Design of Data Acquisition Process and Its Validation Through Statistical Approaches

D. Naga Malleswari, K. Subrahmanyam

Abstract: In SIS framework proposed, data acquisition is playing a vital role. It is acquired for source code analysis, information acquisition, and SWOT processes involved. This paper presents basic factors used in data acquisition process for all the three processes. The obtained data is validated through KMO test, Cronbach alpha test, and KS test for verifying its reliability, normality and validity. The values of data have proven that it has passed KMO test, a test, and normality test. This validated data is further used in SIS Framework.

Index Terms: Cronbach alpha test, Data Acquisition Process, Information Acquisition, KMO test, KS test, SIS Framework, SWOT.

I. INTRODUCTION

SIS Framework

Assessing risk and its treatment is the essence of Risk Management. An error free software always gives competitive advantage for any organization. Risk management measures will help us to build error free software. Identifying risk through analytical approach is always key in the success of risk management process. There are different types of frameworks used to assess the risk but all of them are using either qualitative or quantitative methods and adapted structure is complex and most of them are tool based they are bit inefficient, expensive and not suitable for all the organizations. Hence a hybrid framework SIS (SCARE, Information Acquisition, SWOT) is proposed.

The SIS framework adapted structure is simple and it is using mixed method analysis (both qualitative and quantitative) to assess the risk. SIS framework is divided into 3 phases namely Risk Recognition, Risk Appraisal and Risk Mitigation, in Risk recognition phase the risk was identified with the help of SCARE analysis, Information Acquisition Process and SWOT analysis. Discrimination analysis was done in the Risk Appraisal phase and concluded with the Mitigation rules.
Data Acquisition Process

Based on Work System proposed by Alter (2003) The following 9 factors play a vital role in assessing risk of any system
Participants, Work Practices, Information, Technology, Product & Services, Customers, Environment, Infrastructure, Strategies
For this study we have limited our scope to 5 factors namely
In order to acquire the data for all 3 phases of the framework, only the above 5 factors were considered and the questioners were prepared based on the work proposed by Brace, L. (2004) and were distributed through Google form. Likert Scale proposed by Rensis Likert, Professor University of Michigan Institute for Social Research USA was used to quantify the responses. [1][2]

Sample Questions

Technology
1. Do you have permission to change your password?
2. Can you recover the password in-case you have forgotten?
3. Have you experienced a session time-out when you are working on-site?

Information
1. Has the necessary information been provided in right format?
2. Has content on the Site Been Proofread?
3. Is the Social Media Implementation on the Site Done Correctly?

Environment
1. Error pages are loaded when the site goes offline?
2. Usability aspect is satisfactory?
3. Can be used on different platforms?

Infrastructure
1. Are more no. of user accounts managed properly?
2. Are restrictions of infrastructure affecting usage?
3. Website’s maintenance is good?

For the SCARE Analysis the components of the source code were taken based on the below criteria.
**Participants**  
User Interface which includes placing of labels, buttons, images, and various other controls  

<table>
<thead>
<tr>
<th>Environment</th>
<th>IDE related data with related pages which includes Output Display Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infrastructure</td>
<td>Supporting Programs</td>
</tr>
<tr>
<td>Technology</td>
<td>Impact of Data base, JDBC Connectivity, DAO, RDO</td>
</tr>
<tr>
<td>Information</td>
<td>Database Protection Programs</td>
</tr>
</tbody>
</table>

**Table 1: SCARE Analysis Table**

<table>
<thead>
<tr>
<th>Participants</th>
<th>KMO Test Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Environment</td>
<td>0.750</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.845</td>
</tr>
<tr>
<td>Technology</td>
<td>0.771</td>
</tr>
<tr>
<td>Information</td>
<td>0.847</td>
</tr>
</tbody>
</table>

**Validity Test**  
KMO (Kaiser-Meyer Olkin) measures sampling adequacy which varies between 0 and 1. This measure can be used to validate the intention of the measure. The values above 0.6 is considered as minimum for validation, values <0.6 show the weakness of the measure in the validation. The values close to 1 show the high validation for the measure. The Kaiser-Meyer Olkin (KMO) measure of sampling adequacy was used to examine the appropriateness of Factor Analysis.[9]

Formulas:

\[ KMO_{ij} = \frac{\sum r_{ij}^2}{\sum r_{ij}^2 + \sum a_{ij}^2} \]

Where:  
- \( r_{ij} \) is the correlation matrix and \( a_{ij} \) is the partial covariance matrix.  
- (Invariance of Correlation matrix)

**Reliability Test**  
To identify the consistency of the respondents percentage in a survey the Cronbach’s alpha reliability test will be performed. This test will be able to measure the accuracy of the variable of interest. The values above 0.7 is considered as minimum for reliability, values <0.7 show the weakness of the measure in the reliability. The values close to 1 show the high reliability for the measure.[4]

Formulas:

\[ \alpha = \frac{k}{k-1} \left[ 1 - \frac{S_t^2}{\sum S_i^2} \right] \]

Where:
- \( k \) - no. of items/questions  
- \( S_i^2 \) - variance of scores of item i  
- \( S_t^2 \) - Total variance of overall score

**Normality Test**  
In order to identify the Normal Distribution of the data Kolmogorov-Smirnov Goodness of Fit Test (K-S test). This test doesn’t depend on any parameter. It is also used to check the normality assumption in variance analysis. More specifically, the test compares a known hypothetical probability distribution generated by your data.[3][4]

Formulas:

\[ F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{[-\infty,x]}(X_i) \]

Where:
- \( I(X_i) \) is Indicator function  
- (1 if \( X_i \leq x \) Otherwise 0)

**Results**

<table>
<thead>
<tr>
<th>Validity Test Results</th>
<th>Factor</th>
<th>KMO Test Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>0.690</td>
<td></td>
</tr>
<tr>
<td>Environment</td>
<td>0.750</td>
<td></td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.845</td>
<td></td>
</tr>
<tr>
<td>Technology</td>
<td>0.771</td>
<td></td>
</tr>
<tr>
<td>Information</td>
<td>0.847</td>
<td></td>
</tr>
</tbody>
</table>

It is clearly evident from the results that the data gathered for all the five factors is VALID as the resultant values are all above 0.6.
Design of Data Acquisition Process and Its Validation Through Statistical Approaches

Reliability Test Results

<table>
<thead>
<tr>
<th>Factor</th>
<th>Cronbach’s Alpha Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>0.774</td>
</tr>
<tr>
<td>Environment</td>
<td>0.877</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.899</td>
</tr>
<tr>
<td>Technology</td>
<td>0.888</td>
</tr>
<tr>
<td>Information</td>
<td>0.921</td>
</tr>
</tbody>
</table>

Table 3: Reliability Test

It is clearly evident from the results that the data gathered for all the five factors is RELIABLE as the resultant values are all above 0.7

Normality Test Results

<table>
<thead>
<tr>
<th>Factor</th>
<th>KS Test Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants</td>
<td>0.058</td>
</tr>
<tr>
<td>Environment</td>
<td>0.066</td>
</tr>
<tr>
<td>Infrastructure</td>
<td>0.153</td>
</tr>
<tr>
<td>Technology</td>
<td>0.030</td>
</tr>
<tr>
<td>Information</td>
<td>0.973</td>
</tr>
</tbody>
</table>

Table 4: Normality Test

It is clearly evident from the results that the data gathered for all the five factors is normally distributed.

I. CONCLUSION

The SIS framework has been developed to manage risk in Information Systems using mixed method analysis. This framework requires data acquisition for two processes IA and SWOT. We have successfully identified data acquisition process and implemented its validation through KMO test, Cronbach alpha test, and KS test has validated that the data acquired is normal and reliable in nature. These results will be used in SIS framework for measuring the risk of an Information System. The comparison of the framework with other frameworks will also be carried out in future.

REFERENCES

5. Subrahmanym,K. Ketha, T., 2018 Development of research & amp; development dashboard for an university, International Journal of Engineering and Technology(UAE), 7,60-63
8. 8.Risks Armin Keshushian1, Hasan Rashidi2Computer science Dept, Qazvin Azad UniversityIran, Tehran 2Hasan Rashidi, Qazvin Azad UniversityIran, Tehran
11. Maruf Pasha1, Ghazia Quiser1, Urooj Pasha2 “A Critical Analysis of Software Risk ManagementTechniques in Large Scale Systems” 2169-5356 (c) 2018 IEEE.
13. Development Lifecycle for Successful

First Author D.Naga Malleswari is a research scholar in Computer Science and Engineering of Koneru Lakshmaiah Education Foundation, Vaddeswaram, India. She is a First class M.Tech graduate from JNTU, Kakinada. She has published 10 Scopus indexed papers and her research interest in Software Engineering, Risk Management.

Second Author Dr. Kodukula Subrahmanym, a Gold Medalist from Andhra University (1992-93) is currently working as a Professor (CSE) & Associate Dean in KLEF, Vaddeswaram, Guntur. He is in teaching profession for the past 25 years and prior to joining KLEF he worked as Programme Leader in the School of Engineering, Science & Technology at KDU University, Malaysia for about 10 years. While working in Malaysia, Prof. Subrahmanym has been associated with many reputed Universities like Carnegie Melon University (USA), Murdoch University, Monash University (Australia), Northumbria University (UK) and gained international exposure on various teaching & learning pedagogies. He is the Founder Chairman of ACM Amaravathi Chapter and an active member of other professional societies like CSI, IEEE & CSTA, etc.. He has published more than 80 research papers in both National and international journals/conferences and attended various workshops/conferences in Malaysia, Singapore, USA & India. His research interests include Knowledge Engineering, Software Engineering & Soft Systems Methodologies. He has guided 3 scholars towards their PhD, 100 over students towards their Master’s and Bachelor Dissertations and currently guiding another 8 towards their PhD.