

CASTOR

Software Research Centre

2019 - 2020

Cover page: Maven dependency tree & DepClean



Spoon 8.0.0 [Core]



Direct and used



Transitive and used



Bloated and transitive



Omitted dependency

DepClean

Bloated dependencies are libraries that the build tool packages with the application's compiled code but that are actually not necessary to build and run the application. This phenomenon artificially grows the size of the built binary and increases maintenance effort.

In Castor, we have developed a tool called DepClean, to analyze the presence of bloated dependencies in Maven artifacts. The dependency tree on the cover is the result of analysing *Spoon 8.0.0 [Core]* project.

Spoon <https://github.com/INRIA/spoon>

DepClean <https://github.com/castor-software/depclean>

Credits

Cover design by *Nadia Campo Woytuk* and *Jonathan Ramírez*

Authors: *Amir Payberah, Benoit Baudry, Cyrille Artho, Deepika Tiwari, Didrik Lundberg, Fatemeh Rahimian, Frédéric Loiret, Ingo Sander, Marco Chiesa, Maria Berthelius, Matthias Becker, Musard Balliu.*

Photo credits: *Moa Hörnquist, Long Zhang, CASTOR team.*

Editors: *Benoit Baudry, Frédéric Loiret, Maria Berthelius.*

This document wraps the second year of CASTOR, which has been marked by some key milestones for the construction of a solid, sustainable collaboration between KTH and industry in the area of software research.

The first edition of the CASTOR Software Days marked a key milestone for the visibility of software research at KTH. The three day event attracted more than 180 participants from the Stockholm IT industry, academia and KTH students. Through a mix of keynotes, technical presentations and lightning talks, we heard about state of the art works in DevOps, Security and Distributed Systems. The attendees met the CASTOR PhD students, who showcased their work in a poster session. Several new collaborations between KTH and the Stockholm IT industry have been initiated as a result of these Software Days.

This year, we saw the first batch of CASTOR Seed Projects grow into larger projects, in collaboration with our core industry partners Ericsson and SAAB. Research grants from Vinnova and WASP consolidate advanced software technology at KTH, in collaboration with industry, in critical areas such as embedded systems for avionics, verification for software-defined

networks and artificial intelligence for edge infrastructure.

This second year of CASTOR has been deeply marked by the loss of our colleague Christian Schulte. Christian has been a remarkable scientific leader for the creation of long-lasting research partnerships with Ericsson, which were instrumental for the creation of CASTOR. Christian was a remarkable researcher and professor, a true geek, a great team player, whom is missed by a large international community as well as his colleagues and students at KTH and in industry.

While students are coming back at KTH and Stockholm is slowly reopening after a safe and quiet spring, we start the third year of CASTOR operations with strong motivations for high impact collaborations and scientific results.

Benoît Baudry,
Director of CASTOR

Follow Us



www.castor.kth.se



twitter.com/CastorKth



[CASTOR KTH on YouTube](https://www.youtube.com/castorkth)



[KTH Software Research Meetup](https://www.meetup.com/castorkth)

In memory of Christian Schulte

"Be kind to the students", Christian said. It was his last sentence to me in the hospital, a few days before life let him down.

He died as he lived, kind and caring. Christian was not only a world class scientist with strong impact and top-notch leadership features, but also, and more importantly, a genuinely compassionate and caring person.

These are perhaps a rare combination of characteristics, which were also topped with a quick-wittedness and never ending stream of inspiration and energy.

Christian Schulte was a professor of computer science at the division Software and Computer Systems (SCS) at KTH.

He started his journey in the academic world by getting a diploma in computer science from the University of Karlsruhe, Germany in 1992, and continuing it by acting as a researcher and project leader at the German Research Center for Artificial Intelligence (DFKI) until 1997, and afterwards as a researcher at Saarland University, Germany until 2002.

He obtained a doctoral degree in engineering in 2001, and joined KTH in 2002. Christian also worked as an expert researcher at RISE SICS from 2010 to 2018.

Together with a few colleagues, he established CASTOR Software Research Center at KTH in 2018. He also served as the secretary and elected member on the executive committee of the Association of Constraint Programming since 2019.

As an outstanding researcher, Christian worked mainly on constraint-based compilation and on models, architectures, and implementation techniques for constraint programming systems.

He liked building software systems with impact.

He led the development of Gecode, as an attempt to construct an open,



free, portable, accessible, and efficient environment for developing constraint-based systems and applications. SAP has chosen to leverage Gecode at the core of its configuration engine, since 2017. Unison is another software developed under Christian's supervision, which is a simple and flexible tool for integrating register allocation and instruction scheduling using constraint programming.

Christian was an excellent teacher and a great advocate for all students.

As the program director of the doctoral program "Information and Communication Technology", he was always an amazing support to young researchers and PhD students.

He could always help you find new perspectives, when you were frustrated.

He could open a new door for you, when you found all other doors closed.

He always brought a sense of happiness, excitement and optimism to every conversation. His brilliant mind as an excellent scientist and a great mentor never disappointed those who asked for his advice.

With him around, there was always a sense of liveliness and anticipation of interesting discussions to take place. He could always find something valuable in everyone, and the care and sympathy he demonstrated for people around him was exemplary.

His great contributions will always be remembered and his thoughts and sentiments will always be treasured.

We shall never forget you, Christian.

Amir Payberah & Fatemeh Rahimian

CASTOR Software Days



On October 12-14, 2019, we ran the first edition of the CASTOR Software Days. Software researchers, engineers and students met to present and discuss cutting-edge technology and current challenges in the three areas of DevOps, Security and Large-Scale Distributed Systems. The days were a mix of keynotes, academic and industry talks, demos and lightning talks which were organized by the KTH CASTOR software research centre.

The event attracted approximately 200 guests, a mix of attendants from the industry and KTH. Among the participating companies were Spotify, ABB, Klarna and iZettle as well as the centre's founding partners Saab and Ericsson. Among the speakers were Dimitra Giannakopoulou from NASA, Polhem Award winner Daniel Stenberg from cURL as well as Sahar Asadi who is working with AI at King. ASeveral KTH researchers spoke at the event, such as Laila Bahri, Senior Researcher at KTH, Elena Troubitsyna, Associate Professor of computer science at KTH and Pontus Johnson, Professor of cyber security and software architecture at KTH.

"Initiatives such as the Software Days are important to give researchers and industry the opportunity to share experiences and hear about new processes, tools and insights in software development. It is important to continue to put Sweden on the map as a country that can develop useful software products, such as Swish and Spotify, that are being used across the world. This is fantastic and shows how competent we are in Sweden despite being such a small country" Ingela Persson, DevOps Team Lead, Saab Surveillance

The full program and the presented slides are available:

www.castor-software-days-2019.github.io



DevOps Day

The DevOps day covered a series of topics related to automation of software development. The two keynotes provided first-hand insights into two iconic open source software projects that support automation: Daniel Stenberg opened the day with a keynote about curl, an essential automation technology for data transfer since 1997; Josh Kalderimis gave the second keynote about Travis-CI, the most popular solution for build automation of open source software.

All speakers gave technical insights about the latest works performed within the CASTOR center, both from the academic and the industrial perspective. These insights were complemented by technical presentations from key DevOps players in the Stockholm software industry (e.g., Spotify and Combient).

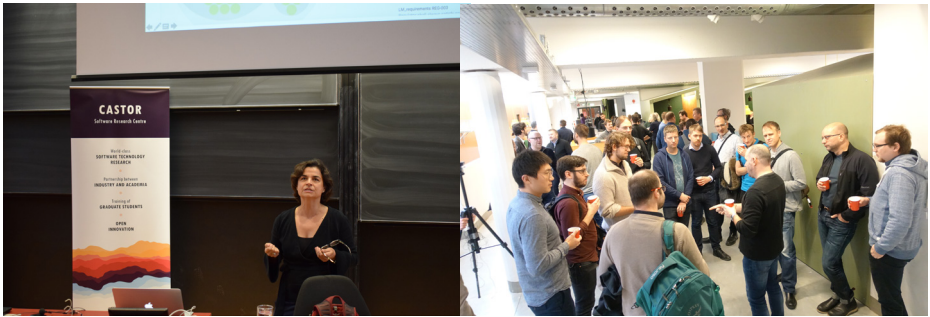


Security & Safety Day

The Security and Safety track covered a wide array of topics ranging from well-oiled areas like security protocols, security-by-design, attack simulations, binary code verification, and safety-critical systems to recently emerging domains like autonomous systems, Internet of Things, and blockchain.

The event benefited from a diverse combination of presentations from world-leading academics and industry experts, including exciting research within our center. Our first keynote speaker, Professor David Basin from ETH Zurich, presented recent advances in formal verification of the protocols driving the security of Internet. Dimitra Giannakopoulou, our second keynote speaker from NASA, discussed security and safety challenges in autonomous systems with compelling applications to autonomous drones and next-generation aircrafts.

Remarkably, the event attracted a large number of undergraduate students, as well as participants with different backgrounds. Active participation, lively discussions, business card exchanges, and happy faces were a clear indicator that event exceeded the expectations. As one undergraduate student put it: *"I joined out of curiosity and left in full inspiration"*. In the spirit of science, inspiration, and creativity, the event closed with a classical electroacoustic concert on transformation of music into data and data into music, by the Royal College of Music.



Distributed Systems Day

The third day of Castor Software Days was dedicated to Distributed Systems and AI. We were honored to host two keynotes and six regular talks. The first keynote in the morning was by Professor Peter Pietzuch from Imperial College London, who gave a talk about his recent research on scaling deep learning over multi GPU servers. His talk was followed by three regular talks by Sahar Asadi from King, Paris Carbone from RISE, and Prem Jonnalagadda from Barefoot Networks.

We resumed the program by the presentation of the second keynote speaker, Ziawasch Abedjan, an Assistant Professor from TU-Berlin. Ziawasch presented an interesting holistic approach for effective error detection. We continued the plan with three regular talks by Sina Sheikholeslami from KTH, Amira Soleiman from RISE, and Wolfgang John from Ericsson.

All in all it was an amazing day full of interesting discussions, which ended by a LLVM meetup organized by Roberto Castañeda Lozano and Diana Picus about keeping code clean with clang-based tools. The whole day was organized by Amir H. Payberah and Philipp Haller.

The Continuous Integration Art Hackathon

On October 14, 2019, we had the first continuous-integration art hackathon in reaktorhallen, at the KTH Royal Institute of Technology, in the decommissioned nuclear reactor R1. The hackathon gathered 60 software students, software developers, artists, designers and curious citizens participated in the first ever performance of software art based on continuous integration data.

From 18h30 to 22h00, 23 teams delved into the extraordinary activity that occurs in a continuous integration server through visual and sound representations. The participants developed their CI-art pieces, with the same constraint for all: use the stream of continuous integration builds of Travis CI. They could use sound, web-based visuals and 24 light fixtures displayed on the back wall of reaktorhallen.

At 22h00, we started an extraordinary series of 18 live performances of CI art pieces. All of them embracing the live, dense stream of continuous integration builds. All teams explored software evolution in creative and unique ways, taking full advantage of the various displays set in reaktorhallen for that special occasion. This resulted in 18 original pieces of software art that unveil the high-frequency activity of the open source communities that build the infrastructure for the world-wide web.

Four prizes were awarded by Josh Kalderimis, CTO of Travis.

An archive of that day is available:
www.hackathon.castor.kth.se/2019



CASTOR Timeline 2019-2020

Here we highlight some of CASTOR's most exciting events and milestones for the year 2019-2020.



1- <https://www.castor.kth.se/project/paus-exchange/>

2- <https://www.ieee-security.org/TC/EuroSP2019/>

3- <https://castor-software-days-2019.github.io/>

4- <https://www.castor.kth.se/2019/11/13/distinguished-artifact-award/>

5- <https://www.chaos.conf.kth.se/>

6- <https://www.svt.se/nyheter/inrikes/de-lyfter-sveriges-ai-satsning>

7- <https://www.vinnova.se/en/p/transform---design-transformation-for-correct-by-construction-design-methodology/>

8- <https://www.castor.kth.se/adl-team/matthias-becker/>

9- <https://www.castor.kth.se/2020/03/17/president-of-kth-sigbritt-karlsson-visited-castor/>

10- <https://www.castor.kth.se/2020/02/21/cooperation-between-castor-and-cdis-centres-on-3-phd-thesis/>

11- <https://www.swedsoft.se/en/2020/02/13/kth-joins-swedsoft/>

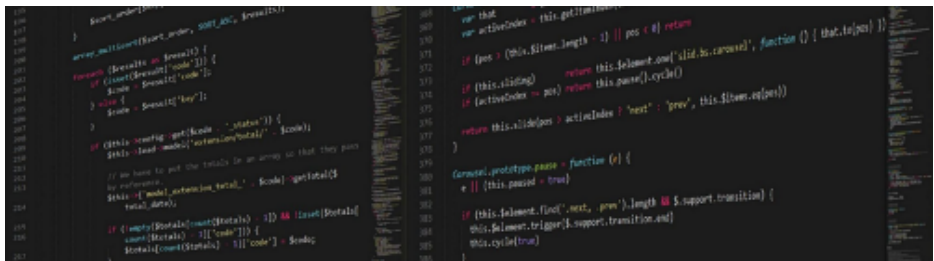
12- <https://www.castor.kth.se/event/online-software-art-exhibition-on-browser-fingerprinting/>

13- <https://www.castor.kth.se/adl-team/marco-chiesa/>

Software technology research is an essential part of WASP, which provides technical foundations to experiment with AI and autonomous systems.

In the last two years, the CASTOR faculty members have successfully acquired WASP funding to consolidate software research at KTH. CASTOR faculty received a total of 14 PhD grants. These grants are evenly distributed between academic and industry PhD grants. The PhD topics span the areas of automated software engineering, programming languages, security and distributed systems and support the research activities of eight faculty members in CASTOR. These grants are provided as part of both WASP-AS and WASP-AI. In addition to PhD grants, four postdocs, as well as two professorships in software technology are supported by WASP.

More about WASP: www.wasp-sweden.org.



The WASP Research Arena – Software

The WARA for Software (WARA-SW) is a research arena that aims to enable software research collaboration between WASP's academic and industrial partners, engaging both senior researchers and WASP PhD students.

The WARA-SW addresses two concerns at the same time – to help PhD students conduct better-evaluated research closer to the needs of the industry, and to shorten the distance between the research conducted at the WASP universities and the software engineering activities at WASP partner companies.

The WASP PhD students in CASTOR contribute to the WARA-SW with tools that result from their research, as well as with datasets resulting from their experiments. WARA-SW shall be a solid platform to collect information about software projects that are of interest for the CASTOR industry partners who are also strongly involved in WASP.

More information

On the WASP website:

www.wasp-sweden.org/research/research-arenas/wara-sw

On Github:

www.github.com/wasp-sweden/wara-sw

News from Seed Projects

The purpose of the seed projects is to conduct pre-studies that have the potential to lead to larger integrated projects part of national or international research projects. All seed projects fall under the CASTOR joint agreement concerning intellectual property rights.

Below are the reports regarding follow-up activities after two seed projects that have been initiated last year.

TRANSFORM

***With Ingo Sander, Dilian Gurov, KTH,
Ingemar Söderquist, Saab***

The new Vinnova NFFP7 project TRANSFORM (2019-2022) with KTH and Saab as partners is a direct result of a previous CASTOR seed project on *Correct-by-Construction design*.

TRANSFORM establishes a new link between research on embedded system design (Ingo Sander, KTH), formal methods (Dilian Gurov, KTH) and avionics industry (Ingemar Söderquist, Saab). The project aims at establishing a transformation-based refinement flow as part of a correct-by-construction design methodology, by transforming an abstract, formal and executable system model by stepwise application of formal transformation rules into a physical implementation on a networked avionics platform.

The project is an excellent example for the close cooperation between KTH and Saab, which currently comprises one international (PANORAMA, ITEA3, 2019-2022), and two national Vinnova projects (TRANSFORM and CORRECT, 2018-2022).

In total four PhD students within the group of Ingo Sander are supported by these projects.

Contact: Ingo Sander, ingo@kth.se

More info: <https://www.vinnova.se/en/p/transform---design-transformation-for-correct-by-construction-design-methodology/>

PAUS Program

***With Cyrille Artho, KTH,
Stefan Hagdahl, Saab***

Creating reliable networked software and protocols has remained a challenge to this day. Traditional network hardware is designed for a fixed topology, with fixed protocols captured in black boxes. New hardware uses so-called software-defined networking (SDN), open design and white-boxing, where the hardware can be configured with different software to allow for different use cases. It is crucial that the network software is correct, especially when the network is used for mission-critical operations.

Different software and protocol analysis approaches exist, each with different trade-offs. The exchange funded by the PAUS program, running from September 2019 through June 2020, was an ideal opportunity to study these trade-offs.

During the exchange, Cyrille Artho (KTH) studied SAAB's needs and tool chains, and collaborated with SAAB on several project proposals and a pilot case study.

The exchange triggered a joint workshop between SAAB and CASTOR, and also resulted in a tutorial given at SAAB by Cyrille Artho and a guest lecture in Kista by a senior SAAB employee.

Because SAAB is both a strategic partner of KTH and also a member of the CASTOR research center, to which Cyrille Art also belongs, this exchange also became the perfect starting point for a long-term research collaboration, with CASTOR extending the initial funding from PAUS (which lasted through February 2020) through June 2020.

Contact: Cyrille Artho, artho@kth.se

More info: <https://www.castor.kth.se/project/paus-exchange/>

CASTOR Portraits

Let's give space for our members who have been actively involved in the CASTOR center since its inception, and for those who have been recently joining us.

Deepika Tiwari | PhD student

Background

I've done my schooling, Bachelor's, and Master's in New Delhi, India, and only just moved to Stockholm last year to join a wonderful group of PhD students, professors, and research engineers at KTH. I started my PhD in January this year, the goal of which is to automatically generate test cases based on the production traces of an application. I'm excited about source-code analysis, code instrumentation, and application monitoring.



Current Research

I'm currently working on static source-code analysis and bytecode analysis to identify a set of methods in an application that meet certain criteria. I will then instrument these methods to add observability to the application. Somewhere down the line, executing these applications with production-like data would help us generate oracles.

Open-source Tools and Technologies

At our lab, we're big on open-science and open-source.

I mostly work with Java, and am currently working a lot more with code analysis and instrumentation libraries like Spoon and Javassist, among others. The tool I'm currently implementing is also completely open-source and can be found here: <https://github.com/castor-software/pankti>.

Didrik Lundberg | New PhD student

Background

My background is right here at KTH, where I have obtained a bachelor's degree in engineering physics as well as a master's degree in computer science (with specialisation in theoretical computer science).

Current Research

In my research, I am specialising in low-level software verification and interactive theorem proving. The collaboration between Saab and KTH consists of the

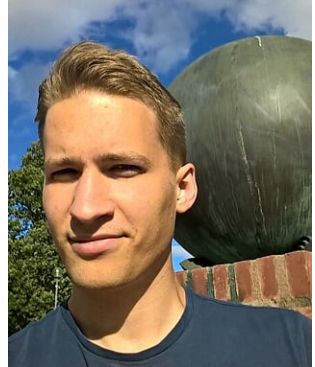
joint funding of my PhD degree (which is a “one-off” in contrast to the WASP programme), which is very similar to that of other Swedish industrial PhDs. 20% of my time - the time regular PhDs spend teaching - is reserved for Saab, which might be allocated to Saab-internal tasks as well as specific research.

In addition to this, the direction of my research as a whole is co-formulated together with Saab.

I don't sit at KTH all of the time like other PhD students, but sometimes I work in Saab's offices, depending on what I'm doing at the moment. I am currently contributing to the HolBA toolbox for formal verification of binary code, and in the near future, I will be working on P4 verification.

During my time at KTH, I have been very engaged in extracurricular activities, such as being the master of ceremonies for the Physics chapter as well as organizing a cruise for 2000 students from various Swedish universities.

Finally, here is my website: <https://csc.kth.se/~didrik/>



Amir H. Payberah | Assistant Professor

Amir H. Payberah is an assistant professor of computer science at the division of Software and Computer System (SCS) of KTH.

He is also a member of the Distributed Computing at KTH (DC@KTH) and the Center on Advanced Software Technology Research (CASTOR).

Prior to that, he was a machine learning scientist at University of Oxford (2017-2018), and a senior researcher at the RISE SICS (2013-2017). He got his PhD from KTH in 2013.



Current Research

Amir's research interests include distributed systems, scalable learning, and data intensive computing platforms.

Amir is responsible for two second-cycle courses at KTH, "Data Intensive Computing", and "Scalable Machine Learning and Deep Learning".

Together with his PhD student, he is currently working on three projects on three different topics including (i) parallelizing deep learning, (ii) multi-modal explainability, and (iii) generalization in reinforcement learning.

Amir's website: <https://payberah.github.io/>

Matthias Becker | New Assistant Professor

Matthias Becker is Assistant Professor at KTH Royal Institute of Technology since November 2019.

Background

He received his Licentiate and PhD degree in Computer Science and Engineering from Mälardalen University in 2015 and 2017 respectively. There he worked on the consolidation of automotive real-time applications on clustered many-core platforms.

During his PhD, Matthias has been a visiting researcher at CISTER - Research Centre in Real-Time and Embedded Computing Systems in Porto, Portugal in 2015 and 2016 (two and three months respectively). Before becoming an Assistant Professor, Matthias has been postdoctoral researcher at KTH during 2018 - 2019.



Current Research

Matthias' research interests are in the area of real-time and embedded systems. For such systems, not only the correctness of computation is important, but also the time at which the computation is performed. With this, they are at the heart of many embedded and safety critical systems all around us. In this context, he is focusing on predictable execution of time-critical applications on multi- and many-core architectures.

As well as on the timing analysis of end-to-end delays, i.e. the delays data experiences when being propagated through software systems.

In this line of research, the tool MECHAniSer (www.mechaniser.com) was developed. MECHAniSer aims to support the development of complex real-time systems that are subject to end-to-end delays by providing insights into the system performance already at early development stages, as well as guidelines

for their later implementation such that end-to-end delay constraints are met. Finally, his online profile can be found here:
<https://www.kth.se/profile/mabecker/>

Marco Chiesa | New Assistant Professor

My name is Marco Chiesa and I am an Assistant Professor in the Network System Lab group of the Communication Systems department at KTH.

My research interests lie on computer networking and, more specifically, on making Internet communication as efficient, robust, and secure as possible.

Prior to joining KTH, I was a postdoctoral researcher in Belgium and Israel as well as a visiting scholar at UC Berkeley (US). I received my Ph.D. from Roma Tre University back in 2014.



I am currently involved in a variety of exciting research projects revolving around the timely problem of building highly performant and resilient Internet. My research builds upon the emerging network programmability paradigms including Software Defined Networking (SDN) and the P4 network programming language. We used such paradigms and related new technologies to recently design and build a high-speed load balancer for Internet services that is 100 times more efficient in processing packets than existing software-based load balancers. We make all our code available to our research community at <https://github.com/cheetahlb/>.

I have recently been fortunate to receive two essential sources of funding for conducting research and building community platforms at the intersection between the areas of computer networks and machine learning. In the SE-CAID Vinnova-funded project, together with other partners, we are building a national digitalization data lab that broadly supports industrial and academic R&D, allowing for industries, SMEs and academia to share a broad range of telecom data and AI models. In the EMERGENCE project, funded by KTH Digital Futures, we will study ways of integrating machine learning techniques with the goal of preventing and mitigating network cybersecurity attacks.

Organization

DIRECTOR & VICE DIRECTOR

Benoit Baudry, KTH
Mads Dam, KTH

BOARD

Catrin Granbom, Ericsson
Stefan Hagdahl, SAAB
Jens Zander, KTH

MANAGEMENT TEAM

Frédéric Loiret, Project Manager, KTH
Maria Berthelius, Service Coordinator, KTH
Åsa Svahn, Finance, KTH
Moa Hörnquist, Communication, KTH
Madeleine Printzsköld, HR, KTH

KTH RESEARCH STAFF MEMBERS

Amir H. Payberah, Assistant Professor
Benoit Baudry, Professor
Cyrille Artho, Associate Professor
Dejan Kostic, Professor
Dilian Gurov, Associate Professor
Douglas Wikström, Associate Professor
Elena Troubitsyna, Associate Professor
Frédéric Loiret, Senior Researcher
Ingo Sander, Professor
Jana Tumova, Assistant Professor
Jim Dowling, Associate Professor
Karl Meinke, Professor
Mads Dam, Professor
Marco Chiesa, Assistant Professor
Martin Monperus, Professor
Musard Balliu, Assistant Professor
Matthias Becker, Assistant Professor
Paris Carbone, Assistant Professor
Philipp Haller, Associate Professor
Roberto Guanciale, Assistant Professor
Rolf Stadler, Professor
Sarunas Girdzijauskas, Associate Professor
Seif Haridi, Professor



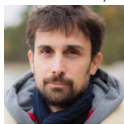
Benoit Baudry



Mads Dam



Jens Zander



Frédéric Loiret



Catrin Granbom



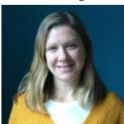
Stefan Hagdahl



Åsa Svahn



Maria Berthelius



Moa Hörnquist



Matthias Becker



Madeleine
Printzsköld



Marco Chiesa



Amir H. Payberah



Cyrille Artho



David Broman



Dejan Kostic



Dilian Gurov



Douglas Wikström



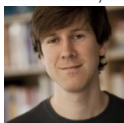
Elena Troubitsyna



Ingo Sander



Jana Tumova



Jim Dowling



Karl Meinke



Martin Monperus



Musard Balliu



Paris Carbone



Philipp Haller



Roberto Guanciale



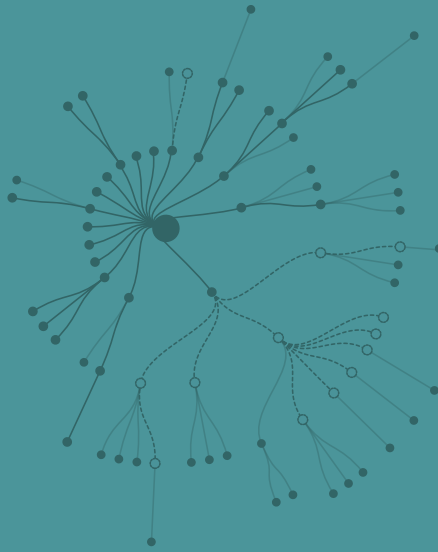
Rolf Stadler



Sarunas
Girdzijauskas



Seif Haridi



CASTOR

Software Research Centre

This document wraps the second year of CASTOR, which has been marked by some key milestones for the construction of a solid, sustainable collaboration between KTH and industry in the area of software research.



SAAB