Digitalizing Teaching Processes – How to Create Usable Data with Minimal Effort

1st Raine Kauppinen[1], 2nd Altti Lagstedt[2], 3rd Juha Lindstedt[3]

[1]1st Haaga-Helia UAS, Ratapihantie 13, 00520 Helsinki, raine.kauppinen@haaga-helia.fi
[2]2nd Haaga-Helia UAS, Ratapihantie 13, 00520 Helsinki, altti.lagstedt@haaga-helia.fi
[3]3rd Haaga-Helia UAS, Hietakummuntie 1A, 00700 Helsinki, juha.lindstedt@haaga-helia.fi

Keywords
Digitalization, data management, teaching processes, thesis

1. ABSTRACT

Having enough timely and correct data is essential for decision making, and information systems (IS) are largely used to collect that data. However, the traditional form-based solutions do not always produce adequate data. Instead, since the systems do not support data collection in users' daily actions, they cause extra work, that may or may not be done depending, for example, on the users' workload. This applies especially in education processes where wrongly designed IS not only causes extra work, but also reduces the autonomy of experts (teachers). This easily leads into the situation where an IS is not used, or it is used only ostensibly to fulfill the given orders. Either way, the data collected is not reliable. In addition, situation like this gives rise to informal procedures and practices: the official and practical ways do not match, and in practical way, no or little data is collected. Since the quality of data collection is emphasized especially in AI-based decision making, and since the process digitalization is affecting education processes as well, we see it important to find out how education processes should be digitalized in order to get timely data with good quality and coverage. Regarding this, we studied and participated on thesis management system development in Haaga-Helia University of Applied Sciences, collected findings and formulated main fundamentals on creating usable data with minimal effort when digitalizing teaching process based on our findings.

2. INTRODUCTION

Digitalization and digital transformation are changing the processes and practices in all kind of business areas (Borg, Olsson, Franke, & Assar 2018), and education is not an exception. Although different kind of e-learning platforms and other learning supporting tools have been in use already several years, not much attention has been paid on digitalizing the teaching processes, and even less on information management point of view in teaching process digitalization.

When processes are digitalized, rather often two aspects are emphasized: process improvements (removing the bottlenecks, speeding up lead times, minimize the resources used etc.) and data collection. But, both aspects can be implemented incorrectly.

Traditionally, based on the ideas of Taylor (1913), the objective of the process improvement is to harmonize and standardize the process in as optimal way as possible. This, in turn, leads easily to mechanistic processes with clear repetitive tasks. However, quite often this is not the optimal target for an organization. Very strict, mechanical processes are not flexible enough if business environment changes, and they do not support expert work where experts have to be able to do independent decisions based on the situation at hand (Davenport 2010). Since teachers, especially in higher education, are experts with a large autonomy, mechanical process harmonization and streamlining is not an optimal approach for their work. Thus, digitalizing teaching processes is not as straightforward as in mechanical industries, since there have to be enough flexibility in digitalized process for the experts to do case specific decisions, and developed information systems (IS) should support that instead of restricting that kind of flexibility.

This affects also methods and practices related to how the digitalization is done: pure top-down approach in expert work digitalization project can be really challenging and may cause the project to fail. There has to be bottom-up approach as well, especially since real operational actions are not always what official process and organizational charts say (Argyris 1977; Brown & Duguid 1991). There are informal ways and organizational structures, which can be beneficial for an organization, although not always understood; developed IS should support this kind of beneficial practices and structures (Brown & Duguid 1991; Wenger & Snyder 2000). Some recommendations exist how the digitalization project should be carried out (see e.g. Lagstedt, Lindstedt, & Kauppinen 2020), but not that much is discussed on the role of information management in teaching digitalization projects.

On the other hand, if we think about data collection, it can fail as well. Normally, the data collected from the process is used for control, steering and development of the process and an organization. The better-quality data, the better tools the management has for making decisions. Because of this, the data model has been in an important role when information system (IS) is developed during process digitalization (Bocij, Greasley, & Hickie 2015; Sommerville 2011). For the management, the formal data collection can be the most important feature of an IS, with less focus on efficient and easy to use tools (if being considered at all). This is easy to see especially in the older applications which rely heavily on the forms different parties must fill. Those systems are more reporting channels causing extra work for users than real tools for the operational level practitioners. The user point of view has often been neglected.

However, although data-centric, and indeed precisely because of it, this approach does not guarantee good quality data. If the application appears to users as mere collection of mandatory forms, and using it is an extra work without any benefits for practitioners (although highly beneficial for management), the motivation to use the system is low. Since in the digitalization project, the IS development is part of the process development, and if the users (process practitioners) are not motivated to use new IS, they easily neglect the new process as well. In this case, the separation between the real actions and reported actions grows and the data inputted in the IS is questionable (see e.g. Argyris 1977). Much better quality data is possible to be collected from systems which are seen as beneficial tools, supporting and easing the work of the practitioners. That kind of systems do not have to be so form-based, since the amount of data inputted manually can be minimized, and still the necessary data can be collected by, for example, deriving it from the user actions. However, it have to be pointed out that data generated automatically into IS logs is not automatically useful or covers all the needs of the digitalized processes, but data collection (also into logs) have to be thoroughly planned when ISs are designed and developed. In addition, in education, quite often the learning supporting information systems (such as e-learning platforms) are emphasised, the teaching and management processes are usually neglected. Not much is studied how should education information systems be developed in order to get useful tools for all parties (teachers, learners and management) supporting their processes and generating meaningful data without causing extra work or other restrictions for users.

To fill this gap and find out how this kind of system should be developed, we formulated the following research question:

RQ1: What kind of fundamentals it is possible to find to guarantee the high-quality data in education digitalization project?

During 2016 – 2018 we conducted a thesis management system development project, where we used the expert oriented digitalization model (EXOD) (Lagstedt et al. 2020). One thing that was heavily stressed in the project was the quality of the data produced with the new system. Especially we wanted a system producing quality and reliable data without extra work from users. The aim was to have as user-friendly system as possible and collect all necessary data from the users' daily actions.

3. BACKGROUND

Digitalization and digital transformation are changing the processes and practices in all kind of business areas (Borg et al. 2018), and education is not an exception. However, in education sector the main emphasis is put on digitalizing learning processes, the teaching processes digitalization is rarely discussed. When teaching processes digitalization is largely neglected, also the quality data collection from the teaching processes has not received sufficient attention, which makes steering and enhancing the processes difficult. Since teaching especially in higher education is expert work, and teachers have

large autonomy and control how they organize their work, new kind models how in these circumstances to create usable systems creating quality and adequate data, as well as good examples are needed (Boehm 1991).

3.1. Motivation in a thesis process

The thesis process is one of the most challenging parts of the studies on higher level education. Not only because of its constructivist nature (see Pritchard & Woollard 2010, 89), but as a project, a thesis is very much a student-driven endeavour. In the thesis project, a student needs help from the supervisor, other teachers, librarians, representatives of commissioning company etc. For some students, the setting may be too challenging, and delays and dropouts occur. Normally, the supervisor is responsible of several thesis students simultaneously, and the work requires good skills to organize the work in order to keep track of the projects. In most cases several tools such as calendar, e-mail, Excel, assessment and other forms are required.

In Self-Determination Theory (SDT) motivation is distinguished between two categories: intrinsic and extrinsic (Ryan & Deci 2000, 55). A thesis can be considered as an example of extrinsic motivation, especially a commissioned one, because it is a clear goal and enables graduation. However, at some point, the thesis project may be halted for different reasons, not always because of the student's own. It is much easier to continue if the intrinsic motivation is high, as well, and Ryan and Deci (ibid., 55) remind that intrinsic motivation leads higher level of learning and creativity. Naturally, high self-discipline or commitment, self-directedness and a strong routine may compensate for the lack of motivation, but with most people, the inner motivation is the driving force.

Keller's (1987) ARCS model of instructional design can be used to operationalize motivation-related ideas. In learning, the first condition is attention (ibid.). Attention is built in the thesis process: students can choose topics from the area of their own interest. The second attribute in the model is relevance (ibid.). This is also covered in a thesis process both because of the topic selection and because of the importance of the thesis for the degree studies. When attention and relevance are more like prerequisites of motivation, it is also important to sustain a high level of motivation as the thesis, especially in a long-lasting project. Confidence (ibid.) is higher in some people who have a higher likelihood of success. This feature related to self-esteem can be supported by proper counselling and feedback. The fourth component, satisfaction, has its origin in behaviouristic reinforcement (ibid.). Successfully completed tasks or phases and positive feedback from the supervisor increase satisfaction.

Even if behavioural learning theories are mostly superseded by cognitive psychology and constructivism, the reinforcement appears in motivation theories (e.g. Keller 1987). Immediate feedback is the most efficient. The challenge of the thesis is that the feedback is often directed to faults and deficits in the report, which sets a need for constructive feedback that does not demotivate the student to continue. Based on the feedback of graduating students (The Ministry of Education and Culture and the Finnish National Agency for Education 2020), some students feel that they do not get constructive feedback or that feedback is given too late, when the project is in its final stage, so not much can be done if the problems are fundamental.

The students' behaviour may vary from a type of student who is highly independent with high self-esteem and is, therefore, not interested in feedback. Some may even get irritated if a supervisor is too keen on giving feedback (see also Keller 1987). Illeris (2009, 16) even mentions mental resistance, which may block or distort learning. In a thesis work, a student may have already put all the effort into the report, and feedback that would require changes may be too much to handle. The other extreme is a student who is unsure about their decisions, so they continuously want feedback on all the details. All students may occasionally be in a situation where the guidelines and knowledge are insufficient to complete a task, so the feedback from the supervisor is vital. Without the response from the supervisor, a student may halt the process. Therefore, it is vital for a supervisor to manage the feedback and keep it at optimal level.

3.2. Data collection in IS development

Traditionally, information systems and their digitalization has often been seen as database centric form-based solutions that are used to input data into database and query (report) the data from the

database as needed (Bocij et al. 2015; Sommerville 2011). Therefore, the focus of data modelling has been from the administrative viewpoint and on creating the necessary forms in order to get the data from different user groups in different phases of the administrative process related to collecting data.

This kind of focus leads often in the situation where the (administrative) information system is separate from the operational work that is either still done by using manual information system or, as nowadays often is the case, by other operational system. This leads to more work on the operational level if the data to the administrative system has to be inputted manually, or to building and maintaining the necessary interfaces between systems for data transfers. This is still a risk, even if more modern product centric approach (e.g. Sommerville 2019) is taken.

While it is possible to maintain different systems for administrative and operational work and data, this also leads to the larger number of separate systems and duplicated data. Also, if the data collection is not integrated to the operational level work and business processes executed on that level, the reliability of the data collected may not be very high. This is because the use of separate administrative system or inputting data for it is often considered as additional workload instead of a part of the daily operational work.

Known possible solutions for these challenges are, for example, systems analysis and master data modelling as a part of an enterprise architecture (Lankhorst 2017; Ross, Weill, & Robertson 2006). Systems analysis focuses on the goals and needs of both the organization (for example, via business process modelling and development) and end users (for example, via behavioural modelling) (Ashrafi & Ashrafi 2014). Master data modelling focuses on identifying the information (data) necessary across the organization and several processes. as well as the information (data) necessary only for a specific part of the organization or for only for a specific process (Lankhorst 2017; Loshin 2010).

In digitalization of teaching processes these approaches should be combined in a feasible way (Ashrafi & Ashrafi 2014; Bocij et al. 2015). This enables digitalization to improve both the administrative and operational level and processes in an organization, as well as the related information systems and data collection. This kind of an approach has potential to also improve the level of utilization of both the information systems and data and to improve the quality of data available.

4. METHODS

Haaga-Helia thesis process was digitalized from 2014 to early 2019 using the expert oriented digitalization model (EXOD (Lagstedt et al. 2020)). To evaluate the results of the digitalization we did a case study research during Spring 2019 (Lagstedt et al. 2020). The EXOD model has relation to both agile development such as Scrum and the systems development lifecycle (SDLC) model and our work in this paper, in addition to the interview focus on the earlier case study, has action design research (ADR) characteristics (see Sein, Henfridsson, Purao, Rossi, & Lindgren 2011).

4.1. Process digitalization

In the process digitalization, the four steps of the EXOD model were applied as follows:

1. Initiation. By the end of 2014, the existing thesis process and all its versions and variations were described. After the existing situation was comprehensibly understood, potential benefits of digitalization were listed and communicated to all stakeholders.

2. Process re-engineering emphasis. During years 2014-2017, the different ways to achieve the benefits listed in step 1 were discussed and new process was developed.

3. IS development emphasis. Starting 2016, the development of an information system (IS) supporting the new process developed in step 2 was started 2016, so the process development and the IS development were partly overlapping. The new system was developed using the agile approach (Beck et al. 2001) and lean principles and it was piloted with the new process in autumn 2018.

4. Stabilization. Staring from 2019, the new process and the IS were taken in full use starting the stabilization phase.

In this case, the IS development followed the agile approach (Beck et al. 2001) and the lean principles, but in general, the IS development in the EXOD model can follow agile methods such as Scrum, but also other development methods are applicable. Also, the EXOD model cover systems development lifecycle more comprehensively than only the IS development that is a part of the lifecycle. For

example, in SDLC model, IS development emphasis of the EXOD model corresponds the system design and system build phases while initiation has similarities to the initiation and feasibility study phases in SDLC. The process re-engineering and stabilization in EXOD correspond the systems analysis and system implementation and changeover phases in the SDLC model. Stabilization in EXOD also has some characteristics of the review and maintenance phase in the SDLC, since changes to the digitalized process and to the supporting IS are made as necessary in EXOD.

4.2. Research method

To evaluate the results of the project we did a case study research during spring of 2019, following the recommendations of Yin (Yin 2009). We used four data collection sources extensively as Yin (Yin 2009) recommends, namely: documentation, archival records, participant-observation, and interviews.

In total, 27 experts were interviewed. The interviewees were chosen based on their above-average activity around Wihi (or Konto, as it is called at Haaga-Helia UAS) at different stages. As the interviewees had extensive experience in the work and roles they were representing, the interviews can be described as expert interviews (Bogner et al. 2009). The experts who were interviewed included the Director of Corporate Planning and IT Services, the Head of Library Services, the Head of Student Services, the Manager of Education Services, degree program directors (2), thesis coordinators (5), thesis supervisors (8), student office secretaries, (2) and students (6).

The interviews were conducted by applying an interview method protocol that was developed by Dahlberg, Hokkanen, and Newman (2016). During the interview, questions were presented to the interviewee, either face to face, or via video call. The interviewer recorded and presented the noted responses to the interviewee immediately before moving to the next question. The act of visible transcription the responses gave the interviewees the ability to validate the typed answers immediately. The method allowed us to easily continue to conduct interviews until data saturation was reached, because we were able to assess saturation after each interview.

In our pervious analysis (Lagstedt et al. 2020), the main emphasis was on the interviews and the other sources were considered complementary. However, because one of the researchers was responsible for the thesis process development and another for the development of the IS (Wihi) supporting it, we had access to the thesis process development, as well as to all Wihi's development documentation (process models, notes, product backlogs, version history, plans, emails, and guidelines). We also utilized Wihi's logs and registers as supporting data to understand the actual usage of the IS. In addition, as supervisors and thesis coordinators, we used and guided the use of the digitalized process and made participant observations during the process.

In this paper, the main emphasis is based on the observations during the thesis process development, as well as Wihi's development based on the aforementioned related material available for us. In addition, we have used the collected information and findings from the interviews from to the extend they apply in creating usable data during the digitalization of the thesis process. It is also worth noticing that since the development has been iterative, especially regarding the development of the IS, but also regarding the thesis process, the method used has characteristics action design research consisting of four stages (Sein et al. 2011) as illustrated in Figure 1.

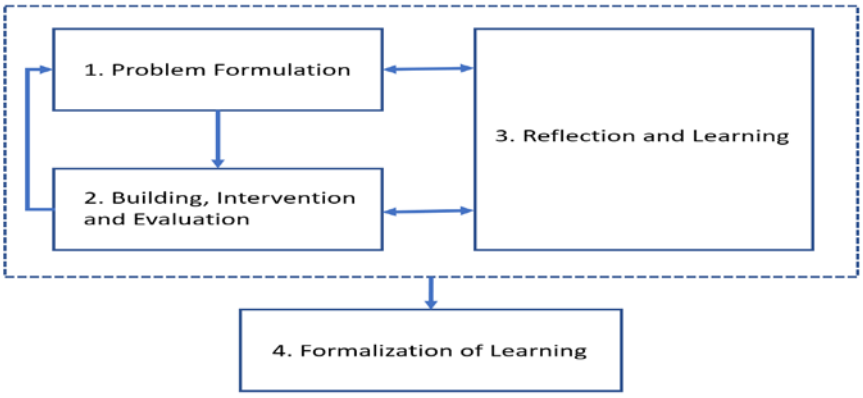

Figure 1: Action Design Research (Sein et al. 2011)

Here, the trigger for the stage one (problem formulation) has been problems perceived in the thesis process, for example related to the creation of useful data during process. Based on stage one, Wihi has been designed interweaving the building of the IS, intervention in the organization, and evaluation in stage two (building, intervention and evaluation). Parallel to these stages, stage three (reflection and learning) has been ongoing continuously throughout the development resulting in general solution concepts based on the originally perceived problems in stage 4 (formalization of learning).

5. FINDINGS

Before the development of formal data collection tools, the thesis coordinators had collected thesis status information by emails, corridor discussions and phone calls. Because the data collection was so laborious, and gotten answers so vague, the real situation was difficult to see. This caused problems to be found too late and resources wrongly allocated.

The development of formal tools can be divided to pre-Wihi development time and Wihi development time. During the pre-Wihi development time rather traditional IS development approach was used: the main purpose was to collect enough data about the progress of thesis writing to support and control decisions. So, the management perspective was emphasized, and the assumption was that it is enough if advisors just have a list of theses they are advising. For this, a simple Access Forms -based solution was developed, and advisors were required to update the status of their theses once a month. Rather soon it was found out that this system does not work. Advisors, especially in the rush time of a period, prioritized other work higher, and filling the form was easily postponed and forgotten. Thesis coordinator had to use quite much time to remind advisors, and still some data was all the time missing. So, the situation did not improve much, and advisors felt that they are forced to do extra work, especially in situations when they were overworked already. Access Forms, although having very simple form for advisors, was also found to be very cumbersome to use.

Based on the findings and feedback of the first version, a mobile phone application was developed. The basic idea was the same than in the first version (data collection approach) but having the application easy to use in phone was planned to lower the threshold to use the application. In a way this was a success: the pilot group found the application easy to use, but they also found that having more features helping advisors to supervise the theses could significantly help advisors' work and improve the motivation to use the application. It was also found out that students should have a user interview as well, so that the tool could be a project management tool for them. However, the technical platform did not support the features desired.

At this point we had several discussion and workshops with different parties involved to thesis supervising, and it was found out that we should have more comprehensive perspective than just data collection. We collected needs and requirements from all parties involved, as well as our experiences from earlier versions, and started Wihi development from a totally new perspective.

The main idea of Wihi development was that only minimal amount of data should be inputted manually, without reducing the amount and quality of the data collected. Instead, the presumption was that if the tool is useful for all parties, it will be used more actively, and quality of data collected will improve. It was found out that it is not only an IS development case, but a thesis process digitalization case as well, so, the process was modified as needed. The development was done in small steps and feedback and new requirements from all parties (advisors, students, thesis coordinators, student office, program managers, assistants and library) were collected continuously during the development, the main question being how the thesis process should be digitalized to help the work as much as possible. In the development, EXOD model was applied, and since it was a novel model, the development was done as a research project, as pointed out in Chapter 4.

From the observations and interviews, which we did during the development and after the implementation, we found out that most of the advisors and management staff are satisfied with the tool. It simplifies advisors' work, especially if they have a big number of students to advice at the same time, by giving one clear view to all projects going on and their status and relevant data (communication, versions, feedback provided). Digitalized thesis process does not force advisors to follow exactly the same workflow every time and advisors do not have to fill out extra forms. Instead, for example, they accept the thesis phases by just clicking one button. It causes some work for advisors to do, but before, they had to do most it most of it other media and tools, so using Wihi is not considered as extra work.

It was carefully analyzed what is the minimum amount of data really needed and how it can be collected with minimum effort. Now the data comes mostly during normal actions of the users such as version saving, comments, phase acceptances. Still, the collected data helps thesis coordinators (and program managers) to, for example, to follow what is the overall situation, whether the theses are in their schedule, is there need for extra arrangements for students having major problems and how many students will finalize their thesis at a given time. Students do not see the use of the IS (Wihi) as extra work either, but are using it as a project management tool, and it helps them to keep track of versions of the thesis documents, tasks and communication with their advisor.

6. DISCUSSION

The developed thesis management system, Wihi (aka Konto), was taken into use in the beginning of 2019. Based on the feedback we collected in the end of the spring semester we found out that the system was considered easy to use, and data collected by it useful. We see that the development concept (EXOD) was good for this case, but more importantly we claim that the basic ideas of the developed system, Wihi, can be utilized also in other teaching process digitalization projects.

The development of the thesis management system was started with clear data-centric approach: the main emphasis was to collect the theses situation from supervisors, and for that a solution with suitable forms was developed. When advisors, although agreeing that this kind of information is important for decision making, were reluctant (or too busy) to use the system, a new, easier-to-use mobile version was developed. However, this did not improve the situation much, and it was found out that the approach has to be changed: instead of developing a data collection tool for management, we should develop a tool for the actors of the process (students, advisors and coordinators) that does not cause extra work for them but helps them to success in their work. Since this approach was found out to be more effective and improving the quality and coverage a collected data, we claim that there is a clear paradox: data-data centric approach seems to produce less and lower quality data than the non-data-centric approach.

Based on this finding, we were curious to find out that if data-centric is not good approach, how should be proceed and what factors should be taken into account when a new education process supporting IS is developed. As an answer to our research question, *RQ1: What kind of fundamentals it is possible to find to guarantee the high-quality data in education digitalization project*?, we formulated the fundamentals presented in Table 1 and consider them to be the most important ones.

Table 1. The main data collecting fundamentals and meanings in education digitalization.

The main fundamentals		Meaning in education digitalization
1	**It has to be understood what data is relevant.**	In expert work, such as teaching and advising, the results are more important than the (mechanical) actions.
2	**The origin of data must be understood.**	For example, teachers should not be required to fill forms when data is derivable from students' actions.
3	**There are limits to who and how much can be affected.**	For example, in some cases it is much easier to get students to do something than the teachers.
4	**No extra work should be caused by collecting data just in case.**	A small amount of data is better than no data, but from process point of view too much data (collection) is worse than no data. Too much efforts used in data collection is away from the other actions, and if quality of process lowers, easily the quality of data lowers as well.
5	**A well-functioning process is the priority.**	For all the user groups, the main emphasis should be facilitating the workflow while inputting the data should be a secondary thing. A well-functioning process generates better quality data than a process which only exist on paper.
6	**The developed IS is a tool for experts, "the tail should not wag the dog".**	Teaching is autonomous, so the teachers must have the control instead of an IS.

Because of the strong result orientation (fundamental 1), the collection of data must be as automated as possible during the process, especially from the result point of view. This requires that the operational level IS collects the data during the process instead of the expert (or someone else) required to input the data separately in a different system. Where possible, the IS should also generate the data instead of requiring manual input. In Wihi certain predefined steps form automatic log data allowing statistical analysis of individual or group performance in thesis process. Based on the data we now know e.g. the mean and variance in length of different type of theses. These ideas have also close relation to the ownership of the data (fundamental 2), since in the educational setting both the teachers and the student own their own data, and the data should be asked from the owner where necessary and derived from the owner's actions whenever possible. It is also important to consider which stakeholder is most likely to provide the data if asked (fundamental 3). In Wihi the data created by routine actions is transparent to students and thesis supervisors, and the data helps students to monitor if they are following their original schedule, and the same information serves supervisors to detect when there is a need for intervention. The same data helps to objectively assess process performance of students when final assessment is done.

Collecting only the essential data (fundamental 4) helps to focus the development of the process and IS regarding the data collection and additionally, makes it more probable that the data needed is actually gotten (coverage) and correct (quality) even if it is necessary to have it inputted manually by the owner. The same idea is behind making inputting the data secondary thing (fundamental 5) compared to actually facilitating the workflow. In addition, flexibility in the workflow is needed meaning that both the process and the IS cannot be too restrictive (e.g. by enforcing excessive amount of data to be manually inputted), because from the expert's point of view this means the loss of control hindering the autonomy (fundamental 6) that is considered to be an essential characteristic in the expert work in educational domain. In Wihi it was carefully considered during EXOD which are the steps and actions that are relevant, i.e. which can produce useful and reliable data, as well as are pedagogically important. The skeleton of the thesis process was kept clear and simple enabling great flexibility, and against some wishes, many details were excluded because the data would have minor usability but had cumulated excessive amount of numbers to be delt with. Since only the order of crucial steps are fixed in Wihi, not the way of work or timing, the IS does not reduce the process autonomy of the users.

It is also worth noticing that while the aforementioned fundamentals minimize the effort in creating usable data when digitalizing teaching processes such as the thesis process, another important aspect is the improvement of the coverage and quality of the data collected. In Wihi the data is never fed to the system for a sake of data, but the essential steps in the process form usable data automatically. This also disables the manipulation of data, I.e. the data is authentic and not prone to intentional or unintentional bias. Good coverage and quality of the available data are important prerequisites for current and emerging ways utilizing data, such as business intelligence (BI) using artificial intelligence (AI) based methods or robotic process automation (RPA).

Based on our findings, we see feasible possibilities to advance into these areas in the future work and research. This is because the current thesis process and the stabilized use of the developed IS result in continuous and automatic creation of usable data of good quality and coverage that does not require extra effort in data collection. Instead, the effort can be focused on utilizing the data in the ways needed, some of which may be novel in the sense that they were not foreseen when developing the current version of the thesis project or the IS.

7. CONCLUSIONS

In this research, we followed and participated to thesis process digitalization in a university of applied sciences. During the development, the emphasis of the development shifted from data-centric approach to process-centric approach. One reason for the shift was a paradox we found: even though traditional data-centric approach emphasizes the quality and coverage of collected data, the approach itself may cause the data collection failure. This is not necessarily because the forms of the IS are not collecting right data from right actors (which they most obviously are), but because causing extra work for all parties resulting in that the IS is not used, or it is used reluctantly in most minimal way when suitable free time slot appears.

We found out that the process-centric approach, paradoxically, produces much better quality, on time and coverage data than the data-centric approach. However, we also found out that developing process-centric data collection is much more challenging than the traditional data-centric approach. The process-centric data collection approach is tightly tied to process digitalization, and to be successful it means that also the process digitalization must be successful. Although these two are highly intertwined, it is possible to see vital data collection fundamentals to be taken into account in education process digitalization. In this study, we found six main data collection fundamentals (see Table 1), and although some of them sound rather self-evident, they are easily forgotten in traditional data-centric development. It is highly important that these fundamentals are not only known but also understood and carefully applied when education processes are digitalized.

We see the found paradox and the fundamentals presented in Table 1 as main contribution of our study. Based on these findings, we have already started a new education process development project. Our next action design research project focuses on digitalizing the work placement processes in Haaga-Haaga University of Applied Sciences, and there we will apply the fundamentals presented in Table 1, as well as EXOD model presented in our earlier work (Lagstedt et al. 2020).

However, we claim that the main ideas behind the presented fundamentals (Table 1) are applicable in other business areas and industries as well, and in future studies it will be interesting to see if these fundamentals, even not written out, are possible to be found from the successful process digitalization projects, or if they are lacking in the failed ones.

In addition, we see that this approach improves the quality and coverage of the collected data, since the data will be collected timely and imperceptibly. Thus, the collected data can be used not only to control day-to-day actions, but in more complex analysis as well. As generally known, the quality of data is one of the biggest obstacles when higher level analyses, such as BI, RPA and AI, are established. Now, using the fundamentals presented (Table 1), it is possible to achieve reliable online data for more sophisticated and advanced education process analysis. A study related to this has been already started, and we are optimistic to find out the new ways to improve the processes and management practices based on the automated and sophisticated AI analysis tools.

8. REFERENCES

Argyris, C. (1977). Organizational Learning and Management Information Systems. *Accounting, Organizations and Society*, *2*(2), 113–123.

Ashrafi, N., & Ashrafi, H. (2014). *Object oriented systems analysis and design*. Harlow: Pearson.

Beck, K., Beedle, M., van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., … Thomas, D. (2001). Manifesto for Agile Software Development.

Bocij, P., Greasley, A., & Hickie, S. (2015). *Business information systems: technology, development and management* (5th edn). Pearson Education.

Boehm, B. W. (1991). Software Risk Management: Principles and Practices. *IEEE Software*, Vol. 8, pp. 32–41. https://doi.org/10.1109/52.62930

Borg, M., Olsson, T., Franke, U., & Assar, S. (2018). Digitalization of Swedish Government Agencies — A Perspective Through the Lens of a Software Development Census. *International Conference on Software Engineering*.

Brown, J. S., & Duguid, P. (1991). Organizational Learning and Toward a Unified View of Working , Learning , and Innovation. *Organization Science*, *2*(1), 40–57.

Dahlberg, T., Hokkanen, P., & Newman, M. (2016). How Business Strategy and Changes to Business Strategy Impact the Role and the Tasks of CIOs: An Evolutionary Model. *Proceedings of the Annual Hawaii International Conference on System Sciences*, *2016-March*(January), 4910–4919. https://doi.org/10.1109/HICSS.2016.609

Davenport, T. H. (2010). Process Management for Knowledge Work. In J. vom Brocke & M. Rosemann (Eds.), *Handbook on Business Process Management 1* (2nd ed., pp. 17–35). Springer Berlin Heidelberg.

Illeris, K. (2009). A comprehensive understanding of human learning. In K. Illeris, P. Jarvis, R. Kegan,

Y. Engeström, B. Elkjaer, & V. Mezirow, J. . . . Stroobants (Eds.), *Contemporary Theories of Learning: Learning Theorists ... in Their Own Words* (pp. 7–20). Routledge: Abingdon.

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, *10*(3), 2–10. https://doi.org/10.1007/BF02905780

Lagstedt, A., Lindstedt, J. P., & Kauppinen, R. (2020). An outcome of expert-oriented digitalization of university processes. *Education and Information Technologies*. https://doi.org/10.1007/s10639-020-10252-x

Lankhorst, M. (2017). *Enterprise Architecture at Work - Modelling, Communication and Analysis*. Berlin: Springer.

Loshin, D. (2010). *Master Data Management*. Burlington, MA: Morgan Kauffman.

Pritchard, A., & Woollard, J. (2010). *Psychology for the Classroom: Constructivism and Social Learning*. Florence: Routledge.

Ross, J. W., Weill, P., & Robertson, D. C. (2006). *Enterprise Architecture as Strategy*. Boston, Massachusetts: Harvard Business School Press.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67.

Sein, M. K., Henfridsson, O., Purao, S., Rossi, M., & Lindgren, R. (2011). Action Design Research. *MIS Quarterly*, *35*(1), 37–56.

Sommerville, I. (2011). *Software Engineering* (9th ed.). https://doi.org/10.1111/j.1365-2362.2005.01463.x

Sommerville, I. (2019). *Engineering software products*. Hoboken, NJ: Pearson.

Taylor, F. W. (1913). *The Principles of Scientific Management*. New York and London: Harper & Brothers Publishers.

The Ministry of Education and Culture and the Finnish National Agency for Education. (2020). *AVOP 2020. The education administration's reporting portal "Vipunen."* Retrieved from https://vipunen.fi/en-gb/

Wenger, E. C., & Snyder, W. M. (2000). Communities of Practice: The Organizational Frontier. *Harvard Business Review*, (January-February), 139–145.

Yin, R. K. (2009). Case Study Research: Design and Methods. In *Essential guide to qualitative methods in organizational research* (4th ed.). https://doi.org/10.1097/FCH.0b013e31822dda9e

9. AUTHORS' BIOGRAPHIES

Raine Kauppinen holds a MSc in computer science. His expertise relates to software development as well as business and ICT co-operation. He has worked at Haaga-Helia as a Lecturer from 2007 to 2012 and since 2018 and has also worked in the public sector in software development and acquisitions, project management and ICT architecture related fields. His current research interests include, for example, digital transformation and information systems development methods.

Altti Lagstedt holds a PhD in information science and an MSc in technology. His expertise relates to information systems development methods, software development and business process digitalization, for which he has written several academic and practical publications. He has worked in Haaga-Helia as a Lecturer from 2004. Before the current position, he has been a software developer and a project manager at a software development company Arts and Minds Ltd. www.linkedin.com/in/altti-lagstedt-0272854

Juha Lindstedt holds a PhD in education and MSc in geography. The expertise areas are theses and thesis process, quality of education and entrance examinations. The current publications are from the same areas. He has worked 10 years at the University of Helsinki, and since 1997 at Haaga-Helia UAS and its predecessors (in total over 22 years). The title has been a Principal Lecturer since 2015, between 2007 and 2015 he worked as the Quality Manager of the university. www.linkedin.com/in/juha-p-lindstedt-b7abb