

A Survival Analysis of Application Life Spans based on Enterprise Architecture Models

Stephan Aier¹, Sabine Buckl², Ulrik Franke³, Bettina Gleichauf¹,
Pontus Johnson³, Per Närman³, Christian M. Schweda², Johan Ullberg³

¹Institute of Information Management
University of St. Gallen
Müller-Friedberg-Str. 8
CH-9000 St. Gallen
{stephan.aier,bettina.gleichauf}@unisg.ch

²Lehrstuhl für Informatik 19 (sebis)
Technische Universität München
Boltzmannstr.3
D-85748 Garching
{buckls,schweda}@in.tum.de

³Industrial Information and Control Systems
Royal Institute of Technology (KTH)
Osqudas väg 12
SE-100 44 Stockholm, Sweden
{ulrikf,pj101,pern,johanu}@ics.kth.se

Abstract: Modern enterprises face the challenge to survive in an ever changing environment. One commonly accepted means to address this challenge and further enhance survivability is enterprise architecture (EA) management, which provides a holistic model-based approach to business/IT alignment. Thereby, the decisions taken in the context of EA management are based on accurate documentation of IT systems and business processes. The maintenance of such documentation causes high investments for enterprises, especially in the absence of information on the change rates of different systems and processes. In this paper we propose a method for gathering and analyzing such information. The method is used to analyze the life spans of the application portfolio of three companies from different industry sectors. Based on the results of the three case studies implications and limitations of the method are discussed.

1 Introduction

The rate of change in the economic environment of enterprises has increased over the past few years [RWR06; Wa05]. Some underlying factors are increased customization of products and services coupled with globalization and a more competitive market situation. Furthermore, regulations like the Sarbanes Oxley Act, Basel II, etc. need to be

met [La05]. Enterprises have to continuously adapt to these environmental changes, by aligning their business, applications, data, and infrastructure to the new requirements. One commonly accepted means to guide such adaptations is enterprise architecture (EA) management, which is a holistic model-based approach to enterprise engineering, specifically addressing business/IT alignment.

The ability to make informed decisions in the complex and highly interdependent area of EA management is closely linked to accurate descriptions and documentations of the EA. An EA model must be both up-to-date [Ci01] and appropriate with respect to the decisions it is to support [JE07; La05]. Planning based on obsolete data or decisions made with insufficient information will almost certainly have an unfavorable influence on the EA, and thus in the long run have a negative business impact.

Due to the continuous changes going on in different parts of any company's structure, EA descriptions will inevitably become obsolete at some point. Therefore, it seems reasonable to establish a maintenance process for EA models, continuously updating the EA documentation. However, such a process is not for free. EA descriptions are extensive, and the cost of collecting information and creating architecture models is considerable. To make this cost-benefit trade-off, enterprises would benefit from a model describing how fast certain parts of the enterprise's structure are likely to change. Based on such a model, an organization could decide not to update the descriptions of selected parts, as their frequency of change is so low that no relevant gain is expected from frequent updating. In a manner of speaking, this amounts to find an optimal EA *sampling rate*.

In this article, we approach the topic of change frequencies for applications as one of the main enterprise artifacts related to business/IT alignment [AW09] with a case study based research paradigm. While change in an EA may include modifications, introduction, and removal of individual elements, we only consider the latter two phenomena in this context. Thus, mere modifications of applications are not within the scope of this article. Rather, the research questions posed are:

- How can life spans of different enterprise artifacts be assessed?
- What are the approximate life spans of such enterprise artifacts?

The rest of the article unfolds as follows. Related work is presented in section 2. There, we discuss the relevant literature on EA model maintenance and prepare our analysis of life spans by giving an overview of appropriate models found in the literature. Based on these discussions, section 3 proposes an analysis model for the decay of artifacts found in EA models. Subsequently, section 4 presents the three case studies, the analysis results of which are presented in section 5. Section 6 contains a concluding discussion on the empirical results, and gives an outlook to further areas of research.

2 Related work

This section elaborates on related work in the field of EA model maintenance (section 2.1) and from the field of methods and models for analyzing decay (section 2.2).

2.1 Enterprise architecture model maintenance

A multitude of methods for EA management has been developed by researchers and practitioners (e. g. [Az05; Az06; BK05; Dv01; Og09; SH93; Wa05]). These methods usually distinguish the following EA management processes: (a) strategic dialogue/architecture visioning, (b) development and maintenance of current-state EA models, (c) development and maintenance of future-state EA models, (d) migration planning, (e) EA implementation, and (f) EA analysis based on EA models.

None of these approaches to EA management pays much attention to specifying maintenance procedures for the EA model in detail. While [Ci01; If99] and [Wa05] mention an EA maintenance process, the corresponding activities are not specified in detail, and neither specific roles nor responsibilities are defined. TOGAF [Og09] introduces the objective of *ensuring that the baseline architectures continue to be fit-for-purpose* as an important aspect of the phase *Architecture Change Management*. However, no details on how to accomplish this objective are given by the framework.

In the academic research community, maintenance procedures for EA models have recently gained increased interest. In an initial paper, [FAW07] developed a systematic decentralized EA maintenance approach describing maintenance processes and roles. [Mo09] present six EA management process patterns related to EA maintenance. Several of these patterns address lifecycles and dynamics of EAs. However, neither [FAW07] nor [Mo09] describe schedules for maintaining EA models. To conclude, requirements on EA maintenance are abundant in the literature, but the concrete solutions are very scarce.

2.2 Methods and models for analyzing decay

Quantitative theories of aging, mortality, and life span have a long history in medicine, demographics, and insurance mathematics. As far as biological systems are concerned, *survival analysis* is used to model the effect of death. Similar analyses can be performed on non-biological systems, e.g. mechanical ones, in the field of reliability theory. This theory aims to analyze, estimate, and predict life span distributions of systems and their components [BPH65; GG01]. Typically, a *reliability function* $S(t)$ is defined as the probability that a system carries out its mission through a time t . Complementing the reliability function, the *hazard function* $\lambda(t)$ is defined as the probability that a failure occurs at a specific time under the condition that it has not occurred up to that point in time.

One of the most popular laws regarding the hazard function was found by Benjamin Gompertz, today known as *Gompertz Law of Mortality* [Go25]. He analyzed the mortality of adult humans in respect to their age, and found that the corresponding hazard func-

tion increases in geometrical progression with the age. Based on this observation, he proposed a mathematical model to explain this behavior – the *Gompertz distribution*.

A special application of survival functions is presented by [FCH83]. The authors investigate the death rates of organizations, by eliciting survival functions and hazard functions for organizations. They adjust the *Makeham extension* of the Gompertz Mortality Law, based on the assumption that the death rate of organizations decreases with age (liability of newness) and taking into account influence factors like size of the organization or environmental changes. By the means of extensive explorative analyses of national labor unions, local newspaper organizations, and semiconductor manufacturing firms, they calculate estimates of organizational death rates.

The Gompertz Mortality Law and its extensions are prominently applied for biological and sociological system. In more technical application areas, failure and reliability analyses give rise to different kinds of laws. For instance, the *Weibull distribution* can be used to model a variety of failure behaviors [Ca03]. The use of the Weibull distribution as well as the Gompertz distribution requires a long period of observation and a complete data set in order to determine the estimators of the distribution's parameters. However, as such data sets are rare in the field of EA, a more flexible method is needed. In particular, information on the age of an EA artifact such as an application is very hard to collect. Most information available is on end-of-life events for EA but little on their creation. Furthermore, the *center of attention* of EA modeling undergoes shifts during the years of observation such that new EA artifacts appear during observation or are lost.

This situation is very similar to some encountered within medical statistics [B100]: In medical statistics, some patients are lost from observation while others are added as the study progresses. Little information is known regarding the start date of e.g. the cancer, but more is known about the end date. For these reasons we will employ methods from medical statistics for our life-span analysis. The next section elaborates on this.

3 Survival data analysis with the life table method

Data constrained in the abovementioned way can be analyzed using the *life table method* [B100], which provides the analyst with a tool with which it is possible to keep track of the probabilities of death and survival of the study subjects for each passing year. To summarize, the life table method does not require knowing actual start dates, allows for losing track of some observed elements and is able to derive knowledge from the observation of elements also before their terminal event. To illustrate the life table method, a generic table is shown below in Table 1, along with a description of its columns.

The terminal event, be it death from cancer or an application taken out of service, is called the *endpoint* and is here referred to as the *death* of the subject. Generally there is also a wish to evaluate the survival well before all the endpoints are known, i.e. while a good fraction of patients or applications are still alive or in use. Data that is only known to be greater than some value is called *right censored* (or just *censored*, for short). This is the kind of data obtained when examining patients or applications that recently entered

the study and have only been observed for a short time and have yet to reach their endpoints. These subjects are called *withdrawn*.

Table 1: Generic life table [BI00].

Period (years, etc.)	Number at start	Withdrawn during period	At risk	Deaths	Prob. of death	Prob. of sur- viving period x	Cumulative prob. of surviving x pe- riods
x	n_x	w_x	r_x	d_x	q_x	p_x	P_x
1	n_1	w_1	$n_1 - \frac{1}{2}w_1$	d_1	d_1/r_1	$1 - d_1/r_1$	p_1
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
T	n_T	w_T	$n_T - \frac{1}{2}w_T$	d_T	d_T/r_T	$1 - d_T/r_T$	$p_T P_{T-1}$

Table 1 depicts important life table concepts. A study period, x , of $[1, T]$ periods is considered, where T corresponds to the longest time period for which there is data available, for instance 9 years if the study covers information between 2000 and 2009. The time period states nothing about the fixed point in time when the observation of a subject commenced, as in for instance the year 2004. It merely describes the amount of time a subject is observed, as in for instance four years. The number at start n_x is the number of elements in the study at the start of *their* corresponding period x under observation. n_1 is the total number of participants in the study and n_x is calculated by taking n_{x-1} and deducting the number of deaths and withdrawals (see below) during period $x-1$.

New subjects that appear after the start date of the survey may also be included in the data analysis. For instance, an application taken into service in year 2006 may be included although the oldest data is from 2000. These new subjects enter the study at time period 1 and can only contribute to the statistics for a limited period of time, since data about their future survival is unavailable. At the relative point in time that a subject becomes unobservable it is declared withdrawn. In the example above applications taken into service 2006 can contribute with survival data for three years until today's date (2009) is reached and thus they are declared withdrawn at the fourth year. w_x represents the amount of withdrawn elements. When reading our statistics it is essential to bear in mind that period 1 is the first period after an application is observed; it is *not* a particular year such as 2001. To summarize; the concept of withdrawn elements is used to reflect the fact that a subject that only can be observed for e.g. four years provides survival information during these years but tells us nothing about the probability of surviving the fifth year.

To get an appropriate number of elements at risk, i.e. the population in which deaths can occur, the number at the start of the period is reduced by half the number of withdrawn elements during the period. This corresponds to saying that the withdrawn cases, on average, contribute half a period of risk to the population. The number of deaths d_x during period x is an observed amount. We consider artifacts that are listed in an EA repository

to be *alive*, i.e. in use. Artifacts not listed in a repository at a given point in time of consideration are declared *dead*, i.e. taken out of service.

The probability of death q_x in period x is simply calculated as the fraction of those at risk that were actually observed to die. The probability p_x of surviving is the complement of q_x . Finally, the cumulative probability P_x of surviving x periods is calculated as the product of the probabilities of surviving the periods. For the first period, $P_1 = p_1$ holds. For the second period, P_2 is the probability of surviving up to the start of period 2, i.e. P_1 multiplied by the probability of surviving period two as well, i.e. p_2 . The same reasoning holds in the general case, so $P_t = P_{t-1} p_t$ for all $t > 1$.

A natural use of the data in a life table is to draw a graph of the cumulative survival probability, usually in a step-wise fashion as to underline the inexactness of the estimation. This is called a *survival curve*. In section 4 we will provide life tables for each case study. The respective survival curves are presented and discussed in section 5.

4 Case studies

Our findings are based on three case studies from different industry sectors, presented in the following. Each case starts with a characterization of the enterprise at hand, including its established EA management approach. Subsequently, characterizations of the used data sets are given, potential errors are discussed, and analysis results are presented.

The data sets used originate from different sources, e.g. application catalogs, MS Excel files, and exports from specialized EA management repositories. As the EA management initiatives analyzed in the three case studies have very disparate backgrounds, including using different tools to store the data, the different data sets of case studies B and C are drawn from a single source, while the data sets of case study A originate from diverse. In order to compare the data sets, a mapping of the different schemas used to document the data about applications had to be performed within each case. During the comparison, two records were regarded to refer to the same application either if they had the same *id* or if they possessed the same *name*. This id or name matching represents a potential source of errors in the data set. To detect and correct such errors, occasional reviews with the respective stakeholders at the industry partners were performed during the conduction of the studies.

4.1 Company A

Company A is one of the principal energy companies in Europe and among the five largest generators of electricity. After the deregulation of several European energy markets in the nineties, company A expanded through acquisitions and has become a big actor in several European countries. As a consequence, the company's IT portfolio has become rather heterogeneous and in the past couple of years there have been several activities to consolidate the application landscape.

Since the generation, transmission, and distribution of electricity and heat are geographically bound activities, company A is organized into a number of geographical business groups each of which is functionally divided into business units according to their contribution to the electricity or heat value chain. The IT infrastructure is centralized in a shared services company but the responsibility for the processes and the applications are distributed amongst the business groups and the business units. Each business group has a coordinating CIO function responsible for the EA work. The group CIO function in turn coordinates and supports local and global EA initiatives within the company.

The empirical data presented here originates from three sources. The first source is the application catalogue that was created in 2000 in order to address the perceived Y2K threat. Secondly, another application catalogue was established in late 2005 and the beginning of 2006 as a part of the first embryo of a group-wide EA program. Finally, the third source is a group-wide application catalogue compiled in 2009, which is used to identify consolidation opportunities within the company.

The coverage of the Y2K catalogue is limited to one business group and encompasses 124 critical technical systems pertaining to different business units. The second application catalogue is more detailed covering over 500 applications. The third application catalogue has an even finer granularity and includes information on which processes are supported as well as information on the age of the applications and their principal integration points. Since the different sources aren't equidistant in time the life table (which contains relative time periods) has four rows while each dataset is however only applicable to at most three of the rows.

Table 2: Life table for applications of Company A

Time period (Year) (x)	Number at start (n_x)	Withdrawn during year (w_x)	At risk (r_x)	Deaths (d_x)	Prob. of death (q_x)	Prob. of surviving period x (p_x)	Cum. prob. of surviving the period x (P_x)
1	577	0	577.0	0	0.00	1.00	1.00
4	577	0	577.0	235	0.41	0.59	0.59
7	342	250	217.0	32	0.15	0.85	0.51
10	60	0	60.0	15	0.25	0.75	0.38

Although the material is bound to contain some errors in minute details, the overall indications from the company is that the data is sufficient to draw conclusions concerning whether applications exist or not at the three points in time. The resulting life table is shown in Table 2 and as seen in the table the cumulative probability of an application to survive 10 years is 38%. A note to the table is that since the granularity of data is three year intervals the rows of the table should be considered to be periods rather than the individual years.

4.2 Company B

The second case is taken from the financial industry. The enterprise under consideration is an internationally operating bank from Germany. The topic of EA management has a long history in this enterprise since a merger in the year 1996. Prior to the merger both companies independently conducted enterprise-wide data modeling endeavors. After the merger, the enterprise-wide data models were maintained, although a change in the focus as well as the reach took place. In certain parts of the enterprise the focus shifted towards a strongly business process centric modeling, while other parts continued to do pure data modeling. In the year 2002 the term EA management makes its first appearance, when a project was launched to increase the business/IT alignment based on a holistic model of the relevant aspects ranging from strategy to infrastructure. In this model architectural information from different parts of the enterprise was consolidated and used to identify fields for action. The first data set analyzed originates from the aforementioned early EA management project. In order to assess the advances made in this field, a similar project was launched in the year 2005. The take-over by an international banking company at the end of 2005 changed the overall make-up of the company significantly. In particular, the IT departments of the formerly independent enterprises, as well as the IT assets developed, operated, and managed by them, were to undergo extensive changes leading to an increased centralization of structures. In 2008, this centralization process has proceeded quite far, such that the final data set from the end of the year 2008 shows the face of the EA according to the new paradigm.

The different data sets analyzed for this case study are several snapshots of the application portfolio of company B. The data sets analyzed cover about 90% of the application portfolio of company B including legacy systems of the company as well as small interim solutions. Table 3 shows the resulting life table for applications of company B, indicating that the cumulative probability to survive 6 years is 49%.

Table 3: Life table for applications of company B

Time period (Year) (x)	Number at start (n_x)	Withdrawn during year (w_x)	At risk (r_x)	Deaths (d_x)	Prob. of death (q_x)	Prob. of surviving period x (p_x)	Cum. prob. of surviving the period x (P_x)
1	1142	16	1134.0	212	0.19	0.81	0.81
2	914	31	898.5	109	0.12	0.88	0.71
3	774	67	740.5	69	0.09	0.91	0.65
4	638	85	595.5	29	0.05	0.95	0.62
5	524	82	483.0	44	0.09	0.91	0.56
6	398	50	373.0	50	0.13	0.87	0.49

Although the data sets gathered and analyzed cover the time period of a merger, a significant impact of the merger on the probability of death of an application system was not observable within the analyzed time period. An interview with the industry partner regarding this expected impact revealed that the impact seems to be delayed about three

years after the actual merger took place. This hypothesis is backed by the planned architecture of company B, which predicts a large number of changes including replacements of applications for the year 2009.

4.3 Company C

Company C is a major financial service provider in Switzerland primarily focusing on standardized retail banking and transaction processing. All EA levels from business architecture to IT infrastructure in [WF06] are managed by broad, defined architecture management processes. An initiative was started to manage business and organizational architecture artifacts by the IS department architecture team as well. However, this has been dropped in favor of managing all business related artifacts by an explicit business architecture management team itself, which forms an organizational unit attached directly to the CEO in the business development department. The alignment of business and IS architectures is explicit and facilitated by a personal interweavement by having former IS architects included in the business architecture unit. Core EA artifacts are captured in an EA repository including timestamps for tracking architectural change. However, only the most requested artifacts are regularly modeled due to the high cost of keeping models up to date.

The data set of company C was extracted from the EA repository via Excel files. It includes applications as well as their clustering in domains. These artifacts are the core elements in the company's EA. For the life table analysis we only considered applications, such as a Data Warehouse, a Card Transaction System, or an SAP system.

The extracted snapshots from the EA repository have multiple timestamps from different dates in the years 2006 through the beginning of 2009. We analyzed six time periods each representing six months of observed time in the years 2006, 2007, and 2008 (cf. Table 4).¹

¹ In order to align the life table with the findings from companies A and B the time periods are denoted on the same scale, resulting in decimal numbers for company C.

Table 4: Life table for applications of company C

Time period (Year) (x)	Number at start (n_x)	Withdrawn during year (w_x)	At risk (r_x)	Deaths (d_x)	Prob. of death (q_x)	Prob. of surviving period x (p_x)	Cum. prob. of surviving the period x (P_x)
1.0	189	0	189.0	28	0.15	0.85	0.85
1.5	161	4	159.0	0	0.00	1.00	0.85
2.0	161	6	158.0	29	0.18	0.82	0.70
2.5	132	1	131.5	6	0.05	0.95	0.66
3.0	126	5	123.5	2	0.02	0.98	0.65
3.5	124	12	118.0	7	0.06	0.94	0.61
4.0	117	10	112.0	8	0.07	0.93	0.57

The overall number of applications captured in the EA remains quite constant over the time period analyzed. However, there are some high losses (d_x) in the periods 1 and 2. This indicates a volatile application landscape at company C concerning the artifacts contained in the EA repository. After the time periods that we analyzed, the cumulative probability of an application to survive the 6 periods was at 0.57, i.e. in average 57% of the applications survive 4 years.

5 Results and discussion

The above case studies demonstrate the feasibility of applying the life table method to assess life spans of EA artifacts via calculating the probability of particular applications to survive a certain number of years. In doing so, data was gathered via EA models within the different companies, presuming that they precisely reflect the survival behavior of the real applications. While observing the life span of applications within an enterprise, one is confronted with similar conditions like in medical studies: subjects enter and leave the study at different times and thus give insights into their life spans. As we were only concerned with externally observable behavior regarding system emergence and decay, we treated the applications as *black boxes*, i.e. we did not analyze their internal changes. Moreover, the findings from the three case studies give quantitative evidence about approximate life spans of the applications listed in EA repositories. Figure 1 summarizes the results from the three case studies by depicting the three *survival curves* that are generated by the calculated cumulative probability of surviving x years. The graphs are drawn stepwise based on the calculations from the life tables. On the one hand, this reflects the gaps in the data sets concerning the time periods analyzed, with abrupt changes in the probabilities. On the other hand, the comparison of the three data sets shows that the results are quite similar for applications in all three companies: The probability to survive for 4 years seems to be around 60% on average.

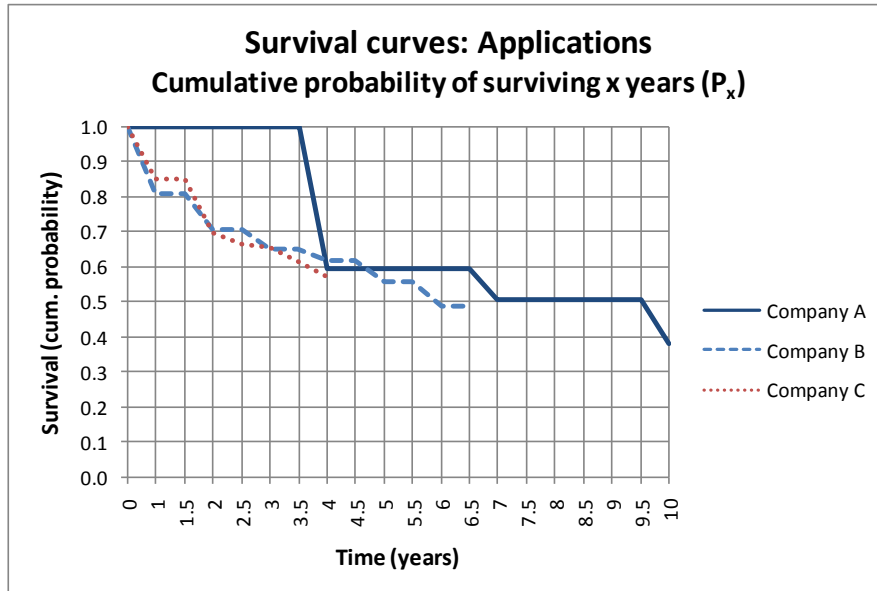


Fig. 1: Survival curves for the case studies

The results provide indications on how to design an EA model maintenance process that is adjusted to the decay rates of applications. A cumulative probability of 60% implies that more than half of the information on applications covered in an EA model becomes obsolete after four years without updates. Consequently, an effective EA model maintenance process will include a sampling rate smaller than four years in order to ensure timeliness of the information and efficiency of the related resource management at the same time.

Drawing more elaborate conclusions from these results is difficult because of several limiting factors. First of all, the limited data sets available hinder a thorough analysis. In the three case studies reliable data was only available for applications. This led to rather exemplary results concerning the survival of EA model artifacts. Also, a detailed analysis is only possible, if the observations underlying the life tables are conducted with equal intervals. While this is often the case in the medical setting from where the life table method has been adopted, it is not necessarily the case in the realm of EA. Indeed, in our investigation, only companies B and C proved mature enough as to have regular and reliable data on a yearly (or even better) basis. Company A could only provide more coarse datasets. Due to this fact, the results of the study necessarily remain on a rather aggregate level.

While our findings result in quantitative figures it is hard to assess these figures without a proper frame of reference, or a baseline. Such a baseline would need to take into account several phenomena, including:

1. Some extensive shifts in the number of applications can be traced back to external events such as mergers and acquisitions (as described above, this will be the case for company B in the near future) or organizational restructurings.
2. The notion of *industry clockspeed*. It has been suggested that different industries have different rhythms in their use of IT [Fi96], and such a phenomenon might be reflected in the data analyzed. At the single point of comparison (4 year period) the analysis results of the different case studies refute the industry clockspeed hypothesis. Nevertheless, the number of companies studied is much too small to draw any final conclusions.

There are many possible factors that add *noise* to the data sets in the present study, or make them hard to interpret. The following discussion aims to address those issues:

The influence of external or contingency factors has not been considered and is thus not reflected in the data. For example, the results are currently not related to the use of specific EA management approaches or methods in the respective companies. Such considerations might increase the commensurability of the data. Other external factors, like mergers and acquisitions, new legislation, or Y2K-phenomena might also influence the findings in detail.

A related uncertainty concerns the previously mentioned question of a baseline. *What is the normal state?* A model aiming to support decisions on optimal sampling rates needs to establish the context in which it is applicable. Another way to phrase this is to ask what is being described. *Is it only changes due to old technology, or also changes due to how organizations work, including the recurring application consolidations?* This means that factors such as the following should be either deliberately included or deliberately excluded:

- Exogenous (as seen from the IT department scope) factors, e.g. mergers and acquisitions, changes in legal regulations
- Endogenous (as seen from the IT department scope) factors, e.g. application consolidation
- Changes in scope of data over time (e.g. models from different time periods might have used different definitions)
- Poor data quality affects model credibility (e.g. very high numbers of withdrawals)
- Reliability of data, i.e. how close to reality the constituent data sets are. For instance, sometimes changes in taxonomies make data sets seem very different, even though the actual landscape they are describing is more or less the same.

In the present studies, no effort has been made to deliberately neither include nor exclude these factors. The construction of reliable statistical procedures for such data cleaning requires, however, more data than is presently available.

6 Conclusions and future work

This paper described a method for analyzing the rate of change of EA artifacts. Specifically, the method has been used to analyze the rate of change of the application portfolios at three different companies. The results were presented in the form of survival curves showing the likelihood of survival of applications in the respective companies. Thereby, we presented a contribution to the field of EA model management and maintenance. The long-term aim is to be able to find optimal sampling rates of model updating, based on the underlying rate of change of the EA artifacts, which the EA models should describe. A final version of such a method would include a trade-off between the benefit of a high sampling rate and the cost of frequent data collection.

In more or less all the cases, longer time series would have been conducive to properly account for contingency factors and find the *normal* state of the rate of change of applications. One indicator of an appropriate length of the time series is the cumulative probability of survival, P_x . When $P_x = 0$ there is obviously no use for a longer time series. While reaching zero is unlikely in most organizations, a low value of P_x could, however, be an indicator that enough data is collected. In this study, the lowest cumulative probability of survival was 37%, which would indicate the need for more data and longer time periods of observation for the case studies.

The presented study focused on investigating the rate of change of applications. To be useful, analyses of change rates need to cover other EA artifacts as well, e.g. business processes or interfaces. A more elaborate analysis method could then be utilized to determine different areas within the EA, which change at different rates and hence require appropriate sampling rates. Such areas might be formed by the types of artifacts, e.g. applications and business processes or different *subtypes* within one artifact type, e.g. different types of applications. Equally such areas could well be defined in other manners, such as for instance all entities connected to a critical business process. Furthermore, an analysis method could be used to investigate tangible EA artifacts, like applications, and their relationships to their costs in order to support investment decisions.

To generalize the analyses of rates of change to other companies and other lines of business there is a need not only for more data from other companies to populate the life tables for different EA artifacts, but probably also a richer model of how contingency factors may affect the change rates of the EA artifacts mentioned above. This would probably entail defining probability distributions in which parameters can be set. Analysis of survivability curves could be a first step towards defining such distributions. Based on the result of more detailed analyses, the process of EA model maintenance could be refined with information on optimal sampling rates for distinct EA artifacts.

References

- [AW09] Aier, S.; Winter, R.: Virtual Decoupling for IT/Business Alignment – Conceptual Foundations, Architecture Design and Implementation Example. In: Business & Information Systems Engineering 51 (2009) 2, pp. 150–163.

- [Az05] Aziz, S. et al.: Enterprise Architecture: A Governance Framework – Part I: Embedding Architecture into the Organization. Infosys Technologies Ltd., 2005.
- [Az06] Aziz, S. et al.: Enterprise Architecture: A Governance Framework – Part II: Making Enterprise Architecture Work within the Organization. Infosys Technologies Ltd., 2006.
- [BK05] Bittler, R. S.; Kreizmann, G.: Gartner Enterprise Architecture Process: Evolution 2005. Gartner Inc., Stamford, CT 2005.
- [BI00] Bland, M.: An Introduction to Medical Statistics. 3 edition, Oxford University Press 2000.
- [BPH65] Barlow, R. E.; Proschan, F.; Hunter, L. C.: Mathematical Theory of Reliability. John Wiley & Sons, New York 1965.
- [Ca03] Carroll, K. J.: On the use and utility of the Weibull model in the analysis of survival data. In: *Controlled Clinical Trials* 24 (2003), pp. 682–701.
- [Ci01] Chief Information Officer Council: A Practical Guide to Federal Enterprise Architecture, Version 1.0. 2001.
- [Dv01] Department of Veterans Affairs: Enterprise Architecture: Strategy, Governance, & Implementation. 2001.
- [FAW07] Fischer, R.; Aier, S.; Winter, R.: A Federated Approach to Enterprise Architecture Model Maintenance. In: *Enterprise Modelling and Information Systems Architectures 2* (2007) 2, pp. 14–22.
- [FCH83] Freeman, J.; Carroll, G. R.; Hannan, M. T.: The Liability of Newness: Age Dependence in Organizational Death Rates. In: *American Sociology Review* 48 (1983) 5, pp. 692–710.
- [Fi96] Fine, C. H.: Industry Clockspeed and Competency Chain Design: An Introductory Essay. In: *Proceedings, 1996 Manufacturing and Service Operations Management Conference* 1996.
- [GG01] Gavrilov, L. A.; Gavrilova, N. S.: The Reliability Theory of Aging and Longevity. In: *Journal of theoretical Biology* 213 (2001), pp. 527–545.
- [Go25] Gompertz, B.: On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. In: *Philosophical Transactions of the Royal Society of London* 115 (1825), pp. 513–585.
- [If99] IFIP-IFAC: GERAM: Generalised Enterprise Reference Architecture and Methodology, Version 1.6.3. IFIP-IFAC Task Force 1999.
- [JE07] Johnson, P.; Ekstedt, M.: Enterprise Architecture: Models and Analyses for Information Systems Decision Making. Studentlitteratur 2007.
- [La05] Lankhorst, M.: Enterprise Architecture at Work: Modelling, Communication and Analysis. Springer, Berlin, Germany 2005.
- [Mo09] Moser, C. et al.: Some Process Patterns for Enterprise Architecture Management. In: *Proceedings, Workshop on Patterns in Enterprise Architecture Management (PEAM2009)*, Bonn 2009.
- [Og09] The Open Group: TOGAF (The Open Group Architecture Framework) Version 9. 2009.
- [RWR06] Ross, J. W.; Weill, P.; Robertson, D. C.: Enterprise Architecture as Strategy. Harvard Business School Press, Boston, Massachusetts, USA 2006.
- [SH93] Spewak, S. H.; Hill, S. C.: Enterprise Architecture Planning – Developing a Blueprint for Data, Applications and Technology. John Wiley & Sons, New York 1993.
- [Wa05] Wager, R. et al.: Dynamic Enterprise Architecture: How to Make It Work. John Wiley & Sons, Hoboken, New Jersey 2005.
- [WF06] Winter, R.; Fischer, R.: Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. In: *Proceedings, EDOC Workshop on Trends in Enterprise Architecture Research (TEAR 2006)* within The Tenth IEEE International EDOC Conference (EDOC 2006), Hong Kong 2006, pp. 1–8.