

In search of well-trodden paths

With an h-index of 85, prof.dr.ir. Wil van der Aalst is the most cited computer scientist of the Netherlands, and belongs to Europe's top 3. And TU/e as a whole is profiting from that: in the Shanghai Ranking for universities, TU/e climbed some seventy spots, all thanks to Van der Aalst's new status as 'ISI Highly-Cited Researcher'. And only recently, the professor of Architecture of Information Systems was the first one ever at TU/e to be elected as a member of Academia Europaea, the European equivalent of KNAW.

The much-cited Van der Aalst is 'home-grown': He graduated and received his PhD in Eindhoven, and has spent most of his working life at TU/e. He's currently studying distributed processes, in which several people or machines influence each other during a process - think of handling a benefit application, an insurance application, or hiring a new employee. "Our research group has three main focuses," says Van der Aalst. "First, there's the mathematical description of process models. TU/e has been doing exceptionally well in that department for

years, and until the mid nineties it's what I was engaged in myself most of the time. After that, my attention shifted to information systems based on these formal process models; you could say it's the engineering aspect of the same design. You're talking work flow systems created by companies like IBM and Oracle, or SAP company software that's used by nearly all major companies."

About ten years ago, Van der Aalst dived into the area for which he and his research group Architecture of Information Systems (AIS) are currently

being applauded for: process mining - the analysis of digital information flows recorded in event logs. Think electronic payments, orders, e-mails, tests in hospitals and exam efforts at TU/e.

"Process mining can also uncover fraud or favoritism"

These information flows show how company processes work in practice and they can be used to improve these processes. "When the dataflow started up, I seized the opportunity to empirically test the theories on company processes. Until then I had only been working with model-based analyses of information systems, and to be honest, I thought they were too subjective. Process mining enables you to see whether process models work properly, and whether people stick to them in practice."

Because according to Van der Aalst, that's the pivot: in many cases, company processes don't run as planned at all. "Based on the data it's possible to form an empirical model that does in fact describe the actual processes. You can demonstrate the differences between what organizations say they do and what they do in reality." He takes out a copy of the book 'Process Mining: Discovery, Conformance and Enhancement of Business Processes' that was published earlier this year and is the very first of its kind. The diagram shows the process of one of the dozens of municipalities Van der Aalst carried out his analyses. Municipalities are known for their high level of conformance - they tend to follow the rules. "Still, it shows that in some respects the actual process differs from the desired process. Cases like these can be analyzed by mapping their social networks in order to unveil fraud or favoritism, for example. Process mining is not exclusively explanatory; it can also be used for checking."

Should the process model that's reconstructed by means of process mining not correspond to the way the process is constructed in theory, it can mean two things, according to Van der Aalst: either reality is faulty - people don't do what they're supposed to - or the procedure's faulty, which may be a reason to change the formal processes based on reality. "It's similar to the desire lanes you see here on campus. They're there because the official route takes you to the exit via a detour. Dwight Eisenhower, who later became president of the United States, had a wonderful solution to that when he had Columbia University campus built. At first, they simply didn't lay any foot paths. It was only after a year that they checked at what locations the ground was trodden bare and that's where they laid the paths."

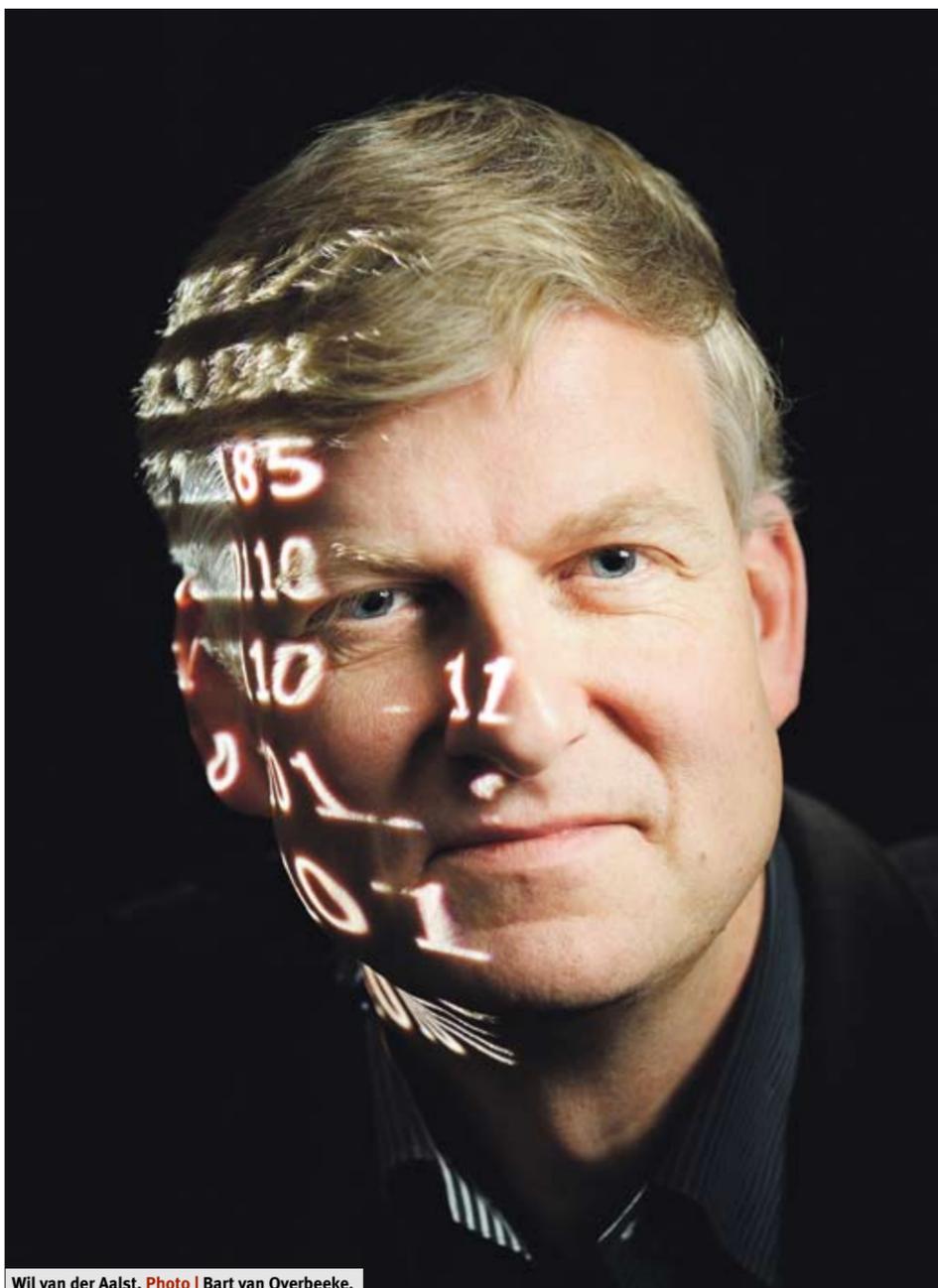
Process mining has rapidly become the third research arm of the AIS group, and it's one that's still maturing.

With his colleagues, 45-year-old Van der Aalst developed open-source software package ProM, which is currently being used by researchers all over the globe. So far, hundreds of tools have been developed for the package by researchers everywhere. "In our field, ProM is by far the most advanced software out there. And because it's open source, any researcher can build on existing technology, and so their contribution can be implemented immediately."

Van der Aalst has also set up two modest spin-offs, Fluxicon and Futura Process Intelligence. The first is a developer of commercial software that enables organizations to analyze their own processes. "The focus here is on user-friendliness", says Van der Aalst. He considers ProM more of a tool for experts, "a kind of Unix" that not just anyone can work with. Future Process Intelligence, the other startup, has been around for a while longer and has been implemented in the Pallas Athena software that's being used by over eighty percent of municipalities right now, for instance. "In five years, I expect there to be at least one hundred commercial tools using our techniques."

"I consider the h-index of 85 my scientific age"

Last month, Van der Aalst was elected as a member of Academia Europaea, and he used to hold a chair at the Royal Holland Society of Sciences and Humanities. What's the secret to his success? "I think my strength lies in the fact that I don't linger until a topic has been flogged to death. As soon as I spot new developments, then that's the direction I take and I make sure I stay the course, even when the road is bumpy." (TJ)



Wil van der Aalst. Photo | Bart van Overbeeke.

The h-index

The h-index, or Hirsch-index, is a way to measure the impact of a scientist's scientific output. The index was proposed in 2005 by physician Jorge E. Hirsch. An h-index of 85 like Van der Aalst's indicates he's published 85 articles that have all been cited at least 85 times. Using the h-index as an indicator for citation impact prevents a researcher from scoring very high with only a single much-cited publication, as well as having someone collect citations by publishing many mediocre articles. The h-index has become quite influential rather quickly, witness the Shanghai ranking for universities. Van der Aalst: "I consider the h-index my scientific age. I can already retire, but my articles on process mining are still too recent to have had any real impact on the index; we'll reap a profit from those in ten years." It seems there's still more to come for the researcher - who's actually 45 years old, so not yet elderly at all.