Design of Framework for Computer Assisted Information Systems Auditing Methodology (CAISAM)

Mario Žgela, M Sc
Croatian National Bank
Trg hrvatskih velikana 3, 10000 Zagreb
Mario.Zgela@hnb.hr

Zdravko Krakar, Ph D
Faculty of Organization and Informatics
University of Zagreb
Pavlinska 2, 42000 Varaždin, Croatia
Zdravko.Krakar@foi.hr

Abstract. *Auditors are compelled to examine not only general information systems controls in order to confirm if information system is in general terms adequate and that certain activities have been undertaken so decisions can be made according to data and reports from information system. Auditors are forced to examine data and processes in details and prove that information system is reasonably resistant to fraudulent activities, omissions and errors. Detailed audit is usually not possible without adequate computer assisted auditing methods. Although these methods are supported by software tools and used by auditors worldwide, complete methodology is still not designed. In this paper meta-model of methodology will be designed. It will be defined what methods should be used, how, in what sequence and what conclusions can be made according to output from certain methods. This paper will show how Benford's Law may be used as a central method of computer assisted information systems auditing methodology (CAISAM).*

Keywords. IS audit, information systems, computer assisted auditing methodology, CAAT tools, Benford's Law, meta model.

1 Present situation on CAISAMs

There are different methods and approaches for overall business and information systems audit developed by certain authorities. Since auditing activity tends to be extremely complicated and covers different business areas, use of auxiliary auditing methods and tools becomes compulsory. Also, as it is shown in [2], [10], [14] and [16] frauds become more and more advanced and resulting in much greater losses. Obviously, existing auditing methods and approaches are not adequate in contemporary business environments.

Although computer assisted auditing is emphasized as important and indispensable by leading organizations in information systems auditing (ISACA), United States banking industry regulation (Federal Reserve Bank), so called „central bank of central banks“ (Bank for International Settlements) and a number of other regulation authorities, widely accepted methodology is still not developed.

The consequence is that auditors use different methods in various sequences with possibly inconsistent input data and results. This is a fact even if the same computer assisted auditing tool is used, and even if the same business process and data are being audited. Also, the same auditor using the same computer assisted auditing tool processing the same business process and corresponding data but in two different actual audits can produce completely different results which can lead to very different

conclusions. Without adequate methodology, auditing processes are hardly comparable. Additionally, inappropriately applied methods, their mutual relationships and possibly wrong conclusions made by auditors result in inadequate audits.

In addition, computer assisted auditing methods are used in existing software tools (Computer Assisted Auditing Tools - CAAT). Methods implemented in CAAT are quite extensive in number and modes of use. However, there is no predefined methodology which could more precisely guide auditor while performing audits. As a consequence, each auditor chooses set of methods for each audit activity, use them in different sequences and makes different conclusions. This approach makes CAATs not effective enough since audit activities based on CAATs without use of methodology are arbitrary (depending on auditor's choice), disparate, not standard and non-comparable.

For that reason development of methodology for computer assisted information systems auditing is of the utmost importance. The development of such methodology must consist of the following:

1. definition of specific methods which will constitute the basis of methodology – defining input data, processing algorithm and output data
2. defining linkage i.e. sequence of methods application depending of output data resulting in adequate meta-model

2 CAISAM Framework

According to previous research, status of auditing methodologies and experience, conclusions about specifically adequate computer assisted auditing methods can be made. Specifically useful auditing method which is used in auditing processes only lately and still quite scarcely is Benford's Law.

Benford's Law, as it is explained in more details in [5], [6], [8], [9], [13] and [14], includes a number of tests. Especially valuable are first digit test, second digit test, first two digits test and first three digits test. These tests proved to be extremely efficient in practical auditing process. For that reason, Benford's Law is the central method in meta-model of methodology designed in this paper. Methods that constitute methodology are:

1. Benford's Law (4 methods)
2. Selection
3. Join
4. Summarization
5. Stratification
6. Duplication
7. Rounded values

2.1 Benford's Law

Benford's Law defines expected digit frequencies in certain number sets. It is noticeable that in sets of numbers from many data sources, certain digits are distributed in a particular way. According to first digit Benford's Law, digit "1" appears as the first digit in number for almost one third of the time, and larger digits appear on the leading number position with lower and lower frequencies. E.g. digit 9 appears as a first digit in slightly more than 4,5% numbers. The basis for Benford's Law lies in the fact that values of real world data sources are often distributed logarithmically, thus the logarithm of this real world data sources is distributed uniformly. Benford's Law may be applied to first, second, first two and first three digits in numbers. That is why there are in fact four methods based on this law: first digit test, second digit test, first two digits test and first three digits test. Benford's Law of first digit i.e. probability V of appearance of digit z1 in number system with base 10 on leftmost position in number can be expressed by the following formula:

$$V(z1)=\log_{10}(1+1/z1),\ z1 \in \{1,2,...,9\} \quad (1)$$

Formulas for probabilities of appearance of second, first two and first three digits in number system with base 10 are:

$$V(z2)=\sum_{z1=1}^{9} \log10(1+1/z1z2),\ z2 \in \{0,1,...,9\} \quad (2)$$

$$V(z1z2)=\log_{10}(1+1/z1z2),\ z1z2 \in \{10,11,...,99\} \quad (3)$$

$$V(z1z2z3)=\log_{10}(1+1/z1z2z3), z1z2z3 \in \{100,...,999\} \quad (4)$$

Each of these methods uses certain numeric attribute as input and counts frequencies certain combination of digits, depending on specific method. Output is list of all digits combinations and their respective frequency.

2.2 Selection

Selection is one of operations of relational algebra which is used for extractions (or selections) of data that holds for defined condition. Selection is very powerful operation which is most commonly used query-based operation in business application software.

In relational algebra, selection (sometimes called restriction or exclusion) is a unary operation written as $\sigma_{xOy}(r)$ or $\sigma_{xOc}(r)$ where:

- x and y are relation's attribute names
- O is a binary operation in the set (greater than, greater or equal than, equal, not equal, lesser than, lesser or equal than). It may also consist of logical boolean operators (AND, OR, NOT).
- c is a constant
- r is a relation

The selection $\sigma_{xOy}(r)$ selects all rows in r for which O holds between the x and the y attribute.

The selection $\sigma_{xOc}(r)$ picks only rows in r for which O holds between the x attribute and the constant value c.

In database management systems, selection relational algebra operation is designated by SELECT

SQL statement. That statement extracts all rows (records, tuples) from a table (term “table” in DBMS terminology corresponds to term “relation” in relational algebra) whose attributes hold against selection criteria:

t2 = select(t1,P)

Table t2 is created out of table t1, and it contains only rows from table t1 which satisfy (or hold against) the predicate P.

A predicate is simply an expression made out of boolean algebra whose operators are the logical boolean expressions (and, or, not) and arithmetic expressions (<, <=, >, >=, =, <>), and whose operands are either attribute names or attribute domain constants. A predicate may also represent another select statement thereby producing relation nesting. In practice, selection is usually used together with projection operation. Projection operation includes only chosen attributes out of table (relation) and may include mathematical functions over included attributes (e.g. maximum, minimum, average).

2.3 Join

Join is operation of relational algebra which is fundamental to the model. It would not be possible to decompose data structure into number of relations if there would not be a method for their composition when necessary. This is precisely what a join operator performs. It combines rows based on the comparison of one or more common fields in both relations in join. It is written as

$$r \underset{x\theta y}{\bowtie} s \quad \text{or} \quad r \underset{x\theta c}{\bowtie} s$$

where r and s are relations, x and y are attribute names of relation. Also, instead of y, constant c may be used.

Join operations can be classified as theta, equi-join, natural and outer (or external) join.

Theta (θ) join means that its result follows the general definition of the join operation. Theta is a predicate which consists of one of the binary operators in the set (greater than, greater or equal than, equal, not equal, lesser than, lesser or equal than) and specifies the join condition. If join condition θ equals "=", the operation is called an equi-join. Natural join is in fact equi-join from which duplicate attributes used for join are excluded. The resulting joined relation contains only one attribute for each pair of attributes holding same data.

Outer join does not require each data row in two joined relations to have a matching record. Resulting joined relation holds each row from at least one of relations even if no corresponding matched row exists, depending on outer join type. There are three types of outer joins: left outer joins, right outer joins, and full outer joins. The resulting relation of a left outer join (or shorter "left join") always contains all rows of the left relation (r), even if the join condition does not find any relating row in the right relation (s). So, left outer join always returns all rows from the left relation (r) together with rows from right relation (s) matching according to θ condition. If there are no matching rows in right relation (s), attributes of right relation (s) will contain only NULL values. If the right relation (s) returns more than one matching row for certain row in the left relation (r), the rows in the left relation will be repeated for each distinct row on the right relation (s).

The opposite case holds for right outer join (shorter "right join"). In case of full outer join, result is union of theta join, left and right joins. One should bear in mind union is a set operator, so all rows that are eventually identical in separate relations which are members in union (theta, left and right) will remain only in one instance in resulting relation.

There is no explicit notation for outer join operations in standard SQL. However, it is easily implemented in all modern databases by combination of expressions based on boolean operators.

2.4 Summarization

Summarization method aggregates amounts of interest according to specific attributes (or columns in SQL notation). Summarization extracts and group all different attribute values chosen by auditor and aggregate according to defined attributes. Auditor must define which attributes have to form unique combination of values (aggregation base attributes), which attributes have to be aggregated on (aggregation value attributes) and what summarization function has to be performed (aggregation function). For each unique extracted row value of aggregation base, summarization is executed for aggregation value attributes according to aggregation function. Notation is as following:

$$S_{AB,AV,f}(r)$$

S stands for summarization method, AB stands for aggregation base attributes, AV denotes aggregation value attributes, f is aggregation function, while r is relation. There are number of aggregation functions that can be applied: average value, sum of amounts, count of rows, minimum value, maximum value etc. Auditor has to remember that each combination of aggregation base attributes shows up in resulting relation only once, and attributes representing aggregation value hold result of aggregation function for all rows with same aggregation base attributes value.

Selection and join methods may be used in order to prepare data i.e. to extract and add additional attributes to final relation ready for summarization.

In SQL, summarization method may be represented with GROUP BY statement. The GROUP BY statement is used in conjunction with the aggregate functions to group the result-set by one or more columns.

2.5 Stratification

Stratification method includes creation of specific data strata (data layers) in accordance with defined value-based categories. As input parameters, method uses relation, attribute and strata limits. So, auditor has to choose which column is of special interest (usually number or date data type), define number of strata and their limits (upper and lower values).

It is not required that auditor should include all values of certain relation's attribute (table column) when applying stratification method (from minimum to maximum value). All rows are included in respective data strata, depending on values and strata limits. After stratification method is performed, auditor may investigate each stratum and make conclusions on data characteristics. This method is especially useful if combined with exception analysis. Auditor may learn about limit values for certain table columns and then additionally analyze column values and rows out of limiting bounds. Notation of stratification method (ς) is:

$$\varsigma_{x,l}(r)$$

where x stands for attribute for which strata should be created while l denotes set of boundaries (lower and upper bound) for each strata.

Method processes each table row and classifies it into corresponding strata, depending on attribute value x. Both totalling and counting functions may be applied on attribute x, and assigned to certain strata which results from pre-defined limit values. Method results in:

- count of rows and totalling of attribute values belonging to each strata
- calculation of percentage of rows in each strata in relation to total number of rows
- counting of all rows with attribute x value lesser than the lowest strata boundary (lowest limit exceptions) and greater than the greatest strata boundary (greatest limit exceptions); percentage of both values in relation to total number of rows; totalling values of attribute x below and above lowest and greatest limit exceptions
- percentage of totalled value for each strata in relation to grand total for all rows

After performing calculations on rows and values of each stratum, auditor may be introduced to a profile of the data in the database. Then, auditor may easier analyze any discrepancies from expected trends.

2.6 Duplication

Duplication method checks for numeric values (amounts) which appear more than once in attribute of interest of certain relation. As input parameters, method uses relation (r) and attributes (x) of interest, and its notation is:

$$D_x(r)$$

Usually, only one attribute for duplication values check is chosen. This method finds all values in chosen attribute that have multiple appearances in relation. Output of these values is sorted according to number of appearances in descending order and value in descending order. The objective of this test from auditing perspective is to find relatively small groups of recurring numbers. As it is shown in [13], when people invent numbers, they often tend to repeat values. This may be of major significance to auditing process. This test is usually performed only on value subsets which have excessive positive deviations in first two and first three Benford's Law tests.

2.7 Rounded values

Rounded values method finds numeric values (amounts) which are rounded to multiples of 5 or 10. Of course, method may be extended to multiples of 100, 1000 etc. which are all multiplies of 10, or to 25, 75 etc. which are all multiplies of 5. Notation is:

$$R_{x,m}(r)$$

where x is numeric attribute, m is multiplication value and r is relation. Method checks values of attribute x for all rows in relation and checks if value is multiplication of m. In fact, method searches for values in x that are divisible by m without remainder. Method also counts recurrences of all rounded values. Method's output is sorted by frequency of rounded values in descending order and rounded value itself in descending order.

This method looks for abnormal repetition of rounded values because such findings may be proof of estimation. Estimation is usually strongly related to omissions, errors and even fraud.

3 CAISAM Meta-model

Although large number of methods and corresponding computer assisted auditing tools are applied in practice, no specific methodology has been developed. A number of auditors used different methods within different procedures, steps and success. However, in this paper certain existing methods are chosen and linked into a new and consistent methodology. Meta-model of methodology is shown in Fig. 1.

Before audit execution, adequate planning step should be undertaken. During planning phase following prerequisites should be realized:

- clear definition of data structures to be processed including quantity, type, format and layout,
- definition of methods to be undertaken together with ways of methods linkage,
- definition of input and output data,
- determination of resource requirements, i.e. personnel, CAATs, processing environment,

- obtained access to the organisation's information system facilities, programs and data, including database definitions.

After planning steps for computer assisted auditing are completed, execution of specific auditing activity should commence.

Crucial prerequisite for computer audit commencement is understanding of data, data structures and appropriate mapping or relating data to business processes and events.

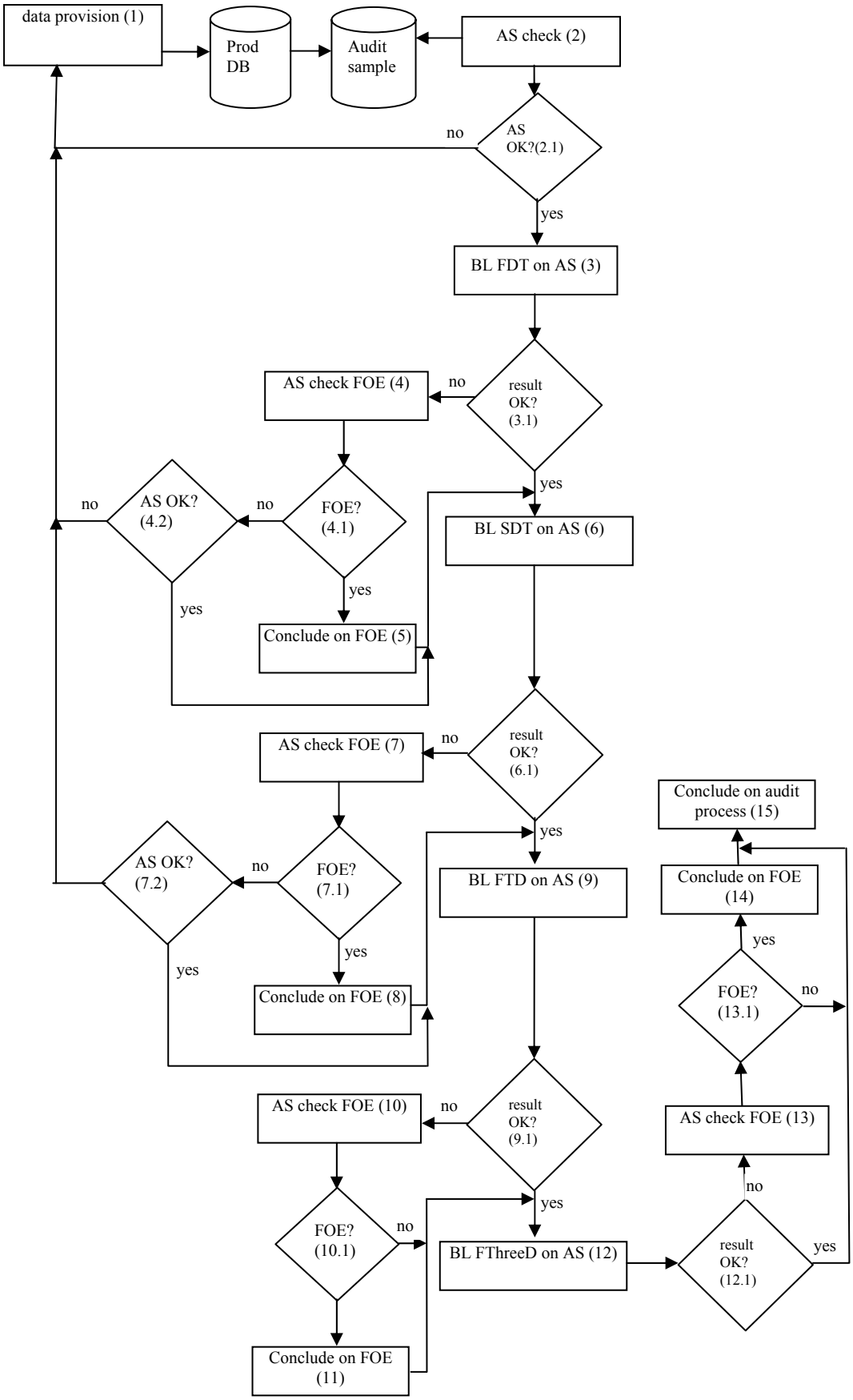

Figure 1. CAISAM Meta-model

First step (step 1 in Fig. 1) in developed computer assisted information systems auditing methodology is data provision. During this step auditor must define ways of taking over data from production databases. It is indispensable that statements for data provision are clearly defined and that conditions are appropriately constructed. As already mentioned, one of major cares in computer assisted auditing is the quality of provisioned data. If data is not appropriately provisioned, the whole auditing process will most certainly fail. The objective of the first step is to ensure that data are ready for auditing i.e. relevant and comprehensive. This step includes usage of following methods:

- selection – in order to declare which data are relevant for audit, to obtain comprehensiveness or completeness of data
- join – in order to combine data from different data tables (relations) or even from different data sources

Typical examples for this step is definition of SQL conditions according to period, types of transactions, certain departments, supplier, customer etc. Auditor should bear in mind that created selections and joins must be compatible with production database systems on which they will be applied. Although almost all modern databases are based on SQL standards, still there are slight differences.

Then abovementioned methods must be applied to production database.

After audit sample (AS) is created out of production database, auditor must check its relevance and completeness (step 2). This step must be executed because of following reasons:

1. It is possible that auditor's knowledge of data and data structures was not sufficient so planning phase and step 1 outcome in inadequate conclusions. Final results are SQL statements which produced incomplete or non relevant data. During this step auditor may correct wrong assumptions and improve his knowledge about data and data structures.
2. In some cases IT personnel will execute SQL queries prepared by the auditor. Since auditor will not be able to monitor process of database log on and queries execution, there is possibility queries may have been changed and/or data have been accessed from non production database (development, testing, data warehousing, temporary).

Of course, if planning phase of computer assisted auditing process was performed correctly and in step 1 data was provisioned according to created queries, then step 2 will confirm relevance and completeness.

Methods that should be used for data relevance and completeness check are selection, summarization (grouping), stratification and join.

For example, selection may be used in order to check if some specific data type (e.g. transaction type) is extracted in certain time period. Summarization may confirm if all data types (e.g. all transaction types) are excluded from production database. This method will clearly show if some data type does not exist in audit sample. Stratification may be set as first required method in data relevance and completeness check. After initial strata are created, auditor may be able to notice discrepancy between expected and actual data set. Stratification will produce strata according to auditor's instructions (e.g. transaction amount) which may show some data subsets were omitted from production database. This may be noticed if some strata lack data i.e. if there are problems with data segments. If that is a case, auditor may question created audit sample. Of course, it is useful to additionally check deviated strata. Join method should be used whenever relevance and completeness check is dependent on relations between audit sample and data in other tables or databases. Except natural or equi-join, auditor should concentrate on left and right outer joins. With outer joins it is possible to check if some data types in audit sample are missing.

Additional selection and summarization together with auditors experience and expectations may be used in questioning data with data and business process owner. Also, following additional checks may be performed:

- comparison of data in audit sample with data subsets from previous audits
- comparison of data in audit sample with data from other databases storing same or similarly organized data (e.g. data warehousing, data backups)
- comparison of data in audit sample with data in some other formats (e.g. paper) and/or with data from other sources (e.g. facility management system, vacation data, accounting data, data, application and event logs, data from corresponding business entities – e.g. suppliers, tax authorities, customers).

If auditor has reason to discard audit sample as incomplete and non relevant, step 1 has to be repeated. Of course, auditor should mitigate reasons that resulted in inappropriate audit sample in the first place. Following actions should be taken in order to get relevant and complete audit sample:

- auditor has to review his knowledge of data, data structure and business being audited
- auditor must work closer with IT personnel in preparation and execution of SQL statements

Steps 1 and 2 have to be repeated until auditor accepts data set as relevant and complete. When and if chosen data set is accepted as relevant and complete, next method that should be executed is test of first digit Benford's Law (step 3). First digit Benford's Law test will result in 9 data subsets (digits 1-9). Quite often, first digit Benford's Law test will not be focused enough and will not give enough proof whether data is fraudulent, erroneous and/or not complete. Even if data subsets will significantly deviate from Benford's Law distribution, usually

auditor will not be able to make conclusions why deviations occurred.

Data distributions should be compared to Benford's Law first test distributions in order to check for conformity. Conformity should be examined by Chi-square test, mean absolute deviation test or Z-statistics test. How conformity tests should be applied is explained in [8] and [9]. Diagrams should be used in order to perform visual inspection of data distributions and their deviations. If data does not follow predefined distributions ("no" result in step 3.1), auditor must check if data is fraudulent, erroneous and/or data resulted from omissions. This check must be performed in step 4.

However, if data satisfies Benford's Law first digit distributions ("yes" result in step 3.1), next Benford's Law test should be applied on data (step 6).

Concerning step 4 (audit sample check on fraud, omissions and errors – "AS check FOE") and 4.1 (conclusion on existence of fraud, omissions and errors – "FOE?") auditor has to understand rules of business conduct especially in relation to limits (e.g. according to authorization for payments, tax obligations and bonus gains), regularly recurring events for certain calculations (e.g. existence of regular interest calculation), rules for calculated database fields (e.g. tax calculations), etc. Limit avoidances resulting in pushing certain amounts bellow or over predefined limits are rather often, so auditor should be aware of their existence. After Benford's Law test shows deviations, additional checks on data should be based on selection, summarization, duplication and rounded values checking methods. Auditor must understand if there are specific reasons why data does not follow Benford's Law. If numbers are clustered around certain limits and clustering can be related to limits set by management than there is high probability of committed fraud. It is important to notice steps 4 and 4.1 have to be repeated after each Benford's Law tests resulting in deviations (steps 7, 7.1, 10, 10.1, 13).

If tests against Benford's Law significantly deviate, auditor may perform some additional tests only for certain digits which deviate from Benford's Law frequencies in order to check against fraudulent or erroneous activity. Examples of use of certain methods in steps 4, 7, 10 and 13 after each Benford's Law test shown discrepancies are:

- finding rounded amounts, amounts starting with multiplies of 10 or 100 (rounded values method)
- counting frequency for each amount – searching for multiple occurrences (number duplication) of the same amounts especially in short period of time (duplication method)
- finding rounded and duplicated amounts on very same day (rounded values, duplication and summarization methods)
- searching for rounded and duplicated crediting amounts for certain accounts in specific period of time (rounded values, duplication and summarization methods)
- rounded and duplicated crediting amounts by detecting specific employees (rounded amounts, duplication and join methods)
- amounts slightly below certain threshold – this checks if payer or payee wanted to evade entering certain tax limits or additional checks by authorities (selection method)
- finding amounts slightly greater than certain threshold – sometimes business entities with performance (sales, investments, contracted amounts etc.) above certain limits may get special status in their relationships with state authorities, banks or other business entities (selection method)
- multiple payments of same type paid same day (or in short time interval) with sum above or below certain threshold (summarization, selection and join methods)

On the contrary, if auditor can not relate clustering, i.e. non conformance to Benford's Law distributions, to evasion of certain business rules, it is possible data set simply does not follow Benford's Law first digit distribution. Auditor may conclude deviations from first digit Benford's Law distributions are result of normal business conduct. As it is already stated, more focused tests should be applied on data. Use of selection, summarization, number duplication and rounded values checks are required. Using these methods auditor can focus on data causing deviations, after selecting and summarizing according to defined criteria. One reasonable example is selection of certain types of transactions committed by certain employees. Then number duplication and rounded values check may be applied on some number subsets. If considerable number of values falling into come clusters is duplicated or rounded, perhaps fraudulent behaviour is on scene.

Also, if application of selection method showed clustering of transactions in certain time period, then summarization of transactions according to certain employee and transaction type in relevant period could be of special audit focus. Additional review based on duplication and rounded values should be also performed.

Nevertheless, even if it is possible to make conclusions about fraud, omissions and/or errors solely on first digit Benford's Law test, it is indispensable to continue with next Benford's Law test. It is second digit Benford's Law test (step 6). This is mandatory, since other types of deviations may be spotted after application of rest of Benford's Law tests.

However, if data are deviated from Benford's Law first digit test and there is no proof of fraud, omissions or errors in business conduct, auditor must again check audit sample data on relevance and completeness (step 4.2 – "AS OK?"). It is possible that first check on relevance and completeness

performed in step 2 was not adequate and that auditor made wrong conclusions. In auditing work it is possible to commit such wrong judgement which is the most often consequence of:

- insufficient knowledge of business system and underlying data (resulting in wrongly performed step 1)
- inadequate application of selection, summarization, join and/or stratification methods (step 2)
- insufficient knowledge of underlying data organization and structure (steps 1 and 2)

If auditor believes there is considerable possibility that data is not fraudulent but not relevant and complete ("no" result in step 4.2), additional checks must be performed. The same methods as in step 1 and 2 may be carried out after auditor additionally improves his knowledge of business system, data organization and data structure. Of course, extensive knowledge on methods in focus is indispensable. So, steps 1 and 2 must be accomplished again not bearing in mind that data was not appropriately sampled and checked in first place.

If, after additional checks, it is concluded data is relevant and complete ("yes" result in step 4.2), auditor should proceed with Benford's Law testing. Second digit test (step 6 – "BL SDT on AS") should be next in sequence. This test is also fairly broad, not very focused on certain data subject and usually will not be enough for final judgement on eventual fraudulent, omitted and/or erroneous data. Second digit test will show distribution of second digits in observed amounts and similar to first digit test it is not very efficient auditing method. It results in only 10 data subsets (digits 0 to 9). Such small number of resulting subsets has consequence of quite large proportions. Any of 10 subsets will often be too large for audit conclusions. However, second digit test may show significant deviations which may be reflection of certain limits evasions resulting in number clustering.

The rest of steps following immediately after Benford's Law second digit test (steps 6.1, 7, 7.1, 7.2 and 8) are in essence coherent with corresponding steps performed after Benford's Law first digit test (steps 3.1, 4, 4.1, 4.2 and 5).

However, it is possible that step 1 resulted in reasonably focused data set i.e. audit sample. If it is so, there is higher probability that already first and/or second digit Benford's Law test and following selections, summarizations, duplication and rounded values checking methods will be enough for conclusions on some fraud, omissions and/or errors in business system. Although in some cases deviations of data set from Benford's Law first and second digit distribution are easily explained after additional focusing by selection, summarization, duplication and rounded values checking methods auditor has to continue with more focused Benford's Law tests, specifically first two digits and first three digits tests. These tests could give additional information and conclusion why deviations occurred. Also, if no deviations where noticed by first and second digit tests it is mandatory to accomplish first two and first three digit Benford's Law tests. As it is already stated, these tests are much more focused and their application result in much focused audit samples so the conclusions about certain deviations may be more precise.

The first two digits Benford's Law test (step 9) has much more audit relevance than first digit and second digit tests. It is more focused since it takes into account 90 combinations of first two left most digits. First two digits test is much more sensible (in theory exactly 9 times) than first digit test. This means auditor will have much more potential in noticing deviations and recovering fraud, omissions and/or mistakes in business activity. After performing this test, auditor can spot finer irregularities or discrepancies and then perform additional analysis with other methods (summarization, selection, rounded amount, duplication, join) belonging to CAISAM methodology as shown in step 4. Each significant deviation on first two digits should be thoroughly checked and compared to certain business rules (steps 10 and 10.1). The difference between first two digits and previous Benford's Law tests performed in steps 3 and 6 is that auditor does not have to recheck on relevance and completeness of audit sample. So, there is no "AS OK?" check which was mandatory if fraud, omissions and/or errors check were not found after noticed discrepancies in first and second digit tests. This is the result of experience - it may be said if audit sample was checked three times (steps 2, 4.2 and 7.2) in worst case scenario that auditor finally extracted appropriate (i.e. complete and relevant) data from database. Furtherly, whatever is outcome of step 10.1 (deviations are result of fraud, omissions and/or errors or not), next and final Benford's Law test must be executed on audit sample.

First three digits test (step 12) is the most specific of all Benford's Law tests. It is highly focused because it gives considerably smaller subsets than any other Benford's Law test. This test covers 900 three leftmost combinations (digits 100 to 999 inclusive). It is 9 times more sensible than first two digits test. Since it has greater precision it will be the best choice for recovering abnormal data duplications. So, duplication method should be performed with more care especially for three digits combinations that are deviated from Benford's Law distribution. Also, this test will be much more beneficial for analysis and conclusions on clusters or so called "positive spikes" on diagram representations resulted from deviations from Benford's Law distribution. Background for explanation why this test is more efficient in cluster analysis is the same as for duplications: test is simply more focused, resulting in larger number of subsets (900) what is effected by smaller number of elements in audit subsets. Apart from that, steps 12.1 to 14 are

already appropriately described in previous paragraph on Benford's Law first two digits test.

After conclusion on existence of fraud, omissions and errors concerning first three digits test (step 14), auditor has to review whole audit process defined by methodology and make final conclusions. This step is necessary as individual tests may show different deviations so auditor may discover various discrepancies resulted from fraud, errors and/or omissions.

4 Conclusion

Business processes are supported and enabled by information systems. As a consequence of constant change and development of business processes, information systems tend to become very complicated and use enormous volumes of data. It brings the conclusion that often it is not possible to audit information system without computer assisted auditing. Although various methods for computer assisted auditing of information systems exist, there is need for development of unified methodology. In this paper, framework and meta-model of such methodology are designed. Methods, steps, input and results of each method together with possible conclusions in practical situations are established. With such methodology, whole auditing process can be improved, become consistent, standard, comparable and executed faster with less resources.

5 Acknowledgement

This contribution as a part of the project No 016-0161199-1718 Developing the ICT management method was supported by Ministry of Science, Education and Sports in the Republic of Croatia.

References

[1] Coderre D. G.: **CAATs and Other BEASTs for Auditors**, Global Audit Publications, Vancouver, Canada, 2001.

[2] Coderre D. G.: **Fraud Detection – A Revealing Look at Fraud**, Ekaros Analytical Inc., Vancouver, Canada, 2003.

[3] DeCeuster M. J. K., Dhaene G., Schatteman T.: **On the hypothesis of psychological barriers in stock markets and Benford's law**, Journal of Empirical Finance 5(3), 1998., pages 263-279.

[4] Doucouliagos C.: **Number preference in Australian stock prices**, School Working Papers – Series 2003, Deakin University, 2003.

[5] Durtschi C., Hillison W., Pacini C.: **The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data**, Journal of Forensic Accounting 1524-5586/Vol. V, 2004., pages 17-34.

[6] Hill T. P.: **The First Digital Phenomenon**, American Scientist, v86 n4, 1998., pages 1-6

[7] Koedijk K. G., Stork P. A.: **Should we care? Psychological barriers in stock markets**, Economic Letters 44, 1994., pages 427-432.

[8] Krakar Z., Žgela M.: **Evaluation of Benford's Law application in stock prices and stock turnover**, Society and Technology 2008, 26th to 28th June 2008, Zadar, Croatia.

[9] Krakar Z., Žgela M.: **Application of Benford's Law in Payment Systems Auditing**, Journal of Informational and Organizational Sciences, Vol. 33, No. 1, Varaždin, Croatia, 2009., pp. 39-51.

[10] Lanza R. B.: **Proactively Detecting Occupational Fraud Using Computer Audit Reports**, The IIA Research Foundation, Altamonte Springs, USA, 2003.

[11] Ley E.: **On the peculiar distribution of the U.S. Stock Indices Digits**, Resources for the Future, Washington DC, USA, 1994.

[12] Matthews R.: **The Power of One**, New Scientist vol. 163 issue 2194, 1999., pages 26-30.

[13] Nigrini M. J.: **Digital Analysis Using Benford's Law**, Global Audit Publications, Vancouver, Canada, 2000.

[14] Nigrini M. J., Mittermaier L. J.: **The Use of Benford's Law as an Aid in Analytical Procedures**, Auditing - A Journal of Practice & Theory, Vol. 16, No 2, 1997., pages 52-67.

[15] Opaska W: **Survey of Accounts Payable Disbursements**, available at http://www.fcc.gov/Bureaus/Inspector_General/Reports/surveymemo.txt, Accessed: 27th March 2008.

[16] Turnbull, C. S.: **Fraud Investigation Using IDEA**, Ekaros Analytical Inc., Vancouver, Canada, 2003.

[17] Wells, J. T.: **Corporate Fraud Handbook - Prevention and Detection**, John Wiley and Sons Inc., Hoboken, USA, 2007.