CALCULABLE-CONTENTS IDENTIFICATION MATRIX: A COMPUTERISABLE KEY FORMAT FOR PLANT IDENTIFICATION

Adepoju Tunde Joseph Ogunkunle

Department of Pure and Applied Biology, Ladoke Akintola University of Technology,
P. M. B. 4000, Ogbomoso, Nigeria

ABSTRACT

Many of the paper-based/printable taxonomic key formats available for plant identification are fraught with inadequacies: fixed sequence of steps, non- readily amenable to automation or computerisation, lack of provision for confirmation of suspected plant identity and indeterminable character states, tedious construction and navigation procedures, and inability of users to ascertain the extent of reliability of the identification process. Aimed at making the practice of plant taxonomy more attractive, less laborious and dreaded, this paper proposes the calculable-contents identification matrix (CCIM), a new key format with structural and functionality attributes to circumvent some of the enumerated challenges. The status and prospects of CCIM are discussed with reference to the inadequacies observed in the dichotomous key format with which most taxonomists are familiar. Based on its features, applicability and potential outputs, CCIM is adjudged to be a useful template for reliable manual and electronic methods of plant diagnosis.

KEYWORDS

Computerised key, Dichotomous key, Expert system, Multi-access key, Single-access key.

1. INTRODUCTION

Correct identification of plants is an important prerequisite to achieving desired goals in health care [1], sustainable food production and housing [2], criminal justice [3], forest resources management [4], environmental protection [5] and biodiversity conservation [6]. However, there are notable constraints on plant identification that are rooted in the structure and functionality attributes [7] of commonly used identification keys including: being tedious to construct [8], having fixed point of entry and daunting path of navigation [7], the problem of 'unanswerable couplet' [9], the associated 'momentary distractions' that can cause a user to forget his or her position in a key [10], being unusable for confirmation of suspected identity, non- readily amenable to automation [11], and inability of users to ascertain the extent of reliability of the identification process. For these reasons, plant misidentification, misrepresentation, adulteration and substitution frequently occur with associated public health [1], social [12], environmental [13], legal [14], and economic [15] burdens. Identification is thus viewed by many practitioners as onerous task, and a huge responsibility on the shoulders of plant taxonomists, who, unfortunately also, have to contend with a number of other challenges including the intricate nature and complexity of plant life, and variability in their characteristics [16], and perceived tediousness of taxonomic practices along with obsolete tools for identification [17]. All these have led to declining interest in plant taxonomy by upcoming students of biology [18].

DOI:10.5121/ijaia.2025.16505

Automation of identification keys appears to be a pragmatic way to enhance the functionality of taxonomic keys, but the format and style of a key are important features that determine whether it can be automated by computerisation. By definition, computerisation is the process of developing, implementing, and using computer systems for activities that were not previously carried out by means of computer [19]. A key, which on account of its design/format allows objective data comparison through numerical computations, is referred to as being programmable or computerisable. When such key format is implemented with appropriate computer language and simulated with diagnostic data on a plant group, it becomes a computerised key [20]. Conceptualisation of this study was informed by the determination to create a new computerisable key format with the prospect to ameliorate some of the enumerated challenges associated with many of the extant key formats.

Basically, a taxonomic key is derived from a data matrix of a given number of ‘objects’ such as plant species. Although it is usually possible to contrive a large number of different keys for one set of objects, the functionality of such keys will not be equal [21]. For this reason, invention of new key formats shall continue to be a welcome development in taxonomy. It is against this background that the present study aimed at making the practice of plant taxonomy more attractive, less laborious and dreaded, and so, the objective is to propose the calculable-entries identification matrix (*CCIM*), a new taxonomic key format with highlights of features, construction procedures and usage that should possibly make it desirable, either as alternative or complementary tool for plant identification.

2. RESEARCH METHOD

2.1. Conceptualisations from Heuristic Approach to Decision Making

In order to actualise the objective of this study, the first step taken was to align with the thoughts of Pankhurst (1970) [21] on the two complementary problems in taxonomy, which are still valid till date. First problem: ‘given a set of objects (e.g. plants), examine their characteristics in order to find a classification i.e. group the objects into subsets (or taxa), and assign names to the subsets’; and second: ‘given a classification and an object, identify that object’. In other words, given a list of the characteristics of named subsets which are known to exist, and an additional object, decide which subset the object belongs (i.e. recognise it, or find its name). Noting the significance of the taxonomists’ diagnostic key as a tool in the process of identification, the second step taken was to undertake a critical examination of the formats and styles of the available ink-on-paper taxonomic key formats along with the challenges associated with their features, construction and application [10]. In doing these, attention was focused mainly, but not limited to the dichotomous key format that is most widely used [22].

In an effort to address some of the inadequacies in some extant key formats, the subject of data structure and ‘decision tree’ in computing science were reviewed, and the concept, types and functionality of key in computer database management systems (*DBMS*) were studied [23]. Consideration was also given to selection criteria for construction of efficient diagnostic keys [24] [25], while efforts made so far on the application of information and communications technology (*ICT*), and computer science in developing identification keys and diagnostic tables were reviewed, including generation of diagnostic keys [21] [26] [27] [20] [28] [29] and development of expert systems [25] [2] [22]. Information obtained from the steps highlighted above, along with those obtainable in some mathematical representations of relationships among objects, such as matrices [30] were integrated into a thought to develop the calculable contents identification matrix (*CCIM*), a new key format with far reaching desirable qualities.

2.2. A Posteriori Determination of the Status of Characters Applicable in the New Key Format

This study derived in part, some strength from the statistics-based conclusions by Adams (1975) [31] that discouraged the use of equal weighting in numerical taxonomy. It therefore agreed with the concept and practice of *a posteriori* characterweighting in the classification process [32], believing that the taxonomic value of a character is increased if the biological significance of the character has been determined. But it has long been established [33] and hitherto believed that the biological significance of many taxonomic characters is unknown or poorly understood. So, for all practical purposes and convenience, the classificatory value, herein referred to as information content of a character is accorded the status of biological importance, at least for the purpose of taxa delimitation [34].

2.3. Procurement of Data for the Purpose of Illustration

Wood anatomical data on five medicinal herbs marketed as plant roots in Ogbomoso township, south western Nigeria were sourced for the purpose of illustration from the 2019 compilation of unpublished results at the medicinal plants research laboratory in the Department of Pure and Applied Biology, Ladoke Akintola University of Technology, Ogbomoso, Nigeria. The data items were obtained in accordance with the standard procedures: tissue sectioning/ maceration [35], staining, dehydration [36], mounting [37], and microscopic observations [38] [39], while the terminology and descriptions of observed features followed those of the International Association of Wood Anatomists (IAWA Committee, 1989) [40]. Staining was done in 1% alcoholic safranin, mounting was carried out in Canada balsam and observations made using Olympus biological microscope CH20iModel with binocular facility.

Twenty-three diagnostic characters (strictly, character states) were collated, and scored ‘present’ or ‘absent’ for each of the five taxa (Table 1). Where a character was scored ‘present’, it was further evaluated quantitatively based on its frequency of occurrence or percent composition by wood tissue volume [41]. Thus, the 23 characters/character states, as defined, including the quantified features were operationally qualitative (present or absent), but each of them was accorded quantitative transformation in line with commonly used descriptive terms as follows: always present/always found (100%); usually found (60-99%); average occurrence (40-59%); sometimes found (10-39%); seldom/rarely found (1-9%); and not found (-100%).

2.4. Design and Features of Calculable-Contents Identification Matrix

Adopting the mathematical definition of matrix as a rectangular array of quantities or expressions in rows and columns that is treated as a single entity, manipulated according to particular rules [30], the *CCIM* was conceptualised as a data matrix with recursively enumerable entries in respect of some objects (in the rows), such that the process of ‘divide and conquer’ can be applied on the characteristics of the objects (in the columns) to come up with precise and quantifiable decisions. Such decisions will lead to recognition/ identification of the rows (i.e. the objects). With this background, the *CCIM* was proposed from the point of view of a number of logical steps as follows:

i. Each character to be adopted in the proposed key format should be unique and clearly defined in such a manner that it can be scored ‘present’ or ‘absent’ for each of the taxa; thus the usual character ‘states’ are elevated to the status of unit characters (see [42]);
ii. In drawing out characters for the construction of the key, consideration should be given to a sizable population and /or observations in each taxon so as to obtain information on

the variations or frequency of occurrence of the characters in the taxon, and across the taxa involved;

iii. Considering each unit character (which is strictly speaking a character state) in the key, a maximum of seven possible quantitative conditions or 'states' are proposed, into which all the taxa can be classified; these states are determinable based on the commonly used descriptive terminology and accordingly assigned weights as follows:

State I: If the character has 100% frequency of observation within the population of plants/specimens observed in a taxon, the state is 'always present (***AP***)' its effective weight being 100%;

State II: If the frequency of observation of the character in the observed population falls between 60-99%, the state is 'usually present (***UP***)' and its effective weight is calculated as the median or arithmetic mean of the range i.e. 80%;

State III: If the frequency of observation of the character in the observed population falls between 40-59%, the state is 'averagely present (***AV***)' and its effective weight is 50%;

State IV: If the frequency of observation of the character in the observed population falls between 10-39%, the state is 'sometimes present (***SP***)' and its effective weight is 25%;

State V: If the frequency of observation of the character in the observed population falls between 1-9%, the state is 'seldom/rarely present (***RP***)' and its effective weight is 5%;

State VI: If the character is absent in the entire observed population of a taxon, the state is 'always absent (***AB***)' and its effective weight is -100%;

State VII: If a situation arises when a unit character had earlier been defined, and another is later being defined as an offshoot of the first, the latter character will be inapplicable to a plant taxon in which the former character was absent. This state is defined as 'not applicable (***NA***)' and its effective weight is 1%.

iv. Where it is not possible to rank the observation of a character in percentage in the observed population of a taxon other than to describe it as present, such character is simply recorded as having 100% positive frequency of observation i.e. ***AP***, and it is treated so;

On the whole, there should be a minimum of two 'conditions' or 'states' per character per taxon, else the character is dropped in the construction of a *CCIM*;

v. The new identification key is proposed as follows:

a. The key consists of a table/data matrix of diagnostic characters in columns; and rows of plant taxa. The body of the matrix consists of 'cells', each of which is labeled*AP, UP, AV, SP, RP, AB, or NA* as appropriate with respect to the observations made on the populations of the taxa during the construction of the key, while the corresponding effective weights assigned thereto i.e. 100, 80, 50%, 25, 5, -100, and 1 respectively are defined as footnotes at the bottom of the table/matrix;

b. From the submission in “a” above , it is possible to calculate a quantity, called the information content or relative importance of each of the characters adopted for the construction of the key in accordance with Bisby (1970) [34] as follows:

$$-H = a \ln a + b \ln b + c \ln c \ldots + n \ln n \qquad (1)$$

where a = the fraction of the plant taxa in a^{th} state; b is the fraction of the taxa in b^{th} state and so on, such that n = the fraction of the taxa in n^{th} state.

The calculated information content of each character is recorded in the first row below the list of the diagnostic characters at the top of the table;

c. The final step in the construction of a *CCIM* is the computation of the sum of absolute weighted information contents (*SAWIC*) which is achievable following the modified procedure for calculating weighted arithmetic mean [43] as follows:

$$SAWIC = \sum_{i=1}^{n} | wiqi | \qquad (2)$$

where

w = effective weight/ value for each character in a taxon
q = the quantity of information/ information content of a diagnostic character
i = counter which ranges from 1 to n where n is the total number of characters used

The *SAWIC* of each taxon recorded for it in the last column of the table.

d. For avoidance of possible ambiguity arising from a large number of taxa and/or characters, the matrix could be fragmented over a number of pages of paper in such a way that in the few taxa (or rows) considered in the first page, all the characters (or columns) should be listed first, possibly running to the next few pages, following which their *SAWIC* and *ASWIC* are recorded; then, consideration is given to the next set of taxa for the same set of characters, and so on (see Table 1).

Table 1: Wood Anatomy-based Calculable –Contents Identification Matrix (CCIM) for Five Medicinal Herbs in Ogbomoso, Nigeria

	Square ray cells in TS	Procumbent ray cells in TS	Uniseriate rays in TLS	Biseriate rays in TLS	Multiseriate rays in TLS	Homocellular rays in TLS	Heterocellular rays in TLS	Linear rays in TLS	Monoconvex rays in TLS	Biconvex rays in TLS	Dumb-bell rays in TLS	Solitary vessels in TS	Apotracheal parenchyma in TS	Paratracheal parenchyma in TS	Paratracheal parenchyma, scanty type	Paratracheal parenchyma, vasicentric type	Paratracheal parenchyma, aliform type	Vessel density, <40/mm²	Vessel density, >70/mm²	Vessel diameter, <70µm	Vessel diameter, 100-250 µm	Parenchyma cells density in TS, < 90/mm²	Parenchyma cells density in TS, >190//mm²	
Species/ Information content (q)	1.333	0.950	1.333	0.950	1.333	1.333	1.610	1.333	0.673	0.673	0.950	1.333	0.501	0.501	0.501	0.501	0.501	0.501	0.501	0.501	0.501	0.673	0.673	*SAWIC*
ARRI	RP	UP	AP	AB	AB	AP	AB	AP	AB	AB	AB	AV	AP	AB	NA	NA	NA	AP	AB	AB	AP	AP	AB	1604.82
CAHA	SP	UP	UP	SP	RP	UP	RP	SP	SP	UP	AB	SP	AP	AB	NA	NA	NA	AB	AP	AP	AB	AB	AP	1030.11
PANI	AV	AV	UP	SP	RP	UP	SP	UP	SP	AB	SP	AV	AP	AB	NA	NA	NA	AP	AB	AB	AP	AB	AP	1115.97
SALA	SP	UP	AV	SP	SP	AB	AP	AB	AB	UP	RP	AP	AP	AB	NA	NA	NA	AP	AB	AB	AP	AP	AB	1356.55
ZAZA	AB	AP	AB	UP	UP	AV	AV	AB	SP	UP	AB	RP	AB	SP	SP	SP	SP	AP	AB	AB	AP	AP	AB	1432.15

SAWIC, Sum of absolute weighted information content; *ARRI*, *Arristolichia . ringens*; *CAHA*, *Calliandra haematocephala*; *PANI*, *Parquetinanigrescens*; *SALA*, *Sarcocephaluslatifolius*; *ZAZA*, *Zanthoxylum zanthoxyloides*; Frequency of observation of characters and their effective weights: AP, Always present (100%); UP, Usually present (80%); AV, Averagely present (50%); SP, Sometimes present (25%); RP, Rarely present (5%); AB, Absent (-100%); NA, Not applicable (1%). The effective weight of 'states' for each character is the arithmetic mean of its range of observations.

2.5. Procedure for Plant Identification using CCIM

A set of rules is being proposed for using a *CCIM* to carry out identification of plant specimens. The process of identification involves the construction of a plant specimen evaluation table (***PSET***) from the features observable on the unknown plant specimen using the identification matrix as a guide. In adopting this procedure for a reliable identification exercise, a sizable number of plant samples should, as much as practicable, be examined and the observations should be quantified as percent occurrence within the samples examined. The proposed rules are as follows:

Step I: Draw a plant specimen evaluation table (*PSET*), which consists of rows, about the number of the taxa in the key (i.e. *CCIM*) to be used, and columns, about the number of the diagnostic characters the user intends to make use for identification;

Step II: Select a character from the *CCIM*, starting with any of the diagnostic features in the list, and write this out at the top of the first character column in the *PSET*;

Step III: Evaluate the specimen(s) of the plant to be identified and determine its status based on the selected character as follows:

a. If the character is absent/ not observable in the plant specimen(s), then check through the list of taxa on the *CCIM* for those having *AB* as character state for the character under consideration. Regarding each of these taxa in the matrix as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*; calculate the respective observed weighted information content (*OWIC*) for the selected taxa as -100q, where q is the information content for the character (as indicated in the *CCIM*) and record same on the *PSET* as appropriate (e.g. see column number 1 in table 2);

b. If the character is present/ observable in the plant specimen (s) for identification, determine first, the frequency of observation of the character in some samples of the plant and use this to calculate the *OWIC* for the affected taxa in the *CCIM* as follows:

  - For 100% frequency of observation, check through the list of all the taxa on *CCIM* which have *AP, UP, AV, SP* and *RP* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where w is the effective weight (100, 80, 50, 25 and 5 respectively) and q is the information content of the character as indicated in the *CCIM*, and record on the *PSET* as appropriate (e.g. see column number 2 in Table 2);
  - For 60- 99% frequency of observation, check through the list of all the taxa on *CCIM* which have *UP, AV, SP* and *RP* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names

into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where w is the effective weight (80, 50. 25 and 5 respectively) and *q* is the information content of the character (as indicated in the *CCIM*);

- For 40- 59% frequency of observation, check through the list of all the taxa on *CCIM* which have *AV*, *SP* and *RP* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where w is the effective weight (50. 25 and 5 respectively) and *q* is the information content of the character as indicated in the *CCIM*, (e.g. see column number 8 in Table 2);

- For 10- 39% frequency of observation, check through the list of all the taxa on *CCIM* which have *SP* and *RP* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where *w* is the effective weight (25 and 5 respectively) and *q* is the information content of the character as indicated in the *CCIM* , (e.g. see column number 9 in Table 2);

- For 1- 9% frequency of observation, check through the list of all the taxa on *CCIM* which have only *RP* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where *w* is the effective weight of 5 and *q* is the information content of the character as indicated in the *CCIM*, (e.g. see column number 7 in Table 2);

c. If the character is present in the plant specimen for identification but it is not possible to evaluate the frequency of its observation (perhaps due to insufficient time, lack of appropriate methodology or equipment or due to unavailability of enough plant material, *etc*.), the observation, for convenience is regarded as being of 100% frequency distribution. It is thus treated in this context for the computation and recording of *OWIC* i.e. using 100, 80, 50, 25 and 5% as effective weights;
d. If the character is inapplicable to the plant specimen, check the list of all the taxa on *CCIM* which haveonly *NA* for the character being considered. Regarding each of these taxa as 'potential identity' of the unknown plant, copy their names into the cells of the selected character column in the *PSET*, calculate and record the *OWIC* for them as *wq* where *w* is the effective weight of 1% and *q* is the information content of the character (e.g. see column number 13 in Table 2;
e. As alternative to copying only the names of taxa which exhibit the features evaluated in the unknown plant specimen into the *PSET*, the names of all the taxa in the key could be copied into the first column of *PSET* from the onset but *OWIC* is calculated for only those taxa that fall in line with the result of specimen evaluation;

Step IV: Select another character from the *CCIM*, write it at the top of the second character column in the *PSET* and repeat step III above, ensuring that each row is exclusively assigned to a taxon, until all the characters determinable by the user on the 'unknown' plant have been exhausted;

Step V: Determine the sum of observed absolute weighted information content (*SOAWIC*) for all the taxa involved in the computation on the *PSET* by adding up the absolute values of weighted observations row by row and recording same at the next column of *PSET*; the name of the taxon with the highest value of *SOAWIC* is taken as the identity of the unknown plant specimen;

Step VI: Compute the reliability of identification (*RID*) for each of the taxa involved in the *PSET* as follows:

$$RID = \frac{SOAWIC}{SAWIC} \times 100 \qquad (3)$$

where *SAWIC* = sum of absolute weighted information content for each taxon as indicated in the last but one column of the *CCIM*.

2.6. Procedure for Confirming Suspected Plant Identity Using CCIM

The procedure for confirming a plant’s identity involves computation of the sum of suspected absolute weighted information content (*SSUAWIC*) for the suspected taxon on the one hand, and all the unsuspected taxa on the other hand, and comparing the magnitudes of these values as follows:

Step I: Prepare a plant specimen evaluation table (*PSET*), which consists of as many rows as the number of taxa in the key (*CCIM*) to be used, and columns as the number of characters the user intends to apply for the confirmation exercise;

Step II: Copy the names of all the taxa from the *CCIM* into the first column of *PSET,* starting with the name of the suspected taxon (e.g. see column number 1 in Table 3);

Step III: Focusing on the suspected taxon first, select and copy one character intended for use into the top of first character column of *PSET* (as shown in column number 2 in Table 3); evaluate the plant specimen whose identity is being confirmed based on the character; if frequency of observation of the character in the suspected taxon falls within the limits recorded from the specimen, compute the suspected absolute weighted information content (*SUAWIC*) as *wq* for the suspected taxon; else, skip the use of the character and proceed to the next, until all the characters evaluated on the specimen have been considered;

Step IV: Select and copy the next character into the next column, and repeat ‘step III’ above until all the characters intended for use in the exercise have been picked;

Step V: Determine the sum of suspected absolute weighted information content (*SSUAWIC*) for the suspected taxon by adding up the *SUAWIC* values; then compute the reliability of confirmation for the taxon (RC_{ST}) as follows and record as appropriate:

$$RC = \frac{SSUAWIC}{SAWIC} \times 100 \qquad (4);$$

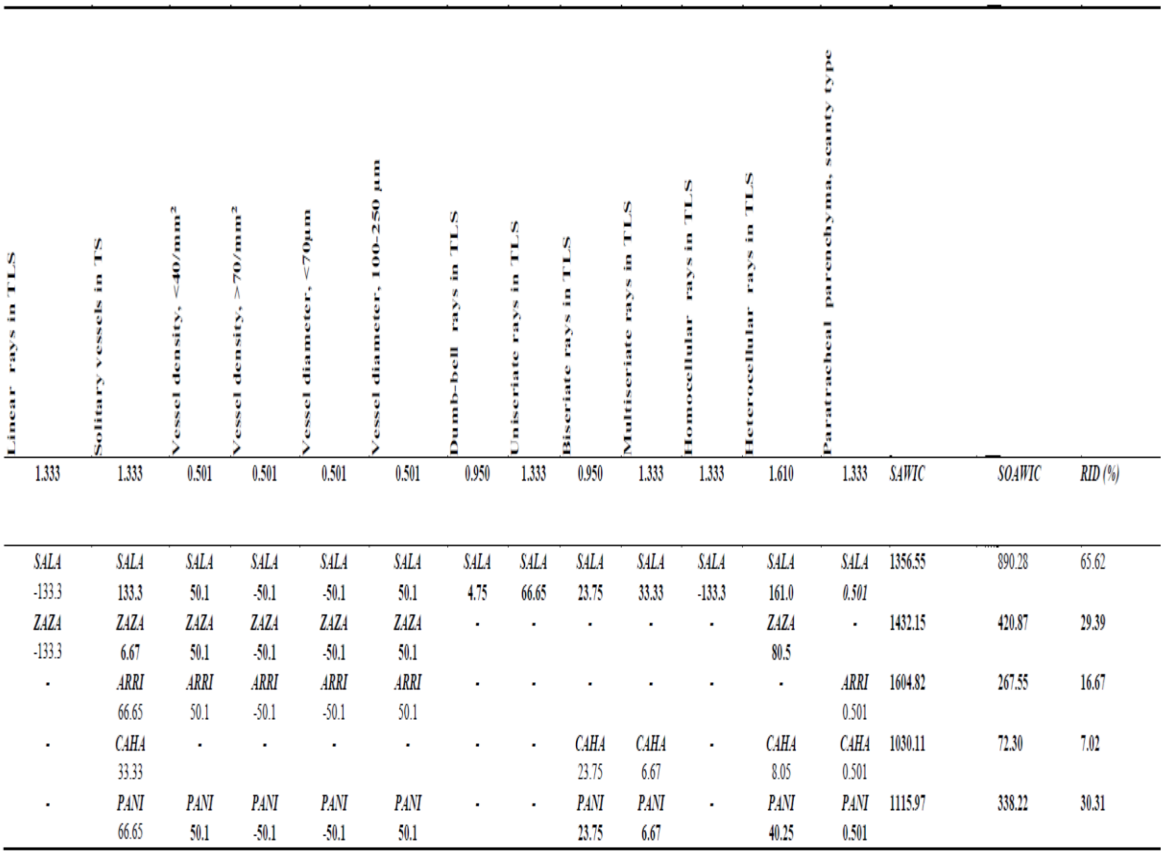
where *SAWIC* = sum of absolute weighted information content for the suspected taxon as indicated in the last column of the *CCIM.*

Step VI: For each of the ‘unsuspected taxa’ in the *PSET*, repeat ‘Step III’ above using the same set of characters;

Step VII: Repeat ‘Step V’ above, and hence compute the reliability of confirmation for each unsuspected taxon (RC_{UT}) as indicated in equation (4) and record as appropriate;

Step VIII: Compare the magnitude of *SSUAWIC* of suspected taxon with those of the other taxa in the *PSET*. If this valueis highest, suspicion is correct, else, it is incorrect.

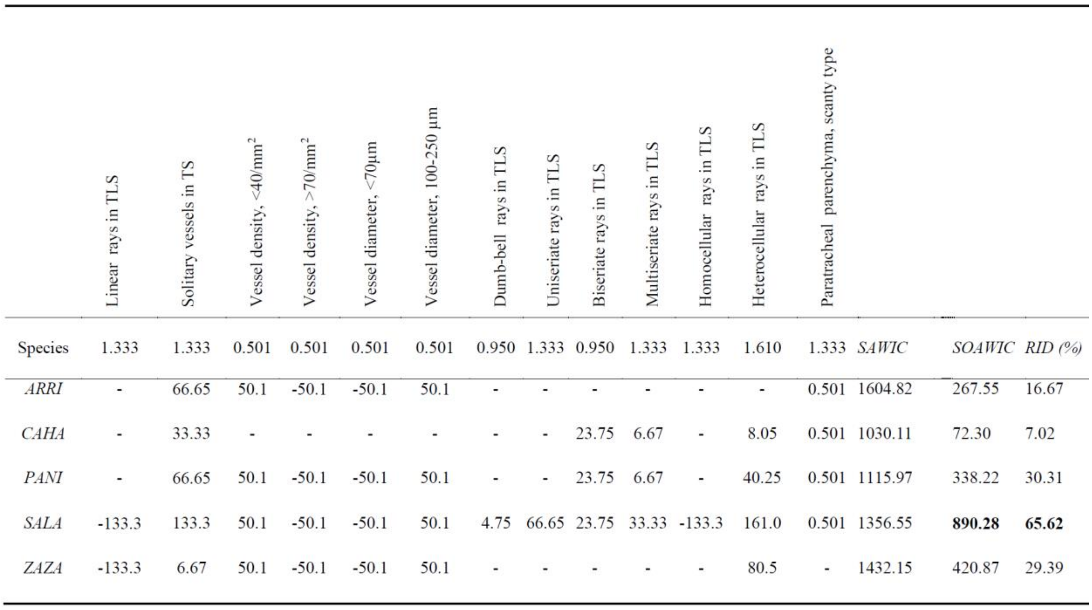
In summary, if a taxon name is suspected for a plant specimen in hand, evaluate the specimen based on a number of observable features listed in the *CCIM*, and use these to calculate *SSUAWIC* for the suspected taxa on the one hand, and all the unsuspected taxon the other. In all the cases, for a character to be used in the computation, the frequency of observation of the character in a taxon should fall within the limits recorded from the specimen. Finally, compare the magnitudes of *SSUAWIC* values in all the taxa involved and draw a conclusion.



Linear rays in TLS	Solitary vessels in TS	Vessel density, <40/mm²	Vessel density, >70/mm²	Vessel diameter, <70μm	Vessel diameter, 100–250 μm	Dumb-bell rays in TLS	Uniseriate rays in TLS	Biseriate rays in TLS	Multiseriate rays in TLS	Homocellular rays in TLS	Heterocellular rays in TLS	Paratracheal parenchyma, scanty type			
1.333	1.333	0.501	0.501	0.501	0.501	0.950	1.333	0.950	1.333	1.333	1.610	1.333	*SAWIC*	*SOAWIC*	*RID (%)*
SALA -133.3	*SALA* 133.3	*SALA* 50.1	*SALA* -50.1	*SALA* -50.1	*SALA* 50.1	*SALA* 4.75	*SALA* 66.65	*SALA* 23.75	*SALA* 33.33	*SALA* -133.3	*SALA* 161.0	*SALA* *0.501*	1356.55	890.28	65.62
ZAZA -133.3	*ZAZA* 6.67	*ZAZA* 50.1	*ZAZA* -50.1	*ZAZA* -50.1	*ZAZA* 50.1	-	-	-	-	-	*ZAZA* 80.5	-	1432.15	420.87	29.39
-	*ARRI* 66.65	*ARRI* 50.1	*ARRI* -50.1	*ARRI* -50.1	*ARRI* 50.1	-	-	-	-	-	-	*ARRI* 0.501	1604.82	267.55	16.67
-	*CAHA* 33.33	-	-	-	-	-	-	*CAHA* 23.75	*CAHA* 6.67	-	*CAHA* 8.05	*CAHA* 0.501	1030.11	72.30	7.02
-	*PANI* 66.65	*PANI* 50.1	*PANI* -50.1	*PANI* -50.1	*PANI* 50.1	-	-	*PANI* 23.75	*PANI* 6.67	-	*PANI* 40.25	*PANI* 0.501	1115.97	338.22	30.31

SAWIC, sum of absolute weighted information content; SOAWIC, Sum of observed absolute weighted information content;RID, reliability of identification; ARRI, Arristolichia . ringens; CAHA, Calliandra haematocephala; PANI, Parquetina nigrescens; SALA, Sarcocephalus latifolius; ZAZA, Zanthoxyllumzanthoxyloides. SALA with the highest SOAWIC (i.e. 890.20) was taken as the identity of the trial herb with reliability of identification (RID) being 65. 62%.

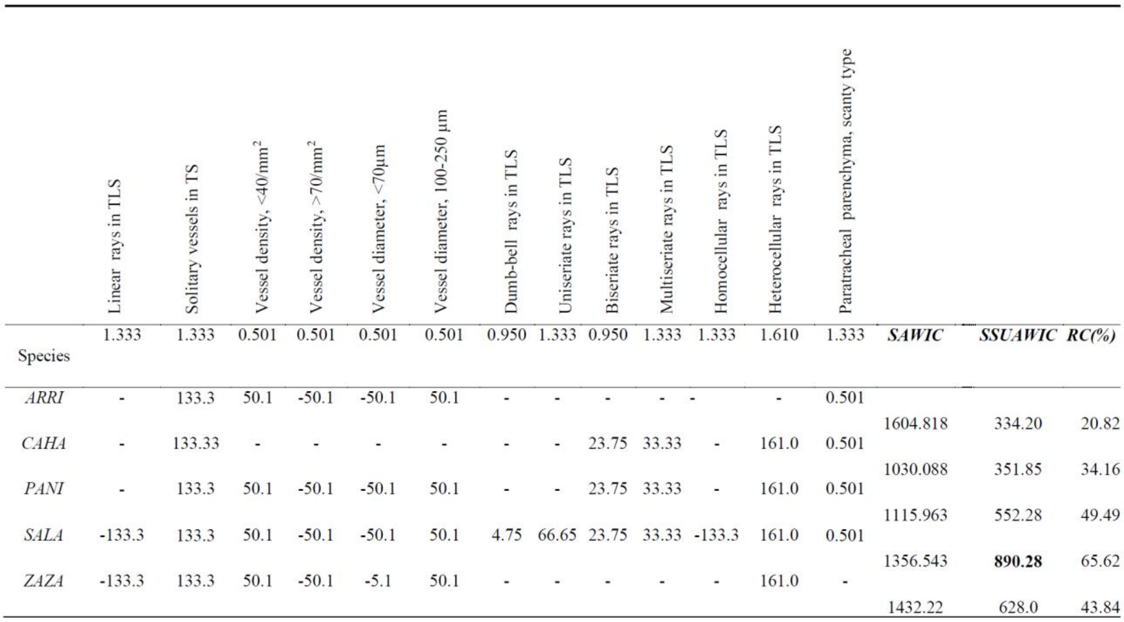
Table 3: Type II plant specimen evaluation table for a manual trial identification exercise using a ***CCIM***



	Linear rays in TLS	Solitary vessels in TS	Vessel density, <40/mm^2	Vessel density, >70/mm^2	Vessel diameter, <70µm	Vessel diameter, 100-250 µm	Dumb-bell rays in TLS	Uniseriate rays in TLS	Biseriate rays in TLS	Multiseriate rays in TLS	Homocellular rays in TLS	Heterocellular rays in TLS	Paratracheal parenchyma, scanty type			
Species	1.333	1.333	0.501	0.501	0.501	0.501	0.950	1.333	0.950	1.333	1.333	1.610	1.333	*SAWIC*	*SOAWIC*	*RID (%)*
ARRI	-	66.65	50.1	-50.1	-50.1	50.1	-	-	-	-	-	-	0.501	1604.82	267.55	16.67
CAHA	-	33.33	-	-	-	-	-	-	23.75	6.67	-	8.05	0.501	1030.11	72.30	7.02
PANI	-	66.65	50.1	-50.1	-50.1	50.1	-	-	23.75	6.67	-	40.25	0.501	1115.97	338.22	30.31
SALA	-133.3	133.3	50.1	-50.1	-50.1	50.1	4.75	66.65	23.75	33.33	-133.3	161.0	0.501	1356.55	**890.28**	**65.62**
ZAZA	-133.3	6.67	50.1	-50.1	-50.1	50.1	-	-	-	-	-	80.5	-	1432.15	420.87	29.39

SAWIC, sum of absolute weighted information content; SOAWIC, Sum of observed absolute weighted information content; RID, reliability of identification; ARRI, Arristolichia . ringens; CAHA, Calliandra haematocephala; PANI, Parquetina nigrescens; SALA, Sarcocephalus latifolius; ZAZA, Zanthoxyllumzanthoxyloides. SALA with the highest SOAWIC (i.e. 890.28) was taken as the identity of the trial herb with reliability of identification (RID) being 65. 62%.

Table 4: Plant specimen evaluation table for manual trial confirmation of Aristolociaringens, a suspected plant identity using CCIM



	Linear rays in TLS	Solitary vessels in TS	Vessel density, <40/mm^2	Vessel density, >70/mm^2	Vessel diameter, <70µm	Vessel diameter, 100-250 µm	Dumb-bell rays in TLS	Uniseriate rays in TLS	Biseriate rays in TLS	Multiseriate rays in TLS	Homocellular rays in TLS	Heterocellular rays in TLS	Paratracheal parenchyma, scanty type			
Species	1.333	1.333	0.501	0.501	0.501	0.501	0.950	1.333	0.950	1.333	1.333	1.610	1.333	***SAWIC***	***SSUAWIC***	***RC(%)***
ARRI	-	133.3	50.1	-50.1	-50.1	50.1	-	-	-	-	-	-	0.501	1604.818	334.20	20.82
CAHA	-	133.33	-	-	-	-	-	-	23.75	33.33	-	161.0	0.501	1030.088	351.85	34.16
PANI	-	133.3	50.1	-50.1	-50.1	50.1	-	-	23.75	33.33	-	161.0	0.501	1115.963	552.28	49.49
SALA	-133.3	133.3	50.1	-50.1	-50.1	50.1	4.75	66.65	23.75	33.33	-133.3	161.0	0.501	1356.543	**890.28**	65.62
ZAZA	-133.3	133.3	50.1	-50.1	-5.1	50.1	-	-	-	-	-	161.0	-	1432.22	628.0	43.84

SAWIC, sum of absolute weighted information content; SSUAWIC, sum of suspected absolute weighted information content; RC, reliability of confirmation ARRI, Arristolichia . ringens;

CAHA, Calliandra haematocephala; PANI, Parquetina nigrescens; SALA, Sarcocephalus latifolius; ZAZA, Zanthoxyllumzanthoxyloides. ARRI the suspected taxon name did not record the highest SSUAWICvalue, so the suspicion was taken as incorrect. SALA with the highest SSUAWIC value(890.28) was the correct identity.

2.7. Illustrative Execution of the Propositions

Execution of the propositions from this study was carried out in three phases: construction, manual trial applications and Microsoft excel-driven multiple trial applications of the key[44] . The output of the first phase (i.e. the constructed key) served as tool for the second (i.e. manual trial identification and identity confirmation) and third phases. Preparatory to the third phase trials, the constructed *CCIM* was reproduced on an excel spreadsheet page in which the descriptive information in the body of the key i.e. *AP*, *UP*, *AV*, *SP*, *RP*, *AB* and *NA* were replaced with their effective weights 100, 80, 50, 25, 5, -100, and 1 respectively as earlier explained. In all the phases, the same set of 13 out of the 23 wood anatomical characters obtained as earlier described were used.

At the third and final phase of execution, a trial identification was carried out for each of the five taxa in the key by asking a research assistant to select a set of data from the *CCIM* pertaining to a taxon whose identity was hidden from the key user. The data so selected were used to implement the relevant provisions of the key before the plant's identity was revealed. Trial identity confirmation was also undertaken for each of the five taxa following a two-stage process with Microsoft excel software [44]. Firstly, the 13-character data set for a known taxon 'A' in the key were deliberately used as standard for computing *SSUAWIC*, an identity confirmation index of another known taxon 'B', whose name was therefore being incorrectly suspected in the exercise. Secondly, the true data set for the incorrectly suspected taxon 'B' as recorded in the key were used for the computation as a form of ensuring correct identity suspicion and confirmation. Thereafter, the two results were compared in order to highlight the applicability of the newly designed key format for the purpose.

3. RESULTS AND ANALYSIS

3.1. Statement of the Results

The results of this study are presented in Tables 1-6. Table 1 is the wood anatomy-based calculable-contents identification matrix (*CCIM*) usable for identification of five medicinal herbs sold as roots in Ogbomoso, Nigeria. Tables 2 and 3 are the outcomes of manual trial identification exercise, indicating that the identity of the trial plant specimen was *Sarcocephalus latifolius,* with the highest sum of observed absolute weighted information content (*SOAWIC*) being 890.20 and reliability of identification (*RID*) being 65. 62%. A trial suspicion that the plant specimen earlier identified as *Sarcocephalus latifolius* was *Aristolochiaringens* yielded Table 4 as the outcome of the manual confirmation exercise. The entries in Tables 5 and 6 are the outputs of multiple trial identification and identity confirmation respectively using Microsoft excel software.

If one examines the entries in Table 4, one discovers that *A. ringens* , the suspected name did not record the highest value of *SSUAWIC*. This result indicates that the suspected plant name was incorrect, and in fact, the correct name of the plant was *Sarcocephalus. Latifolius* as earlier identified, with the highest *SSUAWIC* value of 890.28**.** Applying the same set of data used for the computation in Table 4, if instead of suspecting *A. ringens*, the user had suspected *S. latifolius,* the *SSUAWIC*of *A. ringens* would still have been 334.20 while also, that of *S.*

latifolius would still have been the highest (i.e. 890.28) and the suspicion would have been adjudged accurate (See Table 6).

Table 5: Outputs of multiple trial identification exercises conducted using Microsoft excel software

Trial identification	***SOAWIC* and *RID (%)* values**				
	ARRI	*CAHA*	*PANI*	*SALA*	*ZAZA*
ARRI	**1151.75** (71.77)	373.43 (36.45)	587.47 (52.64)	267.55 (19.72)	368.72 (25.74)
CAHA	0.50 (0.03)	**614.29** (59.64)	244.19 (21.88)	90.90 (6.70)	168.32 (11.75)
PANI	267.55 (16.67)	318.89 (30.96)	**681.89** (61.10)	296.05 (21.82)	273.72 (19.11)
SALA	267.55 (16.67)	72.29 (7.02)	338.22 (30.31)	**890.28** (65.63)	420.87 (29.39)
ZAZA	295.4 (18.41)	133.47 (12.96)	271.07 (24.29)	390.78 (28.81)	**910.98** (63.61)

ARRI, Arristolichia . ringens; CAHA, Calliandra haematocephala; PANI, Parquetina nigrescens; SALA, Sarcocephalus latifolius; ZAZA, Zanthoxyllum zanthoxyloides. SOAWIC, Sum of observed absolute weighted information content; RID, reliability of identification; the RID values are shown in parentheses, while the highest value of SOAWIC in each row identified the taxon in the row

3.2. Structure and Functionality of the Newly Designed Matrix Key System

The most frequently used tool for plant identification is the dichotomous key[45]. This is a type of single-access device which is notable for the various weaknesses earlier enumerated. Random-access or multiple-access key is an identification tool that helps to overcome some of these challenges. The matrix key system is a multi-access or free-access key format, which is associated with some notable merits: It is more flexible than single access keys in that it affords the users the freedom to decide on which characters to choose for scoring, and in which order/sequence preferred. Therefore it allows users to ignore those features that are not clear to them (i.e. unanswerable questions) and still be able to get a reliable diagnosis, or at least a short list of likely identities [2]. This account has only partly described the salient properties of the newly developed *CCIM*, with more desirable functionality attributes. Even at that, for the full potentials of this new key system to be realisable, three conditions are important for compliance: firstly, emphasis shouldbe on the use of characters that can be scored as ‘present’ or ‘absent’; secondly, for each character, a minimum or two states or conditions should apply across the taxa; and lastly, reasonably wide margins in-between the calculated *SAWIC* values of the taxa should be ensured by increasing the number of diagnostic characters as may be deemed necessary to achieve this target.

The use of matrix for identification of living organisms is a practice that has existed for some centuries [46]. The procedure follows the principle of elimination of known taxa with conflicting results on the basis of characters scored for an unknown taxon. In its strict sense, the design of matrix as identification device was not to identify a specimen, but to say what a specimen was not [47]. Its performance can be enhanced bycomputerisation so that as results of the unknown taxon are supplied in a new row of taxa, the number of matrix rows displaying all the known taxa will decrease as those with conflicting information are progressively removed from the list. Identification matrix is therefore best in narrowing down the number of possible identities of an unknown taxon to generate a short-list [48]. Discriminatory features from detailed descriptions of the taxa in question will then have to be sought [49].

Table 6: Outputs of multiple trial identity confirmation exercises conducted using Microsoft excel software

Trial suspicion	**Number of trial**	***SSUAWIC* and *RC* (%) values**					**Conclusion on suspicion**
		ARRI	*CAHA*	*PANI*	*SALA*	*ZAZA*	
ARRI	1st	334.20 (20.82)	351.88 (34.16)	552.2 (49.49)	**890.28** (65.63)	628.0 (43.85)	incorrect
	2nd	**1151.75** (71.77)	562.05 (54.56)	667.45 (59.81)	334.20 (24.64)	628.65 (43.89)	correct
CAHA	1st	295.4 (18.41)	358.14 (34.77)	463.54 (41.54)	516.34 (38.06)	**910.98** (63.61)	incorrect
	2nd	95.50 (5.95)	**614.29** (59.64)	244.19 (21.88)	130.89 (9.65)	139.96 (9.77)	correct
PANI	1st	**1151.75** (71.77)	562.05 (54.56)	667.45 (59.81)	334.20 (24.64)	495.35 (34.59)	incorrect
	2nd	267.55 (16.67)	457.74 (44.44)	**681.89** (61.10)	355.04 (26.17)	373.69 (26.09)	correct
SALA	1st	95.50 (5.95)	**614.29** (59.64)	244.19 (21.88)	130.89 (9.65)	234.97 (16.41)	incorrect
	2nd	334.20 (20.82)	190.88 (18.53)	391.28 (35.06)	**890.28** (65.63)	628.0 (43.85)	correct
ZAZA	1st	267.55 (16.67)	457.74 (44.44)	**681.89** (61.10)	224.65 (16.56)	267.05 (18.65)	incorrect
	2nd	295.4 (18.41)	358.14 (34.77)	463.54 (41.54)	516.34 (38.06)	**910.98** (63.61)	correct

*ARRI, Arristolichiaringens; CAHA, Calliandra haematocephala; PANI, Parquetina nigrescens; SALA, Sarcocephalus latifolius; ZAZA, Zanthoxyllum zanthoxyloides***.** *SSUAWIC,* sum of suspected absolute weighted information content; *RC,* reliability of confirmation (shown in parentheses); the highest value of *SSUAWIC* in the second trial row of each taxon confirmed the identity of that taxon.

Although printable free-access keys are available, they are most suitable for computer-aided identification tools such as DELTA-Intkey, Lucid, Navikeyand Xper [9]. All such computer-aided identification tools have their origin in Database Management System (DBMS) key, which is a set of attributes that help to uniquely identify a row (i.e. tuple or a taxon) in a relation (i.e. table or matrix) by a combination of one or more columns or the diagnostic characters [23]. The newly designed *CCIM* is presentable and usable in printed form, and being matrix-based, the key format possesses those desirable characteristics of multi-access keys earlier enumerated. In addition, it has the potential that its construction and navigation for plant identification, and identity confirmation can be automated. This assertion is premised upon the fact that the algorithms for these three tasks have been articulated in this paper, all of which are in compliance with the stipulated conditions of an executable computer algorithm [50]. With the widespread availability of standard spreadsheet and database programs, the number of taxa and the amount of characters applicable in this key format are limitless [28]. The *CCIM* is therefore a potential addition to the list of extant computer-aided identification tools.

3.3. Applicability of the CCIM

The argument that manual procedures for constructing and browsing the *CCIM* are tedious, boring and time consuming is valid, at least from the point of view of the propositions presented thus far. Practitioners who are also not familiar with numerable exercises can easily be put off; and these will constitute a drawback for this new key format. The cheering news however, is that these exercises are possible and practicable with *CCIM* in the first instance. The fact also remains that the desirable end results of the activities are a justification for the daunting means. This is particularly so with the possibility that these activities can be fully automated in no distant future. Plant identification and identity confirmation are not only possible, the extant of reliability of these exercises can be quantified; the latter being an in-built self-assessment mechanism. It is therefore in order to argue that the *CCIM* is the first taxonomic key format with these laudable features. There is no doubt that matrix-based keys such as the *CCIM*, will require a high initial investment in terms of research into character compilation while single-access keys require less formal investment [9]. But a craftsman is only as good as his tools; so any amount of efforts put into making a good identification key is worthwhile in taxonomy.

In using a *CCIM*, the question of acceptable level of reliability of identification (*RID*) and reliability of confirmation (*RC*) can arise. If it does arise, the relatively low magnitudes of *these* indices should not be interpreted as a limitation to the usability of the key, nor should it constitute a loss of enthusiasm in the user. As it can be deduced from equations 3 and 4, and Tables 2-4, the magnitudes of *SOAWIC*, *RID*, *SSUAWIC*, and *RC* obtained from the trials were determined by both the number of characters evaluated, and the information content of each of those characters so used for identification and confirmation exercises. In this illustration (Tables 2-4), only 13 of the 23 available characters were evaluated, with six of them having information contents below 1.00. The result could not have been the same if up to 20 or more characters had been evaluated. The relatively low values of *RID* of the identified taxa and *RC* of confirmed taxa, which ranged from 59.64% to 71.77% in both cases (Tables 5 and 6) should therefore not be discouraging.

For the sake of standardisation regarding the acceptable values of *SOAWIC*and *RID*on the one hand, and *SSUAWIC*, and *RC* on the other hand, one should be comfortable as long as there is a reasonably wide margin between the highest and the second highest of these indices. For example, after the identification of *S. latifolius* in Table 2, when the *SOAWIC* and *RID* of the other four taxa were calculated, the second highest *RID* of 30.31% was in *P. nigrescens*. Comparing this value with that of *S. latifolius* (i.e. 65.62%), one observes a comfortably wide margin of 35.31%. In event of non-acceptance a *RID* or *RC* value, probably due to their closeness in two or more taxa, the user will still be able to come up with a short list of likely identities of a specimen. Subsequently, it is desirable, as suggested by Platt (1984) [49] and Payne (1988) [25], to check the identification or confirmation of a plant against more detailed descriptions of the taxa in question, if available. This exercise will serve to provide final clearance in such situation.

3.4. Rekindling the Up-Coming Biologists' Interest in Taxonomy

The fact that there is dwindling interest in taxonomy is undeniable [18]. The number of these professionals is shrinking, to the effect that at the moment, most taxonomists are of fairly advanced age [51]. In fact many people consider the field as a dying science [52]. Therefore, the issues surrounding taxonomic practice can be summed as that of 'bountiful harvest but few labourers'. Apart from the thorny issues about the writing and application of keys in taxonomy, another factor that is believed to be contributing to declining interest in taxonomy is the ineffective way in which botany is taught [17]. Among the panacea recommended by Tilling (1987) [16] are development of appealing plant identification resources, making botany relevant to people's lives, and correct use of new teaching aids. Revitalising students' passion for taxonomy is therefore necessary to ensure the 'all-important' field of biodiversity conservation and management will not suffer neglect. The development of *CCIM* in this study is a timely response to this clarion call.

With the advent of the *CCIM*, the procedures and outcomes of key construction, plant identification and identity confirmation are quantifiable, and by implication, can be objectively evaluated for different attempts [53]. These important aspects of taxonomic practice have thus shifted pass the level of subjectivity. The contribution of this paper is therefore an empirical support of the views expressed by Wheeler and Valdecasas (2007) [54] in refuting some myths and misconceptions about taxonomy that seem to contribute to an indefensible lack of respect and support for taxonomists and their collections. It is therefore a morale booster to the upcoming biologists.

In comparison with the most widely used dichotomous key format, the *CCIM* is simple to construct, with some measure of flexibility, reliability, and effectiveness in use. Efficiency will be an added advantage with automation and development into an expert system [55]. With paper-based dichotomous and other types of taxonomic key, the discovery of a new species renders a key incomplete, a development that can be demoralising to an inexperienced user. In contrary, computerised keys are easily updated by adding information for newly discovered species and/or additional diagnostic features, and reposting computer files as appropriate [56]. For an automated *CCIM* in particular, the contents are recalculated, and necessary checks on the three requirements of workability carried out each time a change occurs. It is also interesting to note that the line of demarcation between identification and confirmation of taxa has been sufficiently narrowed with the creation of the *CCIM* to the effect that, for a given key, as identification is being conducted, identity confirmation follows suit, and *vice versa* (see Tables 5 and 6). These features of the new key format are potential encouragers and sustainers of the young minds venturing into taxonomy.

A user of the newly created *CCIM* cannot get lost as is sometimes the case in the use of dichotomous key system; no 'dead ends', and the 'momentary distractions' that can cause a user to forget his or her position in a key [10]. Even though the user may choose to enter and navigate the *CCIM* with a 'guess', the true identity of a taxon will come out provided his scoring of the plant features is not faulty. Therefore, the new key format can in addition, serve as effective tool for training in the act of species identification/confirmation. Another interesting feature of the *CCIM* is that two or more individuals could come up with correct identification/confirmation of a given taxon but with different measures of reliability. This flexibility is a function of the number of characters adopted in the key for the exercise, and the information content of each, calculable at the time of writing or re-writing the key. By design, the *CCIM* is a dynamic key system whereby the contents are re-calculable to accommodate new findings on the plant group concerned. So, working with this key format can turn out to be favourite pastime for specialists and novices alike.

4. CONCLUSION

In this study, a new multi-access key format, the calculable-contents identification matrix (*CCIM*) has been designed and proposed for use in plant taxonomy. With the features and functionality attributes of *CCIM*, the trio activities of key construction, plant identification, and plant identity confirmation are made possible through robust algorithms. These algorithms are adjudged to be in conformity with the principal features of a good/executable computer algorithm: deterministic, general, finite, and with capacity to act on at least one input to produce at least one output. The alternative key format should therefore be programmable for development into an expert system. The *CCIM* is therefore a potential addition to the list of existing computer-aided identification tools, with unprecedented features of being usable for confirming suspected plant identity and ascertaining the reliability of identification and confirmation exercises. Going by its features, applicability and potential outputs, it is clear that the *CCIM* proposed in this paper constitutes a useful template upon which reliable plant diagnostic tools can be based.

ACKNOWLEDGEMENTS

Based I thank my students at the Medicinal Plants Research Laboratory, Ladoke Akintola University of Technology, Ogbomoso, Nigeria: Mrs. Jennifer Ideh and Mr. Gideon Olaniran, for assisting to collate the data used for illustration; and appreciate Prof. Aderemi Okeyinka of Department of Computer Science, Ibrahim Babangida University, Lapai, Niger State, Nigeria, for the helpful comments and pieces of advice on the draft manuscript

.

SUPPLEMENTARY FILES

1. Microsoft excel-constructed CCIM key
2. Excel- assisted trial identification exercises using the constructed CCIM key
3. Excel-assisted trial identity suspicion/confirmation exercises using the constructed CCIM key

REFERENCES

[1] Upton R, Romm A. Guidelines for herbal medicine use, In: Botanical medicine for women's health, Romm, A. Hardy, M. L. and Mills S.(eds.), 2010; Churchill Livingstone, p.75-96. Available at https:// doi.org/10.1016/B978-0-443-07277-2.00004-0.

[2] Amante V D, Norton GA. Developing interactive diagnostic support tools for tropical root crops, Centre for Biological Information Technology, 2003; The University of Queensland, Brisbane Qld 4072, Australia.

[3] Bock JH, Norris DO. Forensic plant taxonomy, In: Forensic plant science, Bock, JH. and Norris, DO, 2016; Academic Press, 2016; p.95-101. Available at https://doi.org/10.1016/C2013-0-19012-5.

[4] Dubey S, Garg M, Sharma K. Bioresources, biodiversity and eco-management. Journal of Experimental Sciences, 2011; 2(10): 62-63. Available at: http://jexpsciences.com/

[5] Panter KE, Welch KD, Gardner DR. Poisonous Plants: biomarkers for diagnosis. In: Biomarkers in Toxicology, Gupta, R. C. (ed.), 2014; Elsevier, Oxford, UK. p. 563-589. https://doi.org/10.1016/C2012-0-01373-7.

[6] Randler C. Teaching species identification- a prerequisite for learning biodiversity and understanding ecology. Eurasia Journal of Mathematics, Science and Technology Education 2008; 4(3): 223-231. Available at: Doi: 10.12973/ejmste/75344.

[7] Jacquemart AL, Lhoir P, Binard F, Descamps C. An interactive multimedia dichotomous key for teaching plant identification. Journal of Biological Education 2016; 50 (4): 442-451. Available at: https://doi.org/10.1080/00219266.2016.1150870.

[8] Lobanov AL. Keys to beetles and biological diagnostics, 2003; Available at: http://www.zin.ru/Animalia/Coleoptera/eng/syst8.htm.

[9] Hagedorn G, Rambold G, Martellos S. Types of identification keys. In: Tools for identifying biodiversity; progress and problems. Nimis PL and Vignes LR (eds.), 2010; EUT, p. 59-64.

[10] Walter DE, Winterton S. Keys and the crisis in taxonomy: extinction or reinvention?. Annual Review of Entomology, 2007; 52: 193-208. Available at: doi:10.1146/annurev.ent.51.110104.151054.

[11] Yin S, Lin X, Liu L, Wei S. Exploiting parallelism of imperfect nested loops on Course-Grained Reconfigurable Architectures. IEEE Transactions on Parallel and Distributed Systems, 2016; 27: 3199- 3213. Available at: Doi: 10.1109/TPDS.2016.2531678.S

[12] Gonzalez JA, Carvalho AM, Vallejo JR, Amich F. Plant-based remedies for wolf bites and rituals against wolves in the Iberian Peninsula: Therapeutic opportunities and cultural values for the conservation of biocultural diversity. Journal of Ethnopharmacology, 2017; 14 (209):124-139. Available at: doi: 10.1016/j.jep.2017.07.038.

[13] Arora NK. Environmental sustainability—necessary for survival. Environmental Sustainability, 2018; 1:1–2. Available at: https://doi.org/10.1007/s42398-018-0013-3

[14] Dukes G . The law and ethics of the pharmaceutical industry, Elsevier Science Publisher, B.V., 2006; The Netherlands. Available at: https://doi.org/10.1016/B978-0-444-51868-2x500-4.

[15] Noble Research Institute.Plant Identification: Is It Worth the Effort?, 2001; Available at: https://www.noble.org/news/publications/ag-news-and-views/2001/may/plant-identification-is-it-worth-the-effort/

[16] Tilling SM. Education and taxonomy: the role of the field studies council and AIDGAP. Biological Journal of the Linnaean Society, 1987; 32(1): 87-96. Available at: https://doi.org/10.1111/j.1095-8312.1987.tb00414.x

[17] Stagg BC, Donkin MC. Teaching botanical identification to adults: Experience of the UK participatory science project open air laboratories. Journal of Biological Education, 2013; 47(2): 104-110. Available at: doi: 10.1080/00219266.2013.764341.

[18] Drew LW. Are we losing the science of taxonomy?.BioScience, 2011; 61(12): 942–946.Available at: doi:10.1525/bio.2011.61.12.4

[19] IaconoS, Kling R. Computerization movements and tales of technological utopianism. In: Computerization and controversy 2nd Edition, Kling R.(ed.), 1996; Academic Press, Inc. p.85-105.

[20] Farr DF. On-line Keys: More than Just Paper on the Web, Taxon, 2006; 56: 589-596.https://doi.org/10.2307/25065636.

[21] Pankhurst RJ. A computer program for generating diagnostic keys. The Computer Journal, 1970; 13 (2): 145-151. Available at: https://doi.org/10.1093/comjnl/13.2.145.

[22] TofilskiA (2018). DKey software for editing and browsing dichotomous keys. Zookeys, 2018;735: 131-140. Available at: doi: 10.3897/zookeys.785.21412.

[23] Javatpoint . DBMS keys, Database management systems, 2018. Available at: www.javatpoint.com

[24] Payne RW. Selection criteria for the construction of efficient diagnostic keys. Journal of Statistical Planning and Inference, 1981; 5 (1): 27-36. Available at: https://doi.org/10.1016/0378-3758(81)90048-3

[25] Payne RW. Identification keys, diagnostic tables and expert systems. In: Edwards D and Raun NE (eds.), 1988; Compstat, Physica-Verlag Heidelberg. Available at: https://doi.org/10.1007/978-3-642-46900-8_27

[26] Morse LE. Specimen identification and key construction with time -sharing computers. Taxon, 1971; 29: 269-282. Available at: https://doi.org/10.2307/1218880.

[27] Bell NL. A Computerized Identification key for 30 Genera of plant parasitic nematodes. New Zealand Plant Protection, 2002; 55: 287-290. DOI: 10.30843/nzpp.2002.55.3954.

[28] Godfray HCJ, Clark BR, Kitching IJ, Mayo SJ, Scoble MJ. The web and the structure of taxonomy. Systematic Biology, 2007; 56 (6): 943-955. Available at: https://doi.org/10.1080/10635150701777521.

[29] Clark B, Charles H, Godfray J, Kitching IJ, Mayo SJ, Scoble MJ. Taxonomy as an eScience, Philosophical Transactions of the Royal Society A, 2008; 367: 953-966. Doi:10.1098/rsta.2008.0190.

[30] Sinha VK. Introduction to Matrix Theory, 2015; Alpha Sciences International Ltd., Oxford, UK.

[31] Adams RP. Statistical character weighting and similarity stability. Brittonia, 1975; 27: 305-316. Available at: https://doi.org/10.2307/2805510.

[32] Vogt L. Testing and weighting characters, Organisms Diversity and Evolution, 2002; 2(4): 319-333. Available at: https://doi.org/10.1078/1439-6092-00051.

[33] Bell CR. Taxonomic characteristics, In: Plant variation and classification. Fundamentals of botany series, 1967; Palgrave, London. Available at: https://doi.org/10.1007/978-1-349-00430-0_6

[34] Bisby FA. The evaluation and selection of characters in angiosperm taxonomy: an example from *Chrotalaria*. New Phytologist, 1970; 69: 1149-1160. Available at: https://doi.org/10.1111/j.1469-8137.1970.tb02495.x

[35] Schoch W, Heller I, Schweingruber FH, Kienast F. Wood anatomy of central European Species. Online version, 2004; Available at: www.woodanatomy.ch

[36] Mota GD, Melo LED, Ribeiro AD, Selvati AO, Pereira H, Mori FA. Ecologic features of wood anatomy of *Caseariasylvestris* SW (Salicaceae) in three Brazilian ecosystems. CERNE, 2017; 23(4). Lavras Available at: https://doi.org/10.1590/01047760201723042387

[37] Arx GV, Crivellaro A, Prendin AL, Cufer K, Carrer, M (2016). Quantitative wood anatomy-practical guidelines. Frontiers in Plant Science 7: 781. Available at: Doi: 10.3389/fpls.2016.00781

[38] De ParniaNE, Miller RB. Adapting the IAWA list of microscopic features for hardwood identification to DELTA. IAWA Bulletin n.s., 1991; 12: 34-50.

[39] Liu Y, Zhou L, Shu Y, Liu S. Anatomical Features and Its Radial Variations among DifferentCatalpabungeiClones. Forests, 2020; 11(824): 1-17 Available at: Doi:10.3390/f11080824

[40] IAWA Committee. International Association of Wood Anatomists (IAWA) list of microscopic features for hardwood identification. IAWA Bulletin. n.s., 1989; 10: 219-332.

[41] Ogunkunle ATJ., Ojo OD, Oni OM. Anatomy and specific gravity of wood samples from six Nigerian tree species in relation to their diagnostic X-ray shielding capabilities. Journal of Natural Sciences Research, 2014; 4(10): 70-77

[42] Freudenstein JV. Characters, States and Homology. Systematic Biology, 2005; 54(6): 965–973. Available at: https://doi.org/10.1080/10635150500354654

[43] Spiegel MR. Theory and problems of statistics: Schaum's outline series, second edition, 1992; McGraw-Hill Book Company, London. p434-463.

[44] Levine DM, Stephan DF, Krehbiel TYC, Berenson ML. Statistics for managers using Microsoft excel, 2008; Prentice Hall, New Jersey 07458.

[45] Sinh NV, Wiemers M, Seettele J. Proposal for an index to evaluate dichotomous keys. Zookeys, 2017; 685: 83-89. Available at:doi: 10.3897/zookeys.685.13625.

[46] Priest FG, Alexander B. A frequency matrix for probabilistic identification of some Bacilli. Journal of General Microbiology, 1988; 134: 3011-3018.

[47] Holmes B, Pinning CA, Dawson CA. A Probability Matrix for the Identification of Gram-negative, Aerobic, Non-fermentative Bacteria that Grow on Nutrient Agar. Journal of General Microbiology, 1986; 132: 1827-1842.

[48] On SLW, Holmes B, Sackin MJ. A Probability Matrix for Identification of Campylobacters, Helicobacters and Allied Taxa. Journal of Applied Microbiology, 1996; 81 (4): 425-432. https;//doi.org/10.1111/j.1365-2672.1996.tb03529.x

[49] Platt HM. (1984). Pictorial taxonomic keys: their construction and use for the identification of free-living marine nematodes. Cahiers De Biologie Marine Tome, 1984; XXV: 83-91.

[50] Okeyinka AE. Introduction to Computer Technology-Lectures in Computer Science Series, 1998; Department of Computer Science and Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria, 68 pp.

[51] Bik HM. Let's rise up to unite taxonomy and technology. PLoS Biol, 2017; 15(8): e2002231. Available at: https://doi.org/10.1371/journal.pbio.2002231

[52] Wagele H, Klussmann-Kolb A, Kuhlmann M, Haszprunar G, Lindberg D, KochA,Wägele JW. The taxonomist - an endangered race, a practical proposal for its survival. Frontiers in Zool., 2011; 8: 25. Available at: https://doi: 10.1186/1742-9994-8-25.

[53] Lavinsky D. The two most important quotes in business. GrowThink, 2020; Availableat: https://www.growthink.com/content/two-most-important-quotes-business.

[54] Wheeler QD, Valdecasas AG. Taxonomy: myths and misconceptions. Anales del JardínBotanico de Madrid, 2007; 64(2): 237-241.

[55] ArabChadegani R, ArabChadegani Z. A review on expert systems and their usage in management. Advances in Environmental Biology, 2013; 7(8): 1460-146.

[56] Dallwitz MJ, Paine TA, Zurcher EJ. Principles of interactive keys delta-intkey.com/, 2018; Available at: https://www.delta-intkey.com/www/interactivekeys.htm