A Quantitative Survey on Software In-house Integration

Rikard Land[1], Peter Thilenius[2], Stig Larsson[1], Ivica Crnković[1]

Mälardalen University
[1]Department of Computer Science and Electronics, [2]School of Business
PO Box 883, SE-721 23 Västerås, Sweden
+46 21 10 70 35

{rikard.land, peter.thilenius, ivica.crnkovic, stig.larsson}@mdh.se, http://www.idt.mdh.se/{~rld, ~icc}

Abstract

As organizations merge or collaborate closely, an important question is how their existing software assets should be handled. If these previously separate organizations are in the same business domain – they might even have been competitors – it is likely that they have developed similar software systems. To rationalize, these existing software assets should be integrated, in the sense that similar features should be implemented only once.

We have previously made qualitative observations on this topic. This report describes the follow-up study, which was performed in the form of a questionnaire aimed at validating and quantifying the previous observations. This report describes the research design, present the questionnaire together with all responses, and make some statistical analyses. This will form a basis for further publications with deeper analyses.

Keywords

Software Evolution, Software Integration, Software In-house Integration, Software Merge, Software Systems Merge.

1. Introduction

Within an organization, from time to time it is realized that two or more in-house developed software systems address similar needs, and there is an overlap in functionality. This may be caused by gradual evolution and extensions to systems that initially addressed different needs; this is also typical when the organization changes through new types of collaborations and mergers. The in-house controlled software may be products of the companies, or some support systems for the core business. In either case, there is typically a desire to rationalize maintenance as well as to business considerations to reduce the number of similar systems offered to customers and users. To study how this situation has been addressed in industry, we have previously performed a qualitative multiple case study where the main data source were interviews.

The previous study was a qualitative and explorative study and can be characterized as grounded theory research [2], the next natural step would be to validate and quantify the proposed generalizations. As research method we chose to construct a questionnaire, which, in the process of construction, was distributed to a number of participants in integration projects. The questionnaire and sufficient observations will enable statistical analyses which can enrich the previous findings. The present paper relates this study to the previous one, describes the current research method in more detail, reprints the questionnaire and all respondent data, and reports some basic analyses performed. We expect to perform deeper analyses and more profound conclusions in subsequent publications, using this report as a reference.

1.1 Acknowledgements
We would like to thank all respondents and their organizations for sharing their experiences and allowing us to publish them. Thanks to Laurens Blankers at Eindhoven University of Technology for previous collaborations in this research field. Thanks to Cecilia Erixon at School of Business at Mälardalen University for help with designing the questionnaire.

1.2 Disposition
Section 2 describes the research method: the purpose, sampling strategy, limitations, and the methods for data collection and analysis. Section 3 goes through the different parts of the questionnaire and presents the statistical analyses performed. Section 4 contains the references. Appendix 1 reprints the questionnaire, and Appendix 2 prints all original respondent data.

2. Research Method
Quantitative studies of a phenomenon are fundamentally conducted for either of two purposes (or sometimes both). One common purpose is to perform a delimited investigation with the aim of making statistically valid statements with a much broader scope. Typically, this means to draw a sample from a population and make statements with a certain certainty, usually defined by statistical measures, on that population based on the information provided by the sample. Accordingly, the discussion on reliability of the study concerns primarily the potential to infer the results of the investigation to the whole population. The other purpose is primarily not to make broad statements on a specific population, but rather to investigate the relation among various aspects of the approached phenomena associated with a population. Often, this purpose calls for multiplex measures of more or less latent aspects and investigation of the correlation among them. Reliability is in this setting a discussion of the measures' qualities in term of scope, accuracy and robustness.

As mentioned, we studied the integration phenomena by means of an exploratory case study [3]. This enabled us to make some qualitative statements about commonalities across the cases, concerning high-level strategies and selection criteria, reuse of implementations and experiences, and process practices [1]. The results were thus suggestions on how to perform in-house integration, and some implications of not following these advices. To refine these results, some of the qualitative statements could be quantified and form the base for developing a measure of in-house integration. This means that the purpose of this study is rather to investigate the relation among aspects of in-house integration as suggested by the qualitative statements about commonalities among the cases than making broad statements on all in-house integration. The measure should therefore consist of a number of interrelated multi-item constructs that together provide the base for more extensive studies of the phenomenon.

In order to develop the measure we chose to construct a questionnaire based on our previous qualitative studies and the outcome from analyses of them. The construction of the questionnaire was made in order to allow for both descriptive analyses of single items as well as multi-variate statistical methods. The construction of a questionnaire initially calls for considerations on face validity, i.e. that the items in a scrutinization appears to be coherent with the studied phenomenon, and content validity, i.e. that the items are adjoining to the appropriate theoretical field of the study. These aspects of validity are also important with the purpose of validation as is the aim of this study. The questionnaire was thus developed in several iterations to ensure validity and consisted eventually of 101 questions and statements. The questions and statements were grouped in four sections covering areas like management, how decisions were reached, integration of systems, reuse and retirement, existing systems and practices relating to in-house integration. More specifically, the phenomena were investigated using the topic areas Integration Strategies: Concept Formation, Decision Making Considerations, Retireability: Quantification, Architectural Compatibility: Quantification and Hypothesis Testing and Quantification of Process Practices and Risk Mitigation Tactics. These five

areas form the base for future further refinement of the questionnaire and must eventually be investigated for construct validity, i.e. their qualities as measures. Construct validity is usually divided into requirements on convergent validity, i.e. that all items forming the construct should be concurrent, and discriminant validity, i.e. that no item should be indicative of more than one construct within in the same setting. Each construct should conform to these requirement and the constructs in the same setting can thereafter be evaluated for nomological validity, i.e. that the constructs in their setting correctly depict the theoretical ideas about the phenomenon.

To achieve high internal data consistency and to facilitate a straightforward way to answer for the respondent were all statements constructed in similar way. They were all assessed using a fixed-alternatives five-grade ordinal scale ranging from "I do not agree at all" to "I agree completely". A few questions were open-ended to allow the respondent to express opinions relating to the statements.

The questionnaire was which was distributed to people representing various roles in several integration projects. The questionnaire is reprinted in Appendix 1, and all respondent data is listed in Appendix 2, and the statistical analyses are presented in Section 3.

2.1 Sampling Method and Sample Size

The cases are chosen with the purpose of validation of the previous results. By returning to the previous cases, some amount of internal validation of our previous interpretations (in terms of theory construction) is ensured. If the respondents would describe the cases in a very different way from what we have done based on the previous interviews, it is a sign that the theory is a bad representation of the reality. By administering the questionnaire to some cases that were not part of the previous study, we get an indication whether the theory extracted from the previous cases makes sense at all.

Ideally, the sample size in any quantitative studies should be considerable (hundreds of companies), even though the relative size of a sample drastically diminishes by increasing size of the population. A large sample is to give statistical confidence in the findings, and enables statistically significant analyses concerning differences between large and small companies, business domains etc. There are databases with e.g. all companies registered in Sweden which are typically used to define a population and retrieve a random sample from for similar kinds of surveys. However, the problem in this case is that it is difficult to formulate the population in terms of the information found in these databases. There are no entries for newly merged companies, so one would have to make some assumptions concerning whether company names and organization numbers are identical or change compared to previous years etc. Also, we are interested in the software development activities within a company, which could be found in virtually any business domain, and the size of the company would not necessarily hint at the size or importance of the software department. In addition, we are not only interested in commercial companies, but other types of organizations as well, such as governmental departments or regional official organizations. Although not necessarily impossible, it would require much research and assumptions only to define a population.

Instead, we chose to use convenience sampling, i.e. by calling and talking to people and organizations we considered were likely to have gone through a significant integration effort. We chose to return to our previous cases as well as pursuing some contacts in other organizations. In the previous cases, we tried to get access to more people than we had interviewed before.

Twelve cases were contacted, with a total of around 25 people, to ensure at least one response from most cases. We received responses from eight cases, nine people. The response rate was thus 2/3 of the cases, and ca 1/3 of the potential respondents. Our conclusions per case are therefore sensitive to individual responses, but conclusions where all responses are summed are less sensitive. We expect to continue distributing the questionnaire to more cases and respondents in the future, and the current data should only be seen as preliminary indications.

The cases and respondents are summarized in Table 1, together with the number of interviews made in our previous study.

Table 1: The Cases and Distribution of Respondents

Case	Organization	System Domain	Number of respondents (previous study)
A	Newly merged international company	Safety-critical systems with embedded software	1 (1)
B	National corporation with many daughter companies	Administration of stock keeping	1 (1)
C	Newly merged international company	Safety-critical systems with embedded software	1 (2)
D	Newly merged international company	Off-line management of power distribution systems	0 (2)
E1	Cooperation defense research institute and industry	Off-line physics simulation	1 (1)
E2	Different parts of Swedish defense	Off-line physics simulation	1 (1)
F1	Newly merged international company	Managing off-line physics simulations	0 (3)
F2	Newly merged international company	Off-line physics simulation	1 (6)
F3	Newly merged international company	Software issue reporting	0 (1)
F4	Newly merged international company	Off-line physics simulation	1 (0)
G	Newly merged international company	Database-centered system	2* (0)
		Total number of respondents:	**9 (18)**

* The respondents are labeled G_a and G_b in this report.

2.2 Data Collection and Analysis

In most cases the questionnaire was distributed via mail or email (after initial contact via telephone and/or email). The data was then organized, coded and stored Microsoft Excel for the basic statistical analyses presented in the paper. Appendix 1 reprints all questions of the questionnaire, and Appendix 2 lists all respondent data.

3. Analysis

This section summarizes the data for each part of the questionnaire. Discussions on reliability and validity as well as deeper analyses and interpretations will be made in subsequent publications.

3.1 Integration Strategies: Concept Formation

In previous publications based on the qualitative data from the multiple case study, we described four high-level integration strategies:

- **No Integration** Do nothing – this requires no extra effort or resources in the short term, but can consequently not give any return on investment.

- **Start from Scratch** Discontinue all existing systems and initiate the implementation of a new system. The new system will likely inherit requirements and architecture from the existing systems.
- **Choose One** Evaluate the existing systems, choose the one that is most satisfactory, and discontinue all others. The chosen system will likely need to be evolved before it can fully replace the other systems.
- **Merge** Take parts from the existing systems and integrate them to form a new system that has the strengths of both and the weaknesses of none.

Although one should not expect any case in reality to fit perfectly within this classification, these strategies show the range of solutions available. They are therefore useful as examples to be discussed.

Based on questions 12-23 and 32-42 in the questionnaire (see Appendix 1), it is possible to see how well the solutions in the cases actually fit these descriptions. We provided strategy profiles, i.e. a set of answers that would signify the perfect match with each strategy (see Table 2). For example, in a case of pure Choose One strategy the statement "All existing systems is (or will be) retired" should yield the response 1 ("I do not agree at all"), while for Start from Scratch the response should be 5 ("I agree completely").

Table 2: The strategy profiles.

Question / Strategy	12	13	14	15	16	17	18	19	22	23
Start from Scratch	1	1	5	5	5	1	(any)	1	5	1
Choose One	1	5	5	1	1	1	5	5	(any)	5
Merge	1	1	1	1	1	5	5	5	(any)	5

Each case was then matched with the profiles by for each statement, summing the absolute difference between the actual response and the strategy profile. There were identical statements concerning management's vision and the actual outcome of the integration project. The exactly same statements used when comparing with the strategy profiles were to be graded both according to what was management's vision for the integration, and what was the actual outcome (or seem to be, if it is not finished yet). Both types of difference are presented in Table 3.

Table 3: The distance of the cases to the strategy profiles (vision/actual outcome); lower means better fit.

Case / Strategy	A	B	C	E1	E2	F2	F4	G_a	G_b
Start from Scratch	14/11	10/7	14/17	10/4	23/18	20/16	25/27	12/11	16/16
Choose One	16/15	16/17	14/16	21/20	13/16	12/25	11/15	25/22	19/20
Merge	20/22	21/21	16/14	19/24	10/15	14/18	8/10	24/25	19/16

3.2 Decision Making Considerations

Questions 24-31 aim at determining who makes strategy decisions, and which considerations are the most important. The responses to each question are summed, and scaled to the percentage of the possible maximum (i.e. if all responses were "5"). The result is shown as a ranked list of statements in Table 4.

Table 4: The relative importance of decision making considerations.

ID	The high-level decision about how to integrate…	Percent

31	...was made by management	84
24	…was based on technical considerations	71
27	…was based on considerations concerning the parallel maintenance and evolution of existing systems	63
30	...was made by technicians	62
26	…was based on considerations for existing users	60
28	…was based on available staff and skills	60
29	...was based on politics	60
25	...was based on considerations on time schedule	38

3.3 Retireability: Quantification

We previously found the discussions concerning the impact of retiring the existing systems to be an important influence when selecting an integration strategy. Questions 44-36 and 52-54 aim at elaborating this notion. The responses to each question are summed, scaled to the percentage of the possible maximum, and the result is shown in Table 5.

Table 5: The responses about retirement.

ID	Statement	Percent
	Management's vision:	
44	None of the existing systems will be retired.	25
45	One or more existing system will be retired.	91
46	There will be a replacement system that covers all the lost functionality of retired system(s).	78
	Personal opinion:	
52	None of the existing systems will be retired.	31
53	One or more existing system will be retired.	87
54	There will be a replacement system that covers all the lost functionality of retired system(s).	73

Questions 47-51 aim at identifying differences between whose opinions the decision on retireability was based on. The responses to each question are summed, scaled to the percentage of the possible maximum, and the result is shown in Table 6. The distribution of responses is fairly large, so the differences cannot be considered important.

Table 6: The relative importance of decision making considerations.

ID	This decision was based on the opinions of...	Percent
51	…management	73
47	…customers	71
49	…developers	69
48	…users	67
50	…marketing people	53

Questions 55-58 aim at ranking between types of whose opinions the decision on retireability was based on. The responses to each question are summed, scaled to the percentage of the possible maximum, and the result is shown in Table 6. The distribution of responses is fairly large, so the differences cannot be considered important.

Table 7: Aspects of backward compatibility.

ID	The future system needs to...	Percent
55	...support the way users currently work.	73
57	...be backwards compatible with existing surrounding tools.	67

56	...be backwards compatible with existing data.	56
58	...be backwards compatible with installations of the existing systems.	51

3.4 Architectural Compatibility: Quantification and Hypothesis Testing

We previously found the architectures of the existing systems to be important when selecting an integration strategy. Questions 59-75 aim at elaborating this notion.

Questions 59-71 consists of statements how similar the existing systems were according to a number of criteria. The responses to each question are summed, scaled to the percentage of the possible maximum, and the result is shown in Table 8. The highest ranked similarity receives a total of only 35 percent, so the conclusion is that no type of similarity is common in general.

Table 8: Ranking of similarities among the systems in the cases.

ID	Group	Statement	Percent
63	Structure	The existing systems contain software parts/components/modules with similar functionality.	78
69	Technology/ Framework	Communication between components/modules/parts in the existing systems is performed through certain interfaces.	66
64	Structure	The hardware topology (networks, nodes) of the systems is similar.	60
65	Data Model	The design of the existing systems is based on the data model.	60
71	Technology/ Framework	The existing software use the same or similar technologies.	56
66	Data Model	The data models in the existing systems are similar.	54
61	Structure	The existing systems interacts with the users in the same way.	53
70	Technology/ Framework	The existing systems use some technology to clearly encapsulate software components/modules/parts.	51
68	Technology/ Framework	The existing systems are written in the same programming language.	51
67	Data Model	The implementations of data handling in the existing systems are similar.	50
62	Structure	The existing systems have similar look-and-feel of the user interface.	49
60	Structure	The parts/components/modules exchange data in the same ways in the existing systems.	36
59	Structure	The software of the existing systems have the same internal structure (architecture).	35

The questionnaire was designed so that each such criterion was considered part either of structure, data model, or technology/framework. The average percent within the groups were 23, 20, and 22, i.e. virtually no differences.

We have previously identified three possible sources of similarities: same time of first implementation, same domain standards, and a common ancestor (the existing systems have evolved from the same system due to previous collaborations). These are plotted with the similarities (average per group). The hypotheses say that a high rank on the y axis should come with a large value on the x axis, and there are some indications of this for same time period and same domain standards (see Figure 1). However, for systems with a common ancestry it is more difficult to claim a link with current similarities (see Figure 2).

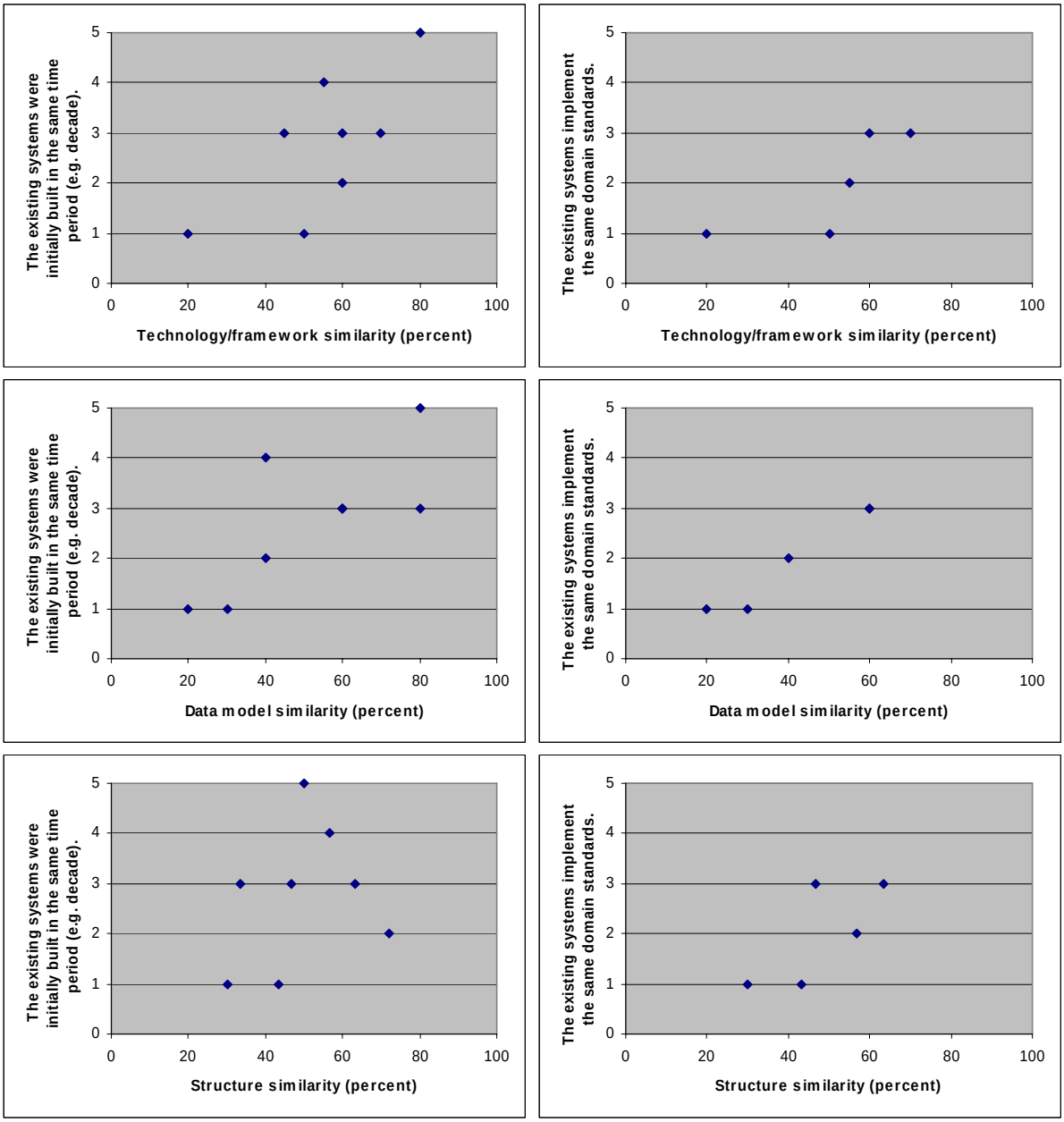

Figure 1: Time period and domain standards (y-axis) plotted with similarities (x-axis).

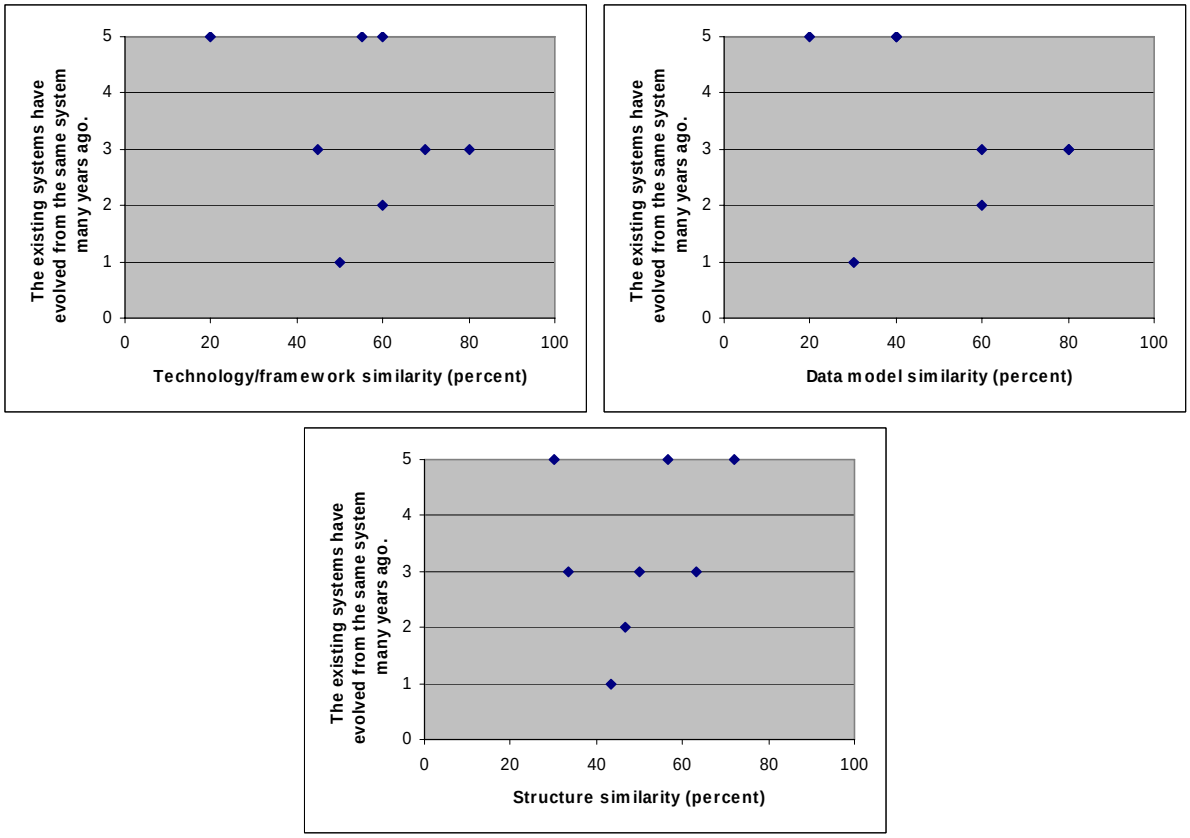

Figure 2: Common ancestry (y-axis) plotted with similarities (x-axis).

In publications based on the qualitative data from the multiple case study, we brought forward the hypothesis that a certain amount of compatibility or similarity between the existing systems is a prerequisite for *Merge*. Table 9 presents the cases in decreasing order of similarity, and we see that the cases that mostly resemble a *Merge* (as described in section 3.1) are at the left.

Table 9: The similarities in the cases.

	F2	E2	A	C	B	E1	G
Average Similarity (Combined)	76	70	64	57	56	53	46
Percentage Structure Similarity	77	50	63	72	47	33	50
Percentage Data Model Similarity	80	80	60	40	60	80	35
Percentage Technology/ framework similarity	70	80	70	60	60	45	53
Strategy (earlier interpretation)	*Merge*	*CO (Merge)*	*SFS*	*CO*	*CO*	*SFS*	*CO*
Strategy (questionnaire interpretation)	*CO Merge SFS*	*CO Merge*	*SFS*	*Merge CO SFS*	*SFS*	*SFS*	*SFS*

3.5 Quantification of Process Practices and Risk Mitigation Tactics

Based on the previous multiple case study data, the questionnaire listed a number of process practices and risk mitigation tactics (questions 76-101). Compared to previous publications, these statements were somewhat shortened (and in some cases divided into several statements) to make them more

straightforward to rank. The respondents were asked to grade: 1) how important this was for project success, and 2) how much attention it was given in the project. Both these grades were summed across the cases (and scaled to the percentage of the possible maximum). In Table 10, the practices are enumerated in decreasing order of importance. The attention given is also listed, as well as the difference between importance and attention. In general, the statements ranked highest were also the ones with less distribution of the answers.

Table 10: Process Practices.

ID	Statement	Importance	Attention	Difference
92	Management needs to show its commitment by allocating enough resources	98	58	40
88	A strong project management is needed	95	68	28
94	The “grassroots” (i.e. the people who will actually do the hard and basic work) must be cooperative, both with management and each other	93	78	16
76	A small group of experts must be assigned early to evaluate the existing systems and describe alternative high-level strategies for the integration.	93	74	18
78	Experience of the existing systems from many points of view must be collected.	91	73	19
90	All stakeholders must be committed to the integration	89	64	25
98	A common development environment is needed	82	70	12
84	Decisions should wait until there is enough basis for making a decision	80	58	22
96	Formal agreements between sites must be made and honored (strictly obeyed)	80	45	35
100	There is a conflict between the integration efforts and other development efforts	77	60	17
86	It is more important that decisions are made in a timely manner, even if there is not enough basis for making a decision	75	60	15
82	The future system must contain more features than the existing systems	58	58	0
80	The future system should be described in terms of the existing systems.	54	50	4

Table 10 also shows that for almost all practices, the attention received in the particular projects was lower than what should be expected. Moreover, the differences tend to be greater for the practices considered most important. This means that there is a danger that these practices are overlooked.

4. References

[1] Land R. and Crnkovic I., “Software Systems In-House Integration: Architecture, Process Practices and Strategy Selection”, In *Information & Software Technology*, volume Accepted for publication, 2006.

[2] Strauss A. and Corbin J. M., *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (2nd edition), ISBN 0803959400, Sage Publications, 1998.

[3] Yin R. K., *Case Study Research : Design and Methods* (3rd edition), ISBN 0-7619-2553-8, Sage Publications, 2003.

Appendix 1: Questionnaire

This appendix reprints the questionnaire. For questions 76-101, the "importance" column is assigned the even question ID and "attention" the odd number; e.g. for the statement "A small group…", "importance" has ID 76 and "attention" 77.

This questionnaire is aimed at studying various aspects of the integration, including how decisions are made, the technical nature of the systems and the integration, and certain practices. Please answer to the best of your knowledge. You do not need to provide any free-text comments, but you are free to communicate anything with us – clarifications, comments on the formulation of questions, or similar.
There are four main sections, labeled A-D, with a total of 101 questions. The questionnaire is expected to take ca 20 minutes to fill. All answers will be treated anonymously and confidentially.
As this questionnaire is distributed to projects in various stages of the integration, we want to clarify that "existing systems" refer to the original systems, that have been or are to be integrated. "Future system" is the system resulting from the integration (it may already exist as well, if the integration is completed.

First we ask you to fill some background information.		
1	Project Name	
2	My experience in software development activities	Years
3	My experience with any of the existing systems	Years
Please mark your role(s) in the current project with "X".		
4	(Technical) architect	[ ]
5	Designer	[ ]
6	Implementer	[ ]
7	Tester	[ ]
8	Project leader	[ ]
9	Line manager	[ ]
10	Product responsible/owner	[ ]
11	Other	[ ]
Comments		

Section A.

You will now be asked some questions concerning management, how decision was reached, and how the existing systems will eventually be integrated.		
The following questions concern what, in your opinion, management's vision is of your project, i.e. the high-level decision about how to integrate.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
12	The existing systems will continue to be maintained, evolved and deployed completely separately.	1 2 3 4 5 NA
13	One of the existing systems is (or will be) evolved into a common system.	1 2 3 4 5 NA
14	One or more systems has been (or will be) retired.	1 2 3 4 5 NA
15	All existing systems is (or will be) retired.	1 2 3 4 5 NA
16	A new generation of this type of systems is (or will be) developed from scratch.	1 2 3 4 5 NA
17	Parts/components/modules of the future system are (or will be) reused from more than one of the existing systems.	1 2 3 4 5 NA
18	Reused parts/components/modules required (or will require) only minor modifications	1 2 3 4 5 NA
19	A significant amount of the existing systems are (or will be) reused in the future system	1 2 3 4 5 NA
20	The functionality of the existing systems are equal.	1 2 3 4 5 NA
21	The quality of the existing systems are equal.	1 2 3 4 5 NA
22	At least some software parts/components/modules is (or will be) completely new	1 2 3 4 5 NA
23	Source code is (or will be) reused from one or more of the existing systems.	1 2 3 4 5 NA

The following questions concern how, in your opinion, this vision was reached.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
The high-level decision about how to integrate…		
24	…was based on technical considerations	1 2 3 4 5 NA
25	...was based on considerations on time schedule	1 2 3 4 5 NA
26	…was based on considerations for existing users	1 2 3 4 5 NA
27	…was based on considerations concerning the parallel maintenance and evolution of existing systems	1 2 3 4 5 NA
28	…was based on available staff and skills	1 2 3 4 5 NA
29	...was based on politics	1 2 3 4 5 NA
30	...was made by technicians	1 2 3 4 5 NA
31	...was made by management	1 2 3 4 5 NA

Now some questions about your personal opinion about what you think will happen (or have

happened) in the project, i.e. how the systems will actually be integrated. This could be identical or different from management's vision/decision.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
32	The existing systems will continue to be maintained, evolved and deployed completely separately.	1 2 3 4 5 NA
33	One of the existing systems is (or will be) evolved into a common system.	1 2 3 4 5 NA
34	One or more systems has been (or will be) retired.	1 2 3 4 5 NA
35	All existing systems is (or will be) retired.	1 2 3 4 5 NA
36	A new generation of this type of systems is (or will be) developed from scratch.	1 2 3 4 5 NA
37	Parts/components/modules of the future system are (or will be) reused from more than one of the existing systems.	1 2 3 4 5 NA
38	Reused parts/components/modules required (or will require) only minor modifications	1 2 3 4 5 NA
39	A significant amount of the existing systems are (or will be) reused in the future system	1 2 3 4 5 NA
40	The functionality of the existing systems are equal.	1 2 3 4 5 NA
41	The quality of the existing systems are equal.	1 2 3 4 5 NA
42	At least some software parts/components/modules is (or will be) completely new	1 2 3 4 5 NA
43	Source code is (or will be) reused from one or more of the existing systems.	1 2 3 4 5 NA

Comments	

Section B. Reuse and retirement

Now follows a number of questions concerning retirement of the existing system and backward compatibility of the final system. (All questions about retiring systems refer to the implementations, not how the systems are named or marketed.)		
The following questions concern what, in your opinion, management's vision is of your project.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
44	None of the existing systems will be retired.	1 2 3 4 5 NA
45	One or more existing system will be retired.	1 2 3 4 5 NA
46	There will be a replacement system that covers all the lost functionality of retired system(s).	1 2 3 4 5 NA
This decision was based on the opinions of...		
47	…customers	1 2 3 4 5 NA
48	…users	1 2 3 4 5 NA
49	…developers	1 2 3 4 5 NA
50	…marketing people	1 2 3 4 5 NA
51	…management	1 2 3 4 5 NA

Now some questions about your personal opinion about what you think will happen (or have happened) in the project, i.e. how the systems will actually be integrated. This could be identical or different from management's vision/decision.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
52	None of the existing systems will be retired.	1 2 3 4 5 NA
53	One or more existing system will be retired.	1 2 3 4 5 NA
54	There will be a replacement system that covers all the lost functionality of retired system(s).	1 2 3 4 5 NA

The following questions concern what, in the project, are (or were) important aspects of backward compatibility.		
Please grade the statements below using the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
The future system needs to...		
55	...support the way users currently work.	1 2 3 4 5 NA
56	...be backwards compatible with existing data.	1 2 3 4 5 NA
57	...be backwards compatible with existing surrounding tools.	1 2 3 4 5 NA
58	...be backwards compatible with installations of the existing systems.	1 2 3 4 5 NA

Comments	

Section C. The existing systems

Now follows a number of questions concerning the existing systems.		
Please grade the statements below according to how well, in your opinion, they describe the existing systems in your project.		
Use the scale 1-5, where 1 means "I do not agree at all" and 5 "I agree completely". NA means "cannot answer".		
59	The software of the existing systems have the same internal structure (architecture).	1 2 3 4 5 NA
60	The parts/components/modules exchange data in the same ways in the existing systems.	1 2 3 4 5 NA
61	The existing systems interacts with the users in the same way.	1 2 3 4 5 NA
62	The existing systems have similar look-and-feel of the user interface.	1 2 3 4 5 NA
63	The existing systems contain software parts/components/modules with similar functionality.	1 2 3 4 5 NA
64	The hardware topology (networks, nodes) of the systems is similar.	1 2 3 4 5 NA
65	The design of the existing systems is based on the data model.	1 2 3 4 5 NA
66	The data models in the existing systems are similar.	1 2 3 4 5 NA
67	The implementations of data handling in the existing systems are similar.	1 2 3 4 5 NA
68	The existing systems are written in the same programming language.	1 2 3 4 5 NA
69	Communication between components/modules/parts in the existing systems is performed through certain interfaces.	1 2 3 4 5 NA
70	The existing systems use some technology to clearly encapsulate software components/modules/parts.	1 2 3 4 5 NA
71	The existing software use the same or similar technologies.	1 2 3 4 5 NA
72	The existing systems implement some domain standards.	1 2 3 4 5 NA
73	The existing systems implement the same domain standards.	1 2 3 4 5 NA
74	The existing systems were initially built in the same time period (e.g. decade).	1 2 3 4 5 NA
75	The existing systems have evolved from the same system many years ago.	1 2 3 4 5 NA

Comments	

Section D. Practices

Now follows a number of questions concerning specific practices.			
For each statement below, please indicate the following: how important it was (or would have been) for your project's success, and how much attention it was given in your project.			
Please use the scale 1-5. For "importance" 1 means "not important at all" and 5 means "essential for success. For "attention", 1 means "no attention was given" and 5 "very much attention was given". The same grade on both "importance" and "attention" means that with respect to importance, enough attention was given but not too much. NA means "cannot answer".			
76	A small group of experts must be assigned early to evaluate the existing systems and describe alternative high-level strategies for the integration.	**Importance** 1 2 3 4 5 NA	**Attention** 1 2 3 4 5 NA
78	Experience of the existing systems from many points of view must be collected.	1 2 3 4 5 NA	1 2 3 4 5 NA
80	The future system should be described in terms of the existing systems.	1 2 3 4 5 NA	1 2 3 4 5 NA
82	The future system must contain more features than the existing systems	1 2 3 4 5 NA	1 2 3 4 5 NA
84	Decisions should wait until there is enough basis for making a decision	1 2 3 4 5 NA	1 2 3 4 5 NA
86	It is more important that decisions are made in a timely manner, even if there is not enough basis for making a decision	1 2 3 4 5 NA	1 2 3 4 5 NA
88	A strong project management is needed	1 2 3 4 5 NA	1 2 3 4 5 NA
90	All stakeholders must be committed to the integration	1 2 3 4 5 NA	1 2 3 4 5 NA
92	Management needs to show its commitment by allocating enough resources	1 2 3 4 5 NA	1 2 3 4 5 NA
94	The "grassroots" (i.e. the people who will actually do the hard and basic work) must be cooperative, both with management and each other	1 2 3 4 5 NA	1 2 3 4 5 NA
96	Formal agreements between sites must be made and honored (strictly obeyed)	1 2 3 4 5 NA	1 2 3 4 5 NA
98	A common development environment is needed	1 2 3 4 5 NA	1 2 3 4 5 NA
100	There is a conflict between the integration efforts and other development efforts	1 2 3 4 5 NA	1 2 3 4 5 NA

Comments	

Thank you for your participation! Your answers will be treated anonymously and confidentially.

Appendix 2: Questionnaire Data

In this appendix, the complete questionnaire data is listed. (Respondent IDs are assigned by the order in which they were received.)

Respondent ID / Question ID	*3*	*1*	*9*	*2*	*4*	*5*	*6*	*7*	*8*
1	A	B	C	E1	E2	F2	F4	G	G
2	23	25	20	12	19	5	20	10	23
3	5	5	10	4	19	0		4	15
4	x							x	
5		x		x		x		x	
6		x		x	x	x		x	
7		x			x			x	
8	x	x	x	x	x			x	
9		x	x				x		
10		x	x						x
11					x				
12	2	2	1	1	3	2	3	4	5
13	5	4	5	1	2	5	4	4	3
14	3	5	5	5	4	5	4	2	5
15	1	1	5	2	1	1	1	5	1
16	2	4	3	4	1	3	1	5	5
17	1	2	4	4	4	4	5	2	4
18	1	3	3	1	2	1	3	2	4
19	1	1	4	3	4	4	4	2	4
20	3	3	4	1	2	3	3	3	4
21	3	1	5	1	2	2	3	3	1
22	5	5	5	4	4	5	5	5	5
23	1	1	3	1	5	4	5	2	1
24	3	4	4	4	4	2	3	3	5
25	1	1	2	1	2	4	1	4	1
26	1	4	4	1	5	3	2	4	3
27	3	4	4	1	5		4	2	2
28	3	3	4	2	2	3	3	3	4
29	3	2	3	1	4	5	3	2	4
30	4	1	4	4	4	2	2	3	4
31	4	4	4	4	5	5	4	3	5
32	2	1	1	1	3	3	4	4	5
33	4	3	5	1	2		4	4	1
34	5	5	3	5	5	2	2	4	5
35	2	2	3	2	1	2		2	1
36	1	4	3	5	4	3	1	5	5
37	1	2	4	1	4	3	4	2	5
38	1	3	2	1	2	1	3	2	5
39	1	1	4	2	4	2	3	1	5
40	3	3	4	1	2	3	3	3	4
41	3	1	5	1	2	2	3	3	1
42	5	5	5	5	4	4	5	5	5
43	1	1	2	1	4	2	5	1	1

Respondent ID / Question ID	3	1	9	2	4	5	6	7	8
44	1	1	1	1	1		2	2	1
45	5	5	5	5	5	4	2	5	5
46	4	4	4	4	2	5	2	5	5
47	3	2	5	4	5	4	2	3	4
48	3	4	5	4	5	4	2	1	2
49	4	4	2	4	4	2	2	4	5
50	4	1	2	NA	1	3	1	5	4
51	4	4	5	4	1	4	2	5	4
52	1	1	1	1	1	4	2	2	1
53	5	5	5	4	5	2	3	5	5
54	4	4	5	4	1	4	1	5	5
55	4	3	5	3	4	3	4	2	5
56	1	5	4	1	3	2	4	2	3
57	4	4	4	1	4	4	3	2	4
58	3	4	3	1	2	2	4	2	2
59	1	2	NA	1	2	4	1	2	1
60	3	1	2	1	1	4	1	2	1
61	4	2	4	2	1	3	1	3	4
62	5	2	4	1	2	3	1	3	1
63	4	4	4	4	5	4	3	3	4
64	2	3	4	1	4	5	2	4	2
65	4	3	NA	NA	4	4	2	NA	1
66	4	3	NA	NA	4	4	1	2	1
67	2	3	2	4	4		1	2	2
68	1	2	4	1	5	3	1	3	3
69	5	5	NA	3	3		1	2	4
70	5	3	NA	3	3		1	2	1
71	3	2	2	2	5	4	1	4	2
72	4	3	NA	NA			3	2	3
73	3	3	NA	NA		NA	1	2	1
74	3	3	2	3	5	NA	1	4	1
75	3	2	5	3	3	NA	5	5	1
76	4	5	5	5		5	5	4	4
77	3	5	5	5		3	2	NA	3
78	5	4	5	5	5	4	4	4	5
79	3	3	4	5	5	4	2	NA	3
80	4	3	4	1			3	2	2
81	3	3	4	1			NA	2	2
82	3	2	3	1	2	3	5	4	3
83	3	4	3	1	2	3	3	NA	4
84	3	4	5	5	4	5	3	5	2
85	3	4	4	5	3	2	2	2	1
86	4	3	4	5	4		4	2	4
87	2	3	4	5	3		3	NA	1
88	5	5	5	4		5	5	4	5
89	4	5	4	4		2	3	3	2
90	5	4	4	5		4	5	4	

Respondent ID / Question ID	*3*	*1*	*9*	*2*	*4*	*5*	*6*	*7*	*8*
91	2	4	5	3		NA	2	NA	
92	5	5	5	4		5	5	5	5
93	2	4	4	4		2	2	3	2
94	5	4	5	5	4	4	5	5	5
95	3	3	4	5		3	4	5	4
96	5	3	5	4	4	5	3	NA	3
97	1	3	4	4	2	2	1	NA	1
98	3	4	4	5	2	5	4	5	5
99	4	4	3	5		1	1	5	5
100	4	4	5	NA	4		5	NA	1
101	4	4	3	NA			NA	NA	1