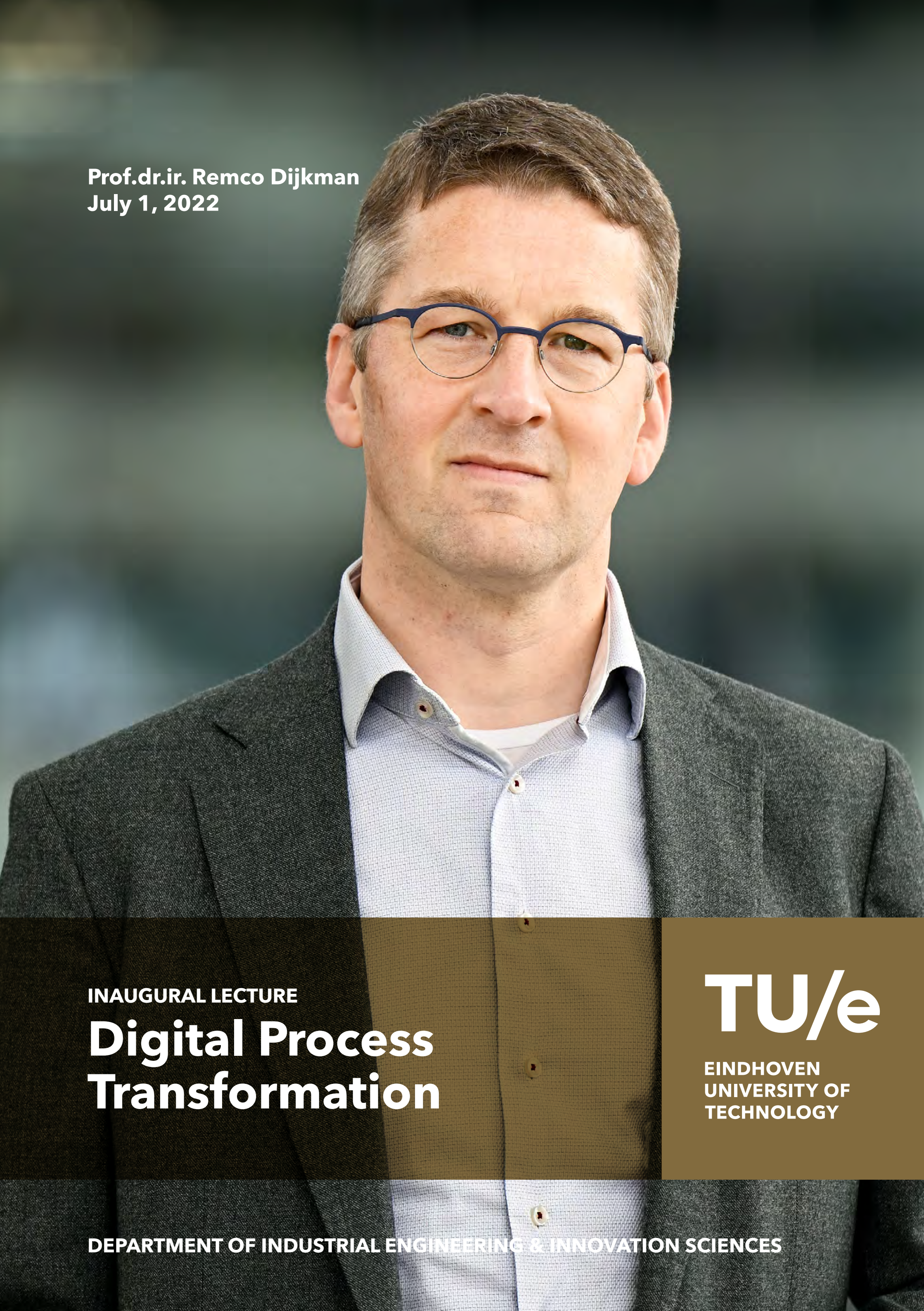
Digital process transformation



Prof.dr.ir. Remco Dijkman
July 1, 2022

INAUGURAL LECTURE

Digital Process Transformation

EINDHOVEN UNIVERSITY OF TECHNOLOGY

DEPARTMENT OF INDUSTRIAL ENGINEERING & INNOVATION SCIENCES

PROF.DR.IR. REMCO DIJKMAN

Digital Process Transformation

Presented on July 1, 2022
at Eindhoven University of Technology

Introduction

The power of computers grows at an exponential rate. This is quantified by the well-known Moore's law (Schaller, 1997), which describes the phenomenon that the number of transistors on a computer chip increases by a factor of two roughly every two years. On a practical level, this means that we can solve problems that require computing power more easily each year and that problems that were impossible to solve within a reasonable timeframe 10 years ago may be easy to solve now, simply due to the increase in computing power.

Related to this, the amount of data that we exchange between computers also increases at an exponential rate. Figure 1 shows both growth curves in a single graph, based on an extrapolation of real-world data (Transistor Count, 2022; Statista, 2022). On a practical level, the growth of the amount of data means that, in addition to computing power, we also have more and more data available to solve problems.

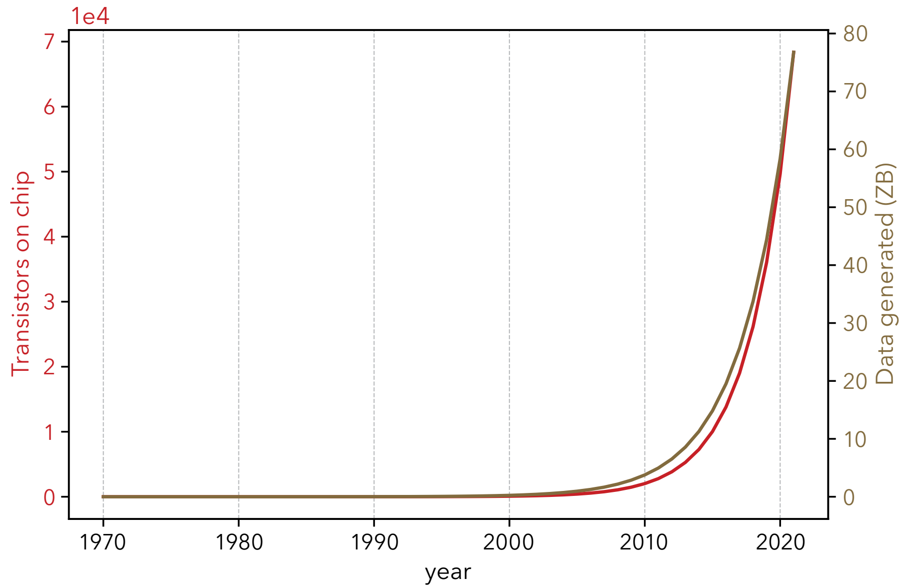


Figure 1. Exponential growth of computing power and data.

Together with the creation of increasingly smart algorithms, these developments have led organizations to 'digitally transform' the ways they operate. Digital transformation is the use of digital technology to change business models and business processes to create more efficiency, more value, or entirely new services. A famous example of a company that transformed an entire industry by going digital is Netflix, which built a business model around delivering movies and series digitally rather than physically. While less famous, existing organizations have also radically changed the way they operate – in other words, they have changed their business processes. Better data exchange has led to more services becoming digital. Think of what this has done to banking and governmental services, which are now largely available online. The availability of more computing power and better algorithms has led to more automated decision-making. This has, for example, led to highly automated warehousing and production processes.

To the greater Eindhoven region, or Brainport, one class of business processes that is of particular interest is that of business planning. For a large high-tech manufacturing company which produces thousands of different products to sell to thousands of different customers, business planning involves accurately planning when to start to produce which batch of products for which customers, being able to deliver them on time and considering the machines, personnel, and raw materials that they need to do so. Of course, to create an accurate plan, they need factual information on how to create each product, such as: what raw materials are needed for the product (also called the bill of materials) as well as the production steps that are necessary, the machines that are required in those steps, and the suppliers from which the raw materials can be procured. In addition to this, they require predicted or estimated information on how many products they expect to sell during a particular period, as well as how long they expect the production steps to take and the delivery time of the ordered raw materials to be. They can estimate that data, they can request it from customers and suppliers directly, and they can predict it based on past results.

In this lecture, I want to show how these processes and similar ones can be transformed into more efficient versions of themselves using data and smarter algorithms. In particular, I want to show that:

1. data helps to optimize processes better; and
2. you need to model the data in a manner that is easy to use for the optimization problem at hand.

I will call this particular form of process transformation 'data-driven business process optimization'. Of course, I will also show how this can be achieved, which will lead to a research agenda for data-driven business process optimization.

The Value of Data

One area in which the value of data is well-recognized is that of demand and sales forecasting. Data about historical demand can be used for forecasting but the forecast can be further improved with external data, for example about the economic situation, the weather, or trends in social media activity. The forecasts that are created with this data are required to make business plans, as described above.

For other parameters of business planning, however, data is used less intensively, while transport times, delivery times of raw materials, delivery time accuracy, production times and the probability that a finished product passes the final quality inspection are also important for creating an accurate plan. For those parameters, estimates are more frequently used than predictions that are computed based on well-prepared historical and external data. Nonetheless, data on these other planning parameters can help improve the plan in multiple dimensions, including costs but also less obvious factors such as CO_2 emissions and nervousness (the extent to which plans change more frequently), which is in turn related to customer satisfaction and employee satisfaction.

For example, in a study where we used predicted arrival times of containers, we were able to show that even relatively inaccurate predictions can help increase the number of containers that are transported by barge rather than by truck, which is both cheaper and leads to less CO_2 emissions (Gumuskaya et al., 2021). I argue that more of such studies should be conducted to show the value of data. Such studies can help justify investments into data aggregation infrastructure and can support business models for companies that want to make a business out of providing data aggregation services.

Optimizing Business Processes with Data

While data is useful for optimization in general, I will focus on a particular form of optimization: business process optimization.

Business processes are collections of related tasks that must be performed to achieve a particular goal. Organizations, including the high-tech manufacturing companies mentioned above, do their work according to business processes. For example, production is a process that relates the various production steps needed to create a product. Purchasing is a process that relates the tasks of getting quotations, placing orders and paying for those orders to obtain the materials that the organization needs to do other work. And there is also an overarching supply process that relates the production process, the purchasing process, and other processes to ensure that these all work together.

Business process optimization concerns taking optimal decisions about the cases that are currently being executed in a business process. Examples of this include choosing which case to handle first such that all customers receive their orders on time and choosing which resource should execute which task of which cases such that overall costs are minimal and all cases are treated fairly.

To illustrate the idea of business process optimization and the role that data plays, Figure 2 shows two strongly simplified processes from the medical domain: one for treating COVID-19 patients and one for treating patients with a heart condition. It shows the tasks that must be executed to treat these patients and the relations between these tasks, modeled in the Business Process Model and Notation (BPMN) (Object Management Group, 2014). At any moment in time, there may be multiple patients that are being treated according to those processes and they may be in different 'states' in those processes: one may be in the nursing ward, another may be in surgery, and yet another may be waiting to be planned for surgery. These processes illustrate optimization problems that we have come to know very well over the past years. At an operational level, we need to take decisions for the individual cases that lead to some optimal overall objective while respecting the constraints that we have. For example, we can take decisions to plan individual patients for heart surgery in such a way that, overall, we keep all patients as healthy as possible while respecting the constraints that we have on the number of doctors, nurses, and beds that we have available. On a tactical level, we take decisions on the design of the process, including the available resources, with some objective in mind and while respecting constraints that we may have, such as the amount of money that is available.

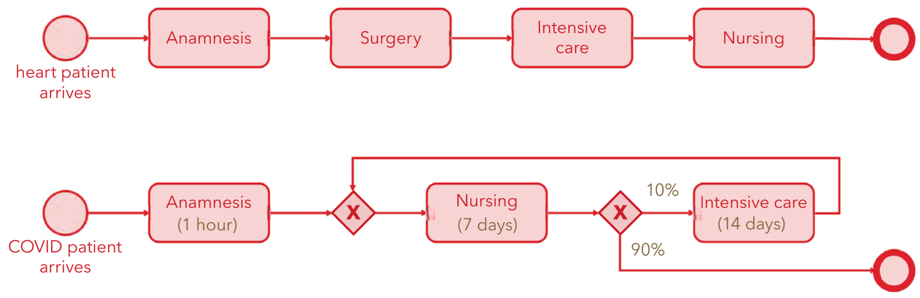


Figure 2. Simplified examples of two medical processes.

To make these decisions, we need to enhance the processes with additional information, such as: the time it typically takes to perform each task, the rate at which new cases arrive, and the probability that a case takes one path or another (in the example, the probability that a COVID patient has to be treated in the ICU). Ideally, this information is based on data about past cases.

A problem regarding the availability of information is that when we are optimizing a process, we are always doing so for some future state that we do not know. For example, we want to make a decision on whether to admit a heart patient in such a way that we have a sufficient number of ICU beds for all of our patients in the future. Consequently, we should be able to predict that future and integrate that prediction into our optimization techniques. It must be noted that these predictions are uncertain. Consequently, if we consider the fact that things may not happen as predicted and make sure that our decision is also good for likely alternative futures, we must also have a prediction of the uncertainty and make our optimization techniques work under uncertainty as well.

Against this background, techniques and (software) tools must be developed in three general areas in order to optimize business processes with data, as illustrated by Figure 3:

1. Techniques and tools must be developed that enhance processes with the information that is required to optimize them. Modeling plays an important role here: (meta)models of the business process must be designed in such a way that they contain the necessary information for doing business process optimization.
2. As business process optimization is always done for some future state of the business process, techniques and tools must be developed that can use this information to predict what that future state will probably look like.
3. Existing optimization tools and techniques must be adapted in such a way that they effectively use those data-driven predictions. From the perspective of the framework, we focus on developing mappings from business process models to existing optimization techniques.

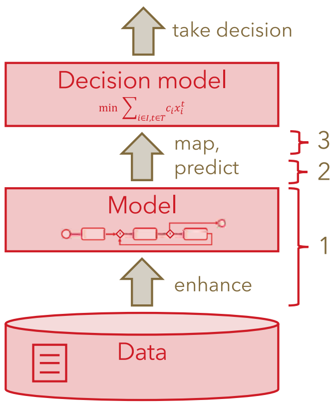


Figure 3. A framework for data-driven business process optimization.

Data Availability

A precondition for using data to optimize your business processes is that it is possible to exchange it on a technical level. Once this is realized, it must be processed and made available conceptually in a form in which it is easy to use for optimization purposes.

TECHNICAL AVAILABILITY

Clearly, a precondition to using data for optimization is that it is made available digitally. This creates specific challenges when data must be exchanged between organizations, and dedicated systems have been built specifically for this purpose. An important aspect of this is standardizing exactly what data is exchanged and how it is exchanged. Over the years, many standards have been developed for these purposes, such as EDI, HL7, and RosettaNet. As these standards are built for specific organizations to interact, they usually have a domain-specific focus and/or a geographical focus. For example, HL7 is used for information exchange between healthcare organizations and RosettaNet primarily by high-tech companies. Examples of systems that serve a specific geographical area are port community systems, which facilitate data exchange around specific ports. Portbase, for example, serves the major Dutch seaports, while APCS serves the Belgian Port of Antwerp.

Strictly speaking, the standards only describe how the information should be exchanged but do not actually do the exchange. A system must still be built according to the standard in order to do the actual exchange. Over the past years, such systems have become known as digital platforms for data exchange. In addition to facilitating the actual data exchange, these platforms provide additional services, including authorization and security, payment functionalities, and functionalities for providing apps and services on top of the data (De Reuver, Sørensen & Basole, 2018). We have been involved in the development of digital platforms for data exchange in the transportation domain, including the GET Service platform (Baumgraß et al., 2015), the IToPP platform (Eryilmaz et al., 2020), and the FENIX Network. From the GET Service platform, the start-up company

Synfioo originated, which provides both data exchange and services on top of the data exchange.

Many of these standards and platforms also have a process component such that they do not only define the messages that are exchanged but also govern the possible sequence in which these messages can be exchanged.

CONCEPTUAL AVAILABILITY

Once the data is technically available, it must be processed in such a way that it can easily be used to optimize business processes. I will call this conceptual availability. Conceptual availability is the availability of data in a manner that is fit for the problem at hand. As an example, consider the bill of materials data structure, which conveniently represents how products are composed of their parts and can be enriched with all sorts of (mined) data about how long the ordering of a part typically takes, the quality of the parts delivered by various suppliers, and how likely a supplier is to deliver on time. When readily available, this information can be used to determine which suppliers to buy which parts from in such a way that the product produced from those parts can be produced on time and with sufficient quality.

Of course, in the area of business process optimization, the business process is an important concept to collect data on so that the data can be used to optimize the business process. The business process management community has a rich history of developing the concepts that can be used to describe business processes and the BPMN has become an important de-facto standard, while a more mathematical conceptual basis is also popular in the scientific community in the form of different variants of Petri nets (van der Aalst, 1997). Of course, the two share a common basis, which is also evident from mappings that have been developed between them (Dijkman, Dumas & Ouyang, 2008) and the fact that the BPMN has Petri net-like execution semantics.

Once a business process has been mapped, it can be enriched with (mined) data. For business processes, this is also known as process enhancement (van der Aalst, 2012). The times and probabilities in Figure 2 are an example of information with which a business process can be enhanced. This information can be aggregated from historical data about patients that went through the processes in the figure.

While enhancing a business process with processing times and the likely next tasks is a good start for doing business process optimization, there are still many open research problems that can be solved in the area of process enhancement. Most notably, in business process management, data is recorded upon the completion of each task. However, mining this data and investigating the way in which it can be used in the further prediction of the future state of the process and optimization of the process has not been studied in a lot of detail. Of particular interest in this context is data that changes value while a task is being executed, such as time-series data or location updates. An important research question is how that data can be used in further analysis and optimization steps that consider tasks as atomic units. Another important open area of research in process enhancement is the behavior of the resources that execute the tasks. Questions that can be asked about research behavior include: which resource is authorized to perform which task, how each resource performs on each task (possibly depending on data that is associated with the task), and when the resource is available to work on the task according to a work or break schedule. While some research exists in this area, this research is primarily focused on mining resource behavior in as much detail as possible. According to the fit-for-purpose goal of conceptual data availability, data mined at a higher level of abstraction may already be helpful for business process optimization. The research question is then at what level of abstraction the process should be enhanced with data to solve which business process optimization problems.

Business processes interact with other data structures in business process optimization problems. Properly conceptualizing these data structures as well as their connection to the business processes and enhancing them with data is also an open research area. One data structure which is important to connect to business processes in industrial engineering is the bill of materials because the bill of materials of a product structures the procurement, production, and possibly logistics processes. The connection between the bill of materials and the optimization of these processes has indeed been and still is an important area of research. Another data structure is process choreography, which is used to represent how processes that are individually managed by different partners cooperate to enable those partners to achieve some goal together (as well as their individual goals). An important challenge in this area is that there is no single controlling party or data source, which makes it harder to determine how tasks from different partners are related to each other (Pourmirza, Dijkman & Grefen, 2017) while, at the same time, information on this is needed when developing a simulation model, for example. A final data structure that I want to mention is the structure of the assignment of resources to tasks. Little structure is provided for such data in the area of business process management while in some areas, such as transport, this is the primary data structure. For example, determining which truck (resource) to assign to which transport (task) is one of the primary problems to solve in transport resource assignment. At the same time, the process perspective can help to represent the fact that transport orders go through different phases (e.g., announced, confirmed, cancelled, in-transport, ...) and have different properties (e.g., a different level of certainty on data such as the arrival time and the pick-up location) in each of these phases. This information can be used to develop a transport plan that is more robust to changes in transport orders that are likely to occur according to the process perspective.

Using Data to Predict the Future, Learn from the Past

Once data is available, we can use it to make better decisions during the execution of a business process. We can use the data to predict what is likely to happen next so that we can optimize for it. We can also use data as experience from the past, looking at what worked well in the past to make better decisions now.

PREDICT THE FUTURE

When optimizing a business process, we are doing so for some future state of that process. Consequently, we need techniques to predict what that future state may be. From the perspective of business process optimization, there are two forms of prediction that are of importance:

1. When new customer cases will arrive.
2. How a customer case will likely progress through the process.

Both these areas are active areas of research. The prediction of case arrival is the subject of the research area of forecasting and the prediction of case progress is the subject of the research area of predictive process monitoring (Metzger, Franklin & Engel, 2012). The latter, in particular, is widely researched by the business process management community and is the main focus of our research.

For a single running customer case, predictive process monitoring can be used to predict how long the current task for that case will still take, what the likely next tasks will be, what the total duration of the case is likely to be, and what the likely outcome of the case will be. These predictions can, for example, be used to determine which collection of tasks must be performed tomorrow such that the optimal staffing for that day can be determined in the optimization step.

To be useful for business process optimization, predictive process monitoring can be extended primarily in two directions. The first extension is predictive process monitoring aimed at predicting multiple likely future states. Currently, predictive process monitoring is primarily aimed at predicting a single value. However, for business process optimization, it is often important to not only have the prediction value but also the probability distribution of likely values around it. This makes it possible to optimize a process not just for a single most likely future state but makes it robust for multiple likely future states. The second optimization-focused extension is predictive process monitoring which can also be focused on collections of cases rather than a single case. The reason for this extension is that business process optimization is always done at an aggregate level over multiple cases and resources. At the same time, predicting the joint state of multiple cases or resources is often easier (i.e., has less uncertainty) than that of a single case or resource. As an example, mining the break schedules of a single employee – and, based on that, predicting when that resource is available tomorrow to pick up a task – is very hard to do. However, the mining of resource scheduling behavior at an aggregate level (Senderovich et al., 2014) can be done with a much higher level of certainty and is sufficient to determine how many tasks of a particular type can be processed per hour.

LEARN FROM THE PAST

When optimizing a business process for some future state, there is one thing that is certain: things will go differently than predicted. Our predictions or even the information that we receive from our customers or partners may be far from reality once it happens. This includes predictions on the estimated time of arrival of sea ships (Gumuskaya et al., 2021) or customer orders (Zeevenhoven, 2020) and estimates on how long activities will take (Claessens, 2016).

These unexpected changes will interfere with our plans and, when they do, we should be able to act efficiently and effectively to keep our customers and employees happy and not disrupt our original plan too much because this will likely be inefficient. Unfortunately, this is still more an art than a science and is not supported very well by (software) tools. There are essentially two ways in which organizations respond to changes to the original plan: (1) they can be dealt with ad-hoc to the best of the ability and knowledge of individual employees who are tasked with solving the problem; or (2) the entire plan can be recomputed. Neither of these options are ideal. If individual employees solve problems in an ad-hoc manner, this takes a lot of time, which is especially problematic in stressful situations where things are not going according to plan. In addition to that, the effectiveness of the solution very much depends on the employee, which may be problematic in terms of knowledge transfer between employees. Recomputing the entire plan, of course, has the problem that not just the problematic customer

cases are reorganized but the non-problematic customer cases as well. The resulting frequent changes in the planning of customer cases are a well-known problem called planning nervousness (De Kok & Inderfurth, 1997) and this is known to lead to customer dissatisfaction as well as frustration for employees.

For these reasons, we need different solutions here. Developing these solutions is very much research in progress, but we are working in two directions.

The first possibility to support adequate responses to changes in the plan is to use solutions that have worked in the past. This is also essentially one of the first things that planners will try to do themselves. They know from experience what has happened in the past, what they did to solve the problem, and how that worked out. We can train a computer to mimic that behavior and then support the planners who must solve the problem by presenting them with the possible solutions that worked best in the past. An important research question is then how we determine which historical problems were similar given a particular planning problem. The essence of this research question is how to determine similarity within the conceptual data structure of the plan that we made. Figure 4 illustrates this with an example. The example shows a plan for a distribution round of a truck in which the truck travels from one place to another to drop off deliveries during the day. A problem that arises now is that the truck has been delayed at one of the drop-off points but still needs to make the deliveries later during the day. This is not a simple matter of moving all deliveries forward because some deliveries can only be delivered during certain time windows. Of course, it is not likely that this exact problem has happened in the past, but it is possible that similar problems happened in the past with a slightly different set of customers. The figure illustrates this with a collection of distribution plans that are similar in that respect. A more detailed research question that is relevant in this context is what information we should encode in the distribution plan that is essential for finding similar distribution plans, such as the geographical location, the name of the customer, the time spent at the location, and the time window within which we should deliver at the location. The second research question is how we compute the similarity between the distribution plans in such a way that it mimics what transport planners consider similar distribution plans. Of course, we can answer these questions in the context of distribution plans, but we can also answer them in other contexts or – preferably – in a more generalizable manner of the more general concepts of business processes.

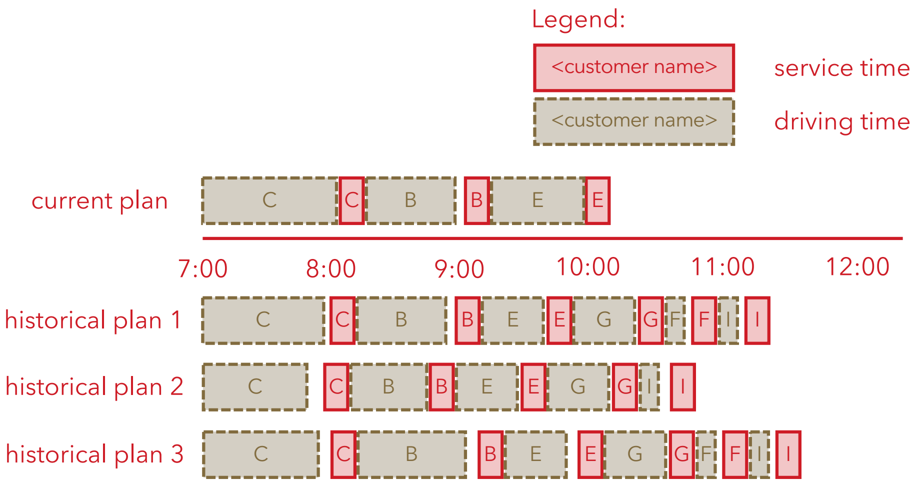


Figure 4. Example of similarity search in a transport plan.

The second possibility to support adequate response to changes in the plan is to automatically identify possible actions and calculate the effect of these possible actions and then select the action with the most desirable effect. The business process simulation models that can be enhanced with mined data, as explained before, are ideally suited for this purpose. When such simulation models are also enhanced with the current state of the business process, they can be used to calculate the effect of performing certain actions from that state. This is also known as a digital twin of the business process.

Business Process Optimization

The techniques presented so far are very suitable for calculating the possible or optimal next actions for a single customer case. However, to determine the optimal course of action for the process as a whole, the interdependencies between the customer cases must be considered as well. When including interdependency constraints between customer cases, we lift the techniques from the area of predictive process monitoring to the area of business process optimization. Business process optimization combines predictive techniques with optimization techniques in a computationally efficient manner, thus contributing to both domains. Business process optimization problems can be solved at a tactical and at an operational level.

At an operational level, constraints primarily exist on the resources that are needed to perform the cases. These resources have limited capacity and availability, which must be shared between the customer cases. In essence, this leads to an assignment or allocation problem in which we must decide which resource to assign to which customer cases during a particular timeframe in such a way that resource constraints are observed (e.g., that an employee does not work on multiple cases at the same time). Note that in the end, resources are assigned to the tasks that must be performed for those customer cases during that timeframe. I will refer to this problem as business process resource allocation.

Resource allocation in general is a well-known problem, for example in the transport domain where we have to assign transport orders to transport resources. Adding the process perspective to resource allocation means acknowledging that cases have a state depending on whether they are more or less likely to need a resource at a particular time in the future, according to which they may be treated differently. Figure 5 illustrates this with a simple example in which transport orders are announced by customers, after which a van can start driving to pick them up. In the meantime, customers can confirm their order and confirmed orders can be loaded and driven to the destination. There is a probability that orders are cancelled, but this differs for confirmed and unconfirmed orders. Consequently, if there is a choice between picking up an order in the announced state and an order in the confirmed state (assuming driving times that are otherwise equal), the order in the confirmed state is preferred. Such distinctions cannot be made if state information on transport orders is not available.

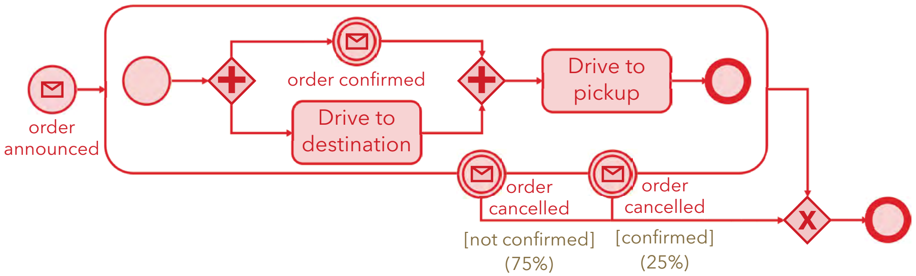


Figure 5. Simple example of transport activities with transport order state information.

Business process resource allocation can benefit from process enhancement and predictions regarding the cases that are likely to arrive, the tasks that must likely be executed in the timeframe for which we are optimizing, and the likely durations of these tasks. It can also benefit from process enhancements and predications regarding the resources that must perform those tasks, including the schedules of these resources and the times that they are likely to be occupied by certain tasks. Once these predictions are calculated, an allocation problem can be solved to assign tasks to resources. If we consider that the tasks have precedence relations with each other, this problem resembles a classical job-shop scheduling problem, which is known to be computationally complex to solve. The computational complexity further increases if we incorporate uncertainty into the prediction and if we consider that the future tasks that must be executed and the speed at which those tasks are executed depends on decisions that are made earlier on. A method that we are looking at to solve business process resource allocation problems is Deep Reinforcement Learning (DRL) and we have several active projects that solve resource allocation problems with this, including FENIX, CERTIF-AI, DynaPlex, and the AI Planner of the Future projects.

At a tactical level, there is the more general problem of how many resources are required during a particular time period in order to process any tasks that need to be executed within given constraints. Those constraints can be constraints on the amount of time within which customer cases must be ready while minimizing the costs of executing those cases. Assuming we have a well-enhanced process model, solving such a tactical business process optimization problem primarily involves simulating that process model for different numbers of resources and different

resource allocation rules. Depending on the complexity of the work allocation rules, this problem can easily become exponential in complexity (Peters, Dijkman & Grefen, 2021). For that reason, efficient algorithms are needed to traverse the space of possible numbers of resources and resource allocation rules and perform simulations for each possible combination.

Improving the Process of Business Process Optimization

The optimization of a business process is also a business process in itself. It can involve multiple tasks to collect data, clean that data, and make predictions based on it. The optimization can also involve multiple tasks in which different aspects of a process are optimized. In between the more technical data processing and optimization tasks, tasks can be included in which the resulting information, forecasts, and decisions are discussed and agreed upon between stakeholders.

There are many examples of optimization processes, especially of processes that optimize business plans. Sales and Operations Planning (Lapide, 2004) is such an example. This process optimizes multiple business plans, including the sales plan, the production plan, and the inventory plan. To create these plans, data must be collected and processed from multiple sources, including the demand forecast. Multiple meetings are held during the execution of this process, including meetings to agree on the correctness of the data and the forecasts and meetings to discuss the plans that are made. Transport planning is typically also a process that involves forecasting the transport demand, planning the transport capacity for a particular period, and then allocating transport orders to capacity. During the execution of the transport, re-planning tasks frequently have to be executed when some unexpected event happens.

These optimization processes can benefit from software support. I envision two types of support. First, scientific workflows (Barker & Hemert, 2007) can be used to compose automated tasks in a flexible manner. For business process optimization, automated tasks include the aggregation and cleaning of data, enhancing business processes with that data, and simulation and prediction tasks. Scientific workflow allows these tasks to be flexibly adapted when a new prediction task is introduced or when the systems from which the data must be aggregated change. Second, classical workflow (Leymann & Roller, 1999) can be used to compose automated tasks and manual tasks in a flexible manner. When manual tasks must be performed, people who must perform these tasks can receive automatic notifications and be supported in performing these tasks. For business process optimization, this facilitates processes in which automatically aggregated data, forecasts, and plans must be approved or adapted by people. The (scientific) challenges of developing these forms of software support are in the clear definition of the types of tasks that must be supported and the interfaces that determine the information that can flow between these tasks. Workflows can only be flexibly constructed and adapted if these interfaces are well-defined.

Impact

When research leads to new developments, these developments must create impact in practice. At the university, we primarily create impact by educating our students and through valorization activities with organizations who can use the research.

EDUCATION

The most impact that we can have as university staff is through our graduates, who enter the job market each year and are educated in the methods, tools, and techniques that are partly developed by us.

Education is, of course, an ongoing activity on which we spend approximately half of our time, including regularly updating courses to include the latest developments. But occasionally, we have to stop and think about the developments in the world around us and whether the way we teach still fits with that world. The university did that a couple of years ago (Eindhoven University of Technology, 2018) and identified trends that I think we see all around us. Some of the most important trends include, and I am paraphrasing, the fact that problems have become more multidisciplinary, that students – and society as a whole – want us to focus on societal topics with sustainability as one of the more important ones, and that information, including information for learning, is publicly and widely (even excessively) available. These trends call for changes to the way we teach, and the university has formulated some ways in which that can be done. Due to its nature, our department is ideally suited to teach in this way: we are, by nature, already multidisciplinary, we already focus on implementing technology in practice and – especially in the Innovation Sciences school – we have a strong societal focus in doing so. However, these trends do require a change in focus and in how we present ourselves.

The way we teach our courses on a daily basis will, in my opinion, be most affected by the third challenge, and I am again paraphrasing: how do we as a university stay relevant in a world where courses are increasingly available online and essentially all of the world's knowledge is in the pockets of the students in the form of their smartphones? A trend that was, of course, amplified by the corona crisis.

This again requires a shift in focus in how we teach our individual courses and, in my opinion, there are two things that we – as lecturers – can rethink: the type of learning goals that we focus on and our own role in teaching.

Not too long ago, courses were mainly set up around lectures in which lecturers taught the students what they felt the students should know. This is exactly the type of information that is now publicly available; students often value YouTube more than they do our lectures. In terms of learning goals, these are the 'after this course the student must know…' learning goals. To stay relevant and to also answer the call for more societally relevant courses, we should shift our focus to 'after this course the students must be able to solve…' type of learning goals. This is related to what the university calls Challenge-Based Learning in which students learn by solving real-world challenges using the techniques that they learn. Using an analogy, we can ask students to build a car (and, in fact, students from Eindhoven University of Technology have[1, 2]) and provide them with the physical tools to do so. Sufficiently motivated by wanting to build the car, they will also study the tools that they need for it. Some courses are already (partly) set up according to Challenge-Based Learning principles and I recognize that this setup encourages students to study the techniques that we provide them with. In education evaluations, we frequently get positive feedback on the assignments that the students do, along the lines of: "The assignment was a nice way to apply the theory in practice" (1BM05 Course Evaluation, 2022). Of course, this does not make us irrelevant as lecturers, but it does require that we rethink our roles, changing from a role where we mainly transfer knowledge into a role where we mainly set the assignments for the students and guide them through the assignment.

While we can provide students with knowledge about techniques, I see teaching students when to use which technique as another important element of Challenge-Based Learning. Therefore, we must also teach methods that explain the steps that students can perform to solve a particular type of challenge and explain which techniques they can use in which steps. The success of helping students to structure problem-solving in this manner is evident from the fact that we frequently see students structure their master's thesis assignments according to the methods

[1] https://solarteameindhoven.nl/
[2] https://www.universityracing.nl/

that we teach, such as CRISP-DM (Wirth & Hipp, 2000) and the Business Process Management Lifecycle (Dumas et al., 2013).

I also see a role for software frameworks in education in general and Challenge-Based Learning specifically. A software framework, properly structured around the steps of a development method, can help students to use and integrate the software tools that they need in the individual steps of their projects. Students can even develop their own tools that can be added to the framework. Examples of software frameworks that we are already using in our education include Torch (Collobert, Bengio & Mariéthoz, 2002) for deep learning and ProM for process mining (van Dongen et al., 2005). I have been developing a framework that can serve this purpose for Data-driven Business Process Optimization (Dijkman, 2021). The framework is built as an extension of the Business Process Management Game (Dijkman & Peters, 2019), which is a serious game in which students optimize a business process of a simulated organization. This framework can be extended to also optimize the business processes of real organizations.

VALORIZATION

As industrial engineers, we do most of our research together with the companies with whom we work together in different forms, from shorter-term collaborations in the form of master's projects to long-term collaborations with companies which have been members of the European Supply Chain Forum (ESCF)[3] for decades.

Related to education, most of our master's students do their master's project in collaboration with a company, where they apply the methods and techniques that they learned to a problem that the company has in projects that take about half a year. Companies see this as a good way to explore the potential impact of novel methods and techniques and, of course, to attract new talent. With some of these companies, we have set up a long-term plan via the ESCF in which we work on solving long-term challenges by identifying subsequent master's projects that build on each other. This form of long-term collaboration is useful for the companies, the university, and the students, and I plan to expand the number of companies with whom we have such collaborations.

[3] https://escf.nl/

Other ongoing valorization activities are the research projects of three to four years that we do together with consortia that usually involve multiple companies. Such research projects create the opportunity for methods, (software) tools, and techniques to be directly applied at the participating companies. One of the more successful projects in this context was the European-funded GET Service project, which has led to implementations at our project partners Portbase, Jan de Rijk Logistics, PTV, and two spin-off companies. One of the challenges of research projects is that they are time-restricted and there is a risk that the developments that are made in the context of the project are abandoned when the project ends. With that in mind, it is important to focus on long-term collaborations in which developments can continue from one project into the next. I also want to explore the possibilities of having an implementation capability at the university that supports companies with the implementation of methods and techniques even after a research project has ended.

As far as long-term collaborations go, the European Supply Chain Forum was founded by Ton de Kok from our department over 25 years ago. It has been incredibly successful in maintaining and expanding long-term collaborations with companies in the area of supply chain management. Through the ESCF, we collaborate with the companies in the form of master's projects, as mentioned before. In addition to this, we organize knowledge exchange activities where we update companies on the latest developments, as well as networking activities where companies can also learn from each other. The ESCF is keeping track of the current topics that companies want to learn more about. One of three current topics in the focus of ESCF companies is that of digital transformation. I will be happy to explore, together with the ESCF team, what digital transformation can mean for the ESCF and its member companies.

Conclusion

The interest of companies in digital transformation takes us back to where we started: the exponential growth of data and computing power and the development of digital platforms and data-processing algorithms which can help organizations to transform their business processes.

The digital transformation has dramatically changed the way organizations work over the past decades; much interaction that was previously done face to face is now done digitally. For better or worse, this has made interactions between organizations and between organizations and their customers much more efficient and has changed entire working cultures. However, when it comes to using the data for purposes other than just to facilitate interaction, we are just starting to scratch the surface. In most business processes, data is aggregated for optimization in an ad-hoc manner and Excel reigns supreme as a data analytics tool. To translate this into customer experience: how often did you have to repeat yourself to an organization on the phone while you were sure that you just gave someone else the exact same information? How often did you stay home to wait for a service engineer who did not show up during the agreed upon timeslot? And how often were you told that what you asked was impossible 'because the system does not allow for it'?

So let me conclude by presenting you with my vision. This vision is of organizations that always have all relevant information available in real time and in an easy-to-interpret manner, end-to-end throughout the entire process, regardless of whether that process is treating a patient, producing a microchip, or transporting a container. Such organizations will never ask you to repeat yourself and will never tell you that they cannot help you because their system does not allow for it. Moreover, they will always have the information that is needed to make the decision that is best for their customers or patients and leads to the lowest production cost, the least waste, and the least CO_2 emissions.

Acknowledgements

I have reached the end of this lecture and, at this moment, want to thank some of the people who have made it possible for me to be where I am right now.

Let me start by thanking the board of the university and of the department for entrusting me with the position and duties of a full professor and group chair. I especially want to thank Ingrid Heynderickx for recognizing that the Information Systems group has some unique characteristics in terms of personnel buildup and history and, accordingly, creating an environment in which we as a group and I as a group chair can perform at our best.

Modern-day science is very much a team effort and I am very thankful for the collaborations that I have built up over the years in Eindhoven, at other universities throughout the Netherlands, and internationally. As the list is too long to thank everyone personally and to avoid the risk of leaving people out, I relied on data to identify the people with whom I have collaborated most over the years: Paul Grefen, Marlon Dumas, Hajo Reijers, Jan Mendling, Boudewijn van Dongen, Mathias Weske, and Dick Quartel. In the last five years, this collaboration was extended to Laura Genga, Yingqian Zhang, Willem van Jaarsveld, and Ivo Adan. Of course, I must also include my PhD students in the list of collaborators. Thank you very much for all the work you put into our research: Heidi, Zhiqiang, Sander, Shaya, Volkan, Amirreza, Jeroen, Mozhgan, Riccardo, Tarkan, and Mark.

As we do most of our research together with industry, I also want to mention the organizations with whom I do much of my work. Again, as the list is too long to mention everyone, data shows that the longest-running collaborations are with Albert Charrel Ernst from Jan de Rijk Logistics, Marten van der Velde from Portbase, and Marcel Huschebeck from PTV, but I certainly also want to mention my co-workers and the partner companies from the European Supply Chain Forum, especially Tom Van Woensel for facilitating this important valorization activity.

The people from the Information Systems group deserve a special mention and I especially want to thank Oktay Turetken and Yingqian Zhang for their collaboration in leading the group as well as Hajo Reijers and Emile Aarts for their advice. And, of course, Emmy Bos for being a fantastic secretary.

To the students of the Department of Industrial Engineering and Innovation Sciences: I thoroughly enjoy teaching you the latest and greatest developments. Thank you for taking my courses. I especially want to thank the master's students who I have supervised over the years for their collaboration.

On a more personal level, I want to thank my parents for teaching me the norms and values that brought me where I am today. The two values that come to mind first – working hard and taking care of others – have helped and are helping me to be a good teacher and group leader.

Martine, over the past decades we have taken on quite some challenges together. Workwise, these challenges have included moving to Eindhoven and picking up our lives temporarily for me to work in Australia and the United States. Thank you, as well as Imre and Lauren, for taking on these challenges together, for listening to me when I get enthusiastic about some research topic, and for helping me reflect on work and on leading a group.

Then I want to thank you all, colleagues, friends, and family, for attending this lecture today and for sharing what is an important moment for me. Let us continue this over drinks.

Ik heb gezegd.

References

Van der Aalst, W. M. P. (1997). Verification of workflow nets. In International Conference on Application and Theory of Petri Nets (pp. 407-426). Springer.

Van der Aalst, W. (2012). Process mining. Communications of the ACM, 55(8), 76-83.

Barker, A., & Hemert, J. V. (2007). Scientific workflow: a survey and research directions. In International Conference on Parallel Processing and Applied Mathematics (pp. 746-753). Springer, Berlin, Heidelberg.

Baumgraß, A., Dijkman, R., Grefen, P., Pourmirza, S., Völzer, H., & Weske, M. (2015). A software architecture for transportation planning and monitoring in a collaborative network. In Proceedings of the Working Conference on Virtual Enterprises (pp. 277-284). Springer.

Claessens, T. R. (2016). Predicting throughput time [Master's thesis, Eindhoven University of Technology].

Collobert, R., Bengio, S., & Mariéthoz, J. (2002). Torch: a modular machine learning software library. Technical report, Idiap.

Dijkman, R. M., Dumas, M., & Ouyang, C. (2008). Semantics and analysis of business process models in BPMN. Information and Software technology, 50(12), 1281-1294.

Dijkman, R.M. (2021). Business Process Optimization Framework. GitHub repository: https://github.com/rmdijkman/bpo-project

Dijkman, R. M., & Peters, S. (2019). The Business Process Management Game. In Proceedings of the Business Process Management Demo Track (pp. 119-123).

Van Dongen, B. F., de Medeiros, A. K. A., Verbeek, H. M. W., Weijters, A. J. M. M., & van Der Aalst, W. M. (2005). The ProM framework: A new era in process mining tool support. In International conference on application and theory of petri nets (pp. 444-454). Springer.

Dumas, M., La Rosa, M., Mendling, J., & Reijers, H. A. (2013). Fundamentals of business process management. Springer.

Eryilmaz, U., Dijkman, R., van Jaarsveld, W., van Dis, W., & Alizadeh, K. (2020). Traceability blockchain prototype for regulated manufacturing industries. In Proceedings of the International Electronics Communication Conference (pp. 9-16).

Gumuskaya, V., van Jaarsveld, W., Dijkman, R., Grefen, P., & Veenstra, A. (2021). Integrating stochastic programs and decision trees in capacitated barge planning with uncertain container arrivals. Transportation Research Part C: Emerging Technologies, 132, 103383.

De Kok, T., & Inderfurth, K. (1997). Nervousness in inventory management: comparison of basic control rules. European Journal of Operational Research, 103(1), 55-82.

Lapide, L. (2004). Sales and operations planning part I: the process. The Journal of business forecasting, 23(3), 17-19.

Leymann, F., & Roller, D. (1999). Production workflow: concepts and techniques. Prentice Hall.

Metzger, A., Franklin, R., and Engel, Y. (2012). Predictive monitoring of heterogeneous service-oriented business networks: The transport and logistics case. In Proceedings of the Annual SRII Global Conference. IEEE Computer Society, 313–322.

Object Management Group. (2014). Business Process Model and Notation (Standard No. formal/2013-12-09). Retrieved from https://www.omg.org/spec/BPMN.

De Reuver, M., Sørensen, C., & Basole, R. C. (2018). The digital platform: a research agenda. Journal of Information Technology, 33(2), 124-135.

Eindhoven University of Technology (2018). TU/e Strategy 2030 – Drivers of Change.

Peters, S. P., Dijkman, R. M., & Grefen, P. W. (2021, October). Resource Optimization in Business Processes. In 2021 IEEE 25th International Enterprise Distributed Object Computing Conference (EDOC) (pp. 104-113). IEEE.

Pourmirza, S., Dijkman, R., & Grefen, P. (2017). Correlation miner: mining business process models and event correlations without case identifiers. International Journal of Cooperative Information Systems, 26(02), 1742002.

Schaller, R. R. (1997). Moore’s law: past, present and future. IEEE spectrum, 34(6), 52-59.

Senderovich, A., Weidlich, M., Gal, A., & Mandelbaum, A. (2014). Mining resource scheduling protocols. In International Conference on Business Process Management (pp. 200-216). Springer.

Statista (2022). Amount of data created, consumed, and stored 2010-2025. Retrieved 29 April 2022, from https://www.statista.com/statistics/871513/worldwide-data-created/.

Transistor Count. (2022). Retrieved 29 April 2022, from https://en.wikipedia.org/wiki/Transistor_count.

Wirth, R., & Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (Vol. 1, pp. 29-40).

Zeevenhoven, M. H. J. (2020). Anomaly detection in order to improve the quality of customer forecast data [Master’s thesis, Eindhoven University of Technology].

Curriculum Vitae

Prof.dr.ir. Remco Dijkman was appointed full-time professor of Data-driven Business Process Optimization at the Department of Industrial Engineering and Innovation Sciences at Eindhoven University of Technology on May 1, 2021.

Remco Dijkman received his MSc (2001) and PhD (2006) in Computer Science from the University of Twente. He then joined Eindhoven University of Technology as an assistant professor in the Department of Industrial Engineering and Innovation Sciences and later became an associate professor (2014). Currently, he leads the research on data-driven business process optimization, with a focus on applications in transport and the supply chain. Remco has published over 100 peer-reviewed papers. His work on business process execution semantics and business process similarity is highly cited and for his work on the latter he has received multiple awards. Remco serves on the editorial board of Information Systems and has been involved in a large number of research projects with industry. He is an enthusiastic teacher and has received multiple education awards.

Colophon

Production
Communication Expertise Center

Cover photography
Bart van Overbeeke Photography, Eindhoven

Design
Grefo Prepress, Eindhoven

Digital version:
www.tue.nl/lectures/

Visiting address
Building 1, Auditorium
Groene Loper, Eindhoven
The Netherlands

Navigation
De Zaale, Eindhoven

Postal address
PO Box 513
5600 MB Eindhoven
The Netherlands
Tel. +31 (0)40 247 9111
www.tue.nl/map

EINDHOVEN
UNIVERSITY OF
TECHNOLOGY