improvement. The cloud reduces development costs of similar systems, but comes with significant adaptation and configuration costs. Automation reduces friction in development and has helped speed up the release of software to users, but automation alone doesn't help determine what system to build or how to design the system.'

The latest development that holds some promise to reduce the complexity of developing software, is the use of large language models to generate software code, design or documentation. Whether or not generation-tools will significantly reduce the effort of building and deploying software, remains to be seen, Murphy said, and she added: 'Or will we just build more complex systems?'

Murphy spent the last part of her keynote lecture arguing for the need to look far more at the whole of software, in its context, rather than continuing to focus on individual components. 'We need to consider the build context, the deployment context and the societal context.'

In the context of building software, supply chains are becoming longer and dependencies can become dangerous, for example leading to security vulnerabilities, Murphy explained. In the deployment context, more research is needed into how developers reason about, explain, grow and verify systems once they are configured and in use. And in the societal context, embedding AI techniques leads to questions around fairness and non-determinism. The use of dynamic instead of fixed data means that software outcomes are different for different users and outcomes will change over time.

Murphy concluded with an appeal to the scientific community as a whole: 'The academic software community and funding agencies need to accept different forms of impact as excellent research, for example by valuing long-term case studies, more study of deployed systems at scale and more integration of research results to solve bigger problems.'

Towards the future

Looking into the future of software engineering, one completely new complexity is on the horizon, which Murphy did not address but was central to another keynote lecture at ICT.OPEN 2023: the complexity that arises when computing is no longer done on classical hardware, using bits, but on quantum hardware, using quantum bits. That was the subject of Harry Buhrman's lecture: "Quantum computing and quantum information". Buhrman leads the Algorithms & Complexity research group at CWI, is founding director of QuSoft, the Dutch research centre for quantum software, and is a professor at the University of Amsterdam.

The state of quantum computing today is comparable to the state of classical computing in the 1960s. Buhrman stated: 'Quantum computing works already at a small scale in research labs around the world, and the first small-scale quantum computers can already be bought commercially.'

But, as Buhrman explained, whether or not the present, rudimentary quantum computers already outperform classical computers on specific computational problems is still being debated. Google made such a claim in 2019 with a 53-qubit system, and Chinese researchers with a 70-qubit photonic system in 2020. However, other researchers said that efficiently configured classical computers can achieve the same.

Buhrman: 'At the moment, we have quantum hardware based on fifty to a hundred qubits, and in the next five years, predictions are a thousand qubits will be reached. That could bring endless new possibilities. However, a big question is for which computational problems

quantum computing leads to a speedup.' One thing is clear: the quantum computer will be a machine for specific computational problems, not a jack-of-all-trades.

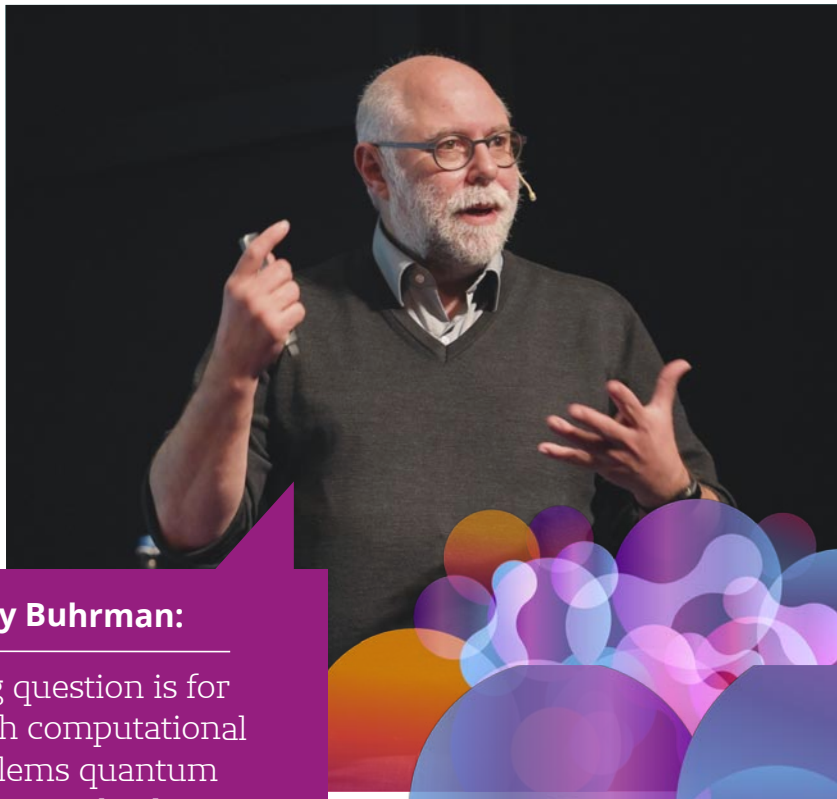
The very first application of quantum computers, discovered by Peter Shor in 1994, is the factorisation of large numbers into their prime factors. That has huge practical implications, because a quantum computer can break globally used digital security systems that are largely based on the assumed hardness of factorising large numbers (RSA cryptography).

On the other hand, Buhrman explained, the laws of quantum physics can also be used to do the opposite: design cryptographic systems that even quantum computers cannot break. And in the field of efficient communication, quantum computing can be used to build a quantum internet which would be much more secure than the classical internet and allow more efficient distributed quantum computing where multiple quantum computers work together to solve complex problems.

Quantum computing also offers exciting possibilities to simulate quantum physical systems and quantum chemistry and thus help design new materials and medicines, optimise logistical problems and perhaps even help accelerate machine learning.

'The question as to which computational problems have a quantum speedup is a difficult and deep one', Buhrman said. 'Roughly speaking, there are three categories: problems with a potential exponential quantum speedup, problems with polynomial quantum speedup and problems with no quantum speedup at all.' Factoring numbers, quantum chemistry and simulations of quantum physical systems belong to the category of problems with a potential exponential quantum speedup. Buhrman: 'These are the first applications of quantum computing that we can expect, as a few hundred quantum bits are already enough.' Computational problems like "backtracking" and "search" fall in the category of polynomial speedup. Due to slow clock speeds, they are not so promising to be solved in the near future. And finally, there are really hard and really easy problems for which the quantum computer gives no speedup at all. Buhrman: 'But the landscape of computational problems is so vast that for many problems, it is not yet known in which category they fall.'

Buhrman concluded that quantum computing is exciting and urgent and holds the promise of a huge economic impact. 'Exciting because as a scientific discipline, it combines two of the greatest revolutions of 20th-century science: quantum physics and computer science. Urgent because medium-sized quantum computers will soon arrive and can be used to address urgent societal needs in drug design, energy and digital security. And economically impactful because quantum computing will fuel a quantum hardware and software industry.'



Harry Buhrman:

'A big question is for which computational problems quantum computing leads to a speedup'

ICT.OPEN 2023

Each year, NWO ICT.OPEN brings together researchers from all ICT research disciplines and industries to learn from each other, exchange ideas, and network. This year ICT.OPEN took place on 19 and 20 April 2023 in the Beatrix Theatre of the Jaarbeurs, Utrecht, and offered a wide range of presentations and demos showcasing the best and most exciting ICT developments. ICT.OPEN 2023 featured four keynote speakers: Gail Murphy from the University of British Columbia (Vancouver, Canada), Harry Buhrman from CWI and QuSoft, Jos De Groot from the Ministry of Economic Affairs & Climate Policy, and Omar Niamut from TNO. Furthermore, there were speed-dating sessions, workshops and, to close the day, an award ceremony.

More information
www.ictopen.nl

Last April, AIVD, TNO and CWI published a handbook for the migration to quantum-secure communications. 'It is vitally important that the Netherlands prepares for the threat of the quantum computer to our secure information and communication. This handbook provides government and industry with important support in this regard', said Alexandra van Huffelen, State Secretary for digitisation, who was handed the first copy.

Making communication more secure

By Sonja Knols

Experts consider the chance small but real that by 2030 quantum computers will be powerful enough to break current cryptographic standards. Therefore, it is important that organisations start migrating to a quantum-safe environment soon. On 4 April 2023, the General Intelligence and Security Service (AIVD), TNO and Centrum Wiskunde & Informatica (CWI) published a handbook for the migration to quantum-secure communications. The handbook is intended for the Dutch government, businesses, vital sectors and knowledge institutions that work with important information that is being encrypted, such as trade secrets.

MIGRATION REQUIRES CUSTOMISATION

The PQC Migration Handbook provides guidance on how to properly prepare and execute the transition to quantum-secure communication using post-quantum cryptography (PQC). PQC is a collection of approved encryption methods that, unlike current encryption, cannot be broken by quantum computers. The handbook helps organisations to identify risks and provides concrete steps to work on a migration strategy. There is not one single strategy for all organisations, as not every organisation has the same interests and ICT structure, for example, because they work with different types of data or have a different degree of confidentiality.

Experts consider the chance small but real that by 2030, quantum computers will be powerful enough to break current cryptographic standards. Therefore, it is important to take appropriate measures for sensitive information in a timely manner. What's more, there are also risks to sensitive information encrypted by current methods. The data currently being sent or stored in encrypted form can be intercepted and deciphered later using a quantum computer. Data that will still be sensitive in 2030 and must remain secret should therefore be encrypted as soon as possible with cryptography that protects against attacks from a quantum computer.

The handbook can be downloaded free of charge from the CWI website: ir.cwi.nl/pub/32988/

Joining defensive forces

'It is great to see that the cooperation between knowledge institutions has led to the publication of this handbook', State Secretary Van Huffelen said when receiving the document. In the field of cybersecurity, collaborations between different actors are far from exceptional. For example, to efficiently and effectively deploy resources from the national government, the business community and knowledge institutions, the various Top Sectors decided to combine their capabilities for cybersecurity innovation in CS4NL. The purpose of CS4NL is to improve cybersecurity awareness and innovation, by bringing together the supply and demand for specific issues, organising an ecosystem that extends across the entire commercialisation chain, and freeing up new financial resources from existing innovation instruments. CS4NL will have a 5-year budget of between 27 and 36 million euros.

More information: dutchdigitaldelta.nl/en/cs4nl

M-GRANT AWARDED TO LUCA LAURENTI

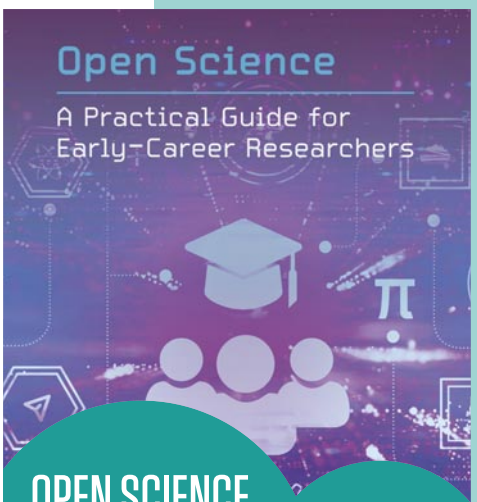
The NWO Science Domain Board has approved 21 grant applications in the Open Competition Domain Science-M programme. One of these is the project from Luca Laurenti (TU Delft), entitled "Formal Verification and Control of Stochastic Dynamical Systems with Bayesian Neural Networks". In this project, Laurenti proposes a shift in deep learning towards Bayesian deep learning to design intelligent systems that use deep learning and have strong correctness guarantees.

M-grants are intended for realising curiosity-driven, fundamental research of high quality and/or scientific urgency. The M-grant offers researchers the possibility to elaborate creative and risky ideas and realise scientific innovations that can form the basis for future research themes.

FOUR NEW COMPUTING TIME PROJECTS

Via the Call Computing Time on National Computer Facilities, researchers can request computing time on the national computer facilities HPC Cloud, Snellius supercomputer, Data Processing (Grid, Spider), and Custom Cloud Solutions. Applications can be submitted on a continuous basis. Researchers can request large amounts of computing time for a project duration of two years via NWO. Small amounts of computing time for a project duration of one year can be requested directly from SURF.

The NWO Committee for the Scientific Use of Supercomputers (WGS) meets six times a year to assess the applications that are complete at that moment. In April, the committee granted four applications, namely:
Dr J. Vreede (University of Amsterdam): 'Rare events in complex molecular systems'
Dr T.A. Hamlin (VU Amsterdam): 'Quantum Theoretical Chemistry'
Dr A.V. Lyulin (Eindhoven University of Technology): 'Nanofiller-enhanced wax for heat storage (WAX+)'
M.A. Khan (TU Delft): 'Numerical modelling of regional-scale wind-farm flow dynamics'



OPEN SCIENCE GUIDE FOR BEGINNERS

What should I pay attention to when it comes to open science? How do I set up my research openly and transparently? Where can I publish? NWO, in collaboration with Universities of the Netherlands, DANS-KNAW and UKB (the partnership of university libraries and KB, National Library of the Netherlands) has published a guide on open science. The guide answers some frequently asked questions that (young) researchers have when getting started with open science.

INTENSIFIED FRENCH/DUTCH COLLABORATION

National Institute for Research in Digital Science and Technology (Inria) and Centrum Wiskunde & Informatica (CWI), the Dutch national research centre for mathematics and computer science, recently signed an agreement to extend their cooperation in the fields of Quantum Computing, Human interaction, Energy, Cryptography, Digital Health, Machine Learning and Software Engineering. New in this collaboration is the creation of European joint research teams.

CWI and Inria have already collaborated successfully over the past decades and have come to know each other as reliable research partners. The more intensive cooperation of both institutes to create a powerful alliance within Europe, as expressed by President Macron during his state visit to the Netherlands, fits in with a joint ambition of both institutes to join forces to cope with major scientific and societal challenges.

A LEARNING ORGANISATION PAR EXCELLENCE

Learning by doing is the central theme in GIPHouse's activities. The student-run company at Radboud University has a spring/bachelor's course for practical code development training and an autumn/master's course aimed at start-up establishment skills. Assistant professor Cynthia Kop of the Software Science group and one of the GIPHouse directors Nick van Oers explain what the company is all about.

By Leendert van der Ent
Image iStock

GIPHouse – Geïntegreerd Practicum (integrated practicum) in House – is a fully student-run, virtual software company. Over 25 years ago, it was the first company of this kind at a Dutch university. GIPHouse places students in a real-life work situation. The company consists of three different courses aimed at separate learning goals. An important aim of the spring double course is to train team effort and code development skills, whereas the autumn course targets establishing a start-up.

The spring course couples five to six bachelor-phase software engineers to one to two master-phase system development managers. Both types of students have their own distinct tasks and educative goals. The team collaborates on a project for a real-life customer, who pays 350 euros for expenses and gets a six months team effort in return.

All customers know they work with students who collaborate on such a project for the first time – and tailor their projects accordingly. Some companies are regular customers that cherish contact with the university. Some of the customers become interested in hiring team members. Cynthia Kop: 'That actually works really well both ways. Certain third-year bachelor students are willing to become part-time employees, while others remember the company once they enter the labour market.'

PROJECT EXAMPLES

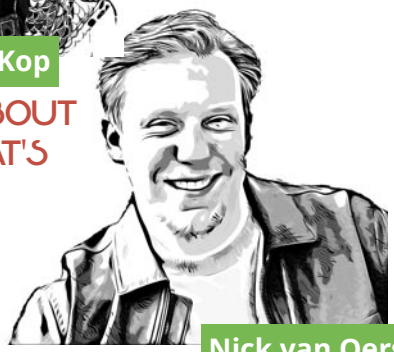
Papua educative app
This GIPHouse project aims to make an app for teaching the language of Bahasa Indonesia to children in West Papua.

Mathematics Scavenger Hunt
Radboud Pre-University College of Science organises a scavenger hunt for candidate science students in Nijmegen city centre, with maths, physics and informatics challenges. The project made a web application to professionalise this scavenger hunt.

Code Flex IT
This project regarded setting up a tracking and tracing system for batches of powder coating. The hardware-software integration with barcode scanners, scales and tablets was an interesting aspect of this case.

Cynthia Kop

'GIPHOUSE IS ABOUT APPLYING WHAT'S LEARNED'



Nick van Oers

'I MAINLY LEARNED TO COLLABORATE WITH COLLEAGUES IN A TEAM'

COLLABORATE, COMMUNICATE

Nick van Oers is one of the current GIPHouse student directors and coaches his fellow software engineering students who take part in the company. Last year he participated in one of the sixteen spring teams as a software engineer: 'We worked on a project for the university itself, and we managed to finish the intended product. It is now being tested as a proof-of-concept. It was exciting to witness what challenges companies experience in practice and what their business is about.' Looking back at the project experience, Van Oers concludes: 'What I mainly learned was to collaborate in a team. The collaboration part was actually more important than the software we built.'

Kop: 'That's exactly how this course is intended. Other courses focus on programming skills; this one is about applying what's learned. Students have to overcome snags, for instance, regarding the Django Python-based web framework. They also have to become familiar with various concepts and work out how they can bring these together in an integrated approach.'

This aspect is combined with collaborating in a scrum setting. 'Communication in a random team is vital for developing an actual product', says Kop. 'Apart from that, it is the process approach that counts. Students learn that a scrum can lead to both efficiency and inefficiency depending on how it is carried out. They learn what works and what doesn't. It is completely different from everything else they learn at university. That's what makes it so useful.'

MASTER'S COURSE

The master's course during the autumn semester revolves around setting up a start-up company. The knowledge transfer department Radboud Innovation is involved in this, for instance, by providing guest lectures and consultancy. The team develops software around a central theme, as a starting point for building a company: How do you build further on a good idea? How can you reach prospects and customers? How do you adapt to comply with interest from the market? Kop: 'I heard that an old GIPHouse team established a start-up last year. There are undoubtedly more cases like that. The course is certainly useful: getting the right image of who customers are and what they want are among the key takeaways. Equally important: real-life conversations with customers. It's also about making beginner's mistakes so that these are already out of the way once a start-up company actually comes into play.'

Looking for smarter ways to test software

By Bennie Mols Images Ivar Pel

GROUP PASSPORT

RESEARCH FIELD

Software testing, automated testing, software quality, education in software testing

INSTITUTION

The Software Testing group is spread over two locations: Open Universiteit (OU) in Heerlen, and Universidad Politécnica de Valencia (UPV).

EMPLOYEES (as of June 2023)

1 professor, 1 assistant professor, 5 students, 1 scientific programmer

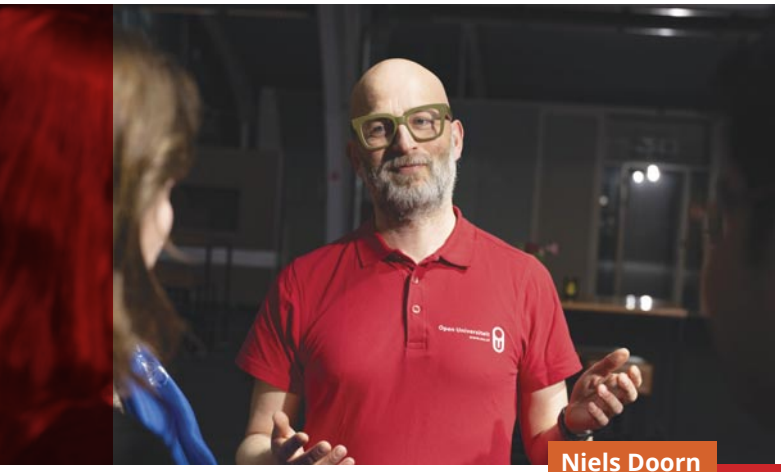
WEBSITES

Tanja Vos: **tanjavos.com**
Open Universiteit (OU): **www.ou.nl/en/onderzoek-informatica**
Universidad Politécnica de Valencia (UPV): **www.upv.es/en**
TESTAR (Automated system testing of desktop, web and mobile applications at the Graphical User Interface Level): **testar.org**
ENACTEST: **enactest-project.eu**

Software testing is paramount for many applications, but because of its difficulty and invisibility, it is not the most popular part of software engineering. Professor Tanja Vos wants to change that by developing more intelligent testing tools and building better educational modules.

Tanja Vos

Beatriz Marín



Niels Doorn

The Software Testing group led by Tanja Vos is unique in spanning two locations in two countries: the Open Universiteit (OU) in Heerlen and the Universidad Politécnica de Valencia (UPV) in Spain. Vos has been an associate professor at UPV since 2003, and in 2016, she also became Professor of Software Engineering at the OU.

'The Open Universiteit has place- and time-independent teaching education at its core', says Vos, 'and then it does not matter whether you collaborate with people working from different locations in the Netherlands or with people working in Valencia. The four PhD students I have at the OU all work from different locations in the Netherlands. In Valencia, there are four people in the group. We communicate with each other in the same way online. We use various chat groups for different purposes and, of course, we meet in person regularly.'

The research group of Vos is specialised in automated testing of desktop, web and mobile applications at the Graphics User Interface (GUI) level. 'Our flagship is a testing tool called TESTAR,' says Vos, 'which we have continuously developed since 2010. TESTAR is a form of scriptless testing in which you do not have to write a test script in advance, but in which TESTAR decides which actions to select to create a test. We construct a very large collection of all possible actions in a given software state, and then we choose one of them to go the next state.'

In its default configuration, TESTAR only does random testing. This may sound strange but in practice, it turns out to be complementary to scripted testing. Vos: 'We have tested software with TESTAR at several companies, including the mobile banking app of ING, and we found that TESTAR covers different parts of the software than human written tests do.'

Still, the question remains whether there are better ways to test than defining a script in advance or choosing random test actions. 'Indeed', says Vos, 'that is why we are investigating if and how we can let TESTAR itself learn the best way to select actions. This involves using techniques from artificial intelligence. Another future research direction is to see whether we can use a ChatGPT type of tool trained on a large number of existing test scripts to predict the best test action to select next. We are constantly looking for smarter ways of testing software.'

Agents and models

Beatriz Marín is one of the senior researchers in the Valencia part of the Software Testing group. She was a professor in Santiago (Chile) for many years, and in 2021, Vos persuaded her to come and join the research group in Valencia as an expert in software quality. 'During the past few years, I have worked on a project that combines the use of intelligent agents and models in testing', says Marín. 'An additional reason Tanja asked me to join the group is that I had experience using serious games to motivate students for software testing. Whether you work in a company or in academia, software testing is often considered difficult, and too few people want to do it. But of course, good testing is extremely important, so our group sees testing and quality as inseparable parts of the education in software engineering.'

Marín has just started working on a new three-year EU-funded project called ENACTEST, to improve the education of testing. The goal of this project is to identify and design seamless teaching materials for testing that are aligned with industry and learning needs. Marín: 'We have four people from our research group working in ENACTEST. Nine different European partners are involved, including universities, vocational centres and small and medium enterprises. In the long run, this project will improve the software quality on which our digitalised society relies.'

Better education

As a PhD student of the Open Universiteit, Niels Doorn is researching how to improve testing education. Doorn works on his doctoral project three days a week; the other two days, he works as a lecturer and researcher in IT and computer science at the NHL Stenden University of Applied Sciences in Emmen. 'The goal of my PhD project is to create effective methods for teaching software testing that are supported by scientific evidence', says Doorn, 'because at present, we know very little about what works well and what not.'

Doorn uses both a theoretical and a practical approach. For example, he compares how test experts test software to how students do that. Doorn: 'What conceptual knowledge and experiences do they use? How can we best integrate testing into programming assignments without compromising the programming concepts we want to teach students? Among other things, we have already created a website with exercises ready for use by teachers.'

Doorn does the day-to-day supervision and most meetings online, but the group members also meet up several times a year. 'In May, we had a computer science day where the group members got together. We conduct workshops for OU students. And we take occasions like the opening of the academic year, an inauguration or a conference to get together for one or two days. Although the group is spread across two countries, the collaborations are going really well.'

Peter M.A. Sloot (1956) was a research professor at the University of Amsterdam and a full professor and director of the Complexity Institute at NTU, Singapore. In 2016, he became the founding scientific director of the Institute of Advanced Study at the University of Amsterdam. He has supervised 57 PhD students and published over 450 research papers.

A REMARKABLE CAREER IN A TOUGH FIELD

After forty years of research on complex adaptive systems, Peter Sloot retired last April. Here, he shares not only a highlight of his career but also his worries about the current hype of using AI for almost every application possible.

By Marysa van den Berg
Image Ivar Pel

How did the research field of complex systems computing emerge?

'Before becoming a professor at the university, I did my PhD at the Netherlands Cancer Institute. There I got interested in complex systems and how to model them. I went to the Santa Fe Institute of Complex Systems in New Mexico (USA), where a lot of the original ideas come from. I think that is the place and time where my colleagues and I kick-started the field of large-scale computing of complex systems.'

You focused on how nature and society process information. Why this topic?

'From all the molecules in your body up to all the people in the world, everything is connected and interacting and thereby processing information. The challenge is that these processes act in non-linear ways. Take the brain as an example. It has billions of neurons interacting. At some point, a thought emerges from all those interactions. There is still no good theory to explain this kind of emerging properties in complex systems. It is both difficult and mind-blowing.'

What are you most proud of?

'I am especially happy with the work we did on HIV infections. It took over a decade, but we managed to build multiscale models that capture many different layers of the disease, from genetic and cellular interactions to transmissions across populations, which allowed us to predict outbursts and to assess potential interventions.'

Where will this kind of research go?

'Predicting the outcome of interventions in complex systems is becoming increasingly important. Think about climate change, geopolitical stress, or the consequences of social inequality. We desperately need to develop tools to study this. I see two main developments. The first is the rise in using AI. There are now top groups that can learn partial differential equations from a black box of data. That is pretty spectacular. Second, there are novel developments in algorithms for individual-based models, that can incorporate elements of human psychology and social behaviour. The merging of the two is the future, I think.'

Do you think the use of AI could also become problematic?

'I am indeed worried about this. In machine learning, one works, by definition, with data about processes that have already happened. But data is only a predictor for the future if the dynamics of the past are predominantly the same as those of the future. This is definitely not the case for interventions in complex systems. Also, most complex systems are open and scale-free systems, meaning they may have infinite variance and zero or infinite averages. So how are you going to train your machine model, then? This is but one of the many challenges ahead that require careful and conscientious research.'

PREDICTING THE EFFECTS OF MEDICATION

By Bennie Mols Image iStock

In her NWO Vidi project "Mechanistic machine learning" Kerstin Bunte combines the explanatory power of expert models with the predictive power of machine learning for applications in personalised medicine.

If you find yourself in the unfortunate circumstance of suffering from a rare disease, then you have no use for modern machine-learning computer programs that need a lot of training data before they can help you. 'It's horrible having to tell people that they can't benefit from computer-aided technology because their disease is too rare', says Kerstin Bunte, associate professor in computer science at the University of Groningen.

This is one of the reasons that in her NWO Vidi project "Mechanistic machine learning", Bunte is investigating ways to combine models that explicitly incorporate expert knowledge with machine learning models. Expert models offer advantages like interpretability, explainability and the ability to handle small datasets, whereas machine learning models offer general adaptability and scalability, but have difficulties with small and dynamic data.

Accumulated knowledge

'Furthermore,' Bunte says, 'purely data-driven, deep-learning approaches often ignore the knowledge accumulated over decades for certain problems. It really just feels wrong to throw away all that knowledge instead of incorporating it in the learning.' Another reason for her approach is to make machine-learning-based technology more transparent.

'Our current focus is the application domain of personalised medicine for treatments of the hormone system,' says Bunte. 'So we are trying to find out what medicine works best for which individual person.' As they are more or less halfway the project, Bunte and her colleagues are preparing two scientific papers for submission.

'As for the machine learning part', says Bunte, 'we use a probabilistic model and not, like is very popular nowadays, a deep-learning model. That is because we use a probability distribution of what is going on in the body at the mechanistic level as input data. Typically not all the data that you ideally need to build a realistic model are available, and you will never even be able to measure some parameters like the total amount of enzyme the body has in the hormone system.'

Good enough

For their papers, Bunte and her colleagues have looked at how far they can come with the mechanistic model, given that you never get all the data that would fully determine the model. Bunte: 'Our conclusion is that you don't need a perfectly fitting mechanistic model, but that it is good enough if the model, for example, can distinguish one disease from the other or if it can help a doctor to choose the best treatment.'

Bunte's research is at the point of formulating the necessary underlying theory. When these techniques will land in the clinic remains to be seen, she says: 'We are working with applied mathematicians and engineers who do biological and pharmacokinetic modelling. They, in turn, work with clinicians. Getting a new tool into clinical applications is a long process, some ten years or so. Meanwhile, we still need to do a lot on the theoretical side at a small scale. But we can already see the big potential. Worldwide the interest of computer science in combining expert knowledge, especially of dynamic systems, with machine learning has grown tremendously in recent years.'

More information **www.cs.rug.nl/~kbunte**

THE NEWS NEEDLE IN THE ONLINE HAYSTACK

By Leendert van der Ent
Images Nadia Bormotova / WAT ontwerpers

Only a remarkably low percentage of the information retrieved by an average person browsing the internet consists of hard news. And it is gathered from a surprisingly wide array of sources. These are some of the outcomes from the Commit2Data project "Inside the filter bubble". Project leader Wouter van Atteveldt from VU Amsterdam and Jisk Attema from the Netherlands eScience Center talk about the project and the added value of their collaboration.

What does modern news consumption in the Netherlands look like? This seemingly simple question is hard to comprehend and also difficult to answer. Project leader Wouter van Atteveldt, Professor of Computational Communication Science & Political Communication at VU Amsterdam, introduces the research topic: 'News consumption has changed dramatically over the last decade. TV news and newspapers used to be the major sources, but now blogs and an extensive array of social media have been added. The technology revolution has changed the role of traditional journalism and created the possibility of filter bubbles and echo chambers to emerge. The very heterogeneity of news sources makes it hard to study the phenomenon.'

Initially, tracking digital traces of the individual news consumption of voluntary test persons seemed the way to go for the project. On second thoughts, although some forms of tracking are allowed, the GDPR regulations together with ethical considerations, were reasons to abandon this idea for now. However, that does not mean that tracking is written off in the future. 'Taking screenshots after a fixed number of minutes could be a viable method,' Van Atteveldt says. 'The ethical issues involved depend on what you do and how you treat the data.'

'News outlets only represent a fraction of the total online news gathering'

The research team decided to go for data donation instead. This leaves test persons in control: they decide which data they wish to share. Doesn't this further increase bias? Van Atteveldt: 'Reckoning with bias is a second nature in social sciences. The major bias here is introduced by the fact that most people with extreme views distrust science and therefore will not participate. We are well aware of this under-representation and other issues that contribute to bias.'

LOW PERCENTAGE

It has recently become apparent that extremists are rarely really caught in a filter bubble. Even they are usually exposed to mainstream news as well. Van Atteveldt: 'People reading no news at all poses a bigger threat to democracy than radicalisation as a consequence of filter bubbles. But that was not the focus of our study. One of the most interesting outcomes is how broad news sources actually are. News outlets in the Netherlands such as NOS and nu.nl only represent a fraction of total news gathering. Hard news represents just a small fraction of people's total browsing experience, but there is a large grey area of news and information on top of that. News also reaches test persons via a wide variety of other apps and websites with blogs, corporate info, opinions and more. It is a much broader phenomenon than was thought beforehand, and that is a positive outcome.'

A follow-up question could be: 'What do people regard as news and do various perceptions create a different bias?' This research question might be answered together with Utrecht University and the University of Amsterdam in a Platform Digital Infrastructure Social Science and Humanities (PDI-SSH) follow-up study.

Besides data donation, an important aspect of the research methodology consisted of automated URL registration plus text analysis through Natural Language Processing approaches designed by Piek Vossen and Anske Fokkens. For software development within the project, Van Atteveldt closely collaborated with Research Software Engineers from the Netherlands eScience Center, established by NWO and SURF.

Van Atteveldt: 'There are several reasons to have a dedicated centre for developing computational software. If software is developed within the framework of a project, there is zero incentive for researchers to make it available to other researchers and thus to support the academic digital infrastructure. The flip side of the coin is equally true: developing this software, however vital it is for a lot of research, doesn't enhance the career perspectives for academic researchers one bit.'

EVER FASTER

Jisk Attema adds: 'At the Netherlands eScience Center, we notice that developing scientific software has become more and more complicated. We enable researchers to make use of the latest technologies. We act as a bridge between scientific users and the software they need, and we encourage them

to get the most out of the digital infrastructure put in place by NWO and SURF.'

SMART USE

Big data has become a household name, but making smart use of it is still a challenge. 'When really big data sets are involved, tools can become very slow,' says Attema from experience. 'Not hardware, but improved and rewritten software can make big data applications exponentially faster overnight. We support researchers with these acceleration capabilities.'

The Netherlands eScience Center's core task is to enable the reuse of research software throughout all science domains, from physics and mathematics to biology, social sciences and the humanities. Attema: 'An example: the Natural Language Processing text analysis tool developed by Vossen and Fokkens can be deployed for legal texts, literature and history. Few people would have guessed it is also used in drug discovery. This shows how the same software can sometimes be applied in wildly different fields.'

The global Findability, Accessibility, Interoperability and Reusability (FAIR) principles for Research Software (FAIR4RS) are valuable to the Netherlands eScience Center. Attema: 'It is not always clear in advance where and how research software might be reused. It is therefore vital to consider the FAIR principles beforehand and enable a broad field of use during development.'

INSIDE THE FILTER BUBBLE

The "Inside the filter bubble" project is being realised under the international Commit2Data programme. This programme focuses on big data research, valorisation and dissemination in areas such as Smart Industry, energy, healthcare and logistics. The project started in 2018 as a cooperation between CWI, University of Amsterdam and VU Amsterdam. It provides a framework for deep semantic analysis of mobile news consumption patterns.

LOGICAL CONNECTIONS

By Sonja Knols

Dirk Heylen

Former head of the Computer Science department and current leader of the Human Media Interaction group at the University of Twente

'At a national level, a series of computer science topics had been chosen that we felt should be reinforced. All of the group leaders from our department came together and discussed which of these topics were most urgent for us. We decided to define positions at the interfaces of existing groups to establish a more coherent research agenda for the department as a whole.

One of the positions we defined aimed to expand our expertise in artificial intelligence, specifically on the knowledge base and reasoning part of it. At the University of Twente, we have many projects related to conversational agents. We put a lot of effort into the graphical design of the user interface, and in collecting data, for example, by automatically capturing and analysing facial expressions and speech to recognise emotions and other social signals. But we had a blank spot when it came to the reasoning component that needs to go in between what an agent recognises and how it should respond.

Birna is now the personal connection between two of our largest projects: the Gravitation programmes "Ethics of Socially Disruptive Technologies" and "Hybrid Intelligence". In addition, she links us to other groups, such as that of one of the other sector plan hires, Giancarlo Guizzardi. It is great to see that our initial idea of using these sector plan funds to create better connections between our different groups is playing out as planned.

The importance of open schemes like the sector plan funds cannot be overstated. At both the national and institutional levels, setting priorities and developing long-term research strategies is crucial. Over the past few years, new appointments were mainly driven by an educational need due to the steep rise in student numbers. And although I think the growth in student numbers for computer science is a positive development, since society is craving for experts in this field, the fact that we had to pull out all the stops to provide our students with the quality of education they deserve has come at a cost. The funds from the second sector plan will hopefully help to reduce the student-staff ratio to more workable proportions so that our scientists can pay more attention to their groundbreaking research.'

The two successive sector plans have resulted in a significant number of new hires at various Dutch universities. Dirk Heylen, former head of the Computer Science department at the University of Twente, reflects on how they decided on where to invest, and associate professor Birna van Riemsdijk explains how her appointment benefits both her own research agenda and that of the university.

M. Birna van Riemsdijk

Associate professor of Intimate Computing at the University of Twente since November 2019

'Ever since I received the Dutch Prize for ICT Research back in 2014, I have been waiting for the perfect occasion to spend the prize money. That time has finally come: the starter pack associated with my appointment here at the University of Twente enabled me to develop an out-of-the-box research idea, which would be fairly impossible to fund by other means. That project called "Designing intimate technology through dance" creates a dance performance with intimate technologies to study questions about what it means to be human in a society where people and digital technologies become intimately connected.

When I received the prize, I was working on digital technologies that take personal values into account. Over the years, that research evolved into an entirely new research area which I call intimate computing. With that term, I refer to digital technologies that, in some sense, come close to us in our daily lives. You can think of things like wearables, high-tech clothing, social media, and smart homes. These technologies collect all kinds of information about you, and by doing so, create new human vulnerabilities. How does being confronted with very personal things like your stress levels or sharing those data with, for example, healthcare workers influence you, your behaviour and your relationships with others? How can we identify these possible vulnerabilities at an early stage of technology development and take them into account in our design? How can we represent users with all their personal and individual features in our computational model in such a way that we do justice to them?

Before coming to Twente, I worked at TU Delft, where the 4TU.NIRICT institute offered support to people who wanted to engage in exchange projects between the four technical universities. Here at the University of Twente, there has always been a strong focus on the interconnectedness of humans and technology. I initially decided to come to Twente for one day a week, but now I work here full-time. Topics like trust between human and machine, ethical aspects of technology development, and cyber-physical systems are well developed here, which means this is the ideal place for me to further develop my research vision. Ultimately, I hope that with my work, I can help build a caring and inspiring digital society that also encourages self-expression.'

DIGITAL HUMANS ARE HERE TO STAY

By Marysa van den Berg Image Bram Saeys

Technology for creating life-like virtual representations is being developed as we speak. And so we can best invest our time and energy in the beneficial applications of this technology for social good, according to Zerrin Yumak, assistant professor at Utrecht University.

'My research is about believable virtual humans and social robots. I use both motion capture and AI techniques to generate facial expressions, gestures and gaze movements using data-driven methods. We aim to push the boundaries of realistic human appearance and expression representation in the 3D digital world, and we are analysing the effect of these representations on the perception of users. As part of this research, we have set up a new AI lab called "Embodied AI Lab for Social Good".

There are worries about the fast development of AI. There was even a call for a six-month pause of Generative AI, particularly for large language models such as ChatGPT. Although I understand the concerns, I do not think such a pause will help solve the problem. Instead, we should focus on improving the regulations and develop new models to catch up with the fast pace of AI developments.

If digital characters and digital worlds are too life-like, that might elicit concerns for certain groups of people, like children and the elderly. We therefore need to carefully analyse and weigh up the pros and cons for each application. For example, the realism requirements for a children's educational game might not need to be that high to achieve the learning goals. On the other hand, a social skills training environment for business might be more effective when facial expressions and behaviours are modelled realistically.'

MEANINGFUL BENEFITS

'There are also discussions about the metaverse, the so-called next version of the internet where we all live and work in a connected 3D digital environment. I do not see one single metaverse platform emerging in the near future or even at all. However, I do think that 3D digital content will be increasingly used in daily life as it has a lot of expressive power. These digital humans can do a lot of the mundane work people are doing at the moment. The benefit is that they are less costly, more scalable and available anytime and anywhere.

Studies have even shown that some people prefer talking to digital humans more than to a real person, since they feel they are not judged and there is no bias. This means we need to think about where the technology could be beneficial and positively impact our lives. And in an EU and Dutch context, we should consider investing more in Generative AI technology and talent development instead of just focusing on the regulatory aspects.'

Zerrin Yumak is an assistant professor at the Human-Centered Computing Group, Department of Information and Computing Sciences at Utrecht University. She obtained her PhD in Computer Science from MIRALab, University of Geneva, and was a scientific collaborator at the HCI Group in EPFL, Switzerland. Subsequently, she was a research fellow at the Institute for Media Innovation, Nanyang Technological University in Singapore.