Data Intensive Applications (DIA)—Project Plan

Vision: Smarter industrial Cyber-Physical Systems will be seamlessly integrated with system engineering in a perpetual process. *Mission*: Striving toward this vision, DIA advances the state of the art and the industry capacities by organizing research and graduate education in relevant smarter technologies.

1 The focus of DIA and the addressed scientific and industry-related needs

1.1 Research Field

This section describes the research field that DIA is addressing and the Smarter Systems environment's position within this field from a national and international perspective.

Cyber-Physical Systems (CPS) are interconnected systems of physical assets and computational capabilities [36]. As observed by Lee et al. [22], recent developments have resulted in higher availability and affordability of sensors, data acquisition systems, and computer networks. The consequent ubiquity of sensor, computing, communication, and storage technologies together with networked machines gives us access to previously unprecedented amounts of data—Big Data [25, 38]. CPS can exploit data driven and related technologies converting Big Data into actionable and machine processible knowledge to reach the goal of resilient, self-aware and self-adaptive systems—*Smarter Systems*. It is understood that integrating CPS with production, logistics, and services in the current industrial practices has significant economic potential [23, 24]. Such an integration would not only disruptively change industry [37],[1] but also the development, maintenance, and operation of CPS *software* as more and more functionality will reside in the software part of CPS.

1.1.1 Smarter Industrial CPS

Smarter systems learn from experience and examples, adapt themselves accordingly, and combine human with artificial intelligence. DIA is the research school of the *Smarter Systems* knowledge environment exploiting *data driven technologies*, self-adaptation, and visual analytics to smarten up industrial CPS and their development, maintenance, and operation. It comprises four areas of computer science: Cyber-Physical Systems (CPS), Self-Adaptation (SA), Data Driven Technologies (DD), and Visual Analytics (VA), see Figure 1.

Cyber-Physical Systems (CPS) monitor data and information from all relevant subsystems and components and synchronize between the physical and the cyber-computational parts [22]. Modern CPS are connected to the Internet referred to as the Internet of Things (IoT) [11] and, specifically in our focus, the Industrial IoT (IIoT). *Proactive Maintenance* of the physical CPS subsystems and components aims at avoiding failure and unscheduled repairs through self-adaptive and self-predictive methods. This is enabled, among other things, by *Digital Twins* that simulate physical subsystems and components of CPS based on abstract system models and data retrieved online via monitoring using sensors and the (I)IoT. In the future, CPS will contain general services, e.g., for monitoring, maintenance, upgrading, and adaptation of functionality.

Data Driven Technologies (DD) enable systems to automatically learn from experience and examples. They automatically transform (big) data into information and further into actionable knowledge (value) while managing challenging quantities (volume, velocity) and qualities (variety) of data. For handling the fundamental challenges: *Scalable computing* handles volume and velocity of Big Data, and *Code and data transformation* handles the variety of heterogeneous data sources. For transforming data into actionable knowledge: *Data-mining* and natural language processing (NLP) extract information from Big Data including human language data; *Machine learning* (ML) discovers classes, clusters, trends, and patterns in Big Data. *Simulations* operate abstract models of processes or systems to understand their behavior. *Digital Twins* are special simulations using online data of system components and make systems self-aware and -predictive.

[1] This change is referred to as the "Fourth Industrial Revolution", conceptualized in "Digitalization" and marketed as "Industry 4.0".

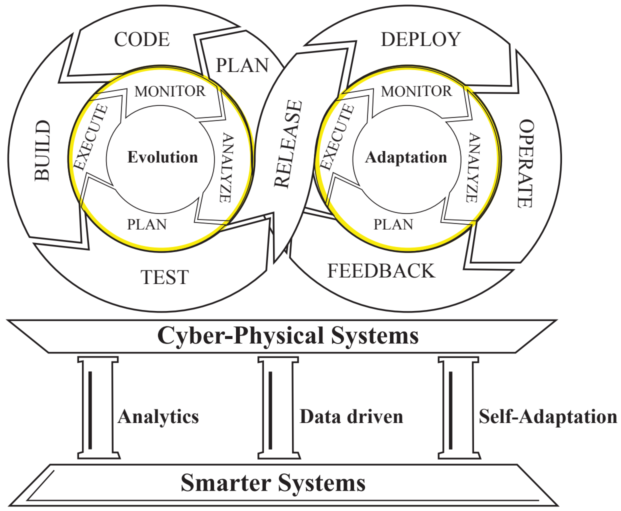

Figure 1: Smarter industrial CPS and system engineering in a perpetual process with relevant technologies.

Self-Adaptive Systems (SA) constantly monitor and analyze data and information of a managed system and its environment in order to plan and execute adaptations while building up system and adaptation knowledge at the same time [16]. Self-adaptive CPS requires the managed system to contain physical assets. Self-adaptive systems subsume self-configuring, -organizing, and -optimizing systems. Related *Self-** technologies, partially overlapping with self-adaptivity, make systems self-aware, -explainable, and -predictive. These self-* technologies make CPS smarter. Smarter CPS are open and tightly integrate human stakeholders, software services, and physical devices. *Software ecosystems* define the interplay for principal and third-party actors that develop, maintain, and market products and services sharing a common platform. Smarter ecosystems address this interplay when products, services, and platforms evolve and adapt as smarter systems.

Visual Analytics (VA) complements data driven technologies by combining the strengths of human and computational data processing and is defined as the science of analytical reasoning facilitated by interactive visual interfaces [40]. It uses appropriate visual metaphors and user interaction to convey information to human stakeholders to create knowledge and, conversely, to feedback insights into automated analysis processes.

1.1.2 Contributions of the Smarter Systems Environment to Smarter Industrial CPS

Cyber-Physical Systems and Proactive Maintenance. In the CPS group[2], we apply data driven technologies for on-line cyber-threat detection with a focus on quantitative vulnerability and risk assessment using probabilistic approaches considering inherent uncertainties [34]. Together with the department of mechanical engineering, we also research on sustainable production with focus on specific aspects concerning today's and future production and products, such as enhancing product development processes, productivity, flexibility, reliability, maintainability, availability, working environment, and prolonging the product's lifetime in a cost-effective way. In collaboration with the Centre for Cost-effective Industrial Asset Management (CeIAM)[3] our research in maintenance management and quality systems addresses questions of how products or systems are used or controlled during their lifetime, in order to meet their specifications at the lowest total cost. This research applies data driven technologies, such as condition monitoring, and signal processing. It organizes quality systems for proactive maintenance.

[2] `https://lnu.se/en/research/searchresearch/cyber-physical-systems-cps`
[3] `https://lnu.se/forskning/sok-forskning/centre-for-cost-effective-industrial-asset-management`

Data Driven Technologies. The Data Intensive Sciences and Applications (DISTA) group[4] contributed to fundamental and applied research in data driven *software* technology improving the engineering efficiency of development, maintenance, and operation of systems. In *Scalable Computing*, we advanced performance portability across various heterogeneous many-core computing systems [4] and introduced high-level parallel programming for large-scale computing systems [6]. We invented context-aware composition [2] and self-adaptive concurrent components [32] to increase the efficiency of parallel [18], concurrent [30], and service-oriented [20] programs. We invented the first ever block-free garbage collector [31]. We applied these technologies for optimizing machine learning [1]. In *Code and data transformation*, we solved heterogeneity and adaptation problems in the analysis of programs [39] and technical documentations [9] using meta-modelling and code generation. In *Data-mining*, we developed a generator of information retrieval components by adapting NLP parsing techniques [29]. In *Machine learning*, we contributed with a unifying theory of different ML approaches [8] while our main focus is on applying ML, in e.g. adaptive robots [15]. For *Simulations and Digital Twins*, we developed an executable model for Timed Automata [13] and co-simulations for cyber- and physical subsystems [7].

Self-Adaptive Systems. In the AdaptWise group,[5] we study the engineering of self-adaptive software systems focusing on assurances of the qualities of the systems and their engineering processes. The group has led the research endeavor on perpetual assurances [44] that moves the provision of assurances into the runtime and has contributed a set of formally specified patterns to design and verify feedback loops for such systems [10]. We defined a novel approach of so-called active feedback loop models to realize self-adaptation [14], which relies on the direct execution of formally verified feedback loop models to realize self-adaptation. Our research on smarter ecosystems focuses on models and process support [35] and applications to CPS [3].

Visual Analytics. The ISOVIS group[6] has contributed to the fundamental research in VA, e.g., to multivariate network visualization [17, 43, 21]. We have also contributed to the applications of these technologies, e.g., in engineering [28, 41]. The latter uses VA to visually and interactively adapt and eventually verify the data driven approach to aggregate multi-dimensional probability based scores [42].

1.1.3 The Smarter Systems Environment's position in the field, nationally and internationally

Cyber-Physical Systems and Proactive Maintenance. The CPS group is a leading center of excellence in smart transportation and intelligent railways, together with the Italian National Inter-University Consortium for Informatics,[7] Delft University of Technology[8] and the University of Leeds.[9] Our CeIAM group is setting the European research frontiers in Proactive Maintenance by leading the H2020 project on Predictive Cognitive Maintenance Decision Support system (PreCoM).[10] PreCoM brings together top competences from 17 organizations, including Bosch[11] and the Technische Universität München,[12] from six European countries.

Data Driven Technologies. Around the ideas of context-aware composition, we organized a Dagstuhl seminar [19] and two Software Composition conferences [27, 5]. Our block-free garbage collector made it into the latest Java OpenJDK version. On a national level, the Wallenberg AI, Autonomous Systems, and Software Program *WASP*[13] is a major national strategic research initiative. We collaborate with the Profs. Jan Bosch (Chalmers) and Per Runeson (Lund

[4] https://lnu.se/en/research/searchresearch/dista
[5] https://lnu.se/en/research/searchresearch/adaptwise
[6] http://cs.lnu.se/isovis
[7] https://www.consorzio-cini.it/index.php/en
[8] https://www.tudelft.nl/citg/over-faculteit/afdelingen/transport-planning/research/labs/digital-rail-traffic-lab/
[9] https://environment.leeds.ac.uk/transport
[10] https://www.precom-project.eu
[11] https://www.boschrexroth.com/en/xc
[12] http://www.iwb.mw.tum.de
[13] https://wasp-sweden.org

University) from the cluster "Software Engineering for Smart Systems" in supporting the Swedish software industry. The Blekinge Institute of Technology, BTH, has the KKS research profile *BigData@BTH*.[14] We collaborate with Prof. Håkan Grahn in common research projects and graduate education. Linköping University, LiU, has an internationally leading Programming Environment Laboratory, *PELAB*[15] with research in parallel computing, transformation technologies and simulation. We collaborate with Prof. Christoph Kessler in common research projects and graduate education. On the European level, the Center for Data-intensive Systems *Daisy*[16] is a research center at Aalborg University, Denmark, and internationally leading in the integration of data management support into general-purpose programming languages. Prof. Christian Jensen is the head of DISA's advisory board.

Self-Adaptive Systems. Our AdaptWise group is setting the European research frontiers here and collaborates with other strong teams including Deepse at Politecnico di Milano, Italy,[17] the TASP group at York University, UK,[18] Spirals at Inria and the University of Lille, France.[19] At the global level, the AdaptWise group collaborates with the ABLE Group at Carnegie Mellon University, US,[20] that focuses on Self-Adaptation and CPS. On the national level, the group collaborates with the Systems-of-Systems research group at RISE, Stockholm, on smarter ecosystems.

Visual Analytics. Our ISOVIS group is setting the European research frontiers in VA. Nationally, there are only two further groups with considerable research in VA. The Division of Media and Information Technology at LiU[21] with mainly scientific visualization and spatio-temporal analysis of scientific data, and the Skövde Artificial Intelligence Lab[22] with a scope on data mining and HCI. ISOVIS collaborates with the strongest VA teams including the Data Analysis and Visualization group[23] at the University of Konstanz and the Visualization Research Center at Stuttgart,[24] both in Germany. At a global level, visual analytics research is performed to a substantial extent at the Georgia Tech's School of Interactive Computing[25] and at UC Davis Center for Visualization.[26]

1.2 From Grand Challenges to Research Questions

This section describes and motivates the grand challenges of the field. From these challenges we derive the research questions addressed in the project and how these relate to the state-of-the-art.

The two high-level objectives for the Smarter Systems environment applied to industrial CPS are:

KO1 To assure smarter industrial CPS capabilities in uncertain conditions.

KO2 To devise engineering processes for the perpetual adaptation and evolution of both software and hardware in smarter industrial CPS.

To achieve these key objectives, the three grand challenges below—identified in the research fields contributing to the Smarter Systems environment—are of particular relevance to enable smarter industrial CPS.

A CPS grand challenge (CPS-3) concerns real applications in Industry 4.0. With a few exceptions, most factories feature limited levels of smartness and intelligence to date. That means a huge work needs to be done in order to not only conceptualize but also design and develop infrastructures, protocols, and algorithms enabling smartness [33]. At the research level, it will be essential to build small-scale test beds for laboratory experimentation of new technologies

[14] http://www.bth.se/bigdata
[15] https://www.ida.liu.se/labs/pelab
[16] https://www.daisy.aau.dk
[17] http://deepse.dei.polimi.it
[18] https://www.cs.york.ac.uk/tasp
[19] https://team.inria.fr/spirals
[20] http://www.cs.cmu.edu/~able
[21] https://liu.se/en/research/media-and-information-technology-2
[22] http://www.his.se/en/sail
[23] https://www.vis.uni-konstanz.de/en
[24] https://www.visus.uni-stuttgart.de/en/index.html
[25] https://ic.gatech.edu
[26] https://vis.ucdavis.edu

and integrated systems for, e.g., smarter production and maintenance. The industry cooperation in DIA provides the requirements of real-world applications. Tackling this grand challenge will contribute to both assuring CPS capabilities (KO1) and devising smart engineering processes (KO2) through the validation of research results in real-world settings of DIA's Ph.D. projects.

A grand challenge of data driven technologies (DD-2) is to ensure guarantees for applications based on data driven models—regarding accuracy, performance, response time, safety etc.—persisting over time. While data driven applications often outperform traditional ones in performance and accuracy, it is hard to guarantee their capabilities in the worst cases. It is challenging to validate, e.g., their safety and response time in the first place. Worse, in an ever-changing world, new data input channels constantly appear and others disappear. The data specifications themselves constantly change as the sensor technology develops. Hence, it is difficult to guarantee the capabilities and maintain learned knowledge (machine- learned models) over time. Continuous machine learning adapts to changes in the data sources and the data itself but makes it even harder to ensure guarantees. Also, it adds to the next challenge described below. Tackling the second grand challenge will contribute to assuring industrial CPS capabilities (KO1) through the development of proper data driven technology that provides guarantees for smarter CPS in operation.

A third grand challenge of both data driven technologies and visual analytics (DD-3, VA-3) is the self-explainability of machine learning models. Opening the black box of computational models, such as supervised or unsupervised machine learning, is a challenge [26] both in data driven technologies and visual analytics. Data driven technologies should not only work according to specifications and constraints, but should also generate repeatable actionable engineering knowledge that can be transferred to new application cases and areas. Not only should CPS get smarter based on experience and examples, but so should their human engineers and users. Addressing this challenge increases the trust into data-driven technologies, and will contribute to both assuring product capabilities (KO1) and devising smart engineering processes (KO2). This is done through the development of proper methods and techniques to understand smarter CPS and support operators in concrete real-world CPS settings.

These three grand challenges lead to the following refined overarching research questions addressed in DIA. The individual Ph.D. projects and the graduate school as a whole will conduct research to answers these questions.

Research Questions

RQ-1 How can we utilize abstract digitalization concepts for smarter industrial CPS? How can we build, maintain, and organize open CPS ecosystems?

RQ-2 How can we verify guarantees of smarter industrial CPS built on data driven models regarding accuracy, performance, response time, safety, etc.? How can we assure that they persist over time?

RQ-3 How can we make data driven models that control smarter industrial CPS self-explainable? How can we convey the knowledge from models to human stakeholders?

1.3 Significance

This section describes the DIA's significance for the university's development of research and doctoral education.

The Digital Europe program[27] aims at building strategic digital skills and capacities. It will launch in 2021 and will last for six years, boosting investments in scalable computing (€2.7 billion), artificial intelligence (€2.5 billion), and for ensuring a wide use of digital technologies across the economy and society (€1.3 billion). Linnaeus University (Lnu) has already strategically invested in the relevant smarter system technologies (CPS, SA, DD, VA). They constitute the core of the Linnaeus University Center (Lnuc) on "Data intensive sciences and applications" (DISA),[28] one of six official research excellence centers at Lnu. They are the cornerstones of the knowledge environment "Smarter Systems" and the building blocks of the advanced level of the Master of Science in Software Engineering (MSc Eng)

[27] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60502

[28] https://lnu.se/disa

program, the first and only MSc Eng program at Lnu. DIA will be the graduate school of DISA and the Smarter Systems environment and, this way, complement research and undergrad education in these technologies.

Also, DIA constitutes a natural next step of further developing the formal Ph.D. education in CS, in general. The Swedish Higher Education Authority, has conducted an evaluation of the educations leading to licentiate and Ph.D. degrees in CS and attested the Ph.D. education in CS at Lnu a high quality.[29] However, the education has been mainly conducted with individual courses and in collaboration with other graduate schools, e.g., CUGS in Linköping. Consequently, DISA has started Ph.D. level courses in data driven technologies. DIA provides dedicated resources for the development of a cohesive Ph.D. curriculum and course in smarter technologies. Also, it adds a critical mass of Ph.D. students. Most importantly, DIA sets focus on the industry application needs of research in smarter technologies. DIA will become the crystallization point for a structured, cohesive Ph.D. education, also to the benefit of the other Ph.D. students enrolled in the Smarter Systems environment and DISA now and in the future. Therefore, DIA is planned to continue even after the initial funding applied from KKS. It is our vision to make it a permanent research school at Lnu that will foster collaboration across academic disciplines and with the industry in smarter industrial system technologies.

1.4 Relevance

This section describes DIA's relevance and its focus to the business sector.

"Industry, innovation and infrastructure" is one of the 17 development goals set by the UN on the Agenda 2030 for the global sustainability efforts in coming years[30] and DIA contributes to resilient (smarter) CPS infrastructures, sustainable (smarter) industrial CPS development and faster (smarter) innovation in industry.

Currently, the industrial sector constitutes one fifth of Sweden's total economy, and is a significant growth engine. It accounts for 72% of the business sector's research and development (R&D) efforts. Industry and industrial services generate 77% of the country's export revenue. The development of the industry is also considered to be of great importance for welfare. However, industrial growth has stagnated in Sweden. Between 2000 and 2013, it has decreased by an average of 0.2%, while in the same period it increased by 2.1% per year in countries such as Belgium, Ireland, Austria, Germany and the Netherlands. If Sweden succeeds in switching from stagnation to the same growth as the above countries, the cumulative increased economic value in 2040 would amount to SEK 6,800 billion.[31]

Innovative digitalization concepts for achieving this goal are summarized under buzzwords such as Industry 4.0, Big Data, Artificial Intelligence (AI), Deep Learning, CPS, and the Industrial Internet of Things (IIoT). These concepts and their potential for the industry are well understood, in general. However, they remain fuzzy, unstructured, unrelated, or overlapping, and partially simply too big for an individual company to bring them into operation. DIA derives its *general relevance* from addressing this contemporary challenge of the Swedish industry in a collaboration of a multi-disciplinary research team consisting of computer scientists and engineers and domain experts of partner companies.

Companies prefer to keep tight control over costs and benefits. They would like to conduct beneficial projects with manageable costs that are promising even with limited resources to build up their know-how gradually following a clear strategy. Industrial digitalization innovation ranks high on their strategic agenda. However, concrete pilot-, mid-, and long-term projects are delayed due to the gap between the high level concepts and a convincing digitalization strategy. At the same time there is a bad conscious about this delay and the belief that other market participants have come further in their digitalization.

We understood this industry challenge due to the collaborations with our partner companies over the course of a year (Scania) up to 15 years (Sigma Technologies) in common BSc/MSc thesis projects, research proposals, publications, and research projects. In order to meet this challenge with concrete actions, we conducted common workshops where companies defined and sharpened their own digitalization vision and goals. They analyzed their current products, services, and processes, their improvement potential, as well as the prerequisites necessary for improvements by digitalization. This gap analysis led to R&D needs and benefits and concrete pilot research projects. In condensed

[29] Universitetskanslersämbetet (UKÄ,https://www.uka.se). Full report published 2018-06-26: https://goo.gl/zLeCP5
[30] https://www.un.org/sustainabledevelopment/development-agenda
[31] https://www.nyteknik.se/sponsrad/industri-4-0-en-tillvaxtmotor-for-sverige-6874890

form, the workshop results are documented in Section 5. For the general method and the canvas applied in the workshops and for a detailed documentation of the workshop results for one of our partner companies, we refer to [12].

We aggregated the workshop results to common Industry Objectives (IO-1–IO-3) that the participating companies confirmed. DIA draws its *practical relevance* from addressing these objectives in the individual Ph.D. projects with and for the partner companies and as the graduate school in total.

Industry Objectives

IO-1 To adopt smart technologies, such as digital twin and proactive maintenance, for solving concrete problems in the development, maintenance, and operation of industrial CPS and for improving these processes.

IO-2 To put together individual solutions and improvements to a common digitalization strategy towards smarter industrial CPS. To define and get started with concrete first pilot projects, set expected benefits, and define a structured systematic roadmap towards digitalization.

IO-3 To create new services, platforms, and ecosystems supporting smarter industrial CPS.

In addition to these technical objectives, we even identified the management challenges to acquire the right competences and staff in relevant technologies. DIA's relevance increases as we address even this challenge by helping companies recruit, train, and keep staff with the right competences to address IO-1–IO-3.

The industry objective contribute nicely to our research questions, see Section 3.2, and this did not happen by chance. Instead, we conducted way more workshops with *potential* industry partners than we eventually invited into DIA and we consciously excluded companies whose objectives did not help answering the research questions. Actually, we did this selection in response to the expert reviewers feedback on the first DIA proposal of 2018.

1.5 Contributions

This section describes the participants' contribution to DIA and how these are complementary in terms of competence.

DIA will contribute with structured *research education* with Ph.D. students employed in the industry. This will supply companies with the fundamental competence of conducting research and development in smarter industrial CPS technologies. Also, DIA will be able to enhance the systematic organization of our Ph.D. education and offer courses on contemporary topics brought forward by industry. This way, DIA will lift the Ph.D. education in CS at Lnu to the next quality level.

DIA will also contribute with *applied research* in co-production with the participating companies and across academic research fields. The actual research will be conducted in individual research projects in the participating companies co-supervised by experts at Lnu and the partner companies. It will address the concrete R&D objectives in the companies (IO-1, IO-3). Some of these R&D objectives lead to common overarching questions for several companies and will be further detailed in Section 3.2.

The Ph.D. projects are structured is smaller steps planned in individual study plans (ISPs). This way, DIA will contribute to the operationalization of digitalization in the participating companies (IO-2). Moreover, due to the exchange of information and experiences in the Ph.D. projects, the individual participating professionals and their companies will be able to learn from each other and to benchmark and evaluate their achievements and setbacks in digitalization.

Ph.D. research projects will test and validate research ideas and hypotheses that aim at answering the common research questions (RQ-1–RQ-3) in an industry context. To this end, DIA will contribute to advancing the state of the art in smarter industrial CPS.

Lnu The researchers/supervisors at Lnu contribute with the appropriate technologies and generic scientific tools and methods that the industry Ph.D. students can then apply in a specific context in their respective company to address the industry objectives in a scientific way. Concretely, Lnu will contribute with:

- Teaching in the foundations of smarter technologies: CPS, self-adaptive systems, data driven technologies, and visual analytics. Courses are detailed in Section 3.1.
- Teaching in research methodology and methods, information search and retrieval, writing scientific texts, reviewing etc., cf. Section 3.1.
- A cohesive curriculum for the graduate education in smarter technologies guaranteeing quality and progression of the individual courses, cf. Section 3.1.
- Senior research and supervisor competence and concrete collaboration in the industry Ph.D. projects. Every Ph.D. student will be supervised by at least one supervisor from Lnu. Researchers/supervisors collaborate in the industry Ph.D. projects. They are responsible for the scientific relevance of the individual research questions and the scientific quality of the results. Projects are detailed in cf. Section 5.

Lnu will lead the joint work of defining and developing missing courses and putting them together to a curriculum. This work and the resulting program will be embedded in the academic and administrative infrastructure at Lnu. Lnu will provide the formal third cycle education program and examination rights, the administrative structure for developing, approving, and quality assuring Ph.D. courses and individual study plans, meeting rooms, lecture halls etc.

Companies will contribute with

- Real-world problems and the resources needed to define research projects of industry relevance and to establish them firmly in their organization.
- Resources needed to prototypically implement research project results, define goal criteria, and evaluate them in real-world settings and on real-world data.
- Engineering and management competence in their respective industry. Every Ph.D. student will be supervised by at least one supervisor from the respective industry partner. They are responsible for the industry relevance of the individual project goals and the applicability of the results.
- Thesis topics and guest lectures on the advanced level of the MSc in Software Engineering program.

Industry partners lead the joint work of defining needs for competences and contribute to the joint work of defining and developing missing courses and putting them together to a curriculum. Industry partners contribute with complementary lectures and seminars both as part of the envisioned Ph.D. curriculum and outside reaching students and researches outside DIA and the general society.

2 Expected results and effects

2.1 Goals for the university

This section describes the university's goals for the graduate school, time-specific, and quantitative.

For the doctoral students DIA creates challenging scientific questions in smart technologies and applications in smarter CPS with high relevance for the industry. This way, it aims to contribute to industry centered and research based Ph.D. education with clear links to solutions and innovations in crucial fields of society. DIA aims to contribute with challenging educations and complements them with applied research and societal embedding. It increases the number of available Ph.D. level positions, students, and courses. It increases the employability of graduates in industry and academia and, hence, opens up new career paths for them. It strengthens the individuals in providing direct meaning and value in the application of otherwise theoretic knowledge and results.

Operational goals: 10 Ph.D. students defend a Licentiate thesis by the end of 2022 and a Ph.D. thesis by the end of 2025; no drop-outs of the program.

For the development of the Smarter Systems environment It is the goal of DIA to contribute to the capability building phase of building the Smarter Systems knowledge environment with Ph.D. students and external relevant competences.

It is also the goal of DIA to improve research and education of the Smarter Systems environment at Lnu. A common research education in smarter CPS is clearly in line with the visions and strategy of the Smarter Systems environment, Lnu, and the Faculty of Technology. We identified a need for a cohesive third cycle education curriculum that DIA will contribute with. DIA is set to extend the Ph.D. education programs at DISA and at the CS department. It is also the goal of DIA to provide a research continuation for the 5-year's MSc in Software Engineering program. Moreover, it is the goal of DIA to increase the maturity of researchers at Lnu in collaborating with the industry in long-term commitments to Ph.D. projects and commonly supervised Ph.D. students. This will lead to a better understanding of industry needs and planning horizons.

Operational goals: The Smarter Systems environment has enrolled (more) 10 Ph.D. students by 9/2020. The Smarter Systems environment is recognized as a *Linnaeus Knowledge Environment*, making Smarter Systems a profile of Lnu by the end of 2021. The Faculty of Technology has approved the general study plan (ASP) for doctoral studies in Computer Science with a specialization in Smarter Systems by the end of 2023. DIA is funded and continues as the graduate school of the Smarter Systems environment after 2025.

For scientific development and positioning The consolidated research performed by DIA Ph.D. students will advance the state of the art in smarter CPS. It will open new avenues for future research and raise new research questions. Consequently, it is expected that DIA generates more high-quality scientific publications, more collaborative research projects with high industry relevance and more external funding in the future.

For the fundamental research in Smarter Systems, DIA and its industry projects provide an invaluable source of real-world data and scenarios in order to validate, if necessary adapt, or even reject the theoretical answers to RQ-1–RQ-3 (and beyond) that otherwise only could have been validated under lab conditions. It also leads to closer collaborations between the Computer Science (CS) department, the successful engineering departments in Mechanical Engineering (ME), and Forest and Wood Technology (FWT).

Operational goals: Successful research will lead to one level 2 plus around five of level 1 journal or conference publications for each of the 10 Ph.D. modules by the end of 2025, see Section 3.1, Publication strategy. Each Ph.D. module will generate at least one externally funded research proposal by the end of 2024 and at least 1 MSc thesis per year under 2020–2025.

For co-production By pursuing the industry objectives in the research modules, DIA is set to increase the relevance of Smarter Systems research for our industry partners.

It is also a defined goal of DIA to summarize and formalize industry collaborations of the participating companies with Lnu and to put the numerous smaller activities with them on a more sustainable basis. This way, DIA simplifies the customer relationship management (CRM) with these partner companies. This amplifies the knowledge exchange with and between the participating companies. Moreover, the collaboration gives mutual access to the respective national and international networks.

Additionally, DIA offers courses related to smarter industrial CPS to other Swedish companies. This way, DIA promotes competence development in smarter technologies even beyond the partnership of DIA. This increases the attractiveness of the Smarter Systems environment for other potential industry collaboration partners and lowers the entry bar of other related R&D collaborations.

Operational goals: 5/9 partner companies will continue collaboration in an externally funded research project overlapping with DIA or directly after 2025. The other 4/9 partner companies will continue collaboration by other means. The Smarter Systems environment has signed formal research and education partnership agreements with all 9 companies participating in DIA by the end of 2021. For Combitech, Scania, Volvo, the Smarter Systems environment has concrete research collaborations with their main national-level R&D centers in addition to the DIA collaboration with the units in the Linnaeus region by the end of 2023. We collaborate with Yaskawa's German R&D center by the end of 2025. Ten new companies beyond the DIA partners collaborate with Smarter Systems researchers by 2025.

2.2 Goals for the business sector

This section describes overarching goals for the business sector.

R&D goals Industry partners are facing similar objectives in digitalization IO-1–IO-3, cf. Section 1.4, i.e., in exploiting data driven and related technologies for building smarter CPS. It is the overarching R&D goal of the industry to address these industry objectives individually in the respective Ph.D. projects and to answer specific and common applied research questions. The common research questions are further detailed and discussed in Section 3.2.

Soft goals Beyond these objectives and questions, industry partners expressed in the initial workshops similar overarching soft goals that are not even limited to the research field of DIA. They include:

- Recruitment of staff that is trained on the job and familiar with the company culture. Alternatively, giving new development opportunities to staff that otherwise would seek new challenges elsewhere.
- Access to relevant research and easy adoption (or rejection) of research ideas for the own business.
- Consulting expertise by senior researchers contributing to solutions relevant for the industry. Long-term relationship building with relevant research and education environments.
- Exchange of experiences on common challenges and approaches with other participating companies.
- Better positioning and competence development of the collaborating regional sites in the globally operating group, e.g., better dialogue with and the central research units of a global group and higher relevance of the collaborating regional sites.
- Training of staff with a skill set to tackle highly relevant and challenging end to end problems, from problem definition and solution creation, to realisation and validation.

3 Implementation of the project

This section describes the planned activities in DIA and describes how DIA will, based on the individual Ph.D. student projects, identify and address common research questions and results. It concludes with a risk analysis.

3.1 Planned activities

This section describes the planned activities including recruitment of doctoral students, joint activities, organization of supervision, courses including development of new courses, publication strategy, and internationalization strategy.

Recruitment of doctoral students Most partners also see DIA as an opportunity to recruit competences that they lack today. For these partners and positions, a joint recruitment process has been defined with the HR and R&D divisions of the partners that has already been started for some companies, such as Scania. This process consists of the following steps: (1) Jointly define common qualification criteria applicable to all candidates. A formal criterion is an MSc in Computer Science or equivalent. (2) Starting for the proposed Ph.D. project description, jointly define required competences and formulate additional qualification criteria. (3) Jointly formulate a position profile (official document at Lnu) and a job advertisement. (4) Get a position profile approved by the board of doctoral studies at the Faculty of Technology (Lnu). (5) Publish the job advertisement in the academic and companies' channels. Especially, prominently publish a joint ad of all open DIA related positions in a national newspaper (Dagens Industri) to get extra attention and publicity for DIA. (6) Jointly assess the candidates. This includes interviews and candidate rankings according to the qualification criteria. Involved are HR and supervisor from the partner company and the academic supervisor(s) from Lnu and the project manager of DIA. (7) Employ the Ph.D. student at the partner company. (8) Formulate an Individual Study Plan (ISP) and get it approved by the board of doctoral studies, which formally enrolls the Ph.D. student. Some partner companies, such as Combitech and Volvo, have own employees who shall be enrolled in DIA as part of their competence development. For these students, the process reduces to the steps 6 and 8.

Joint activities

Joint kick-off day with all supervisors from the industry and Lnu, Ph.D. students and company and Lnu representatives to officially start the school and generate a common loyalty towards DIA.

Annual one-day DIA workshop collocated with DISA's annual Big Data conference. At these workshops, Ph.D. students document their achievements both internally and externally. They are an occasion to bring together DIA Ph.D. students and Lnu supervisors, other DISA researchers, company supervisors, and other interested company representatives with researchers and students at Lnu as well as interested companies and organizations in the general society. In December 2019, we invite to the 5th annual Big Data conference, where we welcome the industry supervisor and confirmed Ph.D. students and inform about DIA.

Bi-annual one-day doctoral seminar At these seminars, Ph.D. students defend their achievements and next steps internally in front of the joint supervisor staff both from the industry and Lnu, and their fellow Ph.D. students. These seminars are a central tool for DIA for achieving several objectives. They serve as a platform for DIA internal R&D related discussions. They train the presentation and public speaking skills of the Ph.D. students, preparing them for conferences and the thesis defence. They correct lesser promising suggestions of the supervisor team by making use of the intelligence of the many. Finally, the motivate they Ph.D. students and set deadlines that help structuring the long-term venture of Ph.D. studies.

Ph.D. curriculum and course development are central activities in DIA although there are some courses that will be integrated from the second cycle Master of Science in Software Engineering and in Software Technology programs. There is currently just one compulsory course of the Ph.D. curriculum in CS, a number that should increase. *Course development* is demand driven where industry and Lnu supervisors define needs and opportunities of competence areas, such as Machine Learning and Visual Analytics. Courses are set as compulsory if there is at least a need in 50% of the Ph.D. modules or if a program coordinator suggests and the executive board decides to do so. *Curriculum development* is driven by prerequisites in some courses and the needs for a general progression over the curriculum. Here the program coordinator suggests and the executive board decides; course syllabi and a general study plan (ASP) formalize the decisions.

Ph.D. courses are the main means of formal Ph.D. education and knowledge transfer, but also an opportunity for the Ph.D. students to meet their peers from DIA and related Ph.D. programs.

Bi-annual steering committee meetings are scheduled in connection to a monthly DISA meeting.

Monthly executive board meetings are scheduled in connection to the monthly DISA meetings.

Continuation efforts start from year 4 on and aim at maintaining DIA beyond 2025. We consider co-funding from the Kalmar and Kronoberg regions as part of the establishment of an European Digital Innovation Hub of the EU program Digital Europe.[32] Alternatively, we consider the creation of a European Industrial PhD school as part of an existing program supported by the European Institute of Innovation and Technology.[33]

Weekly supervisor meetings are detailed below.

Organization of supervision Each Ph.D. module consists of a Ph.D. student with a research project, one supervisor from the industry and one supervisor from Lnu. A supervision group of a Ph.D. module may incorporate other experts from Lnu and the industry if needed, e.g., for adding expertise in mechanical engineering and forest and wood technologies. The Ph.D. research project is jointly defined by the two main supervisors to guarantee both the relevance for the companies' R&D and scientific quality.

Weekly meetings with the Ph.D. student and the supervisors are the main instrument to organize supervision. They bring up questions, resolve practical problems, and provide feedback. They check the progress on micro-level keep the participants of the module in sync with the latest development and with each other. Co-publications are discussed and initiated in these meetings.

Ph.D. students meet in common courses to exchange experiences and support and encourage each other. This part of the supervision is informal but nonetheless invaluable. It is expected and encouraged that the Ph.D. students find

[32] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60502
[33] https://eit.europa.eu/our-activities/education/doctoral-programmes

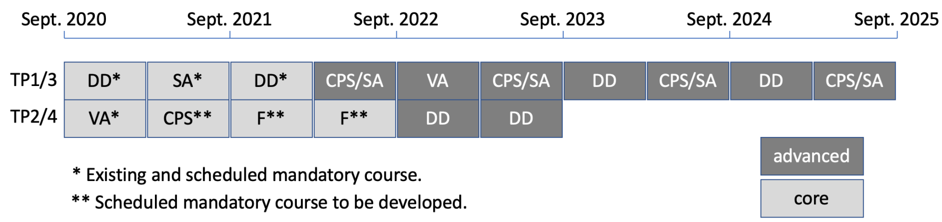

Figure 2: DIA teaching program framework.

more interaction channels, e.g., slack groups and common informal meetings. DIA will fund such activities when needed but, we expect the Ph.D. students to take the initiative.

Ph.D. students and their supervisors also meet in the bi-annual doctoral seminars and at the annual one-day DIA workshops. These activities are an essential part of the official supervision, as they encourage to summarize, present, discuss, and oppose to intermediate results. It puts the focus and sets deadlines on the scientific work progress and prepares the Ph.D. students for conferences and Licentiate and Ph.D. disputations.

Courses including development of new courses The course part in the computer science (CS) research education at Lnu is comprised of at least 90 ECTS credits. To date, there is only one mandatory course on research methodology and research ethics, not provided by nor tailored to the CS department. The curriculum for Ph.D. students in DIA will include more mandatory courses as well as elective courses. The latter are to be decided together with the examiner and the supervisors.

Courses are mandatory on a core level. To guarantee progression, there are also courses on advanced level, mandatory or elective, which is to be decided for the individual course. Courses are grouped into the four areas of computer science that constitute the Smarter Systems environment: Cyber-Physical Systems (CPS), Self-Adaptation (SA), Data Driven Technologies (DD) and Visual Analytics (VA) and a group for foundational courses (F). See Figure 2 for the DIA curriculum framework comprising 80 ECTS credits. Existing course include:

DD Introduction to data analysis and computation using Python, core, 5 ECTS (existing already as a DISA Ph.D. course). This course will be given in TP1, 2020/21.* Introduction to the basics of Python data analysis, including some of the standard modules, such as NumPy, Pandas, and MatPlotLib. The rest of the course is structured around "How do you do X in Python," where X is a topic such as Network Analysis, Text Mining, etc. Each topic will be covered by one or a few overview lectures that cover some of the essential algorithms in detail, how to implement them in Python, and which modules are available to use. The course will also briefly cover how to use the DISA HPCC and how to run Python programs on multicore machines and a cluster of such machines.

VA Applied Information Visualization, core, 5 ECTS (MSc Eng/MSc ST/DISA Ph.D.). This course will be given in TP2, 2020/21.* Information Visualization (InfoVis) focuses on the use of visualization techniques to help people understand and analyze abstract data such as tables or hierarchies. The course includes visual representations, interaction techniques, and visualization systems for: text and documents, network data (graphs), time series, software-related data, SoftVis, WebVis, BioMedVis, and GeoVis.

SA Adaptive Software Systems, core, 5 ECTS (MSc ST). This course will be given in TP3, 2020/21.* The course gives an overview of adaptive software systems and explains the central role of software architecture for adaptive systems. Students will study and critically examine research studies of adaptive software systems. Students will have the opportunity to get hands-on experience from designing and implementing an adaptive software system.

DD Applied Machine Learning, core, 5 ECTS (MSc Eng/MSc ST/DISA Ph.D.). This course will be given in the Teaching Period 1 (TP1), 2021/22.* The goal of the course is to understand ML for bringing meaning to and learning from data. This course mixes theory and practice, with a focus on applied machine learning where we learn what algorithms and approaches to apply on different types of data. The course includes the following: Supervised learning, different types of data and data processing, algorithms for handling text documents, algorithms for handling data with numerical and categorical attributes, neural networks, and deep learning for image recognition.

Other existing courses potentially build into the curriculum are listed below, see the Appendix for course descriptions.

DD: Systems modeling and simulation, advanced, 5 ECTS (MSc Eng/MSc ST).

DD: Code transformation and interpretation, advanced, 5 ECTS (MSc Eng/MSc ST).

DD: Data Mining, advanced, 5 ECTS (MSc Eng/MSc ST).

DD: Parallel computing systems, advanced, 5 ECTS.

DD: Deep learning, advanced, 5 ECTS (MSc Eng/MSc ST).

CPS: Structural dynamics, advanced, 5 ECTS (MSc ME).

DIA needs to develop the mandatory core courses "Foundations of CPS" and "Research Methodology":

*CPS** Foundations of Cyber-Physical Systems, core, 5 ECTS. This course will be given in TP4, 2020/21.* The aim of the course is to make students familiar with the definition, concepts, and issues of CPS. It will introduce concepts that are further detailed in the advanced CPS courses, such as embedded systems, sensors, data acquisition and signal processing, machine diagnostics and condition monitoring. It will introduce the needs of CPS that are detailed in the advanced courses, such as dependability and resilience.

*F** Research methodology in CS, core, 5+5 ECTS. Part 1 of this course will be given in TP2, part 2 in TP4, 2021/22.* The aim of the course is that the student gain knowledge and skills in research methodology. The course consists of modules with the following content: philosophy of science, research ethics, literature searching, quantitative research methodology, oral presentation skills, how to conduct a research project from beginning to end. After the completion of the course, the student should be able to: use scientific databases, formulate a research problem and objectives, design a comprehensive research study that is appropriate to answer the given research questions, perform data analysis and compare the results with the literature and discuss the implications for research and practice, and write a research article including reference other people's work according to recognized academic practice.

DIA needs to develop additional advanced courses, at least 2 in CPS, 2 in SA, 1 in VA, altogether 8–10. Note that we will develop more courses than there are teaching slots in the curriculum framework. This is partially necessary as some courses are inherited from our Master's educations. Ph.D. students who we possibly recruit from our own Master's programs cannot be credited with these courses and must take new courses. The decision on which courses will actually be developed and offered to all Ph.D. students depends on the constraints and demands of the Ph.D. modules. Other courses will only become individual reading courses or will not be developed at all. The list below constitutes the gross list of courses that the Lnu researchers involved in DIA offered to develop on demand, see the Appendix for the course descriptions.

CPS: Predictive Maintenance: Theories and Applications, advanced, 5 ECTS.

CPS: Embedded Systems Dependability, advanced, 5 ECTS.

CPS: Vibration based condition Monitoring for Industry 4.0, advanced, 5 ECTS.

CPS: Sensors, data acquisition, signal processing, and machine diagnostics for Industry 4.0, advanced, 5 ECTS.

SA: Software Product Line Engineering, advanced, 5 ECTS.

SA: Platform-based Software Ecosystems, advanced, 5 ECTS.

DD: Statistical methods for software engineers and computer scientists, advanced, 5 ECTS.

DD: Multivariate analysis, advanced, 5 ECTS.

DD: Data Privacy and Security, advanced, 5 ECTS.

DD: Data Quality and Management, advanced, 5 ECTS.

DD: Designing data driven applications, advanced, 5 ECTS.

VA: Visual Network Analytics, advanced, 5 ECTS.

Common courses will be given on site and on-line in hybrid learning mode, exploiting the experience of developing flexible forms of education, most notably from the program on "Social media and Web Technologies for Innovation and Growth - courses for professionals" funded by the KKS "Expert Competence for Innovation" program. By supporting these courses with modern learning management systems (LMS), we create the foundations for aligning professional and educational activities and for spreading the content even beyond DIA.

Publication strategy Every year, the Faculty of Technology and Lnu assess the individual researchers to set their Faculty funding level for the coming year. The assessment uses the number of publications and the quality of the publication channels according to the Norwegian Register for Scientific Journals, Series and Publishers. In the last five years, the 41 active researchers at the CS department produced a total of 421 publications, i.e., ca. two publications per year and researcher. 6% of the publications are at the highest quality level (level 2), which includes high-impact journals and conferences. 85% of the publications are at level 1. This level includes journals and conferences with somewhat lower impact and even some highly ranked symposia. The remaining level 0 publications, typically in workshop proceeding, account for less than 9% of the total production. The (informal) publication requirement for a Ph.D. thesis to be accepted in CS is one level 2, alternatively two level 1 *journal* publications, plus around five level 1 *conference* publications.

The Smarter Systems environment's goal is to increase the number and relevance of publications: on average ≥ 4 publications per researcher (excl. Ph.D. students) and year, $\geq$20% level 2 publications. DIA aims at contributing to this goal and, hence, will actively work for an increased amount of publications with high impact and relevance by

- Setting individual quantitative publication goals and assessing the number and the quality of publications as part of the individual study plan (ISP) updates,
- Encouraging early publications in the first year even at level 0 and funding travel costs and workshop fees. Discouraging level 0 publications later in the subsequent years and not funding connected costs.
- Adjusting the publication requirement for a Ph.D. thesis to one level 2 journal plus around five level 1 journal or conference publications.

Internationalization strategy Within DIA internationalization is important and gives both the researchers and the partner companies an added value. Researchers understand industry trends, solutions and needs, not only from the perspectives of an individual company but also on national and international level. Our industry partners learn about relevant research and development trends and results in an international context. Both broaden their respective perspectives. Moreover, for researchers, the international context is crucial to keep their research relevant and in the forefront. DIA includes this perspective in their (industrial) Ph.D. students' education from the beginning and throughout their whole Ph.D. program. This internationalization includes

- Researchers/supervisors and Ph.D. students will participate in and present results at international conferences. DIA will actively motivate and financially support their Ph.D. students to present early results at smaller international venues. Later on, DIA encourages publications at the highest quality level, which means in an international context. For more information refer to DIA's publication strategy.
- In collaboration with the Grants and Innovation Office (GIO), DIA will provide seminars, workshops and other activities for researchers, industry partners and Ph.D. students to learn about international funding opportunities and acquire the know-how to write competitive project applications.
- DIA will be active in events of the Artemis Industry Association and other international networks that Lnu is a partner in, to meet and share knowledge with universities and industry from all of Europe and beyond.
- DIA will motivate and financially support the Ph.D. students to attend international courses and summer schools in order to broaden their horizons and build international networks of their own.
- Our researchers/supervisors and Ph.D. students and industry partners will engage in international projects with research addressing and complementing the DIA research questions.
- DIA will make our Smarter Systems research relevant for collaboration with the international R&D centers of our partner companies (in cases where we do not have them in the consortium).

3.2 How to identify and address common research questions and results

How did we initially identify the research questions? We derived common research questions (RQ-1–RQ-3) from the challenges in the relevant fields of the Smarter Systems environment, see Section 1.2 and common industry objectives (IO-1–IO-3) from workshops with our industry partners, see Section 1.4.

How to address the research questions? The individual Ph.D. projects solve concrete R&D problems in the participating companies. Their results contribute to answering the research question (RQ-1–RQ-3) and to achieving the common industry objective (IO-1–IQ-3).

IO-1 adopts digital twin or/and proactive maintenance technologies to specific processes of the participating companies. We generalize the company-specific answers to questions like:

- How can digital twin digital technology be used for proactive maintenance (contributing to RQ-1)?
- How can the level of simulation abstraction in digital twins be set to meet the available data and the expected benefit of proactive maintenance (RQ-1)?
- How to make digital twin and proactive maintenance applications sustainable when the CPS and their environments change (RQ-2)?
- How can we test, validate and trust the predictions (RQ-2, RQ-3)?
- How to extract actionable knowledge for predicting maintenance needs from digital twins (RQ-3)?

Approaching IO-1 contributes to all research questions. Answers and results define concrete real-world smarter CPS solutions, i.e., concrete digital twin and proactive maintenance applications (RQ-1). They provide tested solutions and the sustainability of these solutions over time (RQ-2). Finally, they address the challenge of human control, i.e., understanding, validation, and trust in smarter CPS technologies (RQ-3).

IO-2 aims at a roadmap towards digitalization, IO-3 at creating new services, platforms, and ecosystems exploiting smarter CPS. The research projects identify common reusable artifacts addressing IO-1 along the roadmap defined in the ISPs addressing IQ-2. Some artifacts are worth packaging into general products and services building on common platforms and ecosystems. This allow to transfer the results of IO-1 internally to other divisions or to even create public business offerings and addresses IQ-3. We generalize the company-specific answers to questions like:

- How to package common processes, methods, tools, and techniques for basic data handling and digital twins and proactive maintenance into reusable services?
- How do milestones on the roadmap synergistically contribute to a common ecosystem?
- How to build and maintain an open ecosystem separating common platform services and adaptable application (domain) specific services?
- How to achieve network effects, i.e., how to organize the ecosystem that new services emerge and it becomes more valuable as usage increases?

General answers to RQ-1–RQ-3 are necessary before one can put together the reusable assets of IO-1 to common services, platforms and ecosystems, i.e., approaching IO-2 and IO-3 motivates and gives relevance to RQ-1–RQ-3.

How to handle common results and to adapt the research questions? Answers to these research questions are found in the individual research projects specific to the participating companies. Results are disseminated and discussed in the *Annual one-day DIA workshop as part of DISA's annual Big Data conference* and the *Bi-annual one-day doctoral seminar*, cf. Section 3.1. The bi-annual one-day doctoral seminar will be a forum to discuss results of common interest. It is the responsibility of the participating researchers to abstract, generalize and to publish these results in common publications. The annual one-day DIA workshop will be a forum to realign the common questions and objectives in a similar way as the initial workshops did and to adjust the Ph.D. projects and ISPs accordingly.

Table 1: Risks related to staffing and execution of DIA.

Risk	Probability	Impact	Actions to circumvent or mitigate the risks
Staff of critical importance discontinues his/her work in DIA.	Medium	High	A number of senior researchers are involved, including several full professors, which reduces the impact of this risk. Key researchers are part of operational committees, giving them opportunities to influence the direction of DIA.
Difficult to recruit qualified Ph.D. students.	Medium	Medium	We have already now 6/9 committed companies (7/10 the seats), and several others have expressed a clear interest. 4/10 Ph.D. students are already identified. Active work with our company contacts reduces this risk.
To few relevant Ph.D. courses.	Low	High	There is ongoing work at Lnu to define new courses on 2nd cycle level in the MSc Eng program and within DISA. These will be open for Ph.D. students in general to reduce this risk.
There is a shortage of necessary equipment, software, or other technology required.	Low	Low	The graduate school relies on strong connections to the DISA and Smarter Systems environments, thus we can to a large extent share technical hardware and software infrastructure.

3.3 Risk Analysis

DIA adopts a three-step scale for the probability and impact assessment of risks. In general, a low value indicates that routine procedures need to be identified and followed. A medium value indicates that the risk requires specific monitoring and response procedures. A high value indicates that specific treatment plans must be developed and reported to the project authority. Table 1 and Table 2 on page 17 summarize some of the identified risks associated with the execution of the graduate school.

4 Leadership and organization

This section describes the leadership and the organization of the project: project management, administrative support and steering group. It describes the different parties' roles and responsibilities.

DIA entities have the responsibilities summarized in Table 3 on page 18 and detailed in the subsequent paragraphs.

Steering Committee The strategic governance of DIA is done by a steering committee. The steering committee has the responsibility for the graduate school on a strategic level and focuses on strategy and quality follow ups. The steering committee meets twice a year. It consists of three company representatives, two academic representatives, and one Ph.D. student. In addition, the project manager and the project coordinator are affiliated with the steering committee. It consists of:

- Niklas Malmros, CEO at Sigma Technology Solutions, chairman,
- Margrethe Hallberg, Digitalization Coordinator for Product Introductions at Scania, member,
- Joel Cramsky, Ph.D. student at Volvo CE, member,
- Jesper Andersson, Lnu, Head of Department, Dept. of Computer Science and Media Technology (CS), member,
- Gunnar Bolmsjö, Lnu, Head of Department, Dept. of Mechanical Engineering (ME), member.

Table 2: Risks related to external factors.

Risk	Probability	Impact	Actions to circumvent or mitigate the risks
A company changes focus significantly, causing the research direction to become invalid	Medium	Medium	This is addressed from the start by formulating research questions from long-term objectives of the company. Also, increased data exploitation constitutes a stable strategy for all participating companies, so that a focus change is unlikely to leave the focus area.
Dependency on data or expertise from an industrial partner that discontinues the collaboration.	Medium	Medium	Although the Ph.D. students are connected to individual companies, the graduate school as such has a broader approach and is not dependent on a specific company.
The graduate school fails to meet the expectations of one or more industrial partners.	Low	High	Each Ph.D. student works closely with their company, and will have frequent meetings with their supervisors at the university and the company. During these meetings we will discuss the progress so far and the next steps to monitor progress and plan the studies.
The graduate school fails to meet the expectations of the funding agency.	Low	High	The executive board are responsible for continuously re-assessing the probability and impact of this risk. Mitigated by adhering to the described assessment criteria as well as involving industrial partners early in the development of the application, and by continuously adapting to input from the steering committee and the industrial partners.

Executive Board is responsible for the operational leadership of DIA and for program and research coordination. Also, it runs the daily business, coordinates activities, plans the program, handles formalities, follows up of the Ph.D. students and their projects, supports the participating industries, and handles the administration of DIA. It consist of:

- Welf Löwe,[34] Professor in Computer Science (CS), project manager, (30% of full-time employment),
- Diana Unander,[35] project coordinator, (30%),
- Morgan Ericsson,[36] Assoc. Professor in Computer Science (CS), program coordinator, (15%),
- Johan Bergh,[37] Professor in Forest and Wood Technology (FWT), Pro-Dean of the Faculty of Technology, research coordinator, (15%).

The *project manager* Welf Löwe will lead DIA and the executive board. He is supported by the board members with their respective responsibilities and controlled by the steering committee. He will be ultimately responsible for budget and result, the quality assurance, the scientific coordination of the Ph.D. modules, for successful curriculum and course development, and the operation of the Ph.D. education according to the curriculum. He is the formal point of contact towards KKS and the participating companies.

The *project coordinator* Diana Unander will coordinate activities and stakeholders. She is responsible for scheduling meetings and other activities, requesting and putting together formal reports, relationship management with com-

[34] Welf Löwe is also the head of DISA.

[35] Diana Unander is also the coordinator of DISA.

[36] Morgan Ericsson is the designer of the Master of Science in Software Engineering (MSc Eng, 5 years) and the revised Master of Science in Software Technology (MSc ST, 2 years) programs at Lnu. He is also executive board member of DISA there responsible for research in Data driven Software and Information Quality.

[37] Johan Bergh is also executive board member of DISA, there responsible for Forestry, Wood and Building Technologies.

Table 3: Entities and their responsibilities.

Entity	Responsibility
Steering Committee	Strategic governance, alignment of operation with strategic goals of companies and universities, financial and quality control
Executive board	Reporting to the steering committee, budget planning and financial monitoring, enrollment of modules, quality assurance, KPI definition and monitoring, approval of courses and program *Program coordination*: Definition of the program and courses requirements and approval of syllabi, Scheduling of courses *Research coordination*: Recruitment of Ph.D. students, match making of Ph.D. students and supervisors, progress monitoring, individual study plan approval and adjustment, supervision mentoring for the companies
Researchers responsible for courses	Reporting to the program coordinator, definition of course syllabi, conducting the courses, examination
Supervisors from industry (project lead) and academia	Reporting to the research coordinator, research and co-production in the Ph.D. modules, suggest, discuss and agree on individual study plans and monitor their fulfillment
Industry Ph.D. students	Research and co-production in the Ph.D. modules, suggest, discuss and agree on individual study plans and follow them

panies, supervisors, researchers, and Ph.D. students. She is the point of contact towards the Lnu administration.

The *program coordinator* Morgan Ericsson assures that the curriculum guarantees a progression of skills and knowledge. He defines the requirements for the syllabi so that they fit together in the curriculum of DIA. Syllabi submitted by the course responsible researchers must be checked against the requirements and approved by the committee. Finally, he schedules the courses.

The *research coordinator* Johan Bergh supports the recruitment of Ph.D. students according to academia and industry needs. He also organizes the match making of Ph.D. students and supervisors. For enrolled Ph.D. modules, research coordination monitors the progress of the students, approves and (if necessary) adjusts the ISPs and discusses problems in the executive board. Research coordination also organizes support from academia to industry for supervision mentoring of companies if needed.

Supervisors from the industry and academia have the main responsibility for the Ph.D. students and their individual study plans in close dialogue with industry. They give advice regarding the content of DIA courses and the development of new courses in close cooperation with the industry. A list of potential supervisors and their respective competences are listed Section 8 in Table 5 on page 26.

5 Company participation

This section describes which companies are expected to participate in DIA, their roles, contributions and needs.

Each participating company contributes with one or two Ph.D. modules including a Ph.D. student, a supervisor, a research topic and the required co-funding resources. Each company contributes to the joint activities including but not limited to course and curriculum development. Additionally, Sigma Technology, Scania and Volvo contribute to the formal self-organization of DIA and constitute the Steering Committee.

The subsections below summarize the companies (in lexicographic order) and their specific visions and goals as well as their research needs and benefits derived in individual workshops.

5.1 Combitech—CPS Ecosystems

Combitech has experience in the digitization of the aerospace, defence and telecommunication industries with a focus on creating ecosystems through collaboration, while protecting safety-critical systems and information.

Vision and Goal Combitech's goal is to truly utilize technology and digitalization in the industry. Therefore, business logic needs to change from a traditional model where products are sold through well-established networks to value driven models enabled through a partnership a with the customers and through ecosystem that are shaped together.

Research needs and benefits Digitalization—adopting different technologies, such as data analytics and intelligence, human–machine interaction, CPS—enables a new set of services around traditional products, enhancing their value including extended maintenance and support and insurances and financing. The industry is evolving autonomous solutions and looking for new offerings to their customers, based on the portfolio of traditional products and novel services around them. However, this transformation is yet not accelerating as the industry is too protective trying to manage digitalization changes themselves. Combitech aims at defining the future industry ecosystems (IO-3) facing challenges that need to be researched:

- How will data be shared and who will own the data—conversion to predictability?
- How will data be integrated (machine to machine interoperability)?
- Cyber-security on the higher level: How to secure the entire supply chain process and an entire ecosystem?
- How to govern the ecosystem?
- How will the next generation of industrialization and ecosystem utilize platform economy and how will ecosystems transform supply-chain economy into networks-economy?

Embedding in DIA and the Smarter Systems environment Combitech was a driving force behind the new Master of Science in Engineering program at Lnu. In DIA, the Combitech Ph.D. module will take the top down approach to smarter CPS starting from a business perspective. The Combitech Ph.D. module will contribute to the research *Theme 3 "Smarter Ecosystems"* of the Smarter Systems environment.

5.2 Scania—Product and Process Verification

Scania produces trucks, busses, engines, and related service. DIA collaborates with Scania Cab Production in Oskarshamn, where the cabs for the European market are manufactured.

Vision and Goal The Digital Vision and Goal for Scania's Production and Logistics units are: "a user friendly digital solution that allows Scania to virtually develop, test, simulate, verify and introduce new products and processes, where all deviations are eliminated before production start, with a world class lead time."

Research needs and benefits Scania needs to build up a long term solution for digital product verification according to our the above goal. Scania needs to define gaps and possibilities in earlier digital product/process verification. To this end, they need to establish the prerequisites of extracting data from CPS. Scania's aim is a 100% digital verification of product/process interfaces by adopting digital twin technologies for internal automation, optimization, and general process improvements (IO-1). Expected benefits include:

- Reduced lead-time in Product Development (PD) projects,
- Reduced start up quality issues after product introductions,
- Solutions that are easier to use.

Embedding in DIA and the Smarter Systems environment Scania is the main collaboration partner in a parallel environment initiative on smarter CPS and digital twins, which aims at investigating continuous simulation and analysis of digital twins for assessing and optimizing a production planning system. In particular, the project aims at creating the possibility to predict at runtime how every single activity in the process, e.g., production and material flow influences future steps and thereby opens up for process adaptations that mitigate deviations before they occur. The Ph.D. module will contribute to the research *Theme 1 "Assurances for Unknowns"* of the Smarter Systems environment.

5.3 Sigma Technology—Predictive Maintenance Information Service

Sigma Technology is a global supplier of product information, software, embedded design, and offshore development.

Vision and Goal The product DocFactory will integrate static technical documentation with online CPS data with goal to leverage the change from reactive to predictive maintenance, i.e., with the help of data analysis, to automatically control the optimal time when maintenance and service actions will be performed.

Research needs and benefits Sigma Technology aims at creating a novel products and services in the area of predictive maintenance and addresses IO-3. By implementing proactive maintenance service using Docfactory, Sigma Technology expects to create benefits for their customers:

- Prevent costly equipment failures: Avoid unscheduled downtime by analyzing streaming data to assess conditions, recognize warning signs and trigger preventive actions.
- Maximize uptime: Increase the efficiency of your product by strategically schedule maintenance.
- Learn to improve products: Capture and analyze data and use it to fine-tune processes, maintenance actions and make modifications that improve the product.
- Offer new services to customers: Create new business models and a better customer experience by offering proactive monitoring and predictive maintenance as a service.

Embedding in DIA and the Smarter Systems environment Sigma Technology collaborated with the DD and SA groups in the KKS Synergy project on "Software Technologies for Self-Adaptive Systems" and with the CPS group in a DISA seed project "Smart-Troubleshooting in the Connected Society". Sigma Technology is also the central collaboration partner in the Smarter Systems environment project "Restores" that is currently defined. Based on and in addition to these successful collaborations, Sigma Technology will benefit from joint research in IIoT, CPS and predictive maintenance to reach the above goals and will contribute to the research *Theme 3 "Smarter Ecosystems"* of the Smarter Systems environment.

5.4 Sydved—Strategic Big Data Exploitation in the Forest Industry

Sydved buys wood and offers full forest service to forest owners in southern Sweden.

Vision and goal Sydved takes the perspective of forest owners and their needs when they tailor solutions for active forestry. As a full-service forest company, Sydved focuses on making decisions easy for forest owners. They make sure that forest owners get the most out of their forest - on their terms. The motto Sydved follows is "It builds on trust".

Trust increases with the well-grounded and high-quality advices that Sydved gives to their forest owners. Sydved sees great opportunities in exploiting Big Data to base their consulting on even more actual and updated facts about the complex value chain from production to industrial exploitation of forest material. Additionally, they see a great potential in using Big Data for the automation and optimization of their internal processed.

Research needs and benefits Sydved's long-term goal is to strategically exploit Big Data (addressing IO-2) for

- Automation and optimization of internal processes, production and logistic chains (IO-1).
- Improved decision services for the forest owners and the forest industry allowing them to take well-grounded decisions based on up-to-date information (IO-3).

With this long-term goal in mind Sydved sees a number of challenges to be addressed in a Ph.D. project. These challenges include:

- Exploiting data from different distributed sources most notably but not limited to data from the forest owners and from SDC, the established information hub for the Swedish forest industry.[38]
- Getting trustworthy and real-time data where the data sources are incomplete, sometimes contradicting, and updated in different time intervals. Can prediction models based on probabilities and confidence intervals cope with the fuzzy nature of incoming data and information?
- Regarding ethic and legal aspects of Big Data exploitation for keeping the trust and engagement of forest owners in our operations.

Together with the expert supervisors at Lnu, with Sydved's own experts and interested experts from other companies in the industrial graduate school with similar ambitions and challenges, Sydved's Ph.D. student will research concrete projects exploiting Big Data opportunities for achieving their long-term goals - improving decision making services and self-services and automating and optimizing existing internal processes while addressing the aforementioned challenges.

Embedding in DIA and the Smarter Systems environment Sydved is involved in the DISA seed project on "Data-intensive tools for effective carbon mitigation in forestry". Sydved is specifically interested in the expertise of and the collaboration with Smarter Systems experts in CPS, VA and DD technologies, and wants results to be applied together with the Forest and Wood Technologies researchers. Their Ph.D. module will contribute to the research *Theme 2 "Self-Learning and -Explainability"* and *Theme 3 "Smarter Ecosystems"* of the Smarter Systems environment.

5.5 Södra—Digital Transformation in the Forest Industry

Södra is a member-owned global forest industry group. Södra's assignment from their 51,000 owners is to promote forestry profitability and secure the sales of their members' forest raw materials.

Vision and goal The mission is also to promote a high and valuable forest production with nature and culture conservation, as well as to guard and promote Södra's members' business interests, primarily the ownership interests. Södra's overall goal is that Södra is a more profitable, more competitive, more innovative, and sustainable company with a stronger global market position. Digitalization is a key factor for achieving this goal.

Research needs and benefits Södra will benefit from DIA in its digital transformation roadmap (IO-2). Södra's digitalization needs and benefits could be refined as:

- Forest information made available via a mobile device should be better than information available when standing in the forest. This requires a better use of relevant data in a particular context and an improved access to data and functions for the different stakeholders (IO-3).
- Being smarter, cheaper, faster, and providing better quality by automated business processes. This requires digitization of today's manual business processes and the integration of the various IT systems. (IO-1)

[38] `http://www.sdc.se`

- The forest owner should be able to directly decide and conclude their contracts with Södra without the help of an agent. Currently, the forest owner interacts with the forestry inspector (agent) who meets the owner every 1/2 to 3 years to discuss the operational plan with the owner. Then they decide, face-to-face, on the necessary activities and the Södra services involved. (IO-3)

Achieving these goals poses quite some challenges: *Data ownership* by different stakeholders inside and outside Södra and by Södra members. Ownership of data and access rights and compensation needs to be guaranteed. *Data acquisition* of heterogeneous and incompatible, incomplete, and outdated data sources, e.g., incorporate road information from the government with timely updates. *Data analysis* e.g., to identify relevant attributes to predict volume and quality of a stand to be harvested. *Machine Learning* e.g., to use historical data to predict e.g. customer demands. *Planning and Simulation* e.g., to incorporate road and weather data to improve production planning. *Digital Business* e.g., to develop a marketplace for services, wood and information. The research of a Södra Ph.D. module will address these challenges, suggest approaches and solutions, implement prototypes and evaluate ideas supporting Södra's digital transformation.

Embedding in DIA and the Smarter Systems environment Södra and the researchers of the Smarter Systems environment conducted a series of workshops and smaller joint research projects to map digitalization opportunities and challenges at all steps in the value chain from business development, sales, forest planning, production transport, to industry. Together we jointly suggested ca. 50 valuable digitalization projects and prioritized three to start with. Therefore, Södra wants to collaborate with the all relevant researcher groups in Smarter Systems (CPS, SA, DD, VA) as well as with the researchers in Forest and Wood Technologies and Mechanical Engineering. A Södra Ph.D. module will contribute to the research *Theme 2 "Self-Learning and -Explainability"* and *Theme 3 "Smarter Ecosystems"* of the Smarter Systems environment.

5.6 Volvo CE—Digital Twins for Products and Production (2 Ph.D. Modules)

Volvo Construction Equipment is a leading international manufacturer of premium construction equipment.

Vision and Goal The vision of Volvo CE is to virtualize all designs such that physical prototype verification can completely be avoided. Therefore, Volvo CE aims at predicting and verifying their products' performance.

Research needs and benefits Volvo CE aspires to create a Digital Twin for every machine (Physical Twin) that they sell fed with data from its Physical Twin through sensors and possibly aggregated with edge-based data analytics (IO-1).

Digital Twins are expected to generate benefit for sales and site optimization, requirement management, product development, and after sales. For the latter the focus on predictive maintenance and customer support for optimized operation including fleet management. Each Digital Twin will consider quality deviations from the nominal specification, and thereby define its own capacity. Edge-based analytics is used to process sensor data for each Physical Twin and send it to the Digital Twin for Cloud-based data analytics, where algorithms will calculate remaining component life and plan for maintenance for both machine and fleet level. Therefore, the two Ph.D. modules at Volvo CE will:

- develop and verify virtual models of machine behavior and performance.
- provide plant models for software, function and system development and verification.
- use optimization tools for function and system development.
- use state-of-the-art methods/models for representation of tire and soil properties.
- employ hydraulic models for vehicle simulations.
- use these models for specifying machine and component requirements.
- define scenarios for function development and verification, including functional safety and autonomous driving.
- model the simulation environment for autonomous machines including sensors for autonomous driving in simulation environment.

Embedding in DIA and the Smarter Systems environment Volvo CE has collaborated with researchers of the Smarter Systems environment in the Vinnova project on "Product Development and Open Innovation for Ecosystems of Embedded Systems (EcoFES)".[39] In DIA, they will benefit from experts in DD, such as simulation, image recognition, machine-to-machine communication, edge computing, machine learning, and in predictive maintenance. The Volvo Ph.D. module will contribute to the research *Theme 1 "Assurances for Unknowns"*.

5.7 Yaskawa—AI for Smarter Industrial Robots

Yaskawa is one of the largest manufacturers of industrial robots with 400,000 installed robots worldwide.

Vision and goal Yaskawa has a globally strong commitment to digitalization and aspires to become one of the leading companies within this domain. As IIoT progresses and machinery gets connected, Yaskawa as an equipment manufacturer is working on how to best utilize the collected information and bring value to the customer.

Research needs and benefits Technically, there are several solutions on the market to connect and gather data, and rudimentary data analysis is being done, e.g., stop time analysis, OEE calculation, etc. However, more advanced analysis in order to predict behavior and to give relevant suggestions based on interpreting the data that is gathered with artificial intelligence (AI) is still largely missing.

Industrial robots are used in a number of application areas, such as material handling, loading and unloading of machines, (de-) palletizing, spot and arc welding. They are used in a number of large companies, predominantly in the automotive industry, but also in other industries such as the aerospace industry. Lightweight robots are used to perform complex assembly line tasks, collaborate with humans, and provide real time data into the manufacturing process, creating cognitive factory environments using IoT technology. To this end, Yaskawa aims at exploiting these data streams using digital twin and predictive maintenance technologies (IO-1). In an industrial setting, such solutions offer operators and technicians more control over the management and maintenance of critical machines in the following ways:

- Use a common dashboard across all machines;
- Gain visibility into the health of all machines;
- Monitor the performance of machines in real time using a range of input and output values (operational, conditional, environmental, sensor based, structured) and respond to maintenance issues before they occur;
- Improve the efficiency of the factory's performance.

Value to the customer and for Yaskawa as a supplier also includes increased lifetime of the equipment by preventive maintenance suggestions and increased uptime of the equipment by predicting failures. Another value is being able to turn their internal service organization from a largely responsive organization with low plannability to a more reactive organization that is able to plan the resources for better utilization. By collecting and understanding the data from their connected machines, Yaskawa will also be able to feedback to their R&D and to sharpen their offerings within training to better address the actual pain-points of the customers. Using advanced AI and voice interpretation, one could imagine a more intuitive interaction between operator and machine. To this end, Yaskawa envisions the development of new services around their core products (IO-3).

Embedding in DIA and the Smarter Systems environment Yaskawa jointly worked with researchers of the Smarter Systems environment in the Vinnova EcoFES project and the KKS Synergy project on "Software Technologies for Self-Adaptive Systems". Yaskawa is also involved the KKS Industry research school "ProWood"[40] together with researchers of the Forest and Wood Technologies department at Lnu. In continuation of these collaborations, the Yaskawa module will work with our researchers in CPS, DD, and predictive maintenance. The Ph.D. module will

[39] `https://www.sics.se/projects/ecofes`
[40] `https://prowood.se`

contribute to the research *Theme 2 "Self-Learning and -Explainability"* and *Theme 3 "Smarter Ecosystems"* of the Smarter Systems environment.

6 Timetable

This section describes the overall timetable for the Graduate School.

The graduate school is executed in three phases; a recruitment and startup phase, an operational phase, and finally a graduation phase. In the first phase, the focus is on recruiting the industrial Ph.D. students, setting up the company cooperation and agreements, defining the individual study plans (ISP) for the students, defining the first year of the curriculum, and conducting a joint kick-off for all involved people.

In the operational phase, the main bulk co-production work, research and research education is performed. In this phase, the curriculum is planned, new courses are selected, defined, assigned to researchers, and scheduled. Each year, an annual workshop with all involved partners is organized to increase knowledge exchange among partners. The Lnu supervisors are responsible for an annual followup of the progress of each Ph.D. student. This followup is usually done in January / February and consists of: (i) a check of the progress for the previous year in relation to the goals in the individual study plan, and (ii) an update of the ISP with a detailed plan and goals for the coming year.

Finally, in the graduation phase, the Ph.D. dissertations are defended, and a retrospect activity concludes the graduate school for the first modules. In the parallel to phases two and three, the executive board starts planning activities to make DIA sustainable beyond the project termination. Activities include funding applications and dialogues with industry partners, thematic realignment of the research focus if necessary, etc. Table 4 shows the overall timetable of the DIA graduate school.

Table 4: General timetable for the execution of the DIA graduate school.

Year	**Planned activities**
Year 1 (2020)	*Formal admin*: Sign contracts with KKS and the companies, Recruit and enroll 10 Ph.D. students, Supervisor match-making, Define ISPs. *Organization*: Constitute the boards, first meetings. *Events*: Kick-off for the research school in April, First Annual Workshop in December. *Curriculum development*: Define courses to be developed for 2021/22. *Course development*: Foundations of CPS, Research methodology. *Teaching*: Introduction to Data Analysis, Applied Information Visualization.
Year 2 (2021)	*Formal admin*: Update ISPs, approve syllabi of "Foundations of CPS", "Research methodology" and courses for 2021/2022. *Organization*: Regular board meetings. *Events*: Bi-Annual Doctoral Seminars and Annual Workshop. *Curriculum development*: Define curriculum and courses to be developed. *Course development*: Develop three courses. *Teaching*: Adaptive Software Systems, Foundations of CPS, Applied ML, Research methodology (1).
Year 3 (2022)	**10 licentiate defences** *Formal admin*: Update ISPs, Approve syllabi of courses. *Organization*: Regular board meetings, Conduct internal half-time evaluation to track progress and to identify needed changes and potential improvements. *Events*: Bi-Annual Doctoral Seminars and Annual Workshop. *Curriculum development*: Define general study plan (ASP). *Course development*: Develop three courses. *Teaching*: CPS/SA, Research methodology (2), VA, DD.

Year 4 (2023)	*Formal admin*: Update ISPs, Approve syllabi of missing common courses, Approve ASP. *Organization*: Regular board meetings, Continuation efforts (planning). *Events*: Bi-Annual Doctoral Seminars and Annual Workshop. *Course development*: Develop one common and individual courses. *Teaching*: CPS/SA, 2xDD.
Year 5 (2024)	**5 Ph.D. defences** *Formal admin*: Update ISPs. *Organization*: Regular board meetings. Continuation efforts (proposal writing). *Events*: Bi-Annual Doctoral Seminars and Annual Workshop. *Teaching*: CPS/SA, DD.
Year 6 (2025)	**5 Ph.D. defences** *Formal admin*: Report to KKS. *Organization*: Regular board meetings. Continuation efforts (partnership and agreements). Retrospective to extract knowledge from the execution of the graduate school. *Events*: Bi-Annual Doctoral Seminars and Annual Workshop highlighting the effects of DIA. *Teaching*: CPS/SA.

7 Planned doctoral student projects

This section summarizes planned doctoral student projects.

All Ph.D. students are/will be employed by the partner company. The planned study pace is 80% for all Ph.D. students. All Ph.D. students will be enrolled in the Smarter Systems environment at the Department of Computer Science and Media Technology at Lnu. Their official third cycle subject area will be Computer and Information Science.[41] Details on the planned projects in the table below:

Doctoral student	Project	Company/Employer	Progress at start
Senadin Alisic	CPS Ecosystems	Combitech	0%
Joint recruitment	Product and Process Verification	Scania	0%
Joint recruitment	Predictive Maintenance Information Service	Sigma Technology	0%
Joint recruitment	Strategic Big Data Exploitation in the Forest Industry	Sydved	0%
Anders Gustafsson*	Digital Transformation in the Forest Industry	Södra Skog	0%
Manoranjan Kumar	Digital Twins for Products	Volvo CE	15%
Joel Cramsky	Digital Twins for Production	Volvo CE	15%
Joint recruitment	AI for Smarter Industrial Robots	Yaskawa	0%

* To be confirmed.

8 Project staffing

This section describes and motivates the staffing of the project.

Key personnel includes the project manager Welf Löwe and coordinator Diana Unander (both 30% of full time), the executive board members Morgan Ericsson and Johan Bergh (both 15%), and the steering committee members (each 5%). Table 5 lists 22 potential Ph.D. supervisors from Computer Science and Media Technology (CS), Forest and Wood Technology (FWT), Mechanical Engineering (ME), the Social Sciences (SS) and their individual competences. The 15 CS researchers contribute with competences in the smart technologies (CSP, SA, VA, DD), the three ME

[41] https://lnu.se/en/education/Ph.D.-studies/computer-and-information-science

researchers with competences in proactive maintenance and application knowledge in the heavy vehicle industry, the three FWT/SS researchers with application knowledge in the forest industry and its sustainability. The two main supervisor of each Ph.D module, from Lnu and the respective industry partner, will each dedicate 10% of full time to the project.

Table 5: List of potential supervisors and their respective competences.

Name	**Position, Field**	**Main research area(s)**
Stergios Adamopoulos	Professor, FWT	Forest-based industries, biomaterials, composites, mechanics, recycling, optimisation of manufacturing lines
Basim Al-Najjar	Professor, ME	Condition based maintenance, Condition monitoring, Decision Support System for maintenance, Industry 4.0
Jesper Andersson	Ass. Professor, CS	Software architecture and ecosystems (SA)
Johan Bergh	Professor, FWT	Silviculture, Climate change mitigation and adaptation, forest nutrient dynamics, digitalisation of forestry, Forest production
Gunnar Bolmsjö	Professor, ME	Industrial robotics, modeling - simulation - analysis - programming, reconfigurable systems, Industry 4.0
Giangiacomo Bravo	Professor, SS	Computational social sciences, sustainable development, environmental behavior, natural resource management
Mauro Caporuscio	Assoc. Professor, CS	Distributed Self-adaptive Systems: Architecture, Coordination and Control, Dependability, Simulation (SA, DD)
Morgan Ericsson	Assoc. Professor, CS	Software Data and Information quality, Quality management, Quality assessment (DD)
Francesco Flammini	Assoc.* Professor, CS	Dependable computing, resilient CPS, IoT, infrastructure security (CPS)
Johan Hagelbäck	Ass. Professor, CS	Artificial intelligence, Machine learning, Robotics and Game development (DD)
Lars Håkansson	Professor, ME	Condition monitoring, applied signal processing with focus on sound and vibration analysis, active noise and vibration control
Ilir Jusufi	Ass. Professor, CS	Information Visualization, Visual Analytics, Software Visualization, Data Analysis, IoT, Human-Computer Interaction (VA)
Andreas Kerren	Professor, CS	Information Visualization, Visual Analytics, Software Visualization, Human-Computer Interaction (VA)
Narges Khakpour	Ass. Professor, CS	Security, Formal Methods, Self-* Systems (SA)
Andreas Linderholt	Ass. Professor, ME	Structural dynamics including multibody simulations, finite element analyses and vibrational testing.
Welf Löwe	Professor, CS	Scalable computing and self-optimization; Data, information and software analysis; Composition and self-adaptation (SA, DD)
Marcelo Milrad	Professor, CS	Machine learning, Web engineering, Human Computer Interaction, Learning Analytics (DD)
Raffaela Mirandola	Visiting Professor, CS	Software quality requirements modeling, analysis, verification, Formal methods for (self-)adaptive dependable systems (CPS, SA)
Nuno Otero	Assoc. Professor, CS	Interaction design (CPS)
Diego Perez	Ass. Professor, CS	Self-adaptive systems, Quality of Service, Formal models, Cloud Computing, CPS (CPS, SA)
Sabri Pllana	Ass. Professor, CS	Parallel and Scalable computing, Simulations (DD)
Danny Weyns	Professor, CS	Software engineering, software architecture, self-adaptation, decentralization, formal methods, empirical validation, IoT (SA)

* Formal decision at the Faculty of Technology board meeting on 2019-10-18.

9 Project budget

This section describes and motivates the estimated costs in the budget.

We budget for 600 KSEK salary costs per Ph.D. student and year (altogether 30 MSEK) and 17 KSEK per year for 10% of a supervisor from the companies and Lnu, respectively (altogether 8.5 + 8.5 MSEK).[42] For the DIA management, we account for the actual costs of the Lnu staff ($\approx$ 8 MSEK).

We account for 9 courses and the curriculum to be developed each taking $\approx$ 200 hours corresponding to $\approx$ 12% of a full-time senior researcher with an average annual salary of 170 KSEK. This leads to course and curriculum development costs of $\approx$ 2.04 MSEK. For the operation of the 16 courses, we assume 150 hours per course corresponding to $\approx$ 9% of full-time. With similar salary assumption this leads to costs of $\approx$ 2.15 MSEK.

We account for 15 KSEK annual travel costs per Ph.D. student and for 100 KSEK annul travel costs for the supervisors and the DIA management. For the workshops, seminars, and other communication costs we account for 200 KSEK per calendar year (altogether 1.2 MSEK). Premises costs are based the actual current Lnu overhead on salary ($\approx$ 1.29 MSEK).

We apply for funds of 18 MSEK from KKS of which 4.75 MSEK (26.2%) are operational expenses (management, course and curriculum development, activities).[43] The minimum in-kind co-funding promised by the companies is 450 KSEK per company and year (altogether 22.5 MSEK) which matches the funding requested from KKS.

	DIA costs	**KKS funds**	**Lnu funds**	**Company funds**
Salary	55'018'539 SEK	13'510'000 SEK	15'008'539 SEK	26'500'000 SEK
PhD students	30'000'000 SEK	12'000'000 SEK	-	18'000'000 SEK
Company supervision	8'500'000 SEK	-	-	8'500'000 SEK[42]
Lnu supervision	8'500'000 SEK	-	8'500'000 SEK	-
DIA management	8'018'539 SEK	1'510'000 SEK	6'508'539 SEK	-
Other Expenses	7'932'966 SEK	4'490'000 SEK	3'442'966 SEK	-
Premises	1'292'966 SEK	-	1'292'966 SEK	-
Curriculum/course development	2'040'000 SEK	2'040'000 SEK	-	-
Operation	2'150'000 SEK	-	2'150'000 SEK	-
Communication	1'200'000 SEK	1'200'000 SEK	-	-
Travel	1'250'000 SEK	1'250'000 SEK	-	-
Sum	62'951'506 SEK	18'000'000 SEK	18'451'506 SEK	26'500'000 SEK

[42] The minimum promised co-funding allows to allocate 4.5 MSEK for company supervision corresponding to 5.3% of full-time per Ph.D. student.

[43] If management travels were considered operational expenses, the share of operational support would increase to 5.25 MSEK (29.2%).

Appendix

Short description of existing courses (all advanced, 5 ECTS)

DD: Systems modeling and simulation, (MSc Eng/MSc ST), gives an overview of different deterministic and non-deterministic modeling and simulations techniques. The following topics are covered; introduction to modeling and simulation, understand how models and simulations are used, event driven simulation, queuing theory, input modeling, verification and validation of simulation, and simulation output analysis.

DD: Code transformation and interpretation, (MSc Eng/MSc ST), presents techniques, theories, and tools used to generate code from specifications as in compilation and model driven engineering (MSE). It also discusses how these ideas can be used to define and interpret domain specific languages within MSE. It focuses on the compiler front-end and runtime interpretation of intermediate program representations.

DD: Data Mining, (MSc Eng/MSc ST), introduces data mining and its applications, for example search engines, recommender systems, and web advertisements. It covers the relationship between data mining and machine learning, text processing, clustering, and dimensionality reduction. It also considers data mining from a societal perspective with respect to, e.g., ethical questions, business value, and health.

DD: Parallel computing systems, introduces parallel and distributed computing. It covers how problems can be solved by parallel processing and accelerators (e.g., graphic processors), how problems are decomposed into parallel parts, and how programs can be optimized for different computer and accelerator architectures. Particular focus is put on machine learning applications and cloud deployment.

DD: Deep learning, (MSc Eng/MSc ST), aims at teaching theoretical and practical aspects of deep learning to enable participants to understand deep learning based solutions proposed in their respective fields of study, as well as to enable them to develop own deep learning based solutions. The course covers topics such as: machine learning basics, neural network basics and building blocks, training, common network architectures, learning approaches, and applications.

CPS: Structural dynamics, (MSc ME), addresses the fundamental theory and definitions within Structural dynamics of Single and multiple degree of freedom systems (free vibration, forced vibration, time domain, frequency domain), methods for solving Eigenvalue problems, modal orthogonality and superposition, time-step methods for transient response analyses, reduction of calculation models, state space models, and experimental dynamics.

Short description of courses that will potentially be developed (all advanced, 5 ECTS)

CPS: Predictive Maintenance: Theories and Applications, introduces predictive maintenance (PdM) from theoretical and practical point of views to make students familiar with PdM, its development, application areas, advantages and shortages, comparison with other maintenance strategies. It targets different approaches for PdM and its technical and economic impact on manufacturing machine performance and company business. The students will be able to use relevant models for modelling/developing and applying PdM for manufacturing companies.

CPS: Embedded Systems Dependability, teaches the basic concepts, methods and tools to design and certify dependable systems in critical domains, including avionics, railways, automotive, etc. The course will provide an overview of fault-tolerant computer architectures, model-based risk assessment, CPS security and infrastructure resilience. The course will also address recent AI developments to manage uncertainties in cyber-threat scenarios and to support self-healing in the Internet of Things.

CPS: Vibration based condition Monitoring for Industry 4.0, aims to enable the students to develop basic understanding of robust signal processing methods for extracting reliable information from e.g. sampled vibration time signals about machine's vibration characteristics but also about the health state of rotating machines. The course also aims to provide students with knowledge enabling them to select and apply the signal processing methods for basic information extraction on the vibration properties of a machine and its health state.

CPS: Sensors, data acquisition, signal processing, and machine diagnostics for Industry 4.0, provides students with knowledge on adequate sensors, data acquisition system and fundamental robust signal processing methods for monitoring, etc. relevant for enabling capabilities of Industry 4.0 technologies and techniques. The course also enables the students to exploit machine learning algorithms for classification (diagnose the machine) of ”Vibration features” vibration properties related to the health of rotating machines. Where the vibration properties are extracted with the aid of robust signal processing methods, feature extraction, based on sampled vibration signals measured on e.g. on rotating machines.

SA: Software Product Line Engineering, targets strategic reuse in software development, more precisely the theory and application of Software Product Line Engineering (SPLE). The students will acquire conceptual, theoretical, and practical knowledge in this field, which will prepare them for further research and development. The objective of the course is to introduce students to the foundations of software product-line engineering, including several concrete examples, which will help them identify the research frontier in the field and relate their research area to SPLE and apply this in practice.

SA: Platform-based Software Ecosystems, introduces general software ecosystems and the fundamental theory for ecosystem actors and their interactions. We focus on platform-based ecosystems, which is a backbone for the digital era’s business models. The platform-based ecosystem leverages ideas from software product lines and enables open innovation by third-parties. The course includes several real-world examples and discusses the state-of-practice and state-of-research in software ecosystems. The learning outcomes include fundamental principles and mechanisms for creating open platform-based ecosystems and related topics such as business modeling and open-innovation. Students will create business models and design and develop a platform-based ecosystem in support of that business.

DD: Statistical methods for software engineers and computer scientists, introduces advanced statistical methods in Software Engineers and Computer Scientists and to enable the participants to perform and interpret advanced analyses. The approach of the course is entirely problem-given using several datasets from computer research. After finishing the course, the participants will be familiar with statistical methods useful for software engineering and CS, be able to understand and interpret the results of advanced statistical procedures, have learned how to identify and apply methods that are appropriate for the specific context, and gain the confidence to constructively criticize the application of particular statistical methods.

DD: Multivariate analysis, is based on multivariate statistics, which involves observation and analysis of more than one statistical outcome variable at a time. In design and analysis, the technique is used to perform trade studies across multiple dimensions while considering the effects of all variables on the responses of interest. The course will provide the students with tools to do statistical analysis of complex datasets.

DD: Data Privacy and Security, addresses the concepts of privacy and security in the context of data management systems. It provides students with basic cyber-security definitions and models as well as practical information to achieve information systems certification against international standards and regulations like ISO 27001 and GDPR. The course will cover data classification and encryption standards (e.g. AES), public key cryptography and digital signatures/certificates, redundancy and backup mechanisms, user authentication and biometric identification, risk and vulnerability assessment, threat modeling, malware analysis, network monitoring, intrusion prevention and detection, and physical security.

DD: Data Quality and Management, focuses on data quality and management. Data usually requires preprocessing, cleaning, setting missing values and links before it can be used . The course covers methods to clean data, deal with missing values, but also how to store and maintain large data sets.

DD: Designing data driven applications, focuses on how to build scalable systems for data processing. It mixes distributed systems theory and practical engineering and provides a practical introduction to the different types of distributed systems, e.g., different kinds of databases, event streams, etc. It covers database scalability, replication, data models, and data processing systems architectures. It also covers deployment, monitoring, and fault tolerance.

VA: Visual Network Analytics, is crucial for many analytical problems such as social network analysis or within systems biology. This course introduces the most important techniques and approaches for visualizing networks including special cases like multivariate or dynamic networks. The focus of the course is on algorithms and systems for both network layouts/visualizations as well as on computational methods that are synergistically used together with the layouts/visualizations.

References

[1] Y. Alsouda, S. Pllana, and A. Kurti. Iot-based urban noise identification using machine learning: Performance of svm, knn, bagging, and random forest. In *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, COINS '19, pages 62–67, New York, NY, USA, 2019. ACM.

[2] J. Andersson, M. Ericsson, C. W. Keßler, and W. Löwe. Profile-guided composition. In *Software Composition, 7th International Symposium, SC 2008, Budapest, Hungary, March 29-30, 2008. Proceedings*, pages 157–164, 2008.

[3] J. Axelsson, E. Papatheocharous, and J. Andersson. Characteristics of software ecosystems for federated embedded systems: A case study. *Information and Software Technology*, 56(11):1457 – 1475, 2014. Special issue on Software Ecosystems.

[4] S. Benkner, S. Pllana, J. Larsson Traff, P. Tsigas, U. Dolinsky, C. Augonnet, B. Bachmayer, C. Kessler, D. Moloney, and V. Osipov. Peppher: Efficient and productive usage of hybrid computing systems. *IEEE Micro*, 31(5):28–41, Sep. 2011.

[5] W. Binder, E. Bodden, and W. Löwe, editors. *Software Composition - 12th International Conference, SC 2013, Budapest, Hungary, June 19, 2013. Proceedings*, volume 8088 of *Lecture Notes in Computer Science*. Springer, 2013.

[6] I. Brandic, S. Pllana, and S. Benkner. Specification, planning, and execution of qos-aware grid workflows within the amadeus environment. *Concurrency and Computation: Practice and Experience*, 20:331–345, 03 2008.

[7] M. D'Angelo, A. Napolitano, and M. Caporuscio. Cyphef: A model-driven engineering framework for self-adaptive cyber-physical systems. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceedings*, ICSE '18, pages 101–104, New York, NY, USA, 2018. ACM.

[8] A. Danylenko, J. Lundberg, and W. Löwe. Decisions: Algebra, implementation, and first experiments. *J. UCS*, 20(9):1174–1231, 2014.

[9] M. Ericsson, A. Wingkvist, and W. Löwe. The design and implementation of a software infrastructure for IQ assessment. *IJIQ*, 3(1):49–70, 2012.

[10] D. G. de la Iglesia and D. Weyns. Mape-k formal templates to rigorously design behaviors for self-adaptive systems. *ACM Transactions on Autonomous and Adaptive Systems*, 10(3):15:1–15:31, 2015.

[11] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami. Internet of things (iot): A vision, architectural elements, and future directions. *Future Gener. Comput. Syst.*, 29(7):1645–1660, Sept. 2013.

[12] A. Heberle, W. Löwe, A. Gustafsson, and Ö. Vorrei. Digitalization canvas - towards identifying digitalization use cases and projects. *J. UCS*, 23(11):1070–1097, 2017.

[13] U. Iftikhar, J. Lundberg, and D. Weyns. A model interpreter for timed automata. In *International Symposium On Leveraging Applications of Formal Methods, Verification and Validation (ISoLA 2016)*, 2016.

[14] U. Iftikhar and D. Weyns. Activforms: Active formal models for self-adaptation. In *Software Engineering for Adaptive and Self-Managing Systems*, pages 125–134. ACM, 2014.

[15] P. Jerčić, W. Wen, J. Hagelbäck, and V. Sundstedt. The effect of emotions and social behavior on performance in a collaborative serious game between humans and autonomous robots. *International Journal of Social Robotics*, 10(1):115–129, 2018.

[16] J. O. Kephart and D. M. Chess. The vision of autonomic computing. *Computer*, 36(1):41–50, Jan. 2003.

[17] A. Kerren, H. Purchase, and M. O. Ward. *Multivariate Network Visualization*, volume 8380 of *Lecture Notes in Computer Science*. Springer, 2014.

[18] C. W. Kessler and W. Löwe. Optimized composition of performance-aware parallel components. *Concurrency and Computation: Practice and Experience*, 24(5):481–498, 2012.

[19] C. W. Kessler, W. Löwe, D. A. Padua, and M. Püschel, editors. *Program Composition and Optimization: Autotuning, Scheduling, Metaprogramming and Beyond, 09.05. - 12.05.2010*, volume 10191 of *Dagstuhl Seminar Proceedings*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, Germany, 2010.

[20] J. Kirchner, A. Heberle, and W. Löwe. Service recommendation using machine learning methods based on measured consumer experiences within a service market. *International Journal on Advances in Intelligent Systems*, 8(3,4):347–373, 2015.

[21] J. F. Kruiger, P. E. Rauber, R. M. Martins, A. Kerren, S. Kobourov, and A. C. Telea. Graph layouts by t-sne. *Computer Graphics Forum*, 36(3):283–294, 2017.

[22] J. Lee, B. Bagheri, and H.-A. Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3:18 – 23, 2015.

[23] J. Lee and E. Lapira. Predictive factories: the next transformation. *Manufacturing Leadership Journal*, 2013.

[24] J. Lee, E. Lapira, S. Yang, and A. Kao. Predictive manufacturing system - trends of next-generation production systems. *IFAC Proceedings Volumes*, 46(7):150 – 156, 2013. 11th IFAC Workshop on Intelligent Manufacturing Systems.

[25] J. Lee and D. L. Porretta. Enhancing the motor skills of children with autism spectrum disorders: A pool-based approach. *Journal of Physical Education, Recreation & Dance*, 84(1):41–45, 2013.

[26] S. Liu, X. Wang, M. Liu, and J. Zhu. Towards better analysis of machine learning models: A visual analytics perspective. *Visual Informatics*, 1(1):48 – 56, 2017.

[27] W. Löwe and M. Südholt, editors. *Software Composition, 5th International Symposium, SC 2006, Vienna, Austria, March 25-26, 2006, Revised Papers*, volume 4089 of *Lecture Notes in Computer Science*. Springer, 2006.

[28] H. Meier, M. Schlemmer, C. Wagner, A. Kerren, H. Hagen, E. Kuhl, and P. Steinmann. Visualization of particle interactions in granular media. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1110–1125, 2008.

[29] J. Nilsson, W. Löwe, J. Hall, and J. Nivre. Natural language parsing for fact extraction from source code. In *The 17th IEEE International Conference on Program Comprehension, ICPC 2009, Vancouver, British Columbia, Canada, May 17-19, 2009*, pages 223–227, 2009.

[30] E. Österlund and W. Löwe. Concurrent transformation components using contention context sensors **distinguished paper award**. In I. Crnkovic, M. Chechik, and P. Grünbacher, editors, *ACM/IEEE International Conference on Automated Software Engineering, ASE '14, Vasteras, Sweden - September 15 - 19, 2014*, pages 223–234. ACM, 2014.

[31] E. Österlund and W. Löwe. Block-free concurrent GC: stack scanning and copying. In C. H. Flood and E. Z. Zhang, editors, *Proceedings of the 2016 ACM SIGPLAN International Symposium on Memory Management, Santa Barbara, CA, USA, June 14 - 14, 2016*, pages 1–12. ACM, 2016.

[32] E. Österlund and W. Löwe. Self-adaptive concurrent components. *Autom. Softw. Eng.*, 25(1):47–99, 2018.

[33] J. Papp, D. Tokody, and F. Flammini. From traditional manufacturing and automation systems to holonic intelligent systems. *Procedia Manufacturing*, 22:931 – 935, 2018. 11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, 5-6 October 2017, Tirgu Mures, Romania.

[34] M. Pappaterra and F. Flammini. A review of intelligent cybersecurity with bayesian networks. In *Proceedings of the IEEE SMC Conference 2019*. IEEE Press, 2019.

[35] O. Pettersson, M. Svensson, D. Gil, J. Andersson, and M. Milrad. On the role of software process modeling in software ecosystem design. In *Proceedings of the Fourth European Conference on Software Architecture: Companion Volume*, ECSA '10, pages 103–110, New York, NY, USA, 2010. ACM.

[36] T. Sanislav and L. Miclea. Cyber-physical systems - concept, challenges and research areas. *Control Engineering and Applied Informatics*, 14:28–33, 01 2012.

[37] K. Schwab. The fourth industrial revolution. *Foreign Affairs*, 2015.

[38] J. Shi, J. Wan, H. Yan, and H. Suo. A survey of cyber-physical systems. In *2011 International Conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–6, Nov 2011.

[39] D. Strein, R. Lincke, J. Lundberg, and W. Löwe. An extensible meta-model for program analysis. *IEEE Trans. Software Eng.*, 33(9):592–607, 2007.

[40] J. J. Thomas and K. A. Cook. *Illuminating the Path: The Research and Development Agenda for Visual Analytics*. IEEE CS Press, August 2005.

[41] M. Ulan, S. Hönel, R. M. Martins, M. Ericsson, W. Löwe, A. Wingkvist, and A. Kerren. Quality models inside out: Interactive visualization of software metrics by means of joint probabilities. In *2018 IEEE Working Conference on Software Visualization (VISSOFT)*, pages 65–75, Sep. 2018.

[42] M. Ulan, W. Löwe, M. Ericsson, and A. Wingkvist. Introducing quality models based on joint probabilities. In *Proceedings of the 40th International Conference on Software Engineering: Companion Proceeedings, ICSE 2018, Gothenburg, Sweden, May 27 - June 03, 2018*, pages 216–217, 2018.

[43] T. von Landesberger, F. Brodkorb, P. Roskosch, N. Andrienko, G. Andrienko, and A. Kerren. Mobilitygraphs: Visual analysis of mass mobility dynamics via spatio-temporal graphs and clustering. *IEEE Transactions on Visualization and Computer Graphics*, 22(1):11–20, Jan 2016.

[44] D. Weyns, N. Bencomo, R. Calinescu, J. Camara, C. Ghezzi, V. Grassi, L. Grunske, P. Inverardi, J.-M. Jezequel, S. Malek, R. Mirandola, M. Mori, and G. Tamburrelli. Perpetual assurances for self-adaptive systems. In R. de Lemos, D. Garlan, C. Ghezzi, and H. Giese, editors, *Software Engineering for Self-Adaptive Systems III. Assurances*, pages 31–63, Cham, 2017. Springer International Publishing.