

STRAWBERRY ANALYSIS

WRITING A PAPER-BASED PHD

Wanda Presthus, Westerdals Oslo ACT, prewan@westerdals.no

Bendik Bygstad, University of Oslo, bendikby@ifi.uio.no

ABSTRACT

Many PhD students write paper-based dissertations. This has several advantages; by dividing the PhD project into smaller pieces you get published early, and also reduce the overall risk of running out of time and funding. However, at one point you will need to write a PhD Summary, which essentially should show that all the pieces fit together into one consistent contribution. This is often more difficult and time consuming than expected, and the PhD student – while now well versed in the chosen research method – is often at loss regarding how to do this.

Enter Strawberry Analysis. Building on Carney's Ladder of Analytical Abstraction and Miles & Huberman's data analysis techniques we suggest a framework that guides the PhD student through four steps of analytical abstraction and synthesis, enabling her or him to integrate the contributions of each paper. We illustrate the approach by an example, and discuss the usefulness of the framework. This study should be useful to PhD students writing a paper-based PhD, particularly those who have published a few papers and are in the second half of writing the Summary.

Key words: Strawberries, paper-based PhD, data analysis, data synthesis, Ladder of Analytical Abstraction.

1. INTRODUCTION

Long before he became one of Norway's millionaires and business tycoon, Petter Stordalen was a young boy who tried to sell his strawberries at the market place in a small town of Norway. Frustrated by the poor sales, he finally consulted his father: *The other strawberry vendors have nicer berries than me, and they sell much more!* The father looked at the young boy and said: *Son, you have to sell the strawberries that you have got. They are the only ones you have!* This "strawberry philosophy" became the lifelong lesson for the young boy, who later became a large hotel chain owner (Stordalen, 2009). We will use this metaphor of strawberry selling to frame our investigation in this paper.

In Scandinavia it is common to write a paper-based PhD (or PhD by publications) thesis as opposed to a monograph (Whitley et al., 2004). Quoting Whitley et al.: *In Norway...[...]...The usual PhD thesis is a collection of 5 to 6 published papers and a "kappa" (introductory chapter) of around 100 pages. The papers should preferably be published in peer-reviewed international conferences and journals. Often one paper published at IRIS¹ is tolerated. Given these constraints, the papers in the thesis are normally written in English. A monograph is acceptable but is becoming less customary (p. 323).*

Ideally, a paper-based PhD should be based on a specific research question from the start, and follow a rigid process of data collection and analysis, where the different papers constitute logic parts of a well-planned journey. In practice this is often not the case; on the contrary, the situation facing a PhD student who has chosen a paper-based dissertation is often characterised by the following: Eventually the student has published four, five, or six papers at reasonably strong conferences and/or journals. These papers may address the same topic, but the contributions are often quite different, both regarding type of contribution and level of investigation. The papers have been written at various stages in the PhD journey of the student, hence reflecting growing research maturity. In addition, each paper may have been influenced by

¹ IRIS is an acronym for International Research Seminar in Scandinavia, popular with PhD students.

the various review and publication processes, in the sense that reviewers and editors have influenced the focus and framing of the papers. Some papers may have been rejected, while others have been accepted with substantial revision and reframing.

Consequently, the papers are not internally consistent, nor do they express the same level of maturity of the student (Dowling, Gorman-Murray, Power, & Luzia, 2012). Also, time has passed, normally between three and six years, and the world and the academic field have changed during this period. In a recent study of paper-based PhDs in New Zealand the authors revealed that consistency of the various contributions was an issue of concern (Sharmini, Spronken-Smith, Golding, & Harland, 2014).

The student typically worries about: do I have enough empirical evidence? Is each paper good enough? The answer to these questions is assumedly yes, because they have been published already, usually at peer reviewed conferences or journals. The supervisor is the prime resource for advising which conferences to attend or which journals to submit to. Whether the volume of published research and quality of the papers are sufficient is also the responsibility of the supervisor.

Then the great challenge arises: How do I integrate the content of five rather different papers into one consistent dissertation? A common mistake is “jumping to conclusions”, meaning that the student goes directly from fragmented findings in each paper to one universal conclusion which is supposed to “explain everything”. A useful approach to this challenge may be partly academic analysis and partly rhetoric writing, but how should this be conducted in practice? In this paper we investigate the research question: *Is there a suitable and step-wise approach to guide students in consolidating multiple papers into one PhD Summary?* We here propose such a framework, and we suggest calling it Strawberry Analysis.

We proceed by reviewing the extant literature on paper-based PhD dissertations. Then we present our framework, which builds on Carney’s *Ladder of Analytical Abstraction*. We use one PhD project as empirical evidence and demonstrate how five papers can be analysed and synthesised into a PhD dissertation.

2. RELATED LITERATURE

At the time of our study a search for “*Paper-based PhD*” returned only 10 hits on Google Scholar. A study from 2004 (Whitley et al., 2004) compares the PhD practices from six European countries: Germany, Norway, Italy, the United Kingdom, Spain, and the Netherlands. They investigated choice of topic; the normal time spent completing the dissertation; course workload; and format of the dissertation (monograph or collection of papers). They found that the most visible differences were at the level of the course workload and the format of the dissertation. It is more common in Norway to do a paper-based PhD (rather than a monograph) and it is also a frequent choice in the other countries, with one exception: monograph is preferred in the UK. Whitley et al.’s paper does not investigate possible reasons for this difference. There are, however, some indications that paper-based PhD is currently increasing also in the UK, see for example Lee’s study (Lee, 2010).

Which guidance is available for students that pursue the paper-based approach? A well-known handbook for PhD students and their supervisors, edited by Avison and Pries-Heje, provides rich advice on the research process of a PhD, but does not deal with paper-based PhDs (Avison & Pries-Heje, 2005). We then reviewed the International Journal of Doctoral Studies, an online journal which started in 2006. In 2013, Michael Jones published an article called: “Issues in Doctoral Studies - Forty Years of Journal Discussion: Where have we been and where are we going?” Having examined 995 papers based on doctoral studies from 1971 to 2012, he found six categories: teaching, doctoral program design, writing and research, employment and career, student-supervisor relationship, and the doctoral student experience (Jones, 2013). This paper offers interesting findings, such as why more than one third of PhD students give up during the first year; the reason was due to stress and isolation. However, Jones’ study does not address issues related to paper-based PhDs.

The article “This is not an article” (Sørensen, 2002) provides advice to the following research question: *How do I write a good article which both documents the body of research I have conducted, and which also “sells” the points I am trying to make?* Sørensen’s paper was based on his own experience (hence the title), but being an experienced researcher in the Information Systems field he provides personal ideas

of good craftsmanship. For example: the researcher must have a clear understanding of the research question or problem. If it is blurred, the reader will not remember much afterwards. Sørensen used Brooks's now famous article "No Silver Bullet: Essence and Accidents of Software Engineering" (Brooks, 1987) as a good example of capturing the reader. (If you find our strawberry metaphor strange, let us point to "No silver bullet" where Brooks compares software projects to werewolves, or Dahlbom & Mathiassen (1994, cited in Sørensen, 2002, p. 9) who use the clock as an analogy of the evolution of computing.) While Sørensen's article addresses the issue of selling one's findings he does not deal with paper-based PhDs.

Textbooks, such as David Silverman's *Interpreting Qualitative Data*, will often include extended guidelines beyond data analysis on how to write a good research paper. Silverman guides an academic writer from the research question, via data collection and analysis, to the final report (Silverman, 2011). Another writing guide is *Doing Ethnographies*, by Crang and Cook (2007). This book was written by (former) PhD students for (current) PhD students doing qualitative research. Drawing on their own experiences as ethnographic researchers, Crang and Cook presents a threefold journey of (i) preparation; (ii) collecting data; and (iii) "*Pulling it Together*" (Crang & Cook, 2007, p.129). This third section has two main parts, namely analysing material and writing up the research. While the authors describe three ways of writing styles: "Writing through codes", "Writing autoethnography", and "Writing montage" they do not address a paper-based PhD, but rather a monograph. Still, the style of "Writing montage" address how to create an understanding of fragmented pieces of research. The technique involves "...montaging items together in accordance with established norms of representation so that an audience has an impression of a set of ideas or story being smoothly communicated" (p. 177). This technique of montage has been used by film-makers, photographers, artists, novelists, academics, and musicians since before the 1920s. The montage-maker takes his/her medium and tries out either/or overlaying, overlapping, superimposing, cutting and juxtaposing the fragments with the overall aim of identifying new meanings that one fragment alone cannot provide, but rather by the way they work together.

In an article from 2012 Dowling et al. critically assessed the paper-based PhD. They consider themselves to be pioneer in this field: "...there remains a critical gap in understanding the formation of academic identities through doctoral education in the contemporary era. This paper begins to fill this gap through an exploration of the 'Phd by publication'" (Dowling et al., 2012, p. 293). They found several challenges for students who have chosen a paper-based PhD. Examples include wanting to write the whole thesis in each paper, and feeling vulnerable because they have to publish their individual papers at an early stage. In the same line, a study by Robins and Kanowski compares the pros and cons of a monograph versus a paper-based PhD. They find that the advantages outweigh the disadvantages for a paper-based PhD (Robins & Kanowski, 2008). Amongst the advantages is receiving a continuous assessment in the form of peer reviews when the student publish at conferences or in journals. A related concern in this matter is that publishing is time consuming and requires attendance at the conferences, which again requires a budget for traveling and conference fee.

Summing-up this review, we find the following:

- There is a trend towards more paper-based PhDs (Lee, 2010; Whitley et al., 2004)
- There is a strong need for more research on paper-based PhDs (Dowling et al., 2012; Sharmini et al., 2014)
- In particular, there is a need for more guidance for PhD students on how to write the PhD Summary (Robins & Kanowski, 2008)

From the related literature we see that there are many academic contributions that aim to guide novice researchers on how to get papers published, but there is less on specifically how to write a paper-based PhD and what to do at the final stage in the PhD journey. The aim of our study is to fill this gap and in the next chapter we present our research approach.

3. METHOD AND CASE

This is a conceptual paper, with a case example (Yin, 1994). We started with a problem definition and a research question. Then we reviewed other literature on how to write paper-based PhDs and found that it

was relatively scarce. We chose a well-known framework for analysis (Carney, 1990, cited in Miles and Huberman, 1994, p. 92) and adapted it in order to fit a paper-based PhD project.

As empirical evidence we chose the first author's PhD project since it offered sufficient details and it also illustrated the abstraction process. The PhD project had a paper-based design, as opposed to a monograph. The student had published five papers, the first in spring 2010, and the last in spring 2014. In other words, four years had passed since the first conference paper was published. The PhD project was investigating use of Business Intelligence tools and process, with the following overall research question: Which patterns can be identified for successful use and evolution of Business Intelligence? The five papers are presented in table 1:

<i>Paper</i>	<i>Published</i>	<i>Outlet (Conference or Journal)</i>	<i>Method</i>	<i>Type of data</i>
Paper 1	2010	AMCIS 2010. Proceedings of the American Conference on Information Systems. Association for Information Systems 2010	Interviews	Qualitative
Paper 2	2012	Journal of Information Technology Education: Innovations in Practice	Survey, observation	Qualitative and quantitative
Paper 3	2013	Scandinavian Journal of Information Systems	Data mining, interview	Qualitative and quantitative
Paper 4	2013	NOKOBIT 2013. Norsk konferanse for organisasjoners bruk av informasjonsteknologi, Stavanger. Tapir Akademisk Forlag 2013	Survey, observation	Qualitative and quantitative
Paper 5	2014	ECIS 2014. Proceedings of the European Conference on Information Systems. Association for Information Systems 2014	Interviews	Qualitative

Table 1: The five papers used as empirical data in this study

Although the papers addressed the same overall topic, there were several differences between them, regarding level of investigation, method in use, type of data and findings (see Appendix 1 for more details on findings). Some of the strawberries were mature and juicy, while others were less tasty partly because they were from the early parts of the study.

How does a PhD student proceed from here? In our experience – but not backed up by evidence – most supervisors are reluctant or unable to help at this point, maybe because they believe that solving this issue is a key part of the PhD process.

4. FRAMEWORK

In this section we briefly present Carney's *Ladder of Analytical Abstraction*, on which our framework is founded on. Then we present our framework, which we call strawberry analysis, which we explain in detail by using our empirical evidence.

4.1 Carney's Ladder of Analytical Abstraction

Our framework builds on Carney's (1990) *Ladder of Analytical Abstraction* (figure 1). What Carney (and Miles and Huberman) argue is that research needs to go from a descriptive stage to an explanatory stage, by a step-wise process of abstraction. The first step is summarizing and packaging the data in order to get an overview. The second step is to repackage and aggregate the data from step one. And the third step is to identify explanations for the results from step two. While the ladder can give the impression of a sequential and carefully planned trajectory, Miles and Huberman explain that the journey from bottom to top is iterative and incremental (Miles & Huberman, 1994).

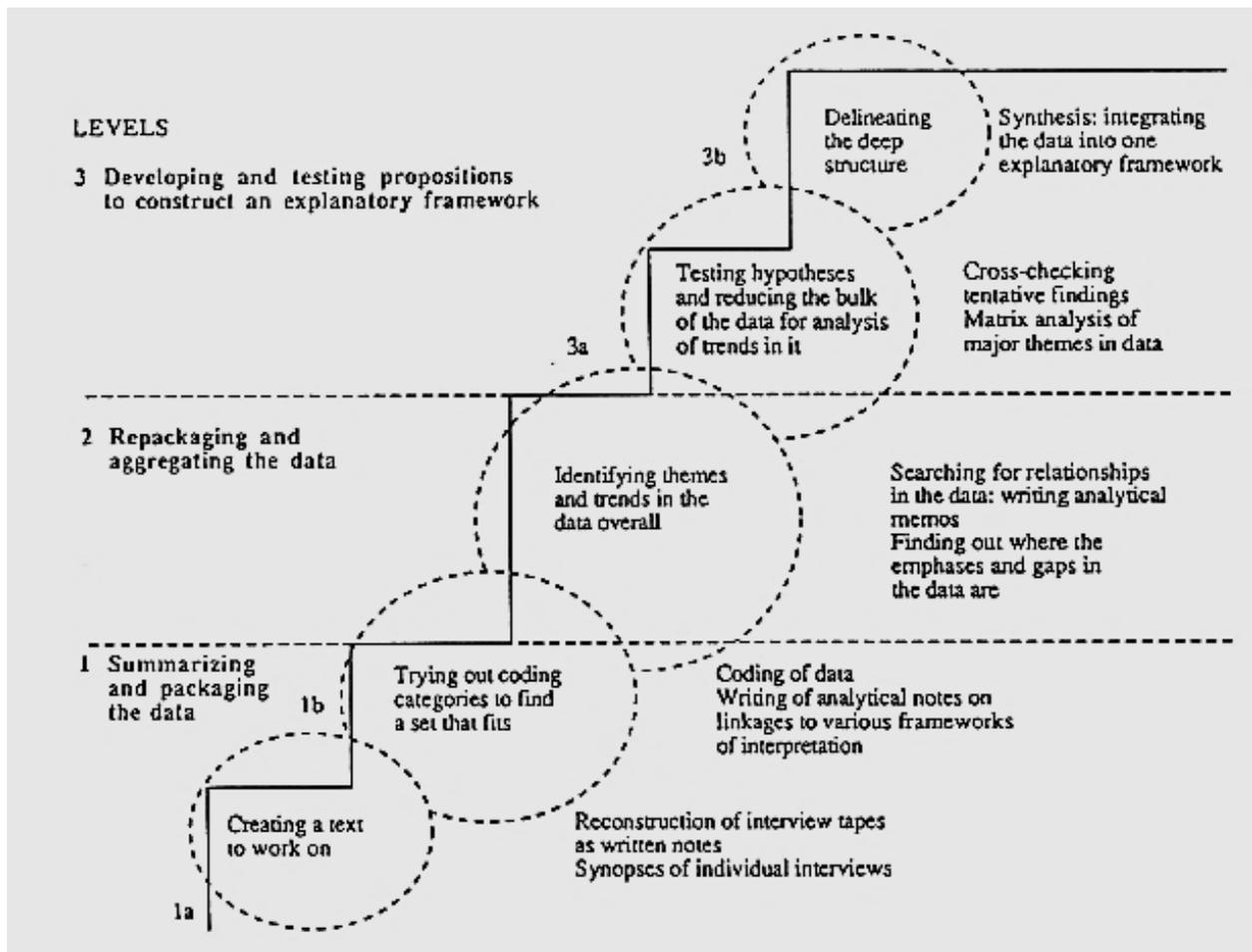


Figure 1: Ladder of Analytical Abstraction by Carney, 1990, cited in Miles & Huberman, 1994, p. 92

4.2 Our framework: Strawberry Analysis

Our framework (figure 2 below) builds directly on Carney's ladder, but differs in two important respects: first, it does not use plain data as level 1, but rather the "strawberries", i.e. we start with the published research papers. Second, we add a fourth level which we call *Selling your strawberries*. We also draw on Miles and Huberman's principle of visual displays: "...you know what you display" (p. 91). Visualisation can for instance be made through matrixes or figures. A key concept in any case is the reduction of the massive amount of data that is gathered. Examples include coding the data or colouring frequent words, either by a computer (see Appendix 2) or manually.

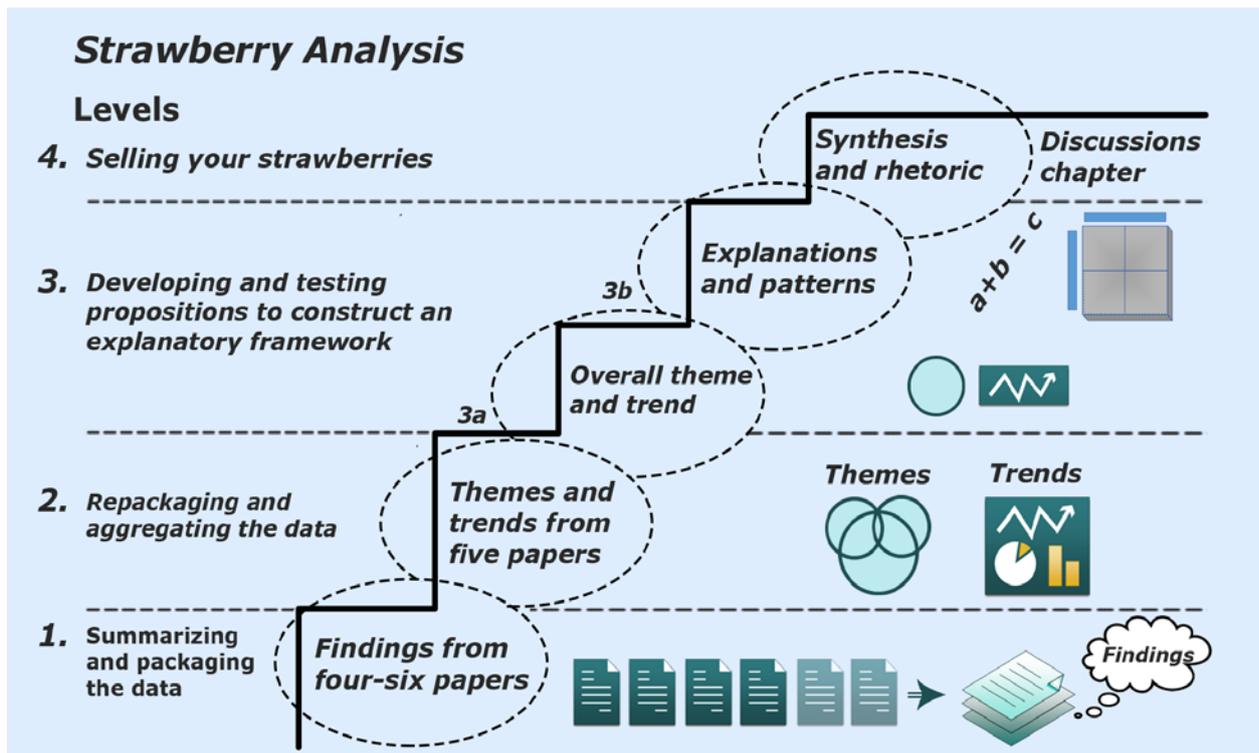


Figure 2: Our framework “Strawberry Analysis” building on Carney (1990)

Level 1: Summarizing and packaging the data. Start by summarizing the key findings from each paper, for example into a matrix, clusters, graphs, and/or main terms. This step is fairly straightforward; you simply extract all findings from your papers, but you may also need to re-consult your raw data materials. The reason we want to take one step backwards is that we wish to reinterpret the findings with fresh eyes. At the time of writing the first papers, the analysis and discussion may have been somewhat limited, in the early stages of the learning process of the PhD student.

Case example: In our case example the five papers were reread again with an open mind. The raw data material, consisting of interview transcripts, print screens, documentation and more, was also studied again. Frequencies of terms were displayed, as illustrated by the Wordle print screen in Appendix 2. All main findings were summarised into a matrix, which was used in the next step. An example of summaries reads *Business Intelligence is useful in identifying customer preferences* and more details can be found in Appendix 1.

Level 2: Repackaging and aggregating the data. At this level you look for themes and trends, while still keeping each paper separated. Or, using our strawberry metaphor, you collect your strawberries and make baskets (themes); for example the large ones in one basket, and the ones from one farmer into another. When it comes to trends you can for example draw timelines based on each paper. This is an iterative process, which means that you will refine your analysis each time. See for instance Crang and Cook’s (2007) advice for *writing montage*. Also, you will expect to go up and down between steps 1 and 2 a few times; is not a setback to take a step down. This process can take several weeks or even months. Preferably, conduct this step together with one or two other people, who should have an interest in your field or your work, and also have some experience within data analysis. Good candidates are fellow PhD students or other researchers at your school. (Addressing the concern in Sharmini et al’s (2014) paper: wait before you involve your supervisor at this stage. This is your chance to prove your individual intellect, especially if you have several co-authors on your publications, which often include your supervisor.) The benefits of involving fellow people (but not you supervisor) are that you become more creative and you can see more connections when other people look at the same whiteboard. Also, fellow people will look at your project from a different viewpoint. Although this step is based on the findings from the papers, it is still a creative process. Regardless of the *packaging* you created in the former stage, visualisation is key in this step. For example, use tables or figures to identify themes and relationships, and use different colours on post-its or draw with different colours on the whiteboard.

Some students might want to introduce a theoretical perspective at this point. There are benefits and disadvantages to this: if you introduce a theoretical perspective at level 2, theme categories and trends can be made based on an established theory. The benefits are a very structured analysis of themes and trends, which will also facilitate the next step of identifying patterns and explanations. However, the downside of this strategy is the risk of forcing the strawberries into square cubes and you lose the richness of your material. If you look for certain theoretical concepts/phenomena, you are very likely to find them. Quoting the economist Ronald Coase: “*If you torture your data long enough, they will confess!*”

Case example: The student started to identify themes and trends. Several attempts had to be made. What criteria have to be present in order for it to qualify to be called a theme? Miles and Huberman’s advice is to use various visualisation techniques and the student chose the *reduction technique* and colour codes. The student consulted fellows at the school; both Master and PhD students could offer useful feedback. As a result, between two and three themes and trends were identified in each paper. In the first paper three themes were identified: *Information quality is important; Users experience a need for a formal BI system; and There are regulations for personalized marketing and privacy.* Amongst trends, the student found: *Increasing use of BI tools, but in symbiosis with existing knowledge; Favour of stand-alone tools (Excel, Google Analytics) but curiosity about traditional technology (SAP); and BI technology is becoming more and more sophisticated and easier to use.* Additional examples are found in Appendix 1.

Level 3: Developing and testing propositions to construct an explanatory framework. Doctoral studies within Information Systems usually aim to make both a practical and a theoretical contribution. At this level, you can identify patterns or guidelines as a practical contribution (Yin, 1994). Your theoretical contribution can be to explain how your findings correspond with, or extend, a theoretical framework or model.

This step is demanding, and will usually require a sub-step, 3a. In this intermediary step you try to synthesize the themes and trends from all papers. For example, you may assemble your themes and trends into one theme and/or one trend on a more abstract level. Let us say that you have published five papers, and four out of these five contain a theme of end-user resistance if they are exposed to a *big bang* implementation of a new information system or lack of training, you may create one theme called “lack of socio-technical perspective”.

Step 3 usually includes introducing a theoretical perspective, for example drawing on the socio-technical literature to develop your overall contribution. The theoretical perspective will help to frame the different themes and trends, and allow for higher-level explanations. These may be constructed in the form of *patterns* (Larman 2004), which often are understood as *solutions of a problem, in a context*. For example, various findings in user adoption of IT could be abstracted into socio-technical patterns.

An alternative approach is *retroduction*, which is a technique associated to critical realism. The point is to start with an observed outcome (such as an event, or a trend), and ask: which underlying mechanism can explain this outcome? For example, Henfridsson and Bygstad investigated why some digital infrastructures are successful, while others are failures. They found three self-reinforcing mechanisms; innovation, adaption and scaling, which consistently could explain the outcomes of 41 documented cases (Henfridsson & Bygstad, 2013). The benefit of retroduction is that it offers a technique for delineating the deep structure of the event, which concurs with Carney’s advice. However it can prove difficult to identify new mechanisms for an inexperienced researcher.

Case example: In our example, the student first tried to theorise using systems theory, building on Ackoff (Sengupta & Ackoff, 1965). It resulted in a causal model that expressed that everything-was-connected-to-everything, and did not really explain the identified themes and trends. The Technology Acceptance Model (Davis, 1989) was the next attempt. Starting to resemble the adventure of *Goldilocks and the three bears*, this theory proved too narrow and consequently it also failed to provide sufficient explanations. Finally, the student tried the concepts *bootstrapping* and *adaptation* which belongs to the framework of the Information Infrastructure Theory (Hanseth & Lyytinen, 2010) which managed to explain the themes and trends.

Some issues arose: When the first paper was published in 2010 the concepts of *bootstrapping* and *adaptation* (Hanseth & Lyytinen, 2010) were not known to the PhD student. They became a part of the

student's knowledge after three years. Nonetheless, these two concepts could more or less explain the themes and trends identified in step 2 (and 3.a) and proved "just right", quoting *Goldilocks*.

Assessing your claims. Can you make sense of your key findings when using the explanations and patterns? If yes, congratulations! If not, we offer two suggestions. First, go back to your raw data, and conduct a new data analysis in the light of your new theoretical perspective. (But don't torture the data.) Second, revise your explanations and patterns, by reconstructing your themes and trends, and go from there to the explanations.

What if you discover data that contradicts your explanations? Should you conveniently dismiss this as an "outlier"? In the PhD example an identified theme was that agile BI tools (such as QlikView) enables bootstrapping. However, in one case a traditional BI tool enabled bootstrapping also. Should you overlook this finding? No! Outliers may be an indicator that there may be something wrong with your explanations, and you should first check this carefully. Also, outliers offer interesting points for discussion, and should generally be valued. In our example, one explanation for this outlier was that agile BI tools could motivate users to explore traditional BI tools at a later time.

Finally, as stated in the introduction: you have to sell the strawberries that you have got. If the final step up the ladder proves difficult, a common mistake for the PhD student is to start contemplating if there is "something wrong" with the collected data, or even wonder if he or she needs to write yet another paper. As Stordalen's father said: *You have to sell the strawberries that you have got, because they are the only berries you have.* Or, as we argue in this paper: you have to concentrate on the papers that you have published at this point in time of the PhD project, and some rhetoric may help you.

Level 4: Selling the strawberries. We recall Sørensen's article on how a researcher has to sell the main point he or she is trying to make by capturing the reader. Level 4 in our Strawberry framework concerns how a student can complete the final part of the PhD thesis, which normally concerns writing the Discussion chapter. We draw on Aristotle's definition of rhetoric as found in Gottweis' book chapter "Rhetoric in Policy Making: Between Logos, Ethos, and Pathos". Gottweis argues that the term rhetoric has suffered a negative association of manipulation; such as covering up something or distracting from the real sequences of events. *"Rhetoric is genuinely linked to the idea of persuasion, but it has also a much neglected performative dimension: in the play of language not only signs are communicated"* (Gottweis, 2007, p. 241). Aristotle suggested *"Let rhetoric be (defined as) an ability, in each (particular) case, to see the available means of persuasion"* (Aristotle (n/d), cited in Gottweis, p. 242).

Rhetoric was divided into three elements by Aristotle: *Logos* addresses rationality and reason. What patterns and connections do you find? *Ethos* points to morale. What is considered right and wrong, as well as legal aspects? *Pathos* addresses the emotional aspect. How do you paint a rainbow over your findings?

From our PhD thesis example, the themes and trends (see Appendix 1 for more details) were finally explained by drawing on two concepts from the Information Infrastructure theory. Subsequently, the PhD student had to convince the reader about the benefits of regarding Business Intelligence as a process with elements from Information Infrastructure. The overall rhetoric was about illustrating how the benefits and value of a Business Intelligence solution would increase recursively as a network effect. This interpretation was somewhat contradictory to the traditional view of Business Intelligence solutions, which tended to focus on technology and disregard the network effect. How should a PhD student start to build such an argument?

Drawing on Aristotle, the student started with logos. The student could build on the steps 1 and 2 from the Strawberry Analysis, and let the strawberries talk for themselves. As Gottweis argues, it is important to keep a fact-based presentation and stay away from any personal opinions at this point. Typical findings included that a Business Intelligence solution would consist of multiple tools from various vendors. Heavy, traditional tools would not scare the user away if he or she had experienced usefulness with agile tools beforehand. It is reasonable to "sell" to the reader that starting out with simple tools will motivate the user to learn more complicated ones.

Presenting these facts was done with cross-matrix as found in step 2, and then the focus was turned to *ethos*. *Ethos* refers to morale and trust, and the student would have to convince the reader that the

research was trustworthy. In our case example each paper had been accepted at peer-reviewed conferences, which gave confidence to the student, knowing that the strawberries had been approved by reviewers. Nonetheless, the findings must be discussed, while finding the right balance between being a humble PhD student and also taking pride and confidence in the work that has been conducted over the years. A simple technique can be differentiating between the work of others and the students' own findings and reflections. In our example, the student drew on concepts from the Information Infrastructure, and used *bootstrapping* and *adaption* to explain how introducing agile tools gradually allowed for successful use of a company's Business Intelligence solution.

Finally, *pathos* has the function to move the reader by drawing on empathy, sympathy, or sensibilities. While emotions have been looked down at modern times, Aristotle saw no contradiction between reason and emotion (Gottweis, 2007). Pathos may be the most difficult aspect for a novice researcher. We recall that logos is about presenting the facts; and ethos is to some extent controlled by the speaker; but pathos lie within each receiver of the message. What moves one person, may be indifferent to another, or even irritate the third person. In our case the doctoral student chose to divide the discussion into two parts; first it was about implications to the Information Systems research and second, to Business Intelligence practitioners in the industry. The reason was that people in these two camps usually have different interests. For the academics, the student tried to illustrate how successful Business Intelligence can be understood by using elements from a theory that had not been associated with Business Intelligence before. The student suggested new opportunities for getting published (such a thought should engage any academic). For the industry, a set of practical guidelines was offered. Practitioners are usually less interested in publications but they do look for new business opportunities. The student pointed to the somewhat surprising and new insights that the case companies had achieved by employing their current BI solutions. The guidelines would hopefully help other companies to explore the gold mine that they are already sitting on.

As explained above rhetoric consists of three elements, and it is possible to move back and forth between the three. A summary of the four steps and the empirical evidence is found in table 2.

<i>Steps in strawberry analysis</i>	<i>Our empirical evidence</i>
1. Summary of five papers	Going back to the source data, five summaries were made
2. Identifying themes and trends	From the summaries, themes and trends were identified by creating matrixes and displays
3.a Abstraction of themes and trends	Using concepts from the Information Infrastructure Theory (Hanseth & Lyytinen, 2010) the themes and trends from step 2 were abstracted into <i>bootstrapping</i> and <i>adaptation</i>
3.b Explaining the themes and trends	A framework based on the Information Theory explained the themes and trends. Four software engineering patterns (Larman, 2005) were identified
4. Selling the strawberries	Inspired by rhetoric's, the student discussed the implications of reframing Business Intelligence for the academics and the practitioners

Table 2: the steps in the strawberry analysis framework and our empirical evidence

5. DISCUSSION

In this section we critically assess the proposed framework.

5.1 Contribution to research

As shown by Whitley et al., (2004) writing a paper-based PhD is a current trend, particularly in Scandinavia. Robins and Kanowski (2008) identified six factors which the student should consider in the decision making process: university requirements, supervisors' attitudes, the research subject matter, intellectual property, capacity and working style, and issues of co-authorship. They also compare monograph and paper-based PhD, and the result of their study is in favour of the paper-based PhD. Currently, there is a strong need for more research on how to guide PhD students who have chosen this form (Dowling et al., 2012). While several benefits have been identified, such as "kick-starting" the study and getting published early, there are also drawbacks such as the pressure of "publish or perish" and the students feeling vulnerable because they doubt that their early publications are good enough. Also, the doctoral students faced the challenge of wanting to cram their whole thesis into each published paper

(Dowling et al., 2012). As mentioned earlier, the alternative to paper-based PhD is the monograph, and some universities give the students the opportunity to choose between the two.

Our key contribution is to fill a methodological gap in addressing the student's need of guidance on how to conduct this type of research. More specifically, we extend the research by Sharmini et al. (2014), who dealt with paper-based PhDs seen from the perspective of supervisors and examination committees. They raise two concerns; the first is the student's challenge of writing a coherent Summary, while other is the examiners' difficulty of differentiating between the student's individual work and that of the co-authors' of the published papers. They call for guidelines for the students on how to write a consistent Summary, and they also raise concern for how these guidelines should be distributed. One of their propositions is a handbook for paper-based PhD students.

Our research addresses much of the same issues as presented above, but we address the particular concerns of the student. What makes our contribution different is that we offer a clear step-wise framework that is particularly created from the student's perspective, addressing the challenges that earlier research has identified. Following all steps of the ladder should facilitate writing a coherent Summary. Our framework, building on rhetoric, demonstrates how you can "sell your strawberries" by synthesizing the publications through the steps in the latter.

5.2 Benefits for the PhD student

Our framework is for a student who has already decided to do a paper-based PhD and finds herself in the process of writing the Summary after the majority of the papers are published. For the student using the Strawberry analysis, the Summary is an opportunity to prove that she or he is worthy of a PhD. This issue addresses Sharmini et al.'s (2014) concern presented above: differentiating between the student's and the co-author's work in the published papers.

Carney's *Ladder of Analytical Abstraction* is a strong and useful tool and we have extended the use of it from "plain" data analysis to writing a paper-based PhD. We have identified three main benefits from using our framework. First, the student is presented with a step-wise framework, which does not allow any shortcuts; it prevents the student from jumping to conclusions or getting lost in the discussion. Second, the framework helps gather all of the existing "loose strings" from the earlier publications, and to stay focused on the actual findings (and not waste time looking for better strawberries). We have illustrated how the loose strings can be synthesized into an overall contribution, and furthermore how to sell this contribution to the reader (or the PhD committee) by use of rhetoric. We believe that by following the framework the student will remain more confident towards the end of the PhD project. Although we have employed an example from the Information Systems field, the Strawberry Analysis can be used as a framework in any case where you conduct qualitative data analysis. Examples are anthropology, history, political science, education, and health care. We also remind that qualitative data is not merely transcripts of interviews. Qualitative data can also include for example pictures and movies (Miles & Huberman, 1994).

5.3 Limitations and concerns

There are limitations to the Strawberry Analysis. First: this is not a recipe, meaning that the framework will provide a line of actions, but not the ingredients. We have tried to provide a few tips, for example using colours and white board when going from the first step to the second. Nonetheless, different persons have different preferences when they do research. Some students are sociable; others have their "Eureka moments" alone. Also, the framework only assists in parts of the dissertation, typically as part of the Methods, Analysis and Discussion section. For example, literature review, research question, or external validity are not part of the framework.

Finally, not all PhD projects may be suitable for this analysis. The Strawberry Analysis may prove less helpful for students with a positivist epistemology, tight research question and quantitative data, because such projects usually require a more stringent and less explorative data analysis.

6. CONCLUSION

In this paper we have investigated the following research question: *Is there a suitable and step-wise approach to guide students in consolidating multiple papers into one PhD Summary?* Building on Carney

and Miles & Huberman, we offer a step-wise framework, which we call Strawberry Analysis. The Strawberry metaphor comes from the father of one of Norway's millionaires to be, who told his young son that he had to "sell the strawberries that he had got" on the market. In our study the strawberries are papers that the PhD student has to merge into one PhD Summary, given that he or she is doing a paper-based PhD project.

We contribute to the methodological literature for guiding the PhD student on how to synthesize the results from the papers into a PhD Summary or "kappa". Our contribution consists of a framework with four main levels and offers three benefits for a PhD student: (i) The step-wise framework allows no short-cuts; it helps gather all existing "strings"; (ii) the framework is an extension of an existing, strong technique; and (iii) Strawberry Analysis also helps the student to sell her strawberries.

ACKNOWLEDGEMENTS

We thank the three anonymous reviewers at NOKOBIT for constructive feedback. In addition we thank Stian Sæthre for graphics.

REFERENCES

- Avison, D., & Pries-Heje, J. (2005). *Research in Information Systems – A Handbook for PhD Supervisors and their Students*: Butterworth Heinemann, Information Systems Series.
- Brooks, F. (1987). No Silver Bullet. Essence and Accident in Software Engineering. *IEEE Computer Society*, pp. 179-203.
- Crang, M., & Cook, I. (2007). *Doing Ethnographies.*: Sage Publications Ltd.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), pp. 319–340.
- Dowling, R., Gorman-Murray, A., Power, E., & Luzia, K. (2012). Critical Reflections on Doctoral Research and Supervision in Human Geography: The 'PhD by Publication'. *Journal of Geography in Higher Education*, 36(2), pp. 293–305.
- Gottweis, H. (2007). Rhetoric in Policy Making: Between Logos, Ethos, and Pathos. In F. Fischer, G. J. Miller & M. S. Sidney (Eds.), *Handbook of Public Policy Analysis: Theory, Politics, and Methods*: CRC Press. Taylor & Francis Group.
- Hanseth, O., & Lyytinen, K. (2010). Design theory for dynamic complexity in information infrastructures: the case of building internet. *Journal of Information Technology*, 25, pp. 1-19.
- Henfridsson, O., & Bygstad, B. (2013). The Generative Mechanisms of Digital Infrastructure Evolution. *MIS Quarterly*, 37(3), pp. 907-931.
- Jones, M. (2013). Issues in Doctoral Studies - Forty Years of Journal Discussion: Where have we been and where are we going? *International Journal of Doctoral Studies*, 8, pp. 83-104.
- Larman, C. (2005). *Applying UML and Patterns: An Introduction to Object-Oriented Analysis and Design and Iterative Development (3rd Edition)*: Pearson Education, Inc.
- Lee, A. (2010). "When the Article is the Dissertation: Pedagogies for a PhD by Publication." In, edited by In C. Aitchison, B. Kamler & A. Lee (Eds.), *Publishing Pedagogies for the Doctorate and beyond* (pp. 12–29). London: Routledge.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis*: Thousand Oaks: Sage Publications.
- Robins, L., & Kanowski, P. (2008). PhD by publication: A student's perspective. *Journal of Research Practice*, 4(2), Article M3. Retrieved May 26th, 2014, from <http://jrp.icaap.org/index.php/jrp/article/view/2136/2154>.
- Sengupta, S. S., & Ackoff, R. L. (1965). Systems Theory from an Operations Research Point of View. *Systems Science and Cybernetics, IEEE Transactions*, 1(1), pp. 9-13.

- Sharmini, S., Spronken-Smith, R., Golding, C., & Harland, T. (2014). Assessing the doctoral thesis when it includes published work. *Assessment & Evaluation in Higher Education*, pp. 1-15.
- Silverman, D. (2011). *Interpreting Qualitative Data* Sage Publications, 4. edition.
- Stordalen, P. (2009). *Jordbærfilosofi. Fra Petter Stordalens blogg.* : Available at:
<http://petterstordalen.no/2009/07/02/jordb%C3%A6rfilosofi/> [accessed February 12, 2014].
- Sørensen, C. (2002). This is Not an Article — Just Some Food for Thoughts on How to Write One. *Department of Information Systems, The London School of Economics and Political Science. No. 121., Working Paper.*
- Whitley, E. A., Sieber, S., Cáliz, C., Darking, M., Frigerio, C., E., Jacucci, E., Nöteberg, A. & Rill, M. (2004). What is it Like To Do an Information Systems PhD in Europe? Diversity in the Practice of IS Research. *Communications of The Association for Information Systems*, 13, pp. 317-335.
- Yin, R. K. (1994). *Case study research. Design and Methods 2nd ed.* (Vol. 5): Sage Publications, Inc.

APPENDIX 1

In this appendix we include two illustrations of using the Strawberry Analysis framework. We choose to present the first and the latest paper of the students' publications.

Paper 1: Presthus, W.; Papazafeiropoulou, A.; Brevik, E. (2010) E-business in entertainment: Insights from the use of Business Intelligence in the Norwegian music industry. *AMCIS 2010 Proceedings*.

LEVELS	
4. Step Selling the strawberries	Implications for the industry Implications for academia
3. Step Explanations	Bootstrapping: The usability of the stand-alone/agile BI tools enables bootstrapping. Adaptation: The perceived usefulness/value from the results of stand-alone/agile BI tools enables further adaptation of BI tools and BI process.
2. Step Identify themes and trends	Themes: Information quality is important. Users experience a need for a formal BI system. There are regulations for personalized marketing and privacy. Trends: Increasing use of BI tools, but in symbiosis with existing knowledge. Favour of stand-alone tools (Excel, Google Analytics) but curiosity about traditional technology (SAP). BI technology is becoming more sophisticated and easier to use.
1. Step Document key findings	BI is useful in identifying customer preferences. Users believe that existing tools are rather sophisticated but difficult to handle while emphasizing the need for a "personal touch".

Paper 5: Presthus, W. (2014) Breakfast at Tiffany's: The Study of a Successful Business Intelligence Solution as an Information Infrastructure. *ECIS 20014 Proceedings*.

LEVELS	
4. Step Selling the strawberries	Implications for the industry Implications for academia
3. Step Explanations	Bootstrapping: Demonstrating the ease of use and usefulness to the end-users enables bootstrapping. Adaptation: The BI competence is kept in-house and is available, which facilitate adaptation. When end-users of BI experience a problem with the BI tool there is an easy and rapid process for a solution.
2. Step Identify themes and trends	Themes: Loosely coupled architecture enabled exchange of BI technology. The BI department is highly competent. The end-users are supporting each other. Trends: Starting with Cognos, QlikView has almost replaced this technology and is becoming a key part of the BI solution.
1. Step Document key findings	The case company's successful BI solution resembles an Information Infrastructure rather than the implementation of technology. Three patterns for BI practitioners.

