Baars *et al. Decis. Anal.* (2016) 3:5
DOI 10.1186/s40165-016-0022-1

Decision Analytics

RESEARCH Open Access

An analytics approach to adaptive maturity models using organizational characteristics

Thijs Baars, Frederik Mijnhardt, Kevin Vlaanderen and Marco Spruit*

*Correspondence:
m.r.spruit@uu.nl
Department of Information and Computing Sciences, Utrecht University, Princetonplein 5, Utrecht, The Netherlands

Abstract

Ever since the first incarnations of maturity models, critics have voiced several concerns with these frameworks. Indeed, a lack of model fit and oversimplification of the real world can be attributed to the rigidity of these models, which assumes that each organization that uses the framework is equal. This research investigates this fundamental rigidity from an analytics perspective, analysing in casu a focus area maturity matrix targeted at information security. The results show that organizational characteristics influence the maturity framework both in parts and as a whole significantly, concluding that current maturity frameworks have a poor model fit and advising that a maturity framework should account for the differences between the characteristics of their target organizations.

Keywords: Information security, SME, Organizational characteristics, ISFAM, Maturity model, Decision analytics

Background

In the past decades, maturity models have become important tools to visualize progress in adopting processes and standards and to benchmark companies in their industry (Becker et al. 2009). Two of the best known and oldest maturity frameworks are the Capability Maturity Model (CMM; Paulk et al. 1993; Watts 1987) and its follow-up, Capability Maturity Model Integration (CMMi; CMMI Product Team 2010). Aside from the aforementioned, over 150 maturity models have been developed (de Bruin et al. 2005), which can be classified in three types of models (Mettler et al. 2010):

1. Maturity grids. A textual description of several levels of maturity.
2. Likert-scale questionnaires. Like maturity grids, these have a few levels of maturity, but instead of a textual description, they let practitioners score "good practice" using a Likert scale.
3. CMM-like models. Based on a formal architecture, which specifies goals and practices to reach a maturity level. These have a greater complexity than the above two.

Furthermore, maturity models can be distinguished in their scope: they are either domain specific, targeting academia and practitioners within the domain, or general,

targeting a wider audience including governments, general management but also practitioners and academia (de Bruin et al. 2005).

Missing from this classification is the focus area maturity matrix (Steenbergen et al. 2010); a recently developed type of maturity framework which consists of a number of "focus areas" which together comprise one or more topics or themes within the framework. Each focus area consists of one or more capabilities that an organization can implement. The maturity of a company is described by the capabilities they have implemented. Focus area maturity matrices extend the classification by Mettler et al. (2010) of CMM-like models in that they have a more formal order in which capabilities should be implemented and express interdependencies between the capabilities making up the maturity levels. A variety of focus area maturity matrices have been developed, implemented, and shown to be effective at displaying progress and benchmarking companies regarding their maturity. Especially for standards adoption, these maturity matrices have shown to be a valuable tool (see for example Bekkers and Spruit 2010; Spruit and Roeling 2014; Spruit and de Boer 2014; Steenbergen et al. 2010).

However, since the development of maturity frameworks there have been critiques on these frameworks for their lack of empirical foundation, lack of documentation (Becker et al. 2009; de Bruin et al. 2005; McCormack et al. 2009), and their focus on predefined end-state, issues and solutions instead of organizational evolution and drivers for change within an organization (Iversen et al. 1999; King and Kraemer 1984). Another often heard criticism is the oversimplification of reality including the "neglect the potential existence of multiple equally advantageous paths" (Pöppelbußand Röglinger (2011), p. 3), a poor model fit, and assumptions that do not pertain to all organizations in the target audience (Bollinger and McGowan 1991). This oversimplification of reality also shows in the rigidity of the models (Iversen et al. 1999; McCormack et al. 2009). The rigidity comes from the maturity model not adjusting to the organization and its environment. Maturity Frameworks thus do not take organizational characteristics into account.

The concept of organizational characteristics is not particularly new. As early as 1972—and possibly earlier—researchers used organizational characteristics in an effort to model risk in organizations (Duncan 1972). In the field of information technology, organizational characteristics have been modeled as well. An example of this is the work by Thong and Yap (1995) who investigated the adoption of IT in SMEs on the basis of organization size, competitiveness of the environment, information density and personal traits of the CEO. Bekkers et al. (2008) more recently took organizational characteristics into account in its investigation of process improvement at software product management.

This work investigates the influence of organizational characteristics on maturity frameworks in an attempt to show that rigid maturity frameworks are ineffective at guiding implementations and change processes at organizations. In the following section the background is described, including the object of research, the ISFAM model. "Methods" section describes the research approach utilized in this research, and "Results and Discussions" section discusses the analysis and results of this research. The results are discussed and concluded in "Results and Discussions" and "Conclusions" sections, respectively.

Spruit and Roeling (2014) developed the Information Security Focus Area Maturity Matrix (ISFAM). The ISFAM covers the complete domain of information security, combining the ISO 2700×, CISSP, Information Security Framework, standards of good practice, and the IBM Security Framework, and applies these in a single model. There are in total 13 focus areas and 12 maturity levels. In these focus areas, a total of 64 capabilities are assigned at the various maturity levels.

Following practices in information security, it divides the capabilities within the maturity matrix in four groups: design, implementation, operational effectiveness, and monitoring. As with all focus area maturity matrices, the lowest implemented capability defines the maturity level reached by a company.

Modern day maturity matrices are rigid. They assume a generalization of the organization in itself. In the case of the ISFAM, a large enterprise is assumed, meaning more than 250 employees and more than €50 Million in revenue. This definition (The Commission of the European Communities 2003) is incredibly broad, featuring large regional companies to multi-national giants. Maturity matrices that target the SME utilize, in general, a definition that limits the organization size at a maximum of 250 employees and up to €50 Million in revenues.

With many security breaches hitting the news (Silveira 2012; The Associated Press 2014), information security (IS), protecting the confidentiality, integrity and availability of information (Guttman and Roback 1995), has attracted a lot of attention in recent years. This makes the ISFAM highly relevant. This is especially true in the Netherlands, where 22% of all small and medium enterprises (SMEs) are hit by cyber-attacks each year, 7% more than the European average (CBS 2011; van Straalen 2011) and costs for the Dutch private sector are estimated at 7.5 Billion Euros annually (Ruiter 2012). Despite all this attention, risk awareness under SMEs is low and risk mitigation is still a low priority. A possible reason for this low priority might lie in the complexity of risk mitigation and the fact that current frameworks mainly target large enterprises. In addition, consultancies mainly target the top companies, while the risks are just as pressing for medium and small sized enterprises.

This also pertains to the ISFAM. It is co-developed with a large consultancy firm, and the standards for information security are targeted at large enterprises. However, earlier research has shown that even large enterprises do not fully adhere to standards as they are too complex and infeasible to implement completely (Baars et al. 2012; Kuilboer and Ashrafi 2000).

These models, and in certain circumstances the accompanying standards, are rigid and do not differentiate on the different characteristics of an organization. This results in certain capabilities within the model not being applicable for specific organizations. The implementation processes in these cases are ineffective, with irrelevant or inapplicable capabilities blocking the implementation. The organization will thus never reach the highest maturity level.

In order to overcome these issues, maturity models like the ISFAM should incorporate organizational characteristics. A previous investigation by Mijnhardt et al. (2014) researched the organizational characteristics in SME information security. Through a literature study and an evaluation with domain-experts, a coherent list of organizational characteristics is delivered. It discusses the operationalization of these organizational

characteristics by identifying themes, specific organizational characteristics and the different means of measuring these organizational characteristics through measurement levels. These organizational characteristics are then applied to the ISFAM, but are applicable to any maturity framework in the domain of SME information security. Mijnhardt et al. identify in total 11 organizational characteristics. These organizational characteristics provide the possibility to distinguish between a wide variety of different organizations measured in both internal as external factors.

Mijnhardt et al. (2014) structure the organizational characteristics in themes. These themes are derived from the literature and the interviews, but are not statistical constructs. They show the topics for which organizational characteristics are applicable within the model's domain, and thus focus on the pertaining dimensions an organization can differ from one another. These themes are (1) General, (2) In- and outsourcing, (3) Dependency on the IT environment, and (4) Complexity of the IT environment. Each organizational characteristic is comprised of a set of measurement levels. These measurement levels factorize the levels to which an organization can differ from one another. An example is the amount of employees employed in the General characteristics theme: (a) 0–9 employees, (b) 10–49 employees and (c) 50–250 employees. In this example, the measurement levels follow the guidelines set forth by the European Commission and the World Bank (Ayyagari et al. 2003; The Commission of the European Communities 2003).

However, (Mijnhardt et al. 2014) did not investigate the relation that the identified organizational characteristics have to the ISFAM, or the influence these organizational characteristics have on information security in a quantitative measure. This research attempts to deploy these organizational characteristics on the ISFAM, and understand the influence the organizational characteristics have on focus areas within the ISFAM. Furthermore it will present an adaptive ISFAM as a proof-of-concept for adaptive maturity matrices.

Methods

In earlier published work (Mijnhardt et al. 2014), organizational characteristics were derived that pertain to SME information security. The research described in this paper follows up on those previous efforts by further evaluating the organizational characteristics and their measurement levels, and how they pertain to the ISFAM maturity matrix through a survey.

Data gathering

The organizational characteristics listed in "Background" section are expected to influence information security implementations at SMEs. However, the relationship, if any, of these influences to the focus areas of the ISFAM are unclear. In order to understand the impact of each organizational characteristic on each focus area within the ISFAM, an impact analysis is performed.

Although there are a variety of ways to perform an impact analysis, this research is largely inspired by the work of Weimer-Jehle (2006, 2007). The method outlined by Weimer-Jehle consists of a large Table where participants can rate, or weight, how each variable influences another. For this research, this method is too expansive. It would result in 1.925.312 corresponding influences to be filled out per participant in case of

bilateral relationships between organizational characteristics. If we assume that there are unilateral relationships, the amount of variables would double in size. (that is: the influence of A on B is not necessarily equal to the influence of B on A.)

The method applied in this research is therefore less extensive, yet captures a perspective of the influence of the measurement level in an organizational characteristic per focus area. Participants are asked to value the importance of focus areas per measurement level by means of a survey, which is detailed below. This still results in a large questionnaire, 48 measurement levels divided over 13 capabilities yielding a total of 624 variables.

Both interviews and a survey were considered as research methods for this project. Although an interview generally provides better data quality and completeness, it would simultaneously severely limit the number of participants able to take part in this study. Furthermore, the questions asked are quite straightforward. In an interview setting this would lead to a structured interview where the questions asked, primarily "how would you rate the importance of these focus areas under setting *x*", which can be answered more efficiently on a computer screen. Therefore, a survey was chosen to be the best method for conducting this research. It allows for a larger participant response, and the usage of an internet survey allowed to use techniques to speed up the process of filling in the survey. These techniques are detailed below in "Survey setup" section, but follow the strategies described by de Leeuw et al. (2009).

Participants are selected from the professional networks of the researchers, and by cooperating with large expert groups in the Netherlands: the Dutch chapters of ISACA, $(ISC)^2$, NOREA, and the PvIB, a platform for information security specialists active in the Netherlands. These expert groups emailed an invitation to their members, and featured a news item concerning our survey on their website. Most experts tend to be part of multiple groups, one becomes a member of $(ISC)^2$ whenever they successfully pass examination of $(ISC)^2$'s certification programs. It is therefore difficult to identify the exact amount of unique experts these emails reached. A conservative estimate made by the directors of ISACA, PvIB and $(ISC)^2$ stated 800 individuals. In an effort for these expert-groups to work along with us, early results of this research have been presented to them at a private gathering for their members.

In order to overcome the issues regarding large surveys, a special survey system is developed. A pre-study is conducted with 5 participants to test the survey system for errors and usability, as well as to test the questions and wording of the survey.

A clear conclusion from this pre-survey revealed that the survey was very long. In response, questions pertaining to the importance of capabilities per organizational characteristic were removed, keeping the survey restricted to questions on the importance of focus areas per measurement level. As a further method of enticing participants to complete the survey, an iPad Mini is raffled to those participants whom completed the survey.

Survey setup

Large surveys have obvious problems: participants might not be willing to invest so much time in a survey, participants might fill in the survey partially, and participants might start inserting patterns or fake data to finish the survey quickly.

To overcome these issues, a survey system is developed to guide participants through the large survey in an unobtrusive and fast method. The main assumption in developing this system is that a selection of the participants will not complete the survey. To account for incomplete surveys, the order of questions that are presented to participants is pseudo-randomly selected from least to most answered. In this situation, each question is answered a (near) equal amount of times, even if a majority of the users would answer only one question. This pseudo-random selection of questions also prevented order-effects (Eisenberg and Barry 1988; Siminski 2006).

Every participant is assigned an individual resume code. In case a participant wants to split the survey up in multiple pieces, they could use that code to resume the survey where the participant left off. Local storage techniques are used to buffer parts of the survey on a participant's computer or other device. This allowed participants to lose their internet connection for a small amount of time, but still resume the survey. In case the time without internet was longer than the amount buffered on a participants device, a participant can fluently resume their survey whenever the internet connection is restored due to session storage on the server side. This provides for a robust survey and minimizes the chance that participants would stop due to internet-connection loss. The survey questions themselves each contained an organizational characteristic per question. Each question is split up in a section per measurement level corresponding to the organizational characteristic at hand. For each measurement level, or section, there are 13 sliders ranging from 0 to 25, which correspond to the 13 focus areas in the ISFAM. A Participant values each slider that represents the amount of importance they give to that capability, higher being more important than lower.

A maximum score of 25 points is selected as it allows participants to give each slider a difference of two points, assuming that this would provide ample space for participants to differentiate the importance of the capabilities.

To improve the consistency of participant's answers, a copy-paste function is provided. This allows participants to copy their ratings from one measurement level, and paste them in another measurement level. This greatly reduced time in answering the questions for participants, since they only need to adjust those sliders which they believe have a difference in importance between the different measurement levels.

Finally, the system is written in Django, a web-framework written in the Python programming language. Django comes with a series of preventive measures against malicious activities. This protected the data against any attacks, and secured the privacy and anonymity of the participants.

The survey is blinded; the researchers are not aware which participants receive the questions in which order, nor are they aware which results belonged to which participant.

Impact analysis

Due to the set-up of the survey, item non-response is anticipated. To assure that this is part of the research design, and not lack of model fitness, Little (1988) test is executed. Little's Test evaluates the non-response of participants for Missing at Completely Random (MCAR), in other words the probability that a missing value doesn't depend on its

value nor on the value of other variables (de Leeuw et al. 2009). MCAR is a prerequisite for multivariate imputations.

Multivariate Imputation by Chained Equations (MICE, see Azur et al. 2011 for an overview) is then applied to address the missing data. MICE is deemed as the proper multiple imputation method as it is specified on a variable-by-variable basis. The uncertainty of the amount of missing values and the actual result set make it hard to predict if the multivariate distribution describes the data, such as is required by Markov Chain Monte Carlo (MCMC) implementations of multiple imputations. MICE overcomes this problem (van Buuren and Groothuis-oudshoorn 2011). The round of imputations will be selected following the method described by Graham et al. (2007). This results in a complete, imputed dataset.

A Shapiro–Wilk test is performed to determine if the data fits a normal distribution. According to these results a Student's t test will be performed, as per the guidelines set forth in Wilcox et al. (2013), Wilcox and Erceg-Hurn (2012).

To gain an overview of the results, especially on a measurement-level, box-plots will be drawn per capability in each measurement level. The box-plots provide a clear overview of the differences per capability at each measurement level, and the differences between measurement levels, but also the consistency of the participants. Following the guidelines by Frigge et al. (1989), Tukey (1977), outliers are marked at ± 1.5 interquartile range (IQR). Aside from box-plots, locally weighted scatter plot smooths (also known as Loess) will be utilized to visualize the results. Loess plots a generalized line over de scatter plot using localized weights (Cleveland and Devlin 1988). This fits our results as it returns lower weights to outliers and does not assume a universal function over the results but examines the polynomial over a local area. This provides freedom to the analysis of the results, which do not have to be linear as is assumed with a linear regression.

Results and Discussions

Twenty-one participants conducted the survey. This data is manually inspected for patterns and abuse, which might occur for participants willing to make a chance to win the iPad Mini that was given away to one of the completed surveys. After this inspection, one participant was removed. Those results showed signs of abuse of the copy-paste function: all capabilities aside from 2, were entered with the value '6'. The data from the remaining 20 participants are used for further statistical analysis.

These twenty participants came from a range of six different industries, where the consultancy business services was the largest sector. All sizes of businesses were represented, but the large enterprises (companies with more than 250 employees) were overrepresented at 55%. This was expected however, as information security is an effort that is uncommon at the smaller companies. Looking at the mean of the answers of each participant, we see that three participants have answered relatively high answers compared to the other respondents. The participant that deviates with 7.45 from the mean also deviates with -2.58 from the standard deviation. This participant indicates that security on all levels should always be high, and deviates little from that in his answers. The analysis shows also the standard deviation of the answers per participant (see Table 1). Besides the aforementioned participant, another participant has a high deviation of 2.5 from the mean standard deviation. He does approach the mean value with his answers,

Table 1 Organizational characteristics and their measurement levels as identified by (Mijnhardt et al. 2014)

Organizational characteristic	Measurement levels
General	
Number of FTE employed	0–9 employees, 10–49 employees, 50–250 employees
Amount of annual revenue	0–2 Million, 2–10 Million, 10–50 Million
Sector the organization participates in	Aerospace and Defense; Agriculture and Foresting; Business Services and Consultancy; Consumer, Media, Leisure, Travel and Entertainment; Finance, Banking and Insurance; Health; IT and Telecom; Industrial Production; Energy, Utilities and Mining; Public, Education and Non-Profit; Transport, Packaging and Logistics
Outsourcing	
% of outsourced development	0–25, 25–50, 50–75, 75–100%
% of outsourced hosting	0–25, 25–50, 50–75, 75–100%
CIA	
Level of availability required by critical data	Low, medium, high
Level of confidentiality required by critical data	Low, medium, high
Level of integrity required by critical data	Low, medium, high
IT Complexity	
FTE in IT	0–1 FTE, 1–2.5 FTE, 2.5–5 FTE, 5–10 FTE,> 10 FTE
IT expense as a percentage of revenue	<1, 1–2.5, 2.5–5, 5–10, >10%
Resilience against IT downtime	<10 min, 10 min to 1 h, 1–24 h, >24 h

which indicates that this participant has made a broad range in answers and probably believes that certain focus areas are clearly less important than others. As removing outliers also removes possible valid data, none of these (marked in a lighter color) outliers in Table 2 have been removed. The darker-indicated outlier has been removed due to false/useless information as aforementioned. The analysis that follows does indicate outliers to assure a proper analysis.

The participants are also asked to identify with which capabilities they felt they were expert in. Except for IT Asset Management, all capabilities are represented. See Table 3 for an overview. Note that participants could select multiple capabilities as their expert knowledge. The sum of participant's expertise therefore outreaches the sum of the participants.

Little's Test is performed on the data to understand if the missing values are predictive of other missing values. If it is not, the missing values are Missing Completely at Random (MCAR, see also "Impact analysis" section). With a Chi-square of 9.366 and DF of 546, and 75 iterations, a significance level of 1.000 is measured. Although the number of variables greatly outranks the number of observations, the data is assumed MCAR since the data does not lend itself for prediction of other missing values. The missing values in the dataset are presumably missing because participants quit the survey as they felt it was too long. This can be inferred because missing data is represented by completely answered questions, or completely unanswered.

This observation, combined with the results of Little's test, satisfy our assumption that the data is MCAR. Because the data is assumed MCAR, Multivariate Imputations by Chained Equations (MICE) are performed to fill the missing data points. The MICE package by Buuren et al. (2013) for the statistical programming environment R is used

Table 2 The participants, the mean of their answers and standard deviation

Sector	Size in FTE	$\bar{x}$	$\Delta\bar{x}/\bar{\bar{x}}$	s	$\Delta s/\bar{s}$
Financial, Banking and Insurance	50–250	17.95	4.92	8.53	1.37
Consultancy/Business Services	50–250	13.17	0.14	7.93	0.77
Education/Non-profit	1–9	13.74	0.72	7.6	0.44
Consultancy/Business Services	1–9	13.36	0.33	8.29	1.13
Consultancy/Business Services	More than 250	20.48	7.45	4.58	−2.58
Consultancy/Business Services	*More than 250*	*5.88*	*−7.14*	*1.99*	*−5.17*
Financial, Banking and Insurance	More than 250	12.73	−0.3	9.66	2.50
Consultancy/Business Services	10–50	12.86	−0.17	8.19	1.03
Financial, Banking and Insurance	More than 250	13.76	0.73	8.14	0.98
Consultancy/Business Services	More than 250	13.61	0.58	8.14	0.98
Consultancy/Business Services	50–250	13.48	0.45	8.2	1.04
Financial, Banking and Insurance	More than 250	13.49	0.47	8.25	1.09
Education/Non-profit	More than 250	12.72	−0.31	7.46	0.29
Industrial Production and Construction	More than 250	13.72	0.69	8.2	1.04
Consultancy/Business Services	10–50	11.46	−1.57	7.23	0.07
Telecommunication	More than 250	16.2	3.17	6.69	−0.47
Energy/Utilities/Mining	50–250	13.86	0.83	8.28	1.12
Consultancy/Business Services	1–9	10.78	−2.25	8.3	1.13
Industrial Production and Construction	More than 250	17.33	4.3	7.56	0.4
Consultancy/Business Services	More than 250	11.91	−1.12	7.86	0.7
Consultancy/Business Services	More than 250	13.62	−0.69	7.16	−0.36

Underline marks potential outliers, italic marks the deleted participant

Table 3 Expertise of the participants

Expertise	#	Expertise	#
Risk management	8	Policy development	9
Organizing information security	14	Human resource security	3
Compliance	9	Identity and access management	9
Secure software development	1	Incident management	4
Business continuity management	9	Change management	3
Physical and environmental security	3	IT asset management	0
Information security architecture	8		

to perform these imputations. As described by Azur et al. (2011), the amount of imputation iterations can be set to eight to get a strong enough result. However, following the guidelines by Rubin (1987), the worst-case scenario is 80% percent missing values, and thus the amount of imputations is set to 25. The amount of iterations is set to 20 after visual inspection of trace line plots (van Buuren and Groothuis-oudshoorn 2011). This assures a strong power without auto-correlation. The imputation method chosen is a Bayesian Linear Regression, as it fits the data best. It copes with a relative small amount of predictive measures and assumes that the residuals are close to normally distributed (van Buuren and Groothuis-oudshoorn 2011).

For further analysis using an ANOVA, the assumption is that the data is normal distributed. As these tests are performed on the measurement level, the variables of the capabilities are combined to create one single variable that describes the measurement

level. Following the comments on normality tests by Wilcox (1998) and Razali and Wah (2011), the Wilk-Shapiro test (Shapiro and Wilk 1965) is performed on the variables.

The influence of measurement levels on maturity frameworks

Each organizational characteristic is comprised of multiple measurement levels (see also Table 1). This section describes the measurement levels and the differences between them, to explain their influence on the complete maturity framework. The cases presented have been manually selected on the basis that they most clearly show the influence a measurement level can have on a maturity framework.

Sectors

The organizational characteristic Sectors describes the sector, or industry, an organization operates in. It is categorized in "general", as it does not pertain necessarily to information security. Note that the focus areas and answers are still described from an SME information security perspective. The sectors Aerospace and Defense, Consumer, Retail, Travel, Leisure, Entertainment and Media (henceforth Consumer), Finance, Banking and Insurance (henceforth Finance), and Business Services and Consultancy are detailed in Fig. 1.

From a visual inspection of the box-plots in Fig. 1, a clear distinction can be seen. Information security within the sectors Finance and Aerospace and Defense is of a higher priority than information security within Business Services and Consultancy and

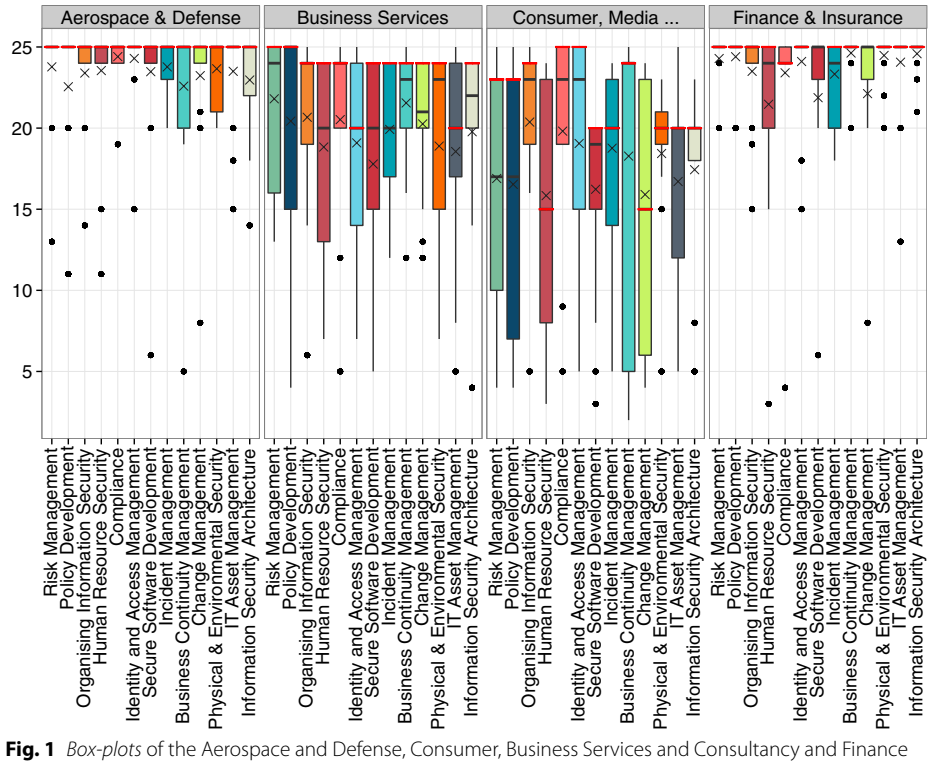

Fig. 1 *Box-plots* of the Aerospace and Defense, Consumer, Business Services and Consultancy and Finance sectors

Consumer. The results within the Finance and Aerospace and Defense sectors are also more consistent between the participants. Their distributions are clearly narrower.

As both the Finance and the Aerospace and Defense sector are not normal distributed (Shapiro–Wilk: $p < 0.001$ and $p < 0.001$ respectively), a Wilcox Robust ANOVA for dependent groups is performed (Wilcox and Erceg-Hurn 2012). The Wilcox Robust ANOVA uses $\hat{\Delta}$ as a measured statistic instead of the *F*-statistic calculated with the parametric ANOVA, and allows for the analysis of specific quantiles without the loss of power but with the advantage of a more detailed analysis. The 10, 20, 75 and 95th quantile are analyzed as the average resides approximately between the 10th and 20th percentiles and the upper quantiles show the largest differences and most modes reside there. These results show (see Table 4) that Finance does significantly differ from Business Services and Consultancy, and from the Consumer sector. Likewise, the ANOVA results from Aerospace and Defense on Consumer and Business Services and Consultancy show significant results over the full range.

These ANOVA results show that measurement levels as a whole influence the maturity framework at hand. By investigating both the full range from the 30th to the 95th percentile, these results are very robust. Not only does the majority of the answers significantly differ, outliers do as well.

Percentage of hosting outsourced

Participants are asked to rate the influence of the focus areas depending of the percentage of software services hosted at a third party. This is broad, Software as a Service

Table 4 ANOVA results

Quantile	$\hat{\Delta}$	Confidence interval		Critical P	*p* value
		Lower	Upper		
Finance and Business Services and Consultancy					
10th	6.457	4.859	9.027	0.05	0*
20th	4.907	3.733	5.282	0.025	0*
75th	0.862	0.208	0.999	0.0167	0*
95th	5.252×10^{-5}	1.49×10^{-8}	0.018	0.0125	0*
Finance and Consumer					
30th	9.229	7.647	9.909	0.025	0*
40th	6.71	5.233	8.321	0.0167	0*
85th	0.543	0.034	1.197	0.0125	0*
95th	1.061×10^{-6}	1.521×10^{-10}	0.002	0.050	0*
Aerospace and Defense and Consumer					
30th	9.144	7.783	9.838	0.05	0*
40th	6.813	5.535	8.532	0.025	0*
85th	0.766	0.111	1.334	0.0167	0*
95th	4.553×10^{-6}	1.062×10^{-9}	0.008	0.0125	0*
Aerospace and Defense and Business Services and Consultancy					
30th	58.263	4.473	8.6	0.05	0*
40th	4.937	4.093	5.120	0.025	0*
85th	0.956	0.417	1	0.01667	0*
95th	1.761×10^{-4}	8.153×10^{-8}	0.061	0.0125	0*

Asterisk signifies significance per quantile. The significance per quantile is identified by the column "Critical P", which is calculated as part of the Robust ANOVA calculation

comes in a variety of ways, from hosted office such as Google's Drive and Microsoft Office365 to traditional web-hosting. Utilizing hosted services implies providing organizational data to a third party and thus affects information security at large.

The box-plot in Fig. 2 shows wide distributions on the results, and although one might infer that the more outsourced hosting services are utilized the more information security becomes important, it is not very clear. The same results have been plotted in the scatter plot shown in Fig. 3. Using a Loess smoothing, as described in "Impact analysis" section, a clear pattern is shown.

The line follows approximately the same trajectory for each of the measurement levels, however the y-intercept per measurement level is higher, leading to the total smooth-function plotted higher in the scatter-plot. This indicates that from the perspective of utilizing hosting services at third party providers the importance of focus areas is always the same independent of the amount of services from the total need for services is outsourced. However, the more services are being outsourced, the more important information security becomes.

Following this visual inspection, an ANOVA is executed to further investigate the differences. A Shapiro–Wilk test shows that the measurement levels are not normal distributed (0–25%: $p < 0.001$, 25–75%: $p < 0.001$, 50–75%: $p < 0.001$, 75–100%: $p < 0.001$), thus a Wilcox robust ANOVA is selected for further analysis. As the Mode (indicated by a red bar) and the average (indicated by an X in the graph) fall both in almost all cases between the 25th and 75th percentile, the ANOVA will investigate those percentiles. This shows that even the two middle measurement levels are significantly different (Table 5).

This analysis reinforces the analysis performed by visual inspection. It shows that across the board each measurement level at each quantile significantly differ from each

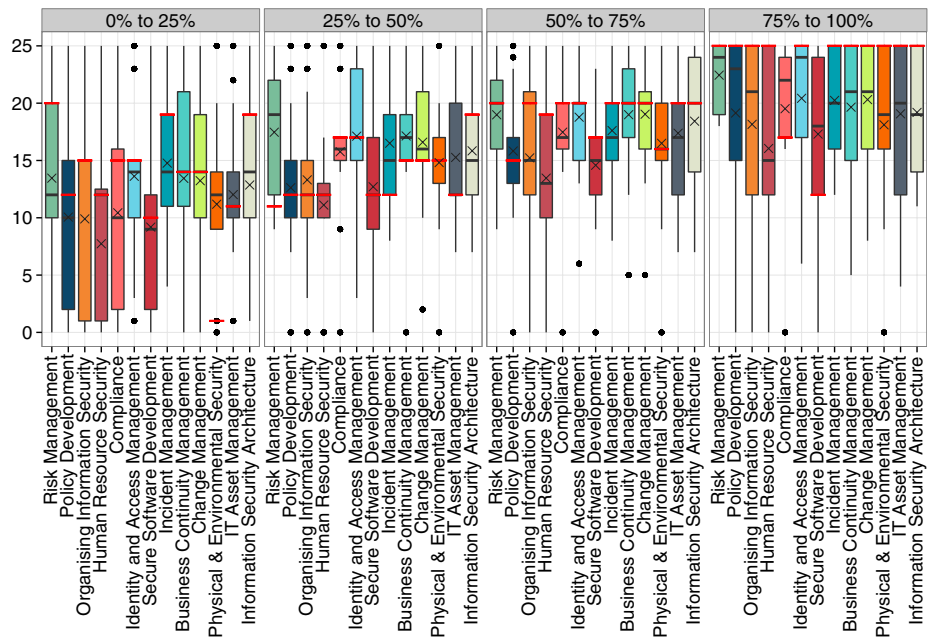

Fig. 2 *Box-plots* of the percentage of hosting outsourced, 0–25, 25–50, 50–75, 75–100%

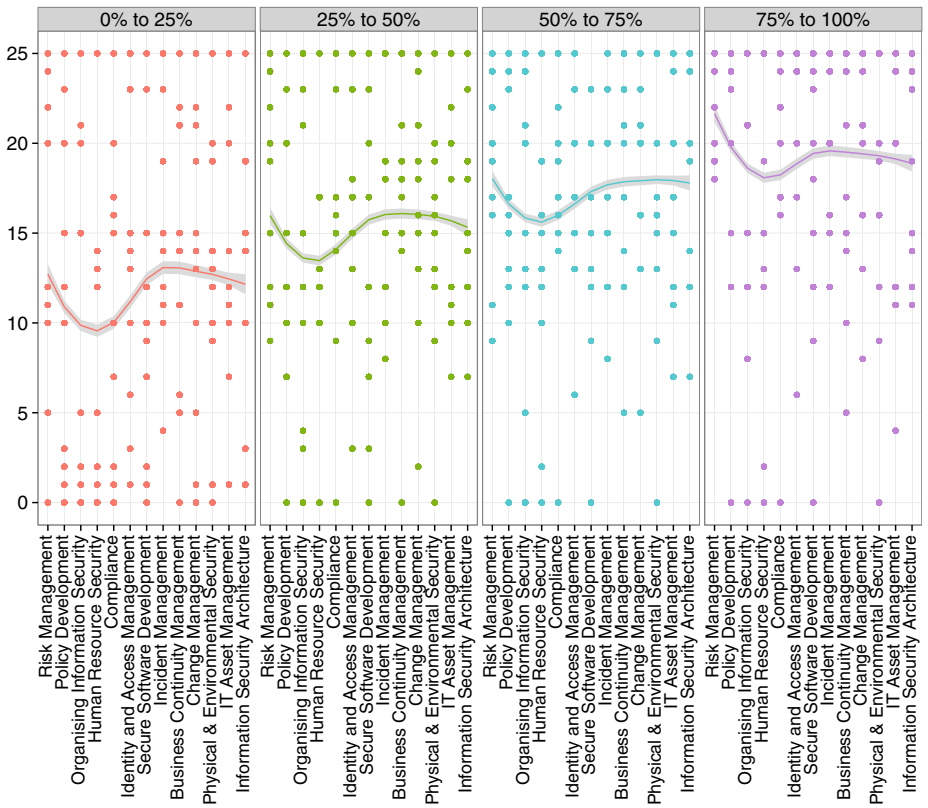

Fig. 3 Scatter-plots of the percentage of hosting outsourced, 0–25, 25–50, 50–75, 75–100%

other. This means that the ISFAM is heavily influenced by the amount of hosting outsourced at organizations.

FTE employed in the IT department

The amount of FTE employed in the IT department of an organization is an indicator of the complexity of the IT as part of organizational operations as well as a measure of dependence on IT by the organization. As part of the complexity measure it assumes that the more FTE employed, the more systems are in place that need people operating it. In case of a dependency measure, if an organization can feasibly employ more personnel in IT, it is inferred that IT is proportionally more important for the organization's operations. Either way, information security becomes more important. As systems become more complex, they are harder to secure and if organizations are more dependent on IT for their operations, the damage done by an intrusion becomes greater thus giving more weight to prevention of such an event.

The scatter-plot shown in Fig. 4 shows a smoothed Loess function over each of the measurement levels. The measurement levels, 0–1 FTE, 1–2.5 FTE, 2.5–5 FTE, 5–10 FTE and more than 10 FTE show a likewise pattern as found at percentage of hosting outsourced. The curve of the Loess smoothing follows the same path among all measurement levels, but the main difference is the height it is plotted at. Although the difference between

Table 5 ANOVA results for percentage of IT hosting outsourced

Quantile	$\hat{\Delta}$	Confidence interval		Critical P	*p* value
		Lower	Upper		
0–25, 25–50%					
25th	−6.731	−9.057	−3.377	0.05	0*
50th	−2.785	−3.588	−1.891	0.025	0*
75th	−3.655	−4.315	−1.815	0.0167	0*
0–25, 50–75%					
25th	−7.988	−10.254	−5.028	0.05	0*
50th	−5.366	−6.973	−4.196	0.025	0*
75th	−4.975	−5.869	−3.296	0.0167	0*
0–25, 75–100%					
25th	−9.496	−11.555	−6.194	0.05	0*
50th	−7.703	−8.542	−6.480	0.025	0*
75th	−9.495	−9.999	−6.909	0.0167	0*
25–50, 50–75%					
25th	−1.258	−2.633	−0.511	0.05	0*
50th	−2.581	−4.230	−15.721	0.025	0*
75th	−1.320	−2.601	−0.609	0.0167	0*
25–50, 75–100%					
25th	−2.765	−3.634	−1.116	0.05	0*
50th	−4.918	−5.659	−3.772	0.025	0*
75th	−5.840	−6.867	−4.732	0.0167	0*
50–75, 75–100%					
25th	−1.508	−2.155	−0.276	0.05	0.016*
50th	−2.337	−3.290	−0.627	0.025	0*
75th	−4.52	−4.915	−3.073	0.0167	0*

Asterisk signifies significance per quantile. The significance per quantile is identified by the column "Critical P", which is calculated as part of the Robust ANOVA calculation

first two measurement levels is not a large difference, the difference between the first and the last is clearly distinguishable.

As the results from FTE employed in IT Department are non-parametric, Shapiro-Wilk returns 0–1 FTE: $p < 0.001$, 1–2.5 FTE: $p < 0.001$, 2.5–5 FTE: $p < 0.001$, 5–10 FTE: $p < 0.001$, > 10 FTE: $p < 0.001$, a Wilcox Robust ANOVA is performed to further investigate the data. The quantiles under investigation are the 25th, 50th, 75th and 95th to understand the consensus between participants and the outlying participants who value information security higher.

Wilcox robust ANOVA reiterates the visual inspection. The quantiles all have significant differences (Table 6).

These results show that the complexity of IT as expressed by the amount of FTE working at the IT department influences maturity frameworks. As all quantiles show significant results it shows that even outliers at the 95th percentile significantly differ. Remarkably, even small changes in the Amount of FTE working at the IT department result in significant changes in the influence of information security as determined by the focus areas.

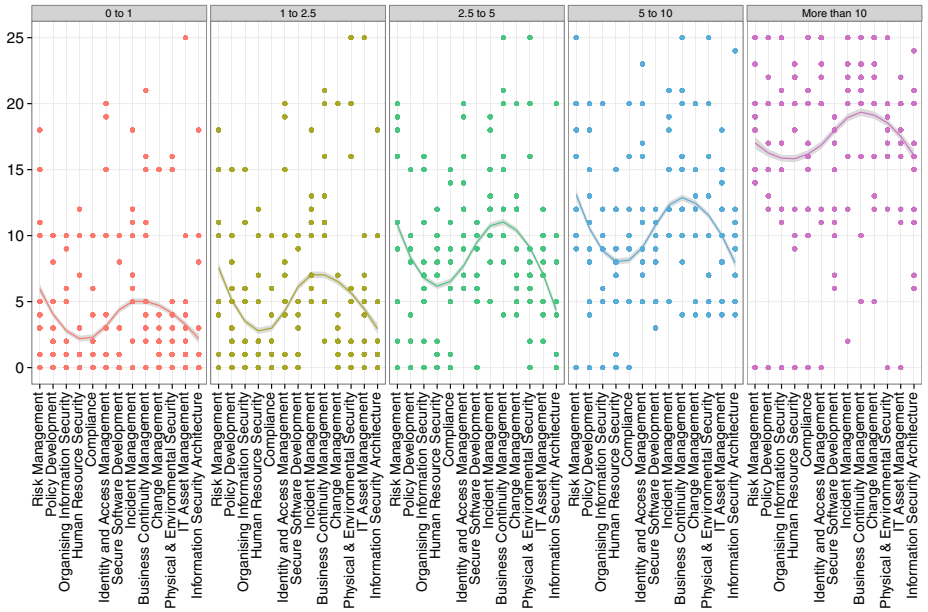

Fig. 4 Scatter-plots of the amount of FTE employed in the IT department, 0–1 FTE, 1–2.5 FTE, 2.5–5 FTE, 5–10 FTE, more than 10 FTE

Table 6 ANOVA results for percentage of IT hosting outsourced

Quantile	$\hat{\Delta}$	Confidence interval		Critical P	*P* value
		Lower	Upper		
0–1 FTE, 1–2.5 FTE					
25th	−0.406	−0.906	−0.012	0.05	0.002*
50th	−2.250	−2.871	−1.276	0.025	0*
75th	−1.813	−3.143	−1.016	0.0167	0*
95th	−1.643	−3.361	−0.623	0.0125	0*
1–2.5 FTE, 2.5–5 FTE					
25th	−3.752	−4.005	−3.008	0.025	0*
50th	−3.715	−4.753	−2.974	0.0167	0*
75th	−3.809	−4.896	−2.317	0.0125	0*
95th	−2.371	−4.063	−0.369	0.05	0.02*
1–2.5 FTE, >10 FTE					
25th	−13.666	−14.418	−11.937	0.05	0*
50th	−14.527	−16.043	−13.418	0.025	0*
75th	−13.977	−15.385	−12.058	0.0167	0*
95th	−8.571	−11.186	−6.046	0.0125	0*
5–10 FTE, >10 FTE					
25th	−8.498	−9.582	−6.791	0.05	0*
50th	−8.615	−9.778	−7.685	0.025	0*
75th	−6.963	−8.6	−5.555	0.0167	0*
95th	−4.913	−5.444	−3.696	0.0125	0*

Asterisk signifies significance per quantile. The significance per quantile is identified by the column "Critical P", which is calculated as part of the Robust ANOVA calculation

The influence of measurement levels on focus areas

Aside from the influence organizational characteristics have on the complete model, as described in 4.1, the results include a granular level of measurement: the influence of organizational characteristics on the focus areas in the ISFAM.

The results discussed in this section have been manually selected to show the influence as clearly as possible.

As the results of interest focuses on the general tendency, an ANOVA with trimmed means is applied. Trimmed means increase the power of the analysis substantially (Cribbie et al. 2012; Wilcox 1995).

Total employees employed

The Total Employees Employed is an organizational characteristic categorized as "General". It distinguishes organizations by the amount of employees that are working at an organization, whether it be full time or part-time. The measurement levels follow the guidelines set forth by the World Bank (Ayyagari et al. 2003) and the European Commission (The Commission of the European Communities 2003), namely: 0–9 employees, 10–49 employees and 50–250 employees.

It uses the number of employees instead of FTE, as employees working 0.1 FTE do need access to systems equally to full time employees, thus a possible risk. In a situation where 10 employees of 0.1 FTE are shown as 1 FTE, it could interfere with the risk assessment of the participants, resulting in wrong answers. The total number employees, regardless of the amount of hours they work, avoids this. Contrary to the results presented in "The influence of measurement levels on maturity frameworks" section, Total Employees Employed does not show a clear difference of the influence per measurement level. Instead, certain focus areas are highly affected, whereas others are not. See Fig. 5.

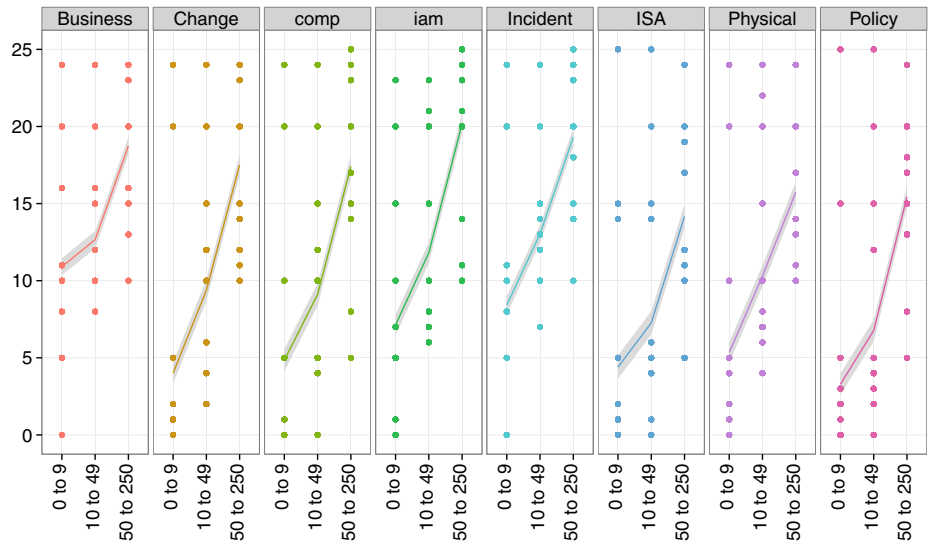

Fig. 5 Scatter-plots per focus area of the ISFAM displaying the Total Employees Employed organizational characteristic

Business Continuity Management shows a stark rise between the measurement levels, so does Organizing Information Security. Compliance and Risk Management are far less affected by the difference in measurement levels and show a near-linear line.

The results from Total Employees Employed is not normal distributed (Wilk–Shapiro test for normality: $p < 0.01$) thus a Wilcox Robust ANOVA is performed. These results underscore the visual inspection, as can be seen in Table 7. This analysis shows what the scatter plot in Fig. 5 also indicates, certain focus areas result in significant differences between the measurement levels, while other do not.

For example, the Information Security Architecture (ISA) is not significant between the levels 0–9 employees and 10–49 employees, but the difference between the other levels are significant. Likewise for Business Continuity Management and the contrary is true for Policy Development where only the difference between 0–9 and 10–49 employees is significant. This shows that under certain circumstances an organization has different needs for information security, implicating that maturity frameworks that do not account for these differences have a poor model fit.

Sectors (IT and Telecom, Agriculture and Forest and Transport and logistics)

Within the Sectors organizational characteristic, the influence of the sectors IT and Telecom, Agriculture and Forest, and Transport and Logistics on focus areas is apparent. Shown in Fig. 6, IT and Telecom clearly has higher information security requirements

Table 7 ANOVA results for total employees employed

Groups	*P* value	Critical P	Confidence interval	
			Lower	Upper
Risk management				
0–9, 10–49	0.045	0.05	−7.314	−1.708
0–9, 50–250	0	0.025	−12.065	−4.812
10–49, 50–250	0.01	0.0167	−6.132	−2.052
Policy development				
0–9, 10–49	0.019	0.025	1.388	11.76
0–9, 50–250	0.098	0.05	−1.012	8.12
10–49, 50–250	0.046	0.017	−5.345	−1.379
Organizing information security				
0–9, 10–49	0	0.05	−5.523	−2.877
0–9, 50–250	0	0.025	−17.1108	−11.791
10–49, 50–250	0	0.017	−13.108	−7.529
Compliance				
0–9, 10–49	0	0.05	−7.212	−3.102
0–9, 50–250	0	0.025	−18.234	−10.735
10–49, 50–250	0	0.017	−13.545	−5.397
Business continuity management				
0–9, 10–49	0.322	0.05	−3.486	0.492
0–9, 50–250	0	0.025	−11.982	−5.037
10–49, 50–250	0	0.017	−9.517	−4.514
Information security architecture				
0–9, 10–49	0.228	0.05	−5.046	−1.659
0–9, 50–250	0	0.025	−14.892	−8.129
10–49, 50–250	0	0.017	−11.443	−5.151

than the other two sectors. However, the results are not identical at all focus areas. Compliance and Incident Management show gradual slopes whereas Policy Development, Information Security Architecture (ISA), and Identity and Access Management (IAM) among others have steep slopes with denoting differences in values between each sector. This shows that the importance of the focus areas is different per measurement level.

The data comprising the sectors IT and Telecom and Logistics and Transportation is non-parametric, Shapiro-Wilk reports a $p < 0.001$ and $p < 0.01$ respectively. Twenty-five percent of the focus areas in the Agriculture and Foresting sector are normally distributed. The averaged p-score of 0.075 returned by the Shapiro-Wilk test does indicate the data is normally distributed, however as for 75% of the focus areas the data has a p-score lower than 0.01, a parametric ANOVA is unsuitable for that part of the data.

Therefore, a Wilcox Robust ANOVA is performed to further investigate the Focus Areas in all the aforementioned sectors. The results are displayed in Table 8.

Table 8 shows that depending on the sector an organization operates in the, certain focus areas are significantly different. The focus area Physical Security for example is significantly different between the IT and Telecom sector and Agriculture and Foresting, whereas the difference in Physical Security between IT and Telecom and Transport and Logistics is not significantly different.

IT and Telecom and Agriculture and Forest differ in more focus areas significantly, such as Change Management and Business Continuity Management. Certain focus areas show significant differences between all sectors tested, whereas other focus areas such as Compliance only shows a significant difference between IT and Telecom and Transport and Logistics.

This shows that the sector an organization operates in significantly influences the information security needs of an organization, but at a fine-grained level.

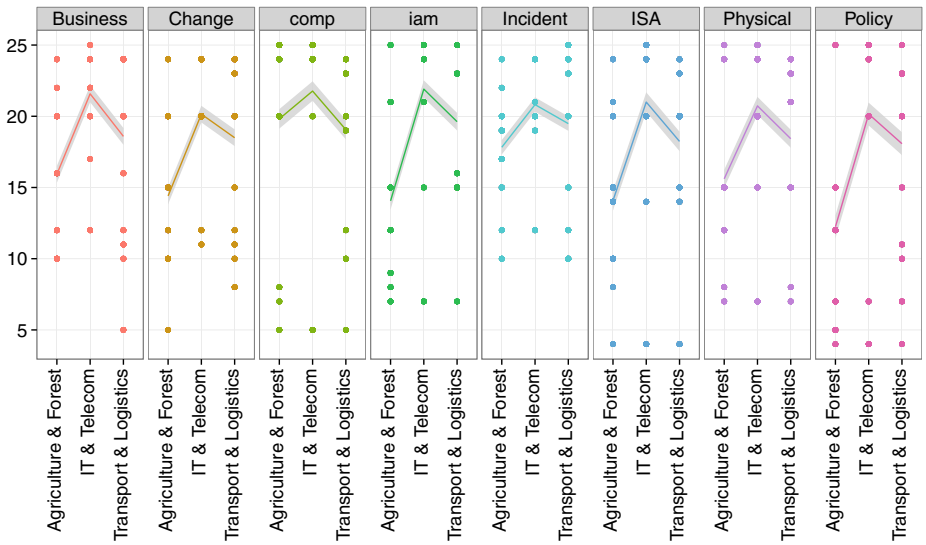

Fig. 6 Scatter-plots per focus area of the ISFAM displaying the sectors IT and Telecom, Transport and Logistics and Agriculture and Forest

Table 8 ANOVA results for IT and Telecom, Agriculture and Forest and Transport and Logistics

Groups	*P* value	Critical P	Confidence interval	
			Lower	Upper
Risk management				
IT, Agriculture	0.001	0.025	4.369	9.206
IT, Transport	0.08	0.05	0.135	1.785
Agric., Transport	0.001	0.017	−8.637	−3.166
Organizing information security				
IT, Agriculture	0	0.05	5.782	10.726
IT, Transport	0.007	0.025	0.782	4.988
Agric., Transport	0.01	0.017	−8.785	−1.837
Compliance				
IT, Agriculture	0.063	0.025	0.095	5.56
IT, Transport	0.002	0.017	1.129	6.212
Agric., Transport	0.506	0.05	−1.806	4.034
Business continuity management				
IT, Agriculture	0	0.0167	3.231	10.274
IT, Transport	0.056	0.025	0.182	6.2
Agric., Transport	0.208	0.05	−8.363	1.305
Change management				
IT, Agriculture	0	0.0147	3.4	10.191
IT, Transport	0.271	0.05	−0.582	4.191
Agric., Transport	0.049	0.025	−9.225	−0.388
Physical security				
IT, Agriculture	0	0.0167	3.062	9.517
IT, Transport	0.05	0.025	0.348	5.615
Agric., Transport	0.121	0.05	−7.219	0.255

Confidentiality, integrity and availability (CIA-triad)

Confidentiality, Integrity and Availability is often used for classifying data (Joint Technical Committee ISO/IEC JTC 1 2008), and is not a tool in itself. The wide distributions showing in these graphs account for that. Participants more likely have different perspectives and experiences with situations concerning the requirements for Confidentiality, Integrity and Availability. Participants were asked how they would rate the importance of the focus areas considering the Confidentiality, Integrity or Availability of their critical data was either high, medium or low. Critical data are defined as data needed for the organization to operate.

At all three organizational characteristics, the focus areas pertaining to those organizational characteristics are ranked more important. Especially availability shows this, as availability is a very clearly defined concept.

Focus areas such as Business Continuity Management pertain fully to Availability, as well as Change Management, which endangers the availability of data and services, is very well defined in the graph.

Although more processes within an organization define the level of Integrity and Confidentiality, there are still focus areas that are greatly influenced by it. In the

Confidentiality graph, Incident Management and Secure Software Development are clear indicators. Risk Management and Incident Management stand out.

The Shapiro–Wilk test returns $p < 0.01$ for Confidentiality $p < 0.01$, and for Availability $p < 0.01$. Following the results from the Shapiro-Wilk test for normality, a Wilcox Robust ANOVA is performed (Table 9).

The analysis performed on the Confidentiality, Integrity and Availability organizational characteristics shows that the needs per organization differ significantly over a majority of the cases tested.

First, it shows that organizational characteristics influence the focus areas of the ISFAM significantly, but that these influences differ per focus area. For instance, Secure Software Development shows significant differences in both Confidentiality and Integrity, but not for the Availability organizational characteristic. Thus, for Availability measurement levels, Secure Software Development is not of a significant influence, whereas for Confidentiality and Integrity it is.

Second, within organizational characteristics, the level of influence differs. This can be seen in Human Resource Security regarding confidentiality, where the difference between the medium to high level is significant, the differences between the other measurement levels are not. Likewise, integrity influences information security architecture only if the measurement levels are low and medium, otherwise there is no significant effect.

Third, it shows that concepts that convey organizational characteristics, such as the aforementioned Business Continuity management in Availability, do indeed differ significantly between the levels. Rigid maturity frameworks that do not account for these differences will thus interfere with a proper implementation of information security.

Analysis

The results presented in "The influence of measurement levels on maturity frameworks" and "The influence of measurement levels on focus areas" sections show that organizational characteristics influence maturity frameworks on two levels.

On the top level, measurement levels within organizational characteristics such as the Percentage Hosting Services Outsourced and the Amount of FTE Employed in the IT Department influence the complete maturity framework. The importance of all focus areas in the maturity framework shift up or down depending on the measurement level of these organizational characteristics. Thus, depending on the organization, a maturity framework can either be too strict and expect capabilities for a maturity level that are infeasible if the importance of focus areas shifts down. If an organizational characteristic determines the opposite, focus areas become more important than maturity levels within the framework, and underestimate the capabilities needed to reach a certain level.

In the analysis described in in "The influence of measurement levels on maturity frameworks" section, different quantiles are analyzed in order to understand the results in more detail. The investigation regarding outliers at the 95th quantile allows for a deeper understanding of participants who feel that security is more important than the average participant.

Even at these quantiles, the results are significant for all measurement levels at the organizational characteristics Amount of FTE Employed at the IT department, and

Table 9 ANOVA results for Confidentiality, Integrity and Availability

Groups	Confidentiality				Integrity				Availability			
	p val.	Crit. P	Conf. interval		*p* val.	Crit. P	Conf. interval		*p* val.	Crit. P	Conf. interval	
			Lower	Upper			Lower	Upper			Lower	Upper
Risk management												
Low, medium	0.224	0.05	−0.806	3.735	0.308	0.05	−1.225	4.585	0.392	0.05	−2.009	3.819
Low, high	0.002	0.025	2.588	12.579	0.034	0.025	1.132	12.394	0.077	0.025	−0.197	9.431
Medium, high	0.001	0.067	2.342	9.923	0.020	0.017	1.222	9.022	0.052	0.017	0.375	7.117
Human resource security												
Low, medium	0.364	0.05	−1.354	4.157	0.143	0.05	−0.157	6.791	0.179	0.05	−0.468	5.394
Low, high	0.049	0.025	0.539	10.625	0.018	0.025	1.874	11.819	0.009	0.025	1.895	11.434
Medium, high	0.007	0.017	1.28	6.929	0.008	0.017	1.123	5.834	0.019	0.017	1.111	7.428
Secure software development												
Low, medium	0.022	0.05	1.338	7.083	0.018	0.05	1.945	6.902	0.462	0.05	−2.098	4.065
Low, high	0.001	0.025	4.105	13.055	0	0.025	6.129	13.831	0.046	0.025	0.015	13.037
Medium, high	0.008	0.017	1.203	7.409	0.002	0.017	2.117	8.911	0.022	0.017	0.785	11.166
Business continuity management												
Low, medium	0.005	0.05	0.942	5.603	0.006	0.05	1.289	6.046	0	0.05	1.529	4.36
Low, high	0.001	0.025	2.065	10.452	0.006	0.025	1.895	9.865	0	0.025	6.505	11.579
Medium, high	0.029	0.017	0.419	5.591	0.283	0.017	−0.871	4.963	0	0.017	3.846	8.462
Information security architecture												
Low, medium	0.014	0.05	1.163	5.197	0.301	0.05	−1.323	5.655	2.449	0.05	0.047	−0.769
Low, high	0.001	0.025	3.582	11.102	0.017	0.025	1.985	11.246	8.215	0.025	0.023	4.08
Medium, high	0.008	0.017	1.606	6.782	0.021	0.017	1.182	7.582	5.766	0.017	0.021	2.105

Underlined are the results that are not significant. Themeasured *p* values, and their respective signicance values are shown in the columns "*p* val." and "Crit. P"

between the sectors Finance, Business Services and Consultancy, Consumer and Aerospace and Defense.

Furthermore, at the lower quantiles, all measurement levels differ significantly. This clearly shows that the different measurement levels have an impact on the focus areas, and therefore the maturity framework at hand. The ISFAM framework under investigation does not account for these measurement levels. The earlier voiced criticism that maturity frameworks like the ISFAM are too rigid is correct. The ISFAM cannot account for the significant differences organizational characteristics present within these models.

On the second level, Organizational characteristics such as Total Employees Employed, certain sectors and the CIA-triad influence specific focus areas, but not necessarily the whole maturity framework. See "The influence of measurement levels on focus areas" section.

These focus areas become more, or less, important in comparison to other focus areas in the framework. This can be explained by organizational characteristics pertaining directly to concepts covered by focus areas. Availability for instance, a part of the CIA-Triad, has a direct link with Business Continuity Management.

Business Continuity Management assures Availability, and this is shown in the results by a significantly higher level of importance than other focus areas in the ISFAM framework.

In other cases, the measurement level is not directly involved with the focus area such as the aforementioned case of Availability, but does influence the focus area significantly. Total Employees Employed shows this, Risk Management, Compliance, and Secure Software Development are not directly a part of the amount of employees working at an organization, but do change significantly per measurement level.

In either case, organizational characteristics influence maturity frameworks on a granular (focus area) level. In the analysis described in "Confidentiality, integrity and availability (CIA-triad)" section, where Confidentiality, Integrity and Availability are investigated, it is shown that both levels of influence occur simultaneously.

Availability does not significantly influence Secure Software Development, but does significantly influence Business Continuity Management. At the measurement level it significantly influences Policy development between the level Low and High, and Compliance only between Low and High, and Medium and High showing that the difference in Compliance between the lower levels is not significantly different. Likewise occurs for Confidentiality and Integrity, albeit at different focus areas.

Current rigid maturity frameworks cannot cope with the differences between organizations such as presented by organizational characteristics. These results show that rigid maturity frameworks are inept at accounting for the differences between organizations and oversimplify organizations at large.

When maturity frameworks are followed to the letter, it can result in grievances at the implementers and the implementation of wrong capabilities or a wrong priority order of implementations. Organizational characteristics show that organizations differ between one another and that these differences influence the maturity framework at multiple design levels: the measurement level, the focus area level, and the overall maturity level.

The results described in section four substantiate earlier claims that maturity frameworks oversimplify reality and lack a good model fit (Bollinger and McGowan 1991;

Pöppelbuß and Röglinger 2011), however there are some minor points that are worth mentioning.

The result set is broad, with 624 variables, but at 20 observations not deep. The discussion surrounding the results has focused on the mode, and the most often returned answer. However, the authors acknowledge that further detailed research into a subset of the results is warranted for a better understanding and development of an adaptive maturity model, further increasing the power of the results.

In "Sectors" section the Sector agriculture comes in significantly different from IT, and the accompanying Fig. 6 shows that overall Agriculture has scored lower than other sectors. No participant active in the Agriculture sector was present in our result set however.

Nonetheless, the survey system developed guaranteed robust data and obviated issues with non-response and the length of the survey. That being said, in hindsight this survey was too long and possible better results pertaining the direct influence of organizational characteristics on information security at small and medium enterprises would have been more robust if a smaller part of information security was addressed.

The results are distinct enough that the authors feel content for the results pertaining to the ISFAM framework. Although these results are generalizable towards information security at SMEs, the organizational characteristics have been developed from the perspective of the ISFAM.

Due to the already many questions in the survey, questions regarding the relations of Organizational characteristics between themselves, and the relations and priority of the focus areas have not been asked. These could provide a baseline for adaptive maturity frameworks.

The dataset does show some results that are not discussed in "Results and Discussions" section, but are interesting nonetheless.

At a 25-point scale, the results show the participant's tendency to select values at round numbers (0, 5, 10, 20, and 25). The authors are unsure why this effect happens, but the expected advantage of 25 points seems does not seem to resonate at the participant's perspective.

There are less outliers at the upper measurement levels than at the lower end of the measurement level scales. This might be because the majority of the participants comes from a large enterprise and their experience is with larger firms as well.

In measurement level scales with four (or more) levels, the middle two levels are more aligned than the first and second and the last and second-to-last. The authors suspect that this is the law of large numbers at work. A normal-distributed scale (16, 34, 34, 16%) could account for this, but in the pre-study participants noted that this was incomprehensible for answering (for example in the outsourcing organizational characteristic).

At certain measurement levels, the distribution of answers fill the full spectrum from null to twenty-five. This could be because certain measurement levels are precise (the fine-grained measurement levels at FTE employed in IT for example) or requires very specific knowledge to return correct answers such as the case is with outsourcing of hosted services. This has greatly changed in the last decade with the rise of cloud computing and Software as a Service models.

Another identifier for these broad ranges of answers is the scope of a measurement level or organizational characteristic. One can outsource nearly every imaginable service or operation such as observed in Baars et al. (2012), and certain sectors such as IT and Telecom and Business Services and Consultancy can be very broad interpreted. Nevertheless, the differences between Sectors, and the amount of Outsourcing used at both development of software and hosting services are astute and evidently underscore that maturity frameworks should account for these differences to provide a better model fit. The organizational characteristics and their respective measurement levels are investigated for their influence of focus areas. Focus Area Maturity Matrices have a lower level object of measurement: the capabilities that comprise one focus area. Broad distributions from the participant responses could be showing that the focus areas in the ISFAM are too broadly scoped. This research, and the investigations by Mijnhardt et al. (2014) do not account for this. Another aspect not analyzed in this investigation is the co-influence of organizational characteristics.

In certain cases it might be that only a sector is not a significant influencing factor on the importance of focus areas, but it gains that significant influence when brought in perspective with another organizational characteristics. For example, a firm developing software for general use web applications will have a different influence on the focus areas then when this firm is developing software for intelligence agencies.

Finally, external aspects such as standards and legislation influence the results. The health sector has seen the implementation of a series of legislation and accompanying standards. These most probably have influenced the results, setting the overall importance of all focus areas higher over sectors that do not feature standards nor legislation.

Although this investigation does not account for these external influences, they do enforce the argument for adaptive maturity frameworks. External influences such as standards and legislation influence maturity frameworks and their model fitness.

Conclusions

One size does not fit all: our data and analyses clearly show that not all SMEs are created equal. Organizations differ from each other, and models that do not account for these differences are oversimplifying reality. Especially maturity matrices could use the input of different organizational characteristics to have a better fit with the organization that is applying them.

The results clearly show a disproportional difference of the importance of capabilities between measurement levels, and between organizational characteristics. Striking are the examples of different industries, see for example Fig. 6 where the IT and Telecom and the Agricultural sector are compared, and the clear difference of the focus area Policy Development and the effects of the amount of FTE employed in an organization (see Fig. 5).

The investigation in "Confidentiality, integrity and availability (CIA-triad)" section on Confidentiality, Integrity and Availability shows clearly how organizational characteristics influence both the maturity framework as the focus areas that it comprises.

The results underline the importance of organizational characteristics in maturity matrices, and presumably so in other benchmark and implementation tools. The differences between organizations are too large to ignore in modelling.

Ignoring organizational characteristics could result in negative consequences for an organization such as unnecessary implementation of capabilities, the wrong order of

priority when implementing capabilities or over-implementing of capabilities. Not only does this bring along unnecessary costs and pressure on an organization, it might also bring along negative impact on organizations in terms of productivity and innovation.

This conclusion is in line with qualitative results from earlier results in Maturity Matrices (Baars et al. 2012). In line with the results, an enhanced version of the ISFAM has been developed which implements a weighted model to account for organizational characteristics. Although a proof-of-concept, it shows the dependencies of capabilities in response to organizational characteristics.

Authors' contributions

TB made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data; has been leading in drafting the manuscript. FM made substantial contributions to conception and design, or analysis and interpretation of data. KV made substantial contributions to the design and implementation of the custom survey system which was used to gather the empirical data. MS initiated and supervised the research; conveived the focus area maturity model design which integrates situational factors, made substantial contributions to conception and design of data analysis and interpretation of data; has been involved in revising the manuscript critically for important intellectual content and co-authored the paper. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the Dutch chapters of Norea, ISACA and (ISC)2 and the PvIB for their cooperation in this research. Funding was provided by Netherlands Enterprise Agency (Grant No. SBIR12C021).

Competing interests

The authors declare that they have no competing interests.

Received: 22 July 2016 Accepted: 1 November 2016
Published online: 10 November 2016

References

Ayyagari M, Beck T, Demirgüç-Kunt A. Small and medium enterprises across the globe: a new database. Policy Research Working Papers. 2003. doi:10.1596/1813-9450-3127.

Azur MJ, Stuart EA, Frangakis C, Leaf PJ. Multiple imputation by chained equations: what is it and how does it work? Int J Methods Psychiatr Res. 2011;20(1):40–9. doi:10.1002/mpr.329.

Baars T, van den Bemd L, Theuns M, van den Akker R, Schönbeck M, Brinkkemper S. Cyber security in smart grid substations. 2012. http://www.cs.uu.nl/research/techreps/repo/CS-2012/2012-017.pdf.

Becker J, Knackstedt R, Pöppelbuÿ J. Developing maturity models for IT management. Bus Inf Syst Eng. 2009;1(3):213–22. doi:10.1007/s12599-009-0044-5.

Bekkers W, Spruit M. The situational assessment method put to the test: improvements based on case studies. In: 4th International workshop on software product management. IEEE. Sydney; 2010. p. 7–16.

Bekkers W, van de Weerd I, Brinkkemper S, Mahieu A. The influence of situational factors in software product management: an empirical study. In: 2008 Second international workshop on software product management IEEE. Barcelona; 2008. p. 41–48. ISBN 978-1-4244-4083-2. doi:10.1109/IWSPM.2008.8.

Bollinger TB, McGowan C. A critical look at software capability evaluations. IEEE Softw. 1991;8(4):25–41.

Buuren SV, Groothuis-Oudshoorn K, Robitzsch A, Vink G, Jolani S, Doove L. Package âĂŸ mice âĂŹ. Technical report, CRAN, 2013. http://cran.rproject.org/web/packages/mice/mice.pdf.

CBS, Share organizations hit by ICT-securityincidents [orig. in Dutch] Amsterdam: Statistics Netherlands (CBS); 2011. http://www.cbs.nl/nl-NL/menu/themas/bedrijven/cijfers/incidenteel/maatwerk/2011-3293-tab.htm.

Cleveland WS, Devlin SJ. Locally weighted regression: an approach to regression analysis by local fitting. J Am Stat Assoc. 1988;83(403):596–610. http://www.jstor.org/stable/2289282.

CMMI Product Team, CMMIÅő for Development, Version 1.3 CMMI-DEV, V1.3, Technical Report November, Software Engineering Institute, Carnegie Mellon University, Pittsburgh; 2010. http://www.sei.cmu.edu/reports/10tr033.pdf.

Cribbie RA, Fiksenbaum L, Keselman HJ, Wilcox RR. Effect of non-normality on test statistics for one-way independent groups designs. Br J Math Stat Psychol. 2012;65(1):56–73. doi:10.1111/j.2044-8317.2011.02014.x.

de Bruin T, Freeze R, Kulkarni U, Rosemann M. Understanding the main phases of developing a maturity assessment model. In: Proceedings of the ACIS 2005. Sidney: AIS; 2005. http://aisel.aisnet.org/acis2005/109.

de Leeuw ED, Hox J, Huisman M. Prevention and treatment of item nonresponse. J Off Stat. 2009;19(2):153–76.

Duncan RB. Characteristics of organizational environments and perceived environmental uncertainty. Adm Sci Q. 1972;17(3):313–27. http://www.jstor.org/stable/2392145.

Eisenberg M, Barry C. Order effects: a study of the possible influence of presentation order on user judgments of document relevance. J Am Soc Inf Sci. 1988;39(5):293–300.

Frigge M, Hoaglin DC, Iglewicz B. Some implementations of the boxplot. Am Stat 1989;43(1):50–4. http://www.jstor.org/stable/2685173.

Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci. 2007;8(3):206–13. doi:10.1007/s11121-007-0070-9.

Guttman B, Roback EA. Special Publication 800-12 an intorduction to computer security: the NIST handbook. Technical Report 800, National Institute for Standards and Technology, Gaithersburg; 1995. 3019753058. http://csrc.nist.gov/publications/nistpubs/800-12/handbook.pdf.
Iversen J, Nielsen PA, Nerbjerg J. Situated assessment of problems in software development. DATA BASE Adv Inf Syst. 1999;30(2):66–82.
Joint Technical Committee ISO, IEC JTC 1, ISO, IEC 27002, Technical Report november 2007, International Organization for Standardization/International Electrotechnical Commission. Geneva; 2008.
King JL, Kraemer KL. Evolution and organizational information systems: an assessment of Nolan's stage model. Commun ACM. 1984;27(5):466–75.
Kuilboer JP, Ashrafi N. Software process and product improvement: an empirical assessment. Inf Softw Technol. 2000;42(1):27–34. doi:10.1016/S0950-5849(99)00054-3.
Little RJA. A test of missing completely at random for multivaraite data with missing values. J Am Stat Assoc. 1988;83(404):1198–202.
McCormack K, Willems J, Bergh JVD, Deschoolmeester D, Willaert P, Temberger MI, Krinjar R, Trkman P, Ladeira MB, Oliveira MPVD, Vuksic VB, Vlahovic N. A global investigation of key turning points in business process maturity. Bus Process Manag J. 2009;15(5):792–815. ISBN 1463715091098. doi:10.1108/14637150910987946.
Mettler T, Rohner P, Winter R. Towards a classification of maturity models in information systems. In: A. D'Atri, M. De Marco, A.M. Braccini, F. Cabiddu, editors Management of the interconnected world. Heidelberg: Physica-Verlag HD; 2010. p. 333–341. ISBN 978- 3-7908-2403-2. doi:10.1007/978-3-7908-2404-9.
Mijnhardt F, Baars T, Spruit M. Organizational characteristics influencing information security maturity. J Comput Inf Syst. 2014;56(2):106–15.
Paulk MC, Curtis B, Chrissis MB, Weber CV. Capability maturity model for software, an analytics approach to adaptive maturity aodels 29 Version 1.1. Technical Report February. Pittsburgh: Carnegie Mellon University; 1993. http://www.sei.cmu.edu/reports/93tr024.pdf.
Pöppelbuß J, Pöppelbuÿ J, Röglinger M. What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in buisiness process management. In: Proceedings of the ECIS 2011. Helsinki; 2011. http://aisel.aisnet.org/ecis2011/28/.
Razali NM, Wah YB. Power comparisons of Shapiro–Wilk, Kolmogorov–Smirnov, Lilliefors and Anderson-Darling tests. J Stat Model Anal. 2011;2(1):21–33.
Rubin DB. Multiple imputation for nonresponse in surveys. 1st ed. New York: Wiley; 1987. ISBN 0-471-08705-X.
Ruiter JT. Cost of cyber crime largely met by businesses, 2012. https://www.tno.nl/content.cfm?context=overtno&content=nieuwsbericht&laag1=37&laag2=2&itemid=2012-04-1011:37:10.0&Taal=2.
Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika 1965;52(3):591–611. http://www.jstor.org/stable/2333709.
Silveira V. Updating your password on LinkedIn and other account security best practices. 2012. http://blog.linkedin.com/2012/06/06/updating-your-password-on-linkedinand-other-account-security-best-practices/.
Siminski P. Order effects in batteries of questions. Qual Quant. 2006;42(4):477–90. doi:10.1007/s11135-006-9054-2.
Spruit M, Roeling M. ISFAM: the information security focus area maturity model. In: Proceedings of the twenty second European conference information system ECIS 2014, Tel Aviv, Israel, 2014.
Spruit M, de Boer T. Business intelligence as a service: a vendor's approach. Int J Bus Intell Res. 2014;5(4):26–43.
Steenbergen M, Bos R, Brinkkemper S, Weerd I, Bekkers W. The design of focus area maturity models. In: Winter R, Zhao JL, Aier S, editors, Global perspectives on design science research, LNCS6105, 6105 edn. St. Gallen: Springer; 2010. p. 319–332. ISBN 978-3-642-13334-3. doi:10.1007/978-3-642-13335-0_22.
The Associated Press, Yahoo Email Account Passwords Stolen, 2014. http://www.npr.org/2014/01/31/269186875/yahoo-email-account-passwords-stolen.
The Commission of the European Communities, COMMISSION RECOMMENDATION of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises. Off J Eur Union 2003;124:36–41. http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:124:0036:0041:en:PDF.
Thong JYL, Yap CS. CEO characteristics, organizational characteristics and information technology adoption in small businesses. Omega Int J Manag Sci. 1995;23(4):429–42. doi:10.1016/0305-0483(95)00017-I.
Tukey JW. Exploratory data analysis. Massachusetts: Addison-Wesley Professional, Reading; 1977. ISBN 0-201-07616-0.
van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations. R J Stat Softw. 2011;45(3). http://www.stefvanbuuren.nl/publications/MICEinR-Draft.pdf.
van Straalen V. Many ICT-security incidents at organizations [original in Dutch], 2011. http://www.cbs.nl/nl-NL/menu/themas/bedrijven/publicaties/digitaleeconomie/artikelen/2011-3293-wm.htm.
Watts HS. Characterizing the software process: a maturity framework. Technical report CMU/SEI-87-TR-11, Carnegie-Mellon University. 1987.
Weimer-Jehle W. Cross-impact balances: a system-theoretical approach to cross-impact analysis. Technol Forecast Soc Change. 2006;73(4):334–61. doi:10.1016/j.techfore.2005.06.005.
Weimer-Jehle W. Cross-impact balances applying pair interaction systems and multi-value Kauffman nets to multidisciplinary systems analysis. Phys Stat Mech Appl. 2007;387:3689–700. doi:10.1016/j.physa.2008.02.006.
Wilcox RR. ANOVA: a paradigm for low power and misleading measures of effect size? Rev Educ Res. 1995;65(1):51–77. doi:10.3102/00346543065001051.
Wilcox RR. How many discoveries have been lost by ignoring modem statistical methods? Am Psychol. 1998;53(3):300–14.
Wilcox RR, Erceg-Hurn DM. Comparing two dependent groups via quantiles. J Appl Stat. 2012;39(12):2655–64. doi:10.1080/02664763.2012.724665.
Wilcox RR, Erceg-Hurn DM, Clark F, Carlson M. Comparing two independent groups via the lower and upper quantiles. J Stat Comput Simul. 2013;84:1–9. doi:10.1080/00949655.2012.754026.