12-7-2010

Evaluating the Applicability of a Use Value-Based File Retention Method

Fons Wijnhoven
University of Twente, a.b.j.m.wijnhoven@utwente.nl

Chintan Amrit
University of Twente, c.amrit@utwente.nl

芽|Sprouts Working Papers on Information Systems ISSN 1535-6078

Evaluating the applicability of a use value-based file retention method

Fons Wijnhoven
University of Twente, Netherlands

Chintan Amrit
University of Twente, Netherlands

Abstract
A well constructed file retention policy can help a company determine the relative value and the corresponding retrieval service level of the different files it owns. Though such a retention policy is useful, the method one can use to arrive at such a policy is under-researched. This paper discusses how one can arrive at a method (based on a systematic literature review) for developing file retention policies based on use values of files. In the case study, we demonstrate how one can develop a file retention policy by testing of causal relations between file retention policy parameters and the use value of files. This case study shows that, contrary to suggestions of previous research, the file type has no significant causal relation with the value of a file and thus should be excluded from a retention policy in this case. The case study also shows that there is a strong causal relation between the position of a user of a file and the value of this file. Furthermore, we have amended an existing subjective file valuation method, namely, the Information Value Questionnaire (IVQ). However, to make file retention methods effective and reliable a substantially more case experiences need to be collected.

Keywords: File retention; use value of data

Permanent URL: http://sprouts.aisnet.org/10-118

Reference: Fons Wijnhoven, Chintan Amrit (2010). "Evaluating the applicability of a use value-based file retention method," Proceedings > Proceedings of SIGSVC Workshop . *Sprouts: Working Papers on Information Systems*, 10(118). http://sprouts.aisnet.org/10-118

EVALUATING THE APPLICABILITY OF A USE VALUE-BASED FILE RETENTION METHOD

1. Introduction

The goal of file retention policies is to store data on the appropriate medium that provides the required service level in the different stages of the data lifecycle [Tanaka et al. 2005]. Here, by file we mean a digital document, and we use the term file and document interchangeably at some places, while retaining the above meaning. The lifecycle of files consists of four stages. The first stage is the creation of new file data or the modification of existing file data. In the second stage, the file is made accessible for other users, by for instance digital, written or verbal communication. The third stage is the actual access and usage of the file. After a period of usage, the file is either archived or deleted. The final stage is called retention. Throughout its lifecycle, the value of a file in general grows after the first stage and declines in the final stage [Tallon and Scannell 2007]. In the final stage, the intensity of usage decreases and the accessibility of the files becomes less important. But, not all types of files have the same value and the evolution of file value over time may differ per file type. Consequently, one of the most important functions of a successful file retention policy is the ability to differentiate files by values in an unbiased manner. We can then understand how the value changes over time, so that decisions can be made on the appropriate storage medium or possible deletion of these documents [Chen 2005]. Hence, what is required is a method to develop a file retention policy where the use value of the files can be relatively easily measured. Such methods are proposed in the literature but we question their applicability. Therefore, our research question therefore is *"How effective is a method for file retention in practice?"* To understand the practical operation of such a method, we first derive an appropriate method from an analysis of literature, and then apply the method in a case study. The paper concludes with a reflection on the findings and suggestions for further research.

2. Literature review

In total nine different data retention policy formation methods have been found in the literature. These methods include the determination of file retention decision parameters (like goals and file attributes) on basis of file valuations. Table 1 gives an overview of these methods.

Table 1: File Policy Retention Determination Methods		
Method	**Goal of data retention policy**	**Important file attributes**
[Chen 2005]	Capture the changing file value throughout the lifecycle and present value differences of files	Frequency of use; Recency of use
[Turczyk et al. 2008]	Determine the probability of future use of files for deciding on the most cost-effective storage medium	Time since last access; Age of file; Number of access; File type
[Bhagwan et al. 2005]	Lay out storage system mechanisms that can ensure high performance and availability	Frequency of use
[Verma et al. 2005]	Optimize storage allocation based on policies	Frequency of use; File type
[Mesnier et al. 2004]	Classify automatically the properties of files to predict their value	Frequency of use; File type; Access mode
[Zadok 2004]	Select files that can be compressed to reduce the rate of storage consumption	Directory; File name; User; Application
[Strange 1992]	Optimize storage in a hierarchal storage management (HSM) solution	Least recently used
[Gibson and Miller 1999]	Reduce storage consumption on primary storage location	Time since last Access
[Shah et al. 2006]	Design a cost efficient data placement plan while allowing efficient access to all important data	Metadata; User input; Policies

A number of criteria for a file retention policy method exist; here we list a few of them based on the findings of our literature review:

The retention policy determination method has to function with little to no human intervention [Chen 2005; Turczyk, Frei, Liebau and Steinmetz 2008],
The method should be based on the subjective use value of files over time in their different life stages [Chen 2005; Turczyk, Frei, Liebau and Steinmetz 2008], and
The method has to use multiple file attributes for the valuation process [Turczyk, Frei, Liebau and Steinmetz 2008].

All the file retention policy determination methods of table 1 can be automated, and thus fulfill the first criterion. In the valuation method of Mesnier [5], the files are only valued at the moment of creation and the value is not measured over time. This method can therefore be excluded as it does not satisfy criterion 2. The method of Verma is excluded for the same reason (criterion 2). The valuation methods of Strange [7], Bhagwan [3] et al. and Gibson & Miller [8] are excluded because they use only one measure for the valuation of the data, and hence do not satisfy the third criterion.

After the evaluation of the literature, only four methods fit the criteria of our research objective; (1) Usage-over-Time Method [Chen 2005]; (2) Probability of Further Use [Turczyk, Frei, Liebau and Steinmetz 2008]; (3) Elastic File Quota System [Zadok 2004]; and (4) the ACE Framework [Shah, Voruganti, Shivam and Alvarez 2006].

Chen's (2005) *usage-over-time* approach to indirectly determine the value of a file, however has as drawback that it does not incorporate the knowledge of administrators and users of the files [Chen 2005; Matthesius and Stelzer 2008]. Furthermore, the method does not take into account that the value of files is not necessarily reflected in their usage. For instance, a trade agreement or contract is of critical value for a business, but the usage count for these types of files can be very low. Developing and adding a classification scheme based on the contents of files could increase the effectiveness of this method.

Turczyk et al's method to determine the *probability of future use* [Turczyk, Frei, Liebau and Steinmetz 2008; Turczyk et al. 2007] has as drawback that all calculations are based on the characteristics and usage of files, while the content and context of a file is not considered in the calculations.

The Elastic Quota File System (EQFS) method developed by Zadok et al. aims to reduce the need for more space on a file system by an intelligent set of policies that allows one user to use the free disk space of another user [Zadok 2004]. The EQFS method uses the experience of data administrators and users to identify the elastic files. When defining the policies for elastic file determination, gaming and politics are unavoidable resulting in subjective allocations of higher service levels (speed of access and disk space) to some actors [Zadok 2004].

We find the ACE framework developed by Shah et al (2006) to be an exemplar method for developing a file retention policy. The framework presents tools and methods for the classification of file and storage locations as well as tools for file placement. The data classification method of ACE is based on metadata (data attributes) and these attributes are compared with predefined policies. In the article of Shah et al. (2006), it is stated that these policies are included in the framework and are based on the consultation of experts [Shah, Voruganti, Shivam and Alvarez 2006]. However, they do not discuss how these policies can be defined. This is remarkable, because a file valuation method can infer priorities of placements. This lacuna makes the usage of ACE framework based on the Shah et al. (2006) article problematic. As without proper guidelines, it is difficult to formulate policies. Table 2 provides an overview of our assessment of the four methods discussed above.

Table 2: Assessment of Methods

Criterion/Method	**Usage over time**	**Future use probability**	**Elastic quota file system**	**ACE: Data classification**
Little human intervention	X	X	X	X
Frequency of use	X	X		X
Measurable metrics	X	X	X	X
Classification of data	X	X	X	X
Knowledge of data managers and users			X	X
Cost reductions	X	X	X	X
System performance				X
Business value of data				X

The ACE framework is the only method which fulfils all the assessment criteria described in Table 2. However, the determination of retention policies by using this framework is problematic. The first problem is that policies should be specified by information *users* not by

system administrators [Tanaka, Ushijima, Ueda, Naitoh, Aizono and Komoda 2005]. An information user is typically a business person, who often has difficulties with understanding metadata attributes. This makes it difficult for a business person to specify policies [Tanaka, Ushijima, Ueda, Naitoh, Aizono and Komoda 2005]. The second problem is that developing file retention policies is a time consuming task [Ohta et al. 2006] and realizing a complete set of policies that cover all files, is too labor intensive [Jin et al. 2008]. Administrators generally use the rules-of-thumb for policy selection, often in anticipation of a certain workload [Short 2006].

A classification of files by attributes of relevance for retention decisions has a key role in ACE. However, ACE does not provide guidelines on how the selection of these parameters (i.e. the development of the retention policy) can be based on the user's valuation of files. Therefore, one of the contributions of this paper is to test the causal relations between file attributes and the (subjective) use value of files.

For subjective valuation of files, Sajko et. al. [2006] developed an information value questionnaire (IVQ) that allows information workers to value the information they use. The IVQ has five dimensions (1) Files Lost, (2) costs of file (Re)building; (3) Market Value; (4) Legislative, and (5) Time as an indicator of obsolescence. The "Lost" dimension measures the impact of information loss on the business operations. This can be anything from "*nothing special"* to "*making wrong decisions with major consequences"*. "(Re)building" measures the cost of replacing the lost information (from "*negligibly small"* to "*intolerably high costs"*). "Market value" measures the consequences if competitors obtain the information (from "*nothing"* to "*competitor gets competitive advantage"*). "Legislative" identifies the obligation to keep the information and the legal consequences if the information is lost (from "*no obligation"* to "*keeping information is obligatory and sanctions are strict"*). The "Time" dimension measures the rate at which the information depreciates in value (from "*very quickly"* to "*does not depreciate at all"*). This questionnaire cannot be automated, and is therefore not directly suitable for an efficient method for determining data retention policy. The measures that are used are subjective; the rankings of different persons are therefore required to create inter-subjective reliability. Because people answer the questionnaire according to their perceptions, the value that is determined is the 'perceived business value'. However, IVQ can be used to align file attributes with use value. The IVQ allows to rank files. By measuring subjective values of information entities, we can combine these values with attributes (like last access date, modification date etc) of files, and then prioritize these attributes to arrive at a decision policy. The approach we hence take is to combine objective observable file attributes with the more subjective IVQ measure. Thus, the most important attributes can be identified and used to be applied to prioritize files over different storage media. We thus can summarize a method for determining a data retention policy, as existing of the following steps:

1. Select a feasible size of representative files and identify their attributes.
2. Let (a sample of) users score the business value of these documents.
3. Correlate value score with the file attributes.
4. Take those file attributes with high correlations with business value as decision parameters in the retention policy. Leave out weakly correlated attributes.
5. Propose the results to users and discuss the applicability of the results.

3. Case Study

It is expected that the behavior of a file has causal relations with its value. Here, by behavior we mean file usage, i.e. the frequency with which the document is accessed or

modified. Based on these causal relations, it is possible to select appropriate file retention parameters. Therefore for data retention we have the following propositions;
Proposition 1: The frequency of access of a file predicts its value.
This proposition is based on the idea that a file is more valuable if it is used more heavily than other files [Chen 2005]. If this correlation is corroborated in this case, file attribute "frequency of access" should be included in the use value-based retention policy as a decision parameter. Unfortunately, the frequency of use is not logged in a Windows file system, and therefore a proxy of "frequency of access" is needed consisting of users' 'perceived frequency of access'. Consequently we have the following proposition,
Proposition 1a: A higher (perceived) frequency of access results in a higher file value.
Gibson and Miller developed a 'file-aging' algorithm based on the assumption that older files are used less and therefore less valuable [Gibson and Miller 1999]. This leads to the following sub-propositions:
Proposition 1b: The older the file the lower the value of the file.
The last modification time of a document refers to the number of days since the file was last updated. If a file is updated recently it implies that people are actively working with the file and therefore the value of the file is higher. Consequently, we have the following proposition:
Proposition 1c: A more recent time of last file modification results in a higher file value.
Turczyk examined the characteristics of different files to find the probability distributions that can be used to determine the probability of future use. He found that the probability distribution depends on the file type of a document [Turczyk, Frei, Liebau and Steinmetz 2008]. This results in proposition 1d:
Proposition 1d: The file type can be used to predict the file value.
The position that a person has in the organization may also influence the file use value (FUV). The reason for this is that the type of files that are used by people in an organization depends on their line of work. Organizational functions (named "grades" in our case study of Capgemini) are used to define the function level of the personnel, resulting in:
Proposition 2: A higher grade of the user results in a higher value for the file they use.
The propositions made above are summarized in the causal model shown in Figure 1. This model displays the observed variables of the files (file age, last modification time, and file type), the behavioral construct of the respondent (user grade) and the perceptual constructs of the respondents (perceived frequency of access and file value). The different constructs are numbered C.1 to C.6.

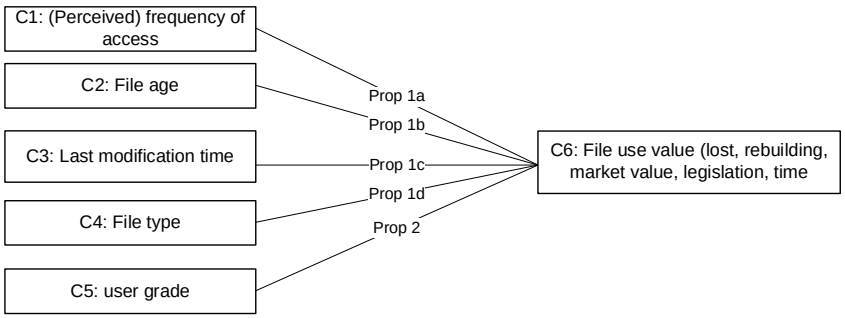

Figure 1: Causalities between file retention parameters (c1-5) and file use value (c6)

Consequently, if the correlations of these propositions are corroborated, frequency of access, file age, last modification time, file type and user grade could be included as decision parameters.

In the case study at Capgemini Netherlands, a dataset was collected to test the propositions. In the dataset the following elements were collected (i) the metadata attributes of a document; (ii) completed IVQ for this document; (iii) the grade of the respondent, and (iv) the perceived frequency of use. The user that completes the questionnaire was asked to indicate his or her current grade at Capgemini. Capgemini uses these grades to indicate the function level of their personnel. The grades are: consultant, senior consultant, managing consultant, principle consultant and business support (secretary). To increase the effectiveness of the questionnaires we applied the following two rules: (1) Each respondent was asked to complete the IVQ for at least 5 files, and (2) Only files of the following file types could be selected; .doc, .xls, .ppt and .pdf. This file type was based on our perception of the file types of important files in Capgemini. An overview of the constructs is presented in table 3.

Table 3: Constructs

Code	Name	Based on	Scale
C.1	Perceived Frequency of Use	Question added to the IVQ regarding perceived frequency of access per time period	Answers are normalized to 'number of accesses per year'
C.2	File Age	'File creation date' in metadata	Number of days since creation date
C.3	Last Modification Time	'File last modification date' in metadata	Number of days since last modification
C.4	File Type	'File type' in metadata	Extension of file (.doc/.ppt/.xls/.pdf)
C.5	User Grade	Question in IVQ	Grade at Capgemini
C.6	File Value	Scores in IVQ	Total score of the five questions in the IVQ, ranging from 0 to 20

For the case study an electronic application was developed. We used this application to collect the valuations of files and the metadata of these files. The application followed the following sequence:

- The respondent manually selected five files that s/he wants to value.
- After selecting five different files, the respondent could progress to the next page. On this page the IVQ was displayed for the first file.
- The IVQ had five multiple choice questions with five possible answers with scores in a range from 0 to 4. We added a sixth question, asking the respondent to give an indication of the number of times s/he uses the files.
- When the IVQ was completed, the respondent was asked to select the employee's current grade at Capgemini.

4. Results

Everyone in the financial service sector of Capgemini Netherlands received an invitation to participate in the case study. In total 654 people were invited, and 77 completed the IVQ, a

response rate of 12%. All respondents were asked to complete the IVQ questions for at least 5 different files. In total the 77 respondents assessed the value of 387 files. A factor analysis was performed to determine if the questions in the IVQ all load on the files' use value (FUV) construct.

Table 4: Factor loadings of File Use Value	
Item	**File Use Value**
Lost	.830
ReBuilding	.800
MarketValue	.742
Legislation	.763
~~Time~~	

The calculated value (0.79) shows that factor analysis is appropriate for the dataset [Field 2005]. The factor analysis indicates that the time-items for the file value construct measurement at Capgemini are not relevant, and thus the time question does not load on the FUV factor. Linear regressions analysis was used to test propositions, 1a, 1b, 1c and 2. To test proposition 1d, one-way independent Analysis of Variance (ANOVA) was used, because an analysis of variance method is more appropriate when a variable (File type) has a nominal nature (Field, 2005). Results for proposition 1a are presented in table 5.

Table 5: *Regression of Perceived Frequency of Access on FUV*				
ANOVA	F (1.385) = 31.82	Sig = 0.000		
Unstandardized Coefficients		Standardized Coefficients		
B	Std. Error	Beta	T	Sig.
.003	0.001	0.276	5.641	0.000

Results for proposition 1b are given in table 6.

Table 6: *Regression Analysis (File Age – FUV)*				
ANOVA	F (1.385) = 8.43	Sig. = 0.004		
Unstandardized Coefficients		Standardized Coefficients		
B	Std. Error	Beta	T	Sig.
- 0.002	0.001	-0.146	-2.903	0.004

Results for proposition 1c are given in table 7.

Table 7:*Regression Analysis (Last Modification Time – FUV)*				
ANOVA	F (1.385) = 9.568	Sig. = 0.002		
Unstandardized Coefficients		Standardized Coefficients		
B	Std. Error	Beta	T	Sig.
-0.002	0.001	-0.156	-3.093	0.002

Proposition 1d is tested by a one-way independent ANOVA test; see Table 8 for results.

Table 8: *One-Way Independent ANOVA (File Type FUV)*		
Levene Statistic	Sig. = 0.515	
ANOVA	F (3.383) = 1.844	Sig = 0.139

Results for proposition 2 are presented in table 9.

Table 9: *Regression Analysis (User Grade – FUV)*				
ANOVA	F (1.385) = 6.81	Sig. = 0.009		
Unstandardized Coefficients		Standardized Coefficients		
B	Std. Error	Beta	T	Sig.
0.363	0.139	0.132	2.610	0.009

Based on the analysis of the dataset from the case study, five of the six propositions are corroborated, see table 10.

Table 10: *Results of the Data Analysis per proposition*		
1a	*"A higher perceived frequency of access results in a higher file value."*	*Corroborated*
1b	*"The older the file the lower the value of the file."*	*Corroborated*
1c	*"A more recent last modification time results in a higher file value."*	*Corroborated*
1d	*"The file type can be used to predict the value of a file."*	*Rejected*
2	*"A higher grade of the user results in a higher value for the files they use."*	*Corroborated*

5. Conclusions

The goal of this research was to answer the following research question: *"How can we develop a method to determine a company's file retention policies?"*. In this research we have described and later demonstrated a method by which a company's file retention policies (based on the use value of the files) can be determined. We have thus also shown that the use value of files can successfully determine useful policy parameters. We have shown that the file behavioral parameters and the context parameter (grade) together can predict the subjective value (FUV) of files. Consequently, these parameters should be part of a file retention policy determination method.

In the case study, we found that the file type of a file has no significant causal relation with the value of a file. File type, contrary to suggestions of others (Verma et al, 2005; Mesnier et al, 2004; Turczyck et al, 2007), is therefore not a usable attribute to specify policies at Capgemini, at the moment we conducted the case study. We also found that a reliable measure of FUV, namely, the IVQ instrument as proposed by Sajko et al. (2006), could be based on 4 instead of 5 factors (depending on the case study). As we have shown in our case study that we can exclude the Time factor. However, our sample is relatively small and only within one firm. Consequently, more research on this is needed.

We have also shown (through the case study), that there is a strong causal relation between the position of the user of a file and the value of the file. We have therefore improved the ACE method by including the position of the user of a file. We have also noted that ACE is a

method which is context dependent. In other words it would provide different results depending on the setting in which it is used. As such, the method is generalizable, but its results are hence probably not generalizable. This implies that a file retention policy should not only include goals and relevant attributes, but also a procedure which guarantees a regular test of attributes in relation to their impact on use value.

The questionnaire helps practitioners in the information life cycle management field to move towards a business oriented approach. We found that people became more aware of the value of their files during the process of the case study. We observed that the employees started discussions about the amount of invaluable data on their own laptops and the data that reside in the different knowledge bases in the organization. With the questionnaire, the business people were stimulated to develop a critical approach towards the files they used and stored. This awareness can be one of the first steps to reducing some of the causes of data proliferation. Furthermore, we think that this study also shows that actual implementation of file retention actions (like file removal or storage to an indirect storage medium) is not just a technical task, or the prime task of a database administrator. Rather, file retention actions are also a task for the owners of the files, which in most cases are their end users. In making file retention decisions, however, end users can be well supported by database administrators, who can take the responsibility for the file retention policies and procedures. The administrators can also advice the end users on basis of research results.

To be useful, the method should contribute in resolving a relevant business problem. To be practical, the method should be workable in an organizational environment. We operationalized usefulness and practicality with the following checklist questions:

- How can this method help you in your project(s)?
- What do you consider to be strong points of this method?
- What do you consider to be the weaker points of this method?
- Can you think of a useful contribution to our method?

We received the following responses from the experts:

- The frequency of issues can depend on the season in a year. If some files become more valuable in a certain season, the accessibility of those files can be increased (during the season). This helps the people that are looking for the files.
- We find that the method designed in this research is not suitable to predict the future behavior and value of files. Consequently, the testing of these propositions must be repeated regularly as part of a policy determination method.
- The value assigned to a file depends on the role and the position of a person in the organization. It can therefore be useful to develop ‘profiles’ of persons. The profile can, for instance, be used to sort the search results of a person. Then, we can place the files with the highest value for the person on top of the search results. The profile can also be used for personalized information on intranet web pages, such as knowledge portals. Files that are assigned a high value by users with the same profile can be presented on the front page of the knowledge portal. The method can determine the moment when a file makes the transition from being directly accessible to being archived or deleted. The designed method can thus be used to select valuable files to publish on a knowledge portal.
- The method can substantially reduce the gap between the work of archivists and the business environment. It can furthermore reduce the workload that is associated with the development of storage policies.

- Finally, depending on the outcome of the IVQ a company can create a file retention policy suitable for the particular company. Table 11 summarizes the findings for Capgemini.

Table 11: Applicability of file retention policy elements in the context of CagGemini	
Qualitative indicators (IVQ)	
Costs of loss	Difficult to assess for each file separately
Cost of rebuilding	Difficult to assess for each file separately
Market value	Difficult to assess for each file separately
Legislative requirements	Easy to assess for each file separately
Time	Difficult to assess for each file separately; probably a redundant item in IVQ
Added: Perceived frequency of use	Difficult to assess for each file separately
File attributes	
Frequency of access	Can be easily assessed; but unclear evidence of correlation with value (FUV)
File age	Can be easily assessed; evidence of correlation with value
Last modification time	Can be easily assessed; evidence of correlation with value
File type	Can be easily assessed; but no evidence of correlation with value
User grade	Can be easily assessed; evidence of correlation with value

Depending on whether the criteria are important, they can be used to create the file retention policy. For example, the file retention policy can state that all files associated with a particular project and accessed at least 5 times in the last week need to be stored in a particular database. A file retention policy can depend on a combination of qualitative indicators and file attributes, but in the case where many files have to be reviewed, a qualitative approach will have to be replaced by a file attribute based approach.

REFERENCES

BHAGWAN, R., DOUGLIS, F., HILDRUM, K., KEPHART, J.O. AND WALSH, W.E. 2005. Time Varying Management of Data Storage. In Workshop on Hot Topics in System Dependability, Yokohama, 222-232.

CHEN, Y. 2005. Information Valuation for Information Lifecycle Management. In Autonomic Computing, 2005. ICAC 2005. Proceedings. Second International Conference on, 135-146.

FIELD, A.P. 2005. Discovering Statistics Using SPSS. Sage, London.

GIBSON, T. AND MILLER, E. 1999. An Improved Long-Term File-Usage Prediction Algorithm. In Annual International Conference on Computer Measurement and Performance (CMG '99), Reno, NV, 639-648.

JIN, H., XIONG, M. AND WU, S. 2008. Information Value Evaluation Model for ILM. In Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, 2008. SNPD '08. Ninth ACIS International Conference on, 543-548.

KAISER, M., SMOLNIK, S. AND RIEMPP, G. 2008. Konzeption eines Information-Lifecycle-Management-Frameworks im Dokumenten-Management-Kontext. In Multikonferenz Wirtschaftsinformatik 2008, Berlin, 483-494.
MATTHESIUS, M. AND STELZER, D. 2008. Analyse und Vergleich von Konzepten zur automatisierten Informationsbewertung im Information Lifecycle Management. In Multikonferenz Wirtschaftsinformatik, München, Germany, 471-481.
MESNIER, M., THERESKA, E., GANGER, G.R. AND ELLARD, D. 2004. File Classification in Self-* Storage Systems. In Proceedings of the Proceedings of the First International Conference on Autonomic Computing2004 IEEE Computer Society, 44-51.
MICROSOFT 2009. Microsoft Visual Studio 2008 - Your development happy place.
MICROSOFT 2009. The Official Microsoft WPF and Windows Forms Site.
MOODY, D. AND WALSH, P. 1999. Measuring The Value Of Information: An Asset Valuation Approach. In Proceedings of the Seventh European Conference on Information Systems, Copenhagen, Business School, Denmark1999, 361-373.
OHTA, K., DAI, K., KOBAYASHI, T., TAGUCHI, R. AND YOKOTA, H. 2006. Treatment of Rules in Individual Metadata of Flexible Contents Management. In Data Engineering Workshops, 2006. Proceedings. 22nd International Conference on, 77-82.
REINER, D., PRESS, G., LENAGHAN, M., BARTA, D. AND URMSTON, R. 2004. Information Lifecycle Management: The EMC Perspective. In 20th International Conference on Data Engineering 10-14.
SAJKO, M., RABUZIN, K. AND BACA, M. 2006. How to calculate information value for effective security risk assessment. Journal of Information and Organizational Sciences 30, 263-278.
SHAH, G., VORUGANTI, K., SHIVAM, P. AND ALVAREZ, M. 2006. ACE: Classification for Information Lifecycle Management. Computer Science IBM Research Report RJ10372, (A0602-0044).
SHORT, J. 2006. ILM Survey: What Storage, IT and Records Managers Say ISIC UCSD Research Report, San Diego.
STRANGE, S. 1992. Analysis of Long-Term UNIX File Access Patterns for Application to Automatic File Migration Strategies University of California, Berkely, California, USA.
TALLON, P.P. AND SCANNELL, R. 2007. Information Lifecycle Management. Communications of the ACM 50, 65-70.
TANAKA, T., USHIJIMA, K., UEDA, R., NAITOH, I., AIZONO, T. AND KOMODA, N. 2005. Proposal and evaluation of policy description for information lifecycle management. In International Conference on Computational Intelligence for Modelling, Control and Automation and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, 261-267.
TURCZYK, L., FREI, C., LIEBAU, N. AND STEINMETZ, R. 2008. Eine Methode zur Wertzuweisung von Dateien in ILM. In Multikonferenz Wirtschaftsinformatik, München, Germany.
TURCZYK, L., GROEPL, M., LIEBAU, N. AND STEINMETZ, R. 2007. A Method for File Valuation in Information Lifecycle Management. In 13th Americas Conference on Information Systems, Keystone, Colorado, 1122-1133.
VERMA, A., PEASE, D., SHARMA, U., KAPLAN, M., RUBAS, J., JAIN, R., DEVARAKONDA, M. AND BEIGI, M. 2005. An Architecture for Lifecycle

Management in Very Large File Systems. In 22nd IEEE / 13th NASA Goddard Conference on Mass Storage Systems and Technologies 160-168.
ZADOK, E., J. OSBORN, A. SHATER, C. WRIGHT, K. MUNISWAMY-REDDY, J. NIEH 2004. Reducing Storage Management Costs via Informed User-Based Policies. In IEEE Conference on Mass Storage Systems and Technologies, Maryland, 101-105.

芽|Sprouts

Working Papers on Information Systems | ISSN 1535-6078

Office:
Sprouts
University of Amsterdam
Roetersstraat 11, Room E 2.74
1018 WB Amsterdam, Netherlands
Email: admin@sprouts.aisnet.org

芽|Sprouts