

# Statistical Analysis of Factors that Influences the Evaluation and Adoption of Multi-Tenant Databases

Olumuyiwa Matthew<sup>1</sup>, Kevan Buckley<sup>2</sup>, Mary Garvey<sup>3</sup>

School of Mathematics & Computer Science, University of Wolverhampton, UK

**Abstract:** Multi-tenant Databases (MTD) are implemented in the deployment of database management services to Information Technology (IT) platform users. A database service provider hosts the Multi-tenant Database Management System (MTDMS) and each tenant subscribes to the service through a standard method such as a web service. Improved groupings of the factors that influence the adoption of MTDs are presented in this paper. A survey is presented here that involves forty one experts from the field of databases. A predictive analytical method called Relative Importance Index (RII) and other statistical tools have been adopted in the analysis. The result has led to the new framework in the adoption of MTDs. The research also considers the direction of decisions about MTDs in situations where two or more factors are combined. A new improved MTD framework is presented that improves the decision making process of MTD adoption.

**Keywords:** Multi-tenant Database, Multi-tenant Database Management System, Predictive Analytical Method, Relative Importance Index, Framework

## I. INTRODUCTION

Multi-tenancy is a concept that evolved from the technology of Software as a Service (SaaS). A Multi-tenant environment is used to help IT organisations to improve the efficiency of their hardware resources by partitioning hardware to provide simultaneous support to multiple applications [1]. This concept is also used in several aspects of Information Technology (IT). One such aspect is in the deployment of database management services to IT platform users. The deployment of Database as a Service (DaaS) involves the implementation of what is called Multi-tenant database (MTD) management system. An MTD refers to a principle where a single instance of the DBMS runs on a server, serving multiple client organisations or tenants [2]. A Multi-tenant database is one which provides database support to a number of separate and distinct groups of users, also referred to as tenants. The multi-tenant data management system amortizes the cost of hardware, software and professional services to a large number of tenants

and thus significantly reduces per-tenant cost by increasing the scale. Therefore, the multi-tenant database system requires having excellent performance, low-space requirement and good scalability [3].

Multi-tenancy can minimize Hardware/Software costs and human costs per tenant. MTD systems have been exploited to store, manage, and retrieve data of tenants. A service provider hosts the MTD and each tenant subscribes to the services by doing necessary configuration, loading data to the data center, and then interacts with the services through a standard method, e.g., web services. Thus, the cost of ownership of database applications and the maintenance costs are transferred from the individual tenant to the service provider [4]. From the point of view of Pengcheng et al [5], clients (tenants) enjoy the desirable features of MTD which include lower upfront investment, pay-as-you-go pricing and reliable performance as specified by the Service Level Agreement (SLA). There is a consolidation at this level, which reduces operational costs without any revenue loss. This is extremely attractive to DaaS providers.

The structure of this paper is as follows: Section II provides the literature review on the different factors that influence the choice of MTD and the initial framework. Section III introduces the methods of analysis done based on the research. Section IV presents the results from the analysis carried out. Section V shows details of findings and discussions based on the results from data analysis and also present the new modified framework and its features. Section VI presents the conclusions and ideas for future work.

## II. REVIEW OF LITERATURES

Multi-tenant technology, a technology making integration of multiple applications to one application system, is a good choice for sharing of resources and economies of scale [6]. MTD architecture is very important for service providers who are based on SaaS. This helps in meeting up with the demands of customers or tenants on that platform of the provider. MTD architecture is very useful when one instance of a database is serving multiple clients. Only one set of hardware resources is needed to fulfil the requirements of all

users. MTD is based on subscriber model, so users have freedom to avail the facility as per business requirement or can turn off.

Jacobs and Aulbach [7] identified three different approaches in implementing MTDs which are: shared machine, shared process, and shared table. These approaches are increasingly better at pooling resources and executing administrative operations in bulk. However, they increasingly break down the isolation between customers, weakening security and increasing contention for resources. Grund et al [8] make comparison based on the three approaches of Jacobs and Aulbachas follows in the shared machine approach each tenants get their own database. The resource sharing is done at machine level. In the shared process approach the tenants share the same physical database process but own different databases. This allows better resource pooling between the tenants but still creates a lot of overheads because the schemas need to be maintained separately for all tenants. The last approach is the shared table approach. Using shared tables the application schema is created once and the different tenants are mapped directly into this schema using different schema mapping techniques.

Schiller et al [9] further explained that, these three approaches vary in the degree of consolidation and that the Shared Machine approach allows consolidating only a few tenants onto one machine due to the large main memory footprint of a database instance. The Shared Process approach consumes less main memory per tenant, yet main memory consumption increases quite fast with the number of tenants, as each tenant obtains a dedicated schema instance. In contrast, the main memory consumption of the Shared Table approach remains constant if the number of tenants increases. Schiller et al concluded that the Shared Table approach seems promising for a provider that targets the long tail because it offers the lowest overhead per tenant and, thus, is suitable for a large number of small tenants, e. g. 1,000 tenants each having less than 50 MB of data and at most 5 concurrent users.

Ying et al [4] explained some of the challenges associated with multi-tenancy database development against the traditional database. The first challenge is the data isolation among tenants. Many tenants can share the same database, but the database must ensure the data of these tenants are isolated from each other and no one can access their data other than themselves. The second challenge is to achieve the economics of scale; the database must have the capability of on-demand scale to support large volume of tenants. This means that irrespective of growth in number of tenants and their demand on the database, it must be capable of meeting the demand. Wood and Anderson [10] argue that complexity through the

different and changing demands and requirement of tenant raises further concerns with regards to maintaining and controlling the system. Because of this changing demand over time, the issue relating to scalability and security must be taken into consideration when deploying a multi-tenant environment. The third challenge is to be transparent to current existing application/skill, that is, the cloud developers can easily deploy the existing applications to on multi-tenant database without a large code change, and the developers can create new multi-tenant application without using new technical knowledge. The forth challenge is to support different isolations for the same application. This means that the use of different application by different tenant should ensure data isolation to each tenant regardless of the number of tenants involved.

This same constraint was also mentioned by Fang and Tong [11] identified the peculiarity need of each tenant, there are problems of; firstly, whether the database can afford the increase of both data and request accompanied with the growth of tenants. Secondly, how the database can meet the specific needs of one tenant efficiently and safely without affecting the others. It seems that the basic challenges associated with this technology remain the same and different models or approaches were proposed to handle each of these challenges.

There are some factors to be considered irrespective of the benefits of adopting MTDs. Matthew et al [12] identified these factors from literature including Keemti [13] and Khan et al [14] and were regrouped into four main factors as follows:

1. Cost.
2. Growth.
3. Security.
4. Regulation.

These were thoroughly analysed with a survey conducted by Matthew et al [2] to reveal the direction that each of these four regrouped factors tend towards. This led to the formation of a framework shown in the Figure 1.

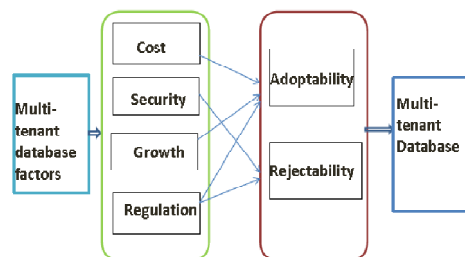


Figure 1- Framework of Multi-tenant Database Adoption [2]

This research will examine each of the above factors and other factors identified by Yaish et al [15] as follows:

1. Flexibility – ability to create multiple tables by tenants.
2. Time- time to build and configure.
3. Regulatory consideration ( UK/EU countries)

Therefore, based on the above other factors, the Matthew et al [16] grouping will now include Time as part of the cost consideration which now makes the grouping as follows:

1. Economic (Time and Cost).
2. Growth.
3. Security.
4. Regulations.

### III. METHODS OF ANALYSIS

This section covers analysis of a survey participated in by database experts. Questionnaires were administered online, responses were received from different continents including America, Europe, Asia, Africa and Australia. There are a total of 41 participants in the survey. The reason for low participation is mainly due to the level of technicality of the questionnaire since the responses were expected from a certain set of experts familiar with the concept of MTDs. Clearly in this part of the survey, time was considered to a very great extent. Hence, bearing this in mind, the questionnaire was constructed of multiple choice and closed questions. The questionnaire was divided into two main sections:

1. Respondent’s general information - This comprises of background questions about the respondents and their organisational information.
2. Investigation of MTD drive – This comprises questions about the concepts of MTD and participants’ opinion about the drive towards them.

This research on MTD is largely quantitative and is concerned with measurement of mainly the nominal and ordinal variables. The data from the survey were coded into SPSS and represented by numerical values. These data were subjected to the following statistical and mathematical tools;

#### A. Percentage Frequency Distribution.

The percentage frequency distribution is a statistical table or diagram that presents a simple analysis of statistical data in terms of percentages. This form of diagram can vary from a bar chart to a pie chart [17]. The frequency distribution is always the total frequency normalised to 100 and the

individual class frequencies expressed in proportion to that figure. The frequency of a particular observation is the number of times the observation occurs in the data. This means that frequency distribution can show either the actual number of observations falling in each range or the percentage of observations. Therefore, percentage frequency distribution is the representation of data that shows both the number of times the observation occurs and the percentage of the occurrence of the observation.

This is used in this research to show percentage distribution of participants in the survey. The questionnaires have been completed by different categories of IT experts all with some knowledge of databases. These ranges from the following group of professions: Database Administrators, IT Consultants, Researchers, IT Engineers, Programmers, Database Users, Systems Administrators, IS Managers, and Software Engineers. The overall number of participants returning valid online questionnaires was 39 out of 41.

TABLE 1 - PERCENTAGE DISTRIBUTION OF RESPONDENT’ PROFESSION

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Unknown	1	2.6	2.6	2.6
	Database Administrator	16	41.0	41.0	43.6
	IT Consultant	4	10.3	10.3	53.8
	Researcher	5	12.8	12.8	66.7
	IT Engineer	3	7.7	7.7	74.4
	Programmer	1	2.6	2.6	76.9
	Database User	2	5.1	5.1	82.1
	Systems Administrator	2	5.1	5.1	87.2
	Information System Manager	2	5.1	5.1	92.3
	Software Engineer	3	7.7	7.7	100.0
	Total	39	100.0	100.0	

Table 1 shows that Database Administrator was the dominant profession (41%) participating in the survey. This was followed by the Researcher profession (12.8%) and the least was those from the Programmer and the Unknown profession (2.6%).

**TABLE 2 - THE PERCENTAGE DISTRIBUTION OF LEVEL OF INVOLVEMENT WITH DATABASE**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Novice(less than 6 Months)	1	2.6	2.6	2.6
	Intermediate(6month -3 years)	9	23.1	23.7	26.3
	Expert ( 3 years above)	28	71.8	73.7	100.0
	Total	38	97.4	100.0	
Missing	System	1	2.6		
Total		39	100.0		

Table 2 shows the percentage distribution of the level of involvement with database systems. This shows that 2.6% of the respondents are Novice, 23.7% are Intermediate those who have just used database system for 6 month to 3 years and 73.7% for the Experts who have used the database system for more than 3 years.

**TABLE 3 - PERCENTAGE DISTRIBUTION OF THE DATABASE ADMINISTRATORS**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	22	56.4	59.5	59.5
	No	15	38.5	40.5	100.0
	Total	37	94.9	100.0	
Missing	System	2	5.1		
Total		39	100.0		

Table 3 shows the percentage distribution of Database Administrators among the respondents. This shows that 37 out the 39 respondents participated in this question where 59.5% represents those that are Database Administrator and 40.5% for those that are not Administrator.

**TABLE 4 - PERCENTAGE DISTRIBUTION OF RESPONDENT'S AWARENESS OF MTD**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	37	94.9	94.9	94.9
	No	2	5.1	5.1	100.0
	Total	39	100.0	100.0	

Table 4 shows the distribution of respondent's awareness of MTD concept. There is 100% participation here where 94.9% for those that aware of MTD while only 5.1% for those not aware of the concept.

**TABLE 5 - PERCENTAGE DISTRIBUTION OF MTD USAGE**

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	26	66.7	70.3	70.3
	No	11	28.2	29.7	100.0
	Total	37	94.9	100.0	
Missing	System	2	5.1		
Total		39	100.0		

Table 5 shows the percentage distribution of the usage of MTD where 37 out of 39 respondents participated in this question. There is a value of 66.7% for those that have used MTD and 28.2% for those that have never used MTD before.

**B. Relative Importance Index (RII).**

Relative Importance Index (RII) is another tool used in this analysis. The use of the term relative importance refers to the contribution a variable makes to the prediction of a criterion variable by itself and in combination with other predictor variables [18]. This definition considers only the relative contribution of a variable to total predictable variance and makes no assumptions about either the statistical significance or practical significance associated with a particular predictor. RII aims of generating an index that can ordinarily arrange those variables being studied in terms of (respondents') perceived agreement, relevance, importance, risk, or other discriminating criterion.

Accordingly to Holt [19] among the literature the RII may be described differently to reflect its application to a particular study – such as “relative agreement index”, “relative importance index”, “risk rating index” and so on. Therefore, for the purpose of this research which is based on measuring the impact of some factor on the adoption of MTD, the concept of RII is adopted and can be referred to as relative impact index.

There are about eight different RII methods identified and discussed by Holt [19], but for the purpose of this study the method adopted is explained below. This method is also used by Sambasi van and Soon [20]; Gunduz et al [21].

$$RII = \frac{\sum W}{A * N} \quad (1)$$

Where W is the weighting given to each factor by the respondents (ranging from 1 to 5), A is the highest weight (i.e. 5 in this case), and N is the total number of respondents.

The RII value had a range from 0 to 1 (0 not inclusive), the higher the value of RII, more important was the cause or effect (i.e. the more the impact in this case).

**C. Cross Tabulation.**

Cross Tabulation (CT) also known as a Contingency Table is a way of summarising the association between variables that have nominal or categorical data [22]. The frequency tables are used to summarise a single categorical variable while CTulation is used to summarise the relationship between two categorical variables. A cross tabulation (or crosstab for short) is a table that depicts the number of times each of the possible category combinations occurred in the sample data [23]. The crosstabs procedure can use numeric or string variables defined as nominal, ordinal, or scale. However, CT should only be used when there are a limited number of categories.

A CT shows the combinations of results of different questions of a survey in a table with the results of one question as the rows and the results of another question as the columns. The frequency data can be represented in a table with the rows as the conditions of one variable and the columns as the conditions of a second variable [22]. This method is therefore adopted in this research to examining the relationship of responses of participants between different variable within the survey. The results of the CT are presented in the next section of this paper.

**IV. RESULTS FROM THE ANALYSIS**

**A. Results from RII**

This section shows the results from the application of the RII formula above on the data collected to show the degree of impact of each factor that influences the move towards the concept of MTD.

Formula (1) above is applied for respondent’s data for Time as a factor

$$RII_{(TIME)} = [(3*4) + (4*20) + (5*4)]/5*28$$

$$= 12 + 80 + 20/140$$

$$= 0.80$$

The results for others are shown in the Table 6.

**TABLE 6 - RELATIVE IMPACT INDEX (RII)**

<b>ECONOMIC</b>						
Factors	Weight/Scale					RII
	1	2	3	4	5	
Time	0	0	4	20	4	0.80
Cost	0	0	4	12	19	0.89
<b>GROWTH</b>						
Factors	Weight/Scale					RII
	1	2	3	4	5	
Size of tenant DB	0	5	14	5	1	0.62
No. of tenants	0	4	15	9	0	0.64
No. of users per tenant	1	11	11	5	0	0.54
Growth rate of No. of tenants	3	7	7	2	1	0.51
Growth rate of tenant DB	3	8	8	5	0	0.53
<b>SECURITY</b>						
Factors	Weight/Scale					RII
	1	2	3	4	5	
Data Isolation	0	0	7	17	5	0.79
Scalability	0	5	15	12	1	0.65
Flexibility	0	9	9	8	1	0.61
Customization	0	2	15	11	2	0.69
<b>MAIN FACTORS</b>						
FACTORS	Weight/Scale					RII
	1	2	3	4	5	
ECONOMIC	0	2	0	8	25	0.92
GROWTH	1	1	25	7	0	0.62
SECURITY	1	0	0	26	7	0.82
REGULATIONS	2	12	6	4	2	0.54

**B. CT Results**

This section shows the results of applying cross tabulation to some of the variables to illustrate the relationships between them. This helps to identify the level of participation of respondents to a particular variable. These variables include the level of involvement with database, database administrator, awareness of MTD and MTD usage. These variables are CT against the four major factors. These are shown in the following tables where A and R represents adoption and rejection respectively.

**TABLE 7 - THE CT OF LEVEL OF INVOLVEMENT AGAINST FOUR FACTORS**

	ECONOMIC			SECURITY			GROWTH			REGULATION		
	A	R	Total	A	R	Total	A	R	Total	A	R	Total
Intermediate	6	2	8	6	2	8	6	2	8	6	2	8
%	17.1	5.7	22.9	17.6	5.9	23.5	17.6	5.9	23.5	17.6	5.9	23.5
Experts	24	3	27	9	17	26	20	6	26	20	6	26
%	68.6	8.6	77.1	26.5	50	76.5	58.8	17.6	76.6	58.8	17.6	76.5
Total	30	5	35	15	19	34	26	8	34	26	8	34
%	85.7	14.3	100	44.1	55.9	100	76.5	23.5	100	76.5	23.5	100

TABLE 8 - THE CT OF DATABASE ADMINISTRATOR AGAINST FOUR FACTORS

	ECONOMIC			SECURITY			GROWTH			REGULATION		
	A	R	Total	A	R	Total	A	R	Total	A	R	Total
YES	18	4	22	9	12	21	19	2	21	16	5	21
%	51.4	11.4	62.9	26.5	35.3	61.8	55.9	5.9	61.8	47.1	14.7	61.8
NO	12	1	13	6	7	13	7	6	13	10	3	13
%	34.3	2.9	37.1	17.6	20.6	38.2	20.6	17.6	38.2	29.4	8.8	38.2
Total	30	5	35	15	19	34	26	8	34	26	8	34
%	85.7	14.3	100	44.1	55.9	100	76.5	23.5	100	76.5	23.5	100

TABLE 9 - THE CT OF AWARENESS OF MTD AGAINST FOUR FACTORS

	ECONOMIC			SECURITY			GROWTH			REGULATION		
	A	R	Total	A	R	Total	A	R	Total	A	R	Total
YES	30	5	35	15	19	34	26	8	34	26	8	34
%	85.7	14.3	100	44.1	55.9	100	76.5	23.5	100	76.5	23.5	100
NO	-	-	-	-	-	-	-	-	-	-	-	-
%	-	-	-	-	-	-	-	-	-	-	-	-
Total	30	5	35	15	19	34	26	8	34	26	8	34
%	85.7	14.3	100	44.1	55.9	100	76.5	23.5	100	76.5	23.5	100

TABLE 10 - THE CT OF MTD USAGE AGAINST FOUR FACTORS

	ECONOMIC			SECURITY			GROWTH			REGULATION		
	A	R	Total	A	R	Total	A	R	Total	A	R	Total
YES	21	3	24	11	13	24	21	3	24	20	4	24
%	60	8.6	68.6	32.4	38.2	70.6	61.8	8.8	70.6	58.8	11.8	70.6
NO	9	2	11	4	6	10	5	5	10	6	4	10
%	25.7	5.7	31.4	11.8	17.6	29.4	14.7	14.7	29.4	17.6	11.8	29.4
Total	30	5	35	15	19	34	26	8	34	26	8	34
%	85.7	14.3	100	44.1	55.9	100	76.5	23.5	100	76.5	23.5	100

Table 7 shows the CT of the level of involvement with database against the four factors. Table 8 shows the CT of the database administrator against the four factors. Table 9 shows the CT of awareness of MTD against the four factors while Table 10 shows the MTD usage against the four factors.

TABLE 11 - CT OF MTD AWARENESS \* ACCEPTANCE LEVEL OF MTD

			Acceptance level of MTD today			Total
			No Difference	Improved	Highly Improved	
Awareness of MTD	Yes	Count	5	25	4	34
		% within Awareness of MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%
Total		Count	5	25	4	34
		% within Awareness of MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%

Table 11 shows the CT between two variables; MTD awareness and the acceptance level of MTD.

This shows the response relationship between the two variables

TABLE 12 - THE CT OF MTD USAGE \* ACCEPTANCE LEVEL OF MTD TODAY

			Acceptance level of MTD today			Total
			No Difference	Improved	Highly Improved	
Have you ever use MTD	Yes	Count	4	16	4	24
		% within Have you ever use MTD	16.7%	66.7%	16.7%	100.0%
		% within Acceptance level of MTD today	80.0%	64.0%	100.0%	70.6%
	% of Total	11.8%	47.1%	11.8%	70.6%	
	No	Count	1	9	0	10
		% within Have you ever use MTD	10.0%	90.0%	0.0%	100.0%
% within Acceptance level of MTD today		20.0%	36.0%	0.0%	29.4%	
% of Total	2.9%	26.5%	0.0%	29.4%		
Total		Count	5	25	4	34
		% within Have you ever use MTD	14.7%	73.5%	11.8%	100.0%
		% within Acceptance level of MTD today	100.0%	100.0%	100.0%	100.0%
		% of Total	14.7%	73.5%	11.8%	100.0%

Table 12 shows the CT between two variables; MTD usage and the acceptance level of MTD. This shows the response relationship between the two variables.

V. DISCUSSION OF FINDINGS AND DEVELOPMENT OF GUIDELINES.

A. Discussion of Findings from PFD

The PFD analysis has helped show the respondents' profession, age range, country, level of database usage, MTD awareness level, MTD usage level etc. For the purpose of this study, concentration is on the data relevant to the aim and objectives of the research.

From Table 2 above, 73.7% are the Experts and 23.7% are Intermediate users. Table 3 also shows that 59.5% are Database Administrators and Table 1 shows that 97.4% are IT related professions. The results in Table 4 further proof a point where 94.9% are aware of MTD concept. Table 5 shows that 70.3% of participants use the concept MTD. This shows that participants are more conversant with the subject of study. All the results from

Tables 1 to 5 show that the outcome and conclusions from this study will be very viable with sound judgement because of the areas of specialization of the participants. This is a proof of the quality of the research.

### **B. Discussion of Findings from RII**

This section of discussion is based on the results from the application of RII formulae used in the analysis section. The discussion is done based on the four factors to help determine the degree of impact of these factors in relation to the MTD concept.

The value of RII for the time and cost as factors shows the relative impact. Time has a value of 0.80 and cost with 0.89 respectively. This means that both factors have a great impact on the decision process towards the concept but with cost having greater impact. These two factors form the economic factor which is considered to have the highest RII of 0.92 from the Table 6 after the regrouping of the factors into four.

There are four sub factors that make up the security which includes Data isolation, scalability, flexibility and customization. These four factors form the security factor which is considered to have second highest RII after economic factor with the value of 0.82.

There are five sub factors here based on the study that make up the growth factor which are the size of tenant's database; number of tenants; the number of users per tenant; growth rate of number of tenants and growth rate of tenant's database. These five factors form the growth factor which has the third highest RII value 0.62 from Table 6.

Regulation as a factor does not have any other sub factor based on the study. The application of RII formulae shows that it does have the least RII value of 0.54 from Table 6. In conclusion, this stage of the analysis method has shown that the economic has the highest impact followed by security, after security we have growth and after growth is regulation.

### **C. Discussion of Findings from CT**

This section of discussion is based on the results from the CT done in the analysis section. The CT uses the four re-grouped factor against some other factors to determine whether the re-grouped factor leads to adoption or rejection of MTD. This shows whether it's positive or negative impact on MTD usage. This section also examining the level of acceptance of MTD in today's IT market using the CT of level of acceptance against two other

variables namely; MTD usage and MTD awareness.

#### **1. CT of level of involvement against Four Factors**

The Table 7 shows the CT of the level of involvement with database against the four factors. For economic factor, putting the two categories together we have 30 adoptions and 5 rejections which represents 85.7% for adoption and 14.3% for rejection. This show that there a very high gap between the adoption and the rejection. This means that the economic factor has a positive impact and it leads to adoption of MTD. This is just based on the level of involvement with database.

The combination of the two categories for regulation factor amounts to 76.5% for adoption and 23.5% for rejection. This also has shown that there is a wide gap between adoption and rejection. This means that regulation factor has a positive impact and it leads to the adoption of the concept.

For the Security factor, we have a total of 15 adoptions and 19 rejections which represents 44.1% and 55.9% respectively. In this result, there is no much gap between adoption and rejection but since more of the experts agreed that security leads to rejection, we can conclude that security will not encourage adoption of MTD. This means that security is more of negative impact than positive.

The combination of the two categories for the Growth factor gives a total result of 26 for adoption with 76.5% while 8 for rejection with 23.5%. This shows that there is a wide gap between the adoption and the rejection. Therefore conclusion can be made that growth factor encourages that adoption of MTD. This also means that growth has a positive impact.

#### **2. CT of Database Administrator against Four Factors**

There is a CT of database administrator against the four factors shown in table 8. We have 30 adoptions and 5 rejections which represents 85.7% for adoption and 14.3% for rejection. This is a wide margin between the adoption and the rejection. Conclusion can therefore be made that economic factor will lead to adoption of MTD and that economic factor has a positive impact on the drive towards the concept.

The combination of these two categories of participants regulation factor gives a total result of 26 for adoption with 76.5% while 8 for rejection with 23.5%. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, regulation can lead to the adoption of the concept.

The combination of these two categories of participants gives a total result of 15 for adoption with 44.1% while 19 for rejection with 55.9% for security factor. In this result, there is no much wide margin between the adoption and the rejection but more of the respondent said security will lead to rejection. Therefore, security will have a negative impact on the adoption of MTD.

For growth factor, there is a total result of 26 for adoption with 76.5% while 8 for rejection with 23.5%. This also has shown a wide margin between adoption and rejection. This is indicative of positive impact toward the drive about MTD. Therefore, growth can lead to the adoption of the concept. The results under database administrator are same with the level of involvement for each factor.

### **3. CT of Awareness of MTD against Four Factors**

There is a CT of awareness of MTD against the four factors shown in Table 9. The result in this table shows that we have 30 adoptions and 5 rejections which represent 85.7% for adoption and 14.3% for rejection under economic factor. This result is a confirmation of the result obtained the CT of database administrator against economic factor. Conclusion can also be made that economic factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

There are a total number of 26 adoptions with percentage value of 76.5% while only 8 rejections with a percentage value of 23.5%. This result is also a confirmation of the result obtained in the CT of database administrator against regulation factor. Conclusion can also be made that regulation factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

Table 9 also shows that there are 15 adoptions with percentage value of 44.1% and 19 rejections with percentage value of 55.9% for security factor. This result is a confirmation of the result obtained in the CT of database administrator against security factor. Conclusion can also be made that security factor has negative impact on the drive towards MTD and that it will lead to the rejection of the concept.

The result in this Table shows that there are a total number of 26 adoptions with a percentage value of 76.5% while only 8 rejections with a percentage value of 23.5% for growth factor. This result is a confirmation of the result obtained the CT of database administrator against growth factor. Conclusion can also be made that growth factor has positive impact on the drive towards MTD and that it will lead to the adoption of the concept.

### **4. CT of MTD Usage against Four Factors**

There is a CT of MTD usage against the four factors shown in Table 10. The result in this Table shows that there are 30 adoptions with 85.7% and 5 rejections with 14.3% for economic factor. This result also confirms other CT against economic factor. Conclusion can therefore be made that economic factor will lead to adoption of MTD and that economic factor has a positive impact on the drive towards the concept.

There are 15 adoptions with 44.1% and 19 rejections with 55.9% for security factor. Therefore, security will have a negative impact on the adoption of MTD. This result also confirms other CT against security factor.

Table 10 also shows 26 adoptions with 76.5% and 8 rejections with 23.5% for growth factor. This indicates a positive impact toward the drive about MTD. Therefore, growth as a factor can lead to the adoption of the concept. This result also confirms other CT against growth factor.

We have a total of 26 adoptions with 76.5% and 8 rejections with 23.5% for regulation factor. This is indicative of positive impact toward the drive about MTD. Therefore, regulation as a factor can lead to the adoption of the concept. This result also confirms other CT against regulation factor.

### **5. CT of acceptance level of MTD against Two Factors**

Table 11 shows the CT of acceptance level of MTD against the MTD awareness. And Table 12 shows the CT of acceptance level of MTD against the MTD usage. There are values of 14.7%, 73.5% and 11.8% for no difference, improved and highly improved respectively in both Tables. These show that there is a wide margin between improved and no difference. And putting together the value of improved and highly improved gives the value of 85.3% which further widen the gap. This shows that MTD acceptance has improved tremendously in the last 2 years. The value for no difference, improved and highly improved are the same in Table 11 and 12 which proves the reliability and consistency in the results of the research.

### **D. Development of Guidelines**

This section presents the guidelines for an intending subscriber of MTD based on the findings from this study. These guidelines are as follows:

G1 – Expert in the field of Database management system should be given the responsibility of heading the project team when considering MTD adoption.



Factor	% Adoption	% Rejection	RII	Final Decision
E	85.7	14.3	0.92	Adopt
S	44.1	55.9	0.82	Reject
G	76.5	23.5	0.62	Adopt
R	76.5	23.5	0.54	Adopt
ES	64.9	35.1	0.87	Adopt
EG	81.1	18.9	0.77	Adopt
ER	81.1	18.9	0.73	Adopt
SG	60.3	39.7	0.72	Adopt
SR	60.3	39.7	0.68	Adopt
GR	76.5	23.5	0.58	Adopt
ESG	68.8	31.2	0.79	Adopt
EGR	79.6	20.4	0.69	Adopt
SGR	65.7	34.3	0.66	Adopt
ESR	68.8	31.2	0.76	Adopt
ESGR	70.7	29.3	0.73	Adopt

G2 – The economic factor must be considered which includes the cost and time factor as explained in earlier sections.

G3 – The level of security you want your MTD to achieve and tolerate should be examining which include data isolation, scalability, flexibility and customization.

G4 – The growth rate of the MTD should also be considered in terms of the size of tenant database, number of tenants, number of users per tenant, growth rate of tenants and growth rate of tenant database.

G5 – The harmonious balance between the regulations governing prospective tenants should be look into by both the tenant and the service provider.

G6 – The choice of the MTD model must be determined by evaluating the features of the three model approaches of the concept which includes share machine, share process and share table.

**E. The Amendment of Framework**

Based on the analysis and discussion sections above, the postulates in Matthew et al [1] are now further proofed to have the following as the final ones. These will now form the amended framework.

Postulate 1  
ECONOMIC ➡ MTD ADOPTION

The economic factor still shows that it will lead to the adoption of MTD following the results of the analysis.

Postulate 2  
SECURITY ➡ MTD REJECTION

The security factor still shows that it will lead to rejection of MTD as a result of the analysis.

Postulate 3  
GROWTH ➡ MTD ADOPTION

The growth factor still shows that it will lead to the adoption of MTD as a result of the analysis.

Postulate 4  
REGULATION ➡ MTD ADOPTION

The regulation factor does not tend towards both directions as indicated Matthew et al [1], after the analysis of the responses from the participants; it shows that regulation now leads to adoption. Putting all these postulate together will form the new framework shown below.

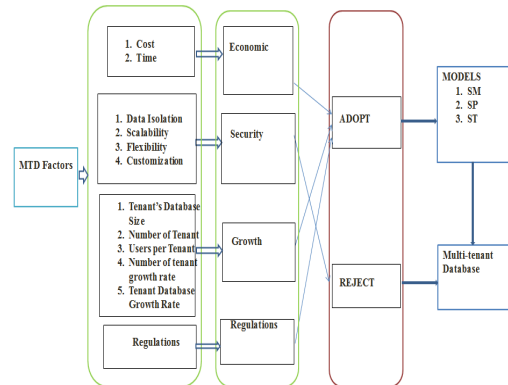


Figure 2 – The New Framework

**F. The Results of the Combination of Two or More Factors**

When an intending subscriber wants to know the evaluation impact of considering two or more factors together, this section takes into consideration these types of scenarios. Using the CT results of MTD usage and the four factors in Table 10, the percentage of adoption and rejection for the four factors can be determined as seen in the Table 13 below. This will also be used in the case of combination of two or more factors by calculating the average cumulative percentage of adoption and rejection for each factor. In the same way, using the results from the RII Table 6, the RII of any combination of the four factors will be determined by the average of RII of those factors that make up the combination. In Table 13, E represent economic, S represent security, G represent growth and R represent regulation and others are combinations of two or more of these factors.

TABLE 13 – THE TABLE OF RESULTS FOR TWO OR COMBINATIONS OF THE FACTORS

In Table 13 above the final direction of any of the combinations still leads to adoption because none of the final percent of rejection is up to the 55.9 as in the case of security that leads to the rejection direction. This means that for the direction to be rejection the percentage rejection must be 55.9 or more. Based on this the framework will now incorporate all these possible combinations of the factors. This is now shown below in Figure 3.

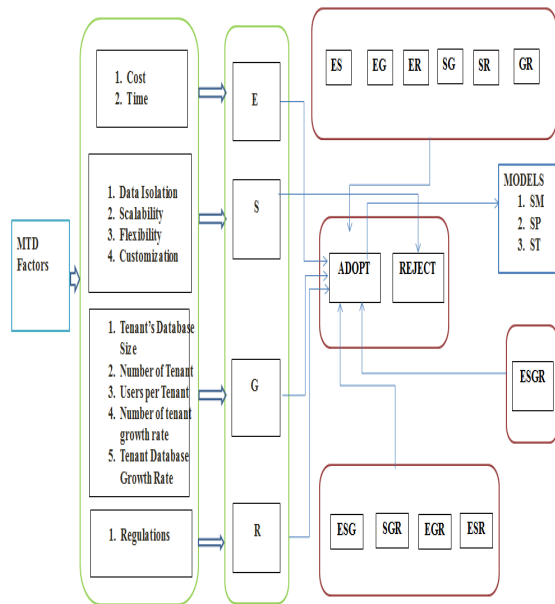


Figure 3 - The Modified New Framework

**G. The Comparison of the Initial and New Frameworks**

There are some differences between the initial framework and the new framework. These differences are as results of the analysis of the data collected and discussion from the findings. Looking at Figures 1 and 2 there are obvious similarities and also differences.

There are more similarities than differences which include economic factors tend towards adoption, security factors tend towards rejection and growth factors also tend towards adoption.

In terms of differences, the regulation as a factor now only tends towards adoption unlike in the initial where it was pointing to both directions. The new framework also considers the MTD model to

be adopted considering different features of each type of MTD models.

The modified new framework in Figure 3 now incorporates the combinations of two or more factors and what these could lead to. This is more comprehensive and user friendly when an intending user wants the benefits of the combination of two or more factors to inform his or her decision about MTD. This shows that all the possible combinations lead to the adoption of MTD.

**VI. CONCLUSION AND FUTURE WORK**

This paper has shown the improvement in the MTD framework and identifies the differences and similarities with the new modified framework and the initial framework. This also identified some other factors and subsequently improves the grouping done in the earlier stage of the research. This research has also shown the benefit of using mathematical and statistical tools in determining the direction of decisions about MTD and the degree of impact each of the factors has on the adoption of MTD. The cumulative impact of two or more factors was also identified and the possible resulting direction of the different combinations. There appears to be a further study required in this aspect to validate the framework.

**REFERENCES**

[1] Chaturvedi, A. and Bhat, Z.A.(2015) Adaptive Resource Scaling Methods for Multi-tenant cloud system. International Journal of Computer Trends and Technology (IJCTT) [online] 30(2) pp93-97

[2] Matthew, O., Garvey, M., Buckley, K (2016) A Framework for Multi-Tenant Database Adoption Based On the Influencing Factors. International Journal of Information Technology and Computer Science 8(3), pp.1-9.

[3] Ni, J., Li, G., Zhang, J., Li, L. and Feng, J. (2012) Adapt: adaptive database schema design for multi-tenant applications Proceedings of the 21st ACM international conference on Information and knowledge management. [online]. Maui, Hawaii, USA New York, NY, USA: ACM, pp.2199-2203.

[4] Ying, H., Wang, Q., Wang, Z., and Wang, N. (2011) DB2MMT: A Massive Multi-tenant Database Platform for Cloud Computing e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on. [online]. pp.335-340

[5] PengchengXiong, Yun Chi, Shenghuo Zhu, Hyun Jin Moon, Pu, C. and Hacgumus, H. (2015) SmartSLA: Cost-Sensitive Management of Virtualized Resources for CPU-Bound Database Services. Parallel and Distributed Systems, IEEE Transactions on [online], 26(5), pp. 1441-1451 .

[6] Sang, C., Li, Q. and Kong, L. (2012) Tenant Oriented Lock Concurrency Control in the Shared Storage Multi-tenant Database Enterprise Distributed Object Computing Conference Workshops (EDOCW), 2012 IEEE 16th International. [online]. pp.179-189.

[7] Jacobs, D. and Aulbach, S. (2007) Ruminations on multi-tenant databases. BTW Proceedings [online], 103pp. 514-521 .

[8] Grund, M., Schapranow, M., Krueger, J., Schaffner, J. and Bog, A. (2008) Shared Table Access Pattern Analysis for Multi-Tenant Applications Advanced Management of Information for Globalized Enterprises, 2008. AMIGE 2008. IEEE Symposium on. [online]. pp.1-5.

[9] Schiller, O., Schiller, B., Brodt, A. and Mitschang, B. (2011) Native support of multi-tenancy in RDBMS for software as a service Proceedings of the 14th International Conference on Extending Database Technology. [online]. Uppsala, Sweden New York, NY, USA: ACM, pp.117-128.

[10] Wood, K. and Anderson, M. (2011) Understanding the Complexity Surrounding Multitenancy in Cloud Computing e-Business Engineering (ICEBE), 2011 IEEE 8th International Conference on. [online]. pp.119-124.

[11] Fang, S. and Tong, Q. (2011) A comparison of multi-tenant data storage solutions for Software-as-a-Service Computer Science and Education (ICCSE), 2011 6th International Conference on. [online]. pp.95-98.

[12] Matthew, O., Dudley, C. and Moreton, R. (2014) A Review Of Multi-Tenant Database And Factors That Influence Its Adoption. UKAIS 2014 Conference [online].

[13] Keemti, P. (2010) Multi-tenant Database Architecture [online]. [Accessed 23 August 2013]. Available at <[http://www.msdn.microsoft.com/eus/library/aa479086.aspx#mltntda\\_topic1](http://www.msdn.microsoft.com/eus/library/aa479086.aspx#mltntda_topic1)>

[14] Khan, M.F. and Ullah, M.A. (2012) An Approach Towards Customized Multi-Tenancy. International Journal of Modern Education and Computer Science (IJMECS) [online], 4(9), pp. 39.

[15] Yaish, H., Goyal, M. and Feuerlicht, G. (2013) Proxy service for multi-tenant database access. in Availability, Reliability, and Security in Information Systems and HCI. Springer, .

[16] Matthew, O., Buckley, K. and Garvey, M (2015) Predicting The Impact Of The Factors That Influence The Adoption Of Multi-Tenant Databases. International Conference on Computer and Information Science and Technology (Cist'15); 05/2015 University of Ottawa, Canada

[17] Dodge, Y. (2006) The Oxford dictionary of statistical terms. [online] Oxford University Press.

[18] Tonidandel, S. and LeBreton, J. (2011) Relative importance analysis: A useful supplement to regression analysis. Journal of Business and Psychology [online], 26(1), pp. 1-9 .

[19] Holt, G. (2013) Asking questions, analysing answers: relative importance revisited. Construction Innovation [online], 14(1), pp. 2-16 .

[20] Sambasivan, M. and Soon, Y.W. (2007) Causes and effects of delays in Malaysian construction industry. International Journal of Project Management [online], 25(5), pp. 517-526 Available at:<<http://www.sciencedirect.com/science/article/pii/S0263786306001700>>.

[21] Gündüz, M., Nielsen, Y. and Özdemir, M. (2012) Quantification of delay factors using the relative importance index method for construction projects in Turkey. Journal of Management in Engineering [online], 29(2), pp. 133-139

[22] Hinton, P.R., McMurray, I. and Brownlow, C. (2014) SPSS explained. [online] Routledge.

[23] Miller, R. and Acton, C. (2009) SPSS for social scientists. [online] Palgrave Macmillan.

## Authors' Profiles



**Olumuyiwa Matthew** received B.Tech and MSc degrees in Computer Science and Information Technology respectively. He is currently Management working towards the PhD degree with the School of Mathematic & Computer Science University of Wolverhampton, UK. His current research interest is issues relating to the adoption of Multi-tenant database. He is an associate member of BCS the Chartered Institute for IT and a member of Computer Professionals of Nigeria (CPN)



**Dr Kevan Buckley** is the Head of School of Mathematics and Computer Science at the University of Wolverhampton so has diverse interests across the subjects. In particular he is interested in the application of Machine Learning and Optimisation techniques to data from a range of disciplines from Social Media to Forensic Science. Techniques for data capture, processing and storage underpin this research



**Dr Mary Garvey** is a senior lecturer at the University of Wolverhampton, specialising in teaching databases at all levels and web programming. She has several academic publications to her credit. She is a Member of the BCS CITP and is the Chair of the Wolverhampton Branch of the BCS.