A Data Mining Process In An Indigenous Knowledge Ethno Medicinal Database

Audrey Masizana, Gobusamang Oscar, Harriet Okatch, Barbara N. Ngwenya, Keitseng N. Monyatsi, Mbaki Muzila, Kerstin Andrae-Marobela

Abstract: Botswana is a landlocked country with diverse ethnic groups amounting to a population of around 2 million. Batswana like all the other African nations have a strong sense of culture which is expressed and strongly felt through the language, Setswana, traditional food, traditional healing and the music which form part of various indigenous knowledge systems (IKS). However, the way indigenous knowledge is or should be documented properly is subject of intense debate. Hence research projects are coming up with various methods and tools to contribute to the documentation. The Ethnosurvey Research project at University of Botswana's Centre for Scientific Research Indigenous Knowledge and Innovation (CesrIKi) set out to contribute to IKS documentation by collecting and documenting the country's traditional medicinal plants in Botswana. This paper presents the findings of conducting a data mining process on the collected data to uncover patterns and trends emerging from the data through a data mining technique.

Index Terms: Apriori, Association, Confidence, Data Mining, Indigenous Knowledge Systems, Medicinal Plants, Support, Weka.

1 INTRODUCTION

THIS research project is motivated by the University of Botswana Centre for Scientific Research Indigenous Knowledge and Innovation (CESRIKI). CESRIKI is an interdisciplinary and science-focused center of study at the University of Botswana. The primary objective of the center is to be an intellectual and cultural centre that draws upon the nation's indigenous knowledge base to promote Botswana's social and cultural heritage. Currently there is inadequate documentation of indigenous knowledge in Botswana, however there are some efforts made by the Botswana National Museum and other archiving associations like libraries.

- *Dr. Audrey Masizana, Phone: 00267 355 2131, Fax: 00267 355 4465, e-mail: masizana@mopipi.ub.bw, Center for Scientific Research, Indigenous Knowledge and Innovation (CesrIKi)*
- *Department of Computer Science, University of Botswana, Gaborone, Botswana, Block 235/Office 251, University of Botswana*
- *Oscar Gobusamang (oscar.gobusamang@gmail.com) Department of Computer Science, University of Botswana, Gaborone, Botswana*
- *Dr. Kerstin Andrae-Marobela (marobelak@mopipi.ub.bw) Center for Scientific Research, Indigenous Knowledge and Innovation (CesrIKi), University of Botswana, PO Box 70237, Gaborone, Botswana*
- *Dr. Harriet Okatch (okatchh@mopipi.ub.bw) Center for Scientific Research, Indigenous Knowledge and Innovation (CesrIKi), University of Botswana, PO Box 70237, Gaborone, Botswana*
- *Dr. Barabara Ngwenya (bntombi@mopipi.ub.bw) Okavango Research Institute (ORI), University of Botswana Maun Campus, P/Bag 285, Maun, Botswana*
- *Center for Scientific Research, Indigenous Knowledge and Innovation (CesrIKi), University of Botswana, PO Box 70237, Gaborone, Botswana*
- *Mbaki Muzila (muzilam@mopipi.ub.bw), University of Botswana Maun Campus, P/Bag 285, Maun, Botswana*
- *Keitseng N. Monyatsi (kmonyatsi@aripo.org) African Regional Industrial Property Organization (ARIPO), PO Box 4228 Natal Road, Harare, Zimbabwe*

CESRIKI's existence intends to contribute and promote documentation of IKs in various forms. One such initiative is through the ethno survey research that CESRIKI has established which was funded by UNDP in 2007, to document medicinal plants in Botswana. This activity gave CESRIKI an opportunity to interact with communities in six different regions in Botswana, being North East Botswana Central Botswana, Kgalagadi, Mmankgodi, Ghanzi and Maun. The result of this survey was the development of a database containing knowledge captured on the plants and their medicinal uses together with relevant information. However the ethnosurvey project generated diverse and large data sets from different districts and ecological zones in which different traditional healers practice in the country. There were inherent inconsistencies, overlaps and redundancies in the data. In addition, large quantities of data were collected which made it difficult to manually analyze and reconcile these anomalies/inconsistencies. There was a need therefore to conduct a comprehensive analysis of the patterns and trends emerging from the data as a foundation for improving data quality and credibility developments in future and also improve on the technical representation of the data. This paper applies a data mining technical approach to identify and analyze emerging patterns and trends in order to improve understanding of the ethno-survey data. The data mining, as a concept is generally applied to uncover patterns of data from large data samples. The aim is to

1) Discover patterns emerging from the database in order to make informed decisions about the ethno medicinal data as the database grows phenomenally. The process is expected to produce results that reveal commonalities, that reduce anomalies and inconsistencies, and that improve usefulness and accuracy of the CesrIKi indigenous medicinal plants database.

2) Secondly, it is aimed to indicate better pointers for further research and development of other types of IKS databases.

The data mining process is conducted through two approaches;

a) Data mining using manual searches. Using existing search methods within the DBMS (Database Management System) and sieve through the data

to get any visible trend that might be of use in achieving the set objectives.

b) Data Mining using computing algorithms. The reason for the second objective is to improve on the first objective which can be limited in terms of human capability and human error when going through large chunks of data manually.

2 INDIGENOUS KNOWLEDGE OF MEDICINAL PLANTS

Botswana, like many African countries has a wealth of Indigenous Knowledge (IK), and has taken a major step forward in the recognition of the legitimacy of systems of knowledge and reflects a growing shift in consciousness among academics, policy-makers and practitioners in a number of fields as to the value and the legitimacy of Indigenous Knowledge Systems. Such steps include the current involvement of CESRIKI in documenting Indigenous Knowledge. Indigenous Knowledge as defined in [1] is variably referred to as traditional knowledge, local knowledge or traditional environmental knowledge that an indigenous (local) community accumulates over generations of living in a particular environment. Indigenous Knowledge as described in [2] is vital information which is diminishing at an alarming rate and as such there is an urgent need to collect it before it is irretrievably lost. In terms of medicinal plant use in Botswana the knowledge holders are traditional doctors '*Dingaka(plural*) or ngaka(singular) *tsa Setswana'* who practice *Bongaka jwa Setswana,* a form of diagnosis and prescription of traditional medicine. This dates as far back as the origin of Batswana and it is passed from one generation to the next. Hedberg and Staugård [3], in their compilation of plants that are used as traditional medicine in Botswana, made an observation that about 2/3 of Batswana have undergone traditional healing or have used traditional medicines at some point in their lives. Traditional doctors are consulted for a wide variety of reasons, ranging from sexually related diseases, lack of child bearing, Aids, tuberculosis, sore feet, broken body parts, and a lot more *Setswana* culture related diseases. Indigenous Knowledge has been researched around the world and some surveys are as follows: An Ethno medicinal survey and documentation of medicinal folklore food of Samahni valley [4] showing the need and importance of Indigenous Knowledge systems. An ethno botanical analysis was conducted in order to document the traditional medicinal uses of plants, particularly medicinally important folklore food of Samahni valley, Azad Kashmir (Pakistan). Another example is an Ethno botanical survey in Canhane village, district of Massingir, Mozambique which was focusing on medicinal plants and traditional knowledge. This study in [5] shows the social importance of the floristic richness in the Canhane village, particularly regarding the significance of medicinal plants in primary healthcare. This is reflected in the great diversity of plants used for medical purposes as well as in the wide range of their applications ancd associated procedures. An investigation on South African plants as a source of new antimalarial drugs [6] has been researched on. Based on the historical success of natural products as antimalarial drugs and the urgent need for new antimalarials, a number of South African medicinal plants have been evaluated for their antimalarial properties. This paper reviews the major studies conducted and their findings. Also Research has been done on the Ethno medicinal plant knowledge and practice of the Oromo ethnic group in southwestern Ethiopia, [7]**.** A study was conducted to document the indigenous medicinal plant knowledge and use by traditional healers in south western Ethiopia. The research finding shows that ethno medicinal plant species used by healers are under serious threat due to several factors, which indicates the need for urgent attention towards their conservation and sustainable utilization through Indigenous Knowledge Systems.The following shows a sample of Indigenous Knowledge System of medicinal plants. The Prelude Medicinal Plants database is reported by [8] to have been initially created at the Catholic University of Louvain (Louvain-la-Neuve, Belgium) but in order to insure better conservation and technical management it was transferred to the Royal Museum for Central Africa on the Metafro Web site. It contains medicinal plants, scientific names, its family, symptoms and other attributes. Another contribution in [9] shows a database with a variety of Amazon rainforest plants, their main medicinal properties, and in what form they are ingested as well as their medicinal properties. Indigenous knowledge data is collected and documented for knowledge reference and future research which can be of importance to the country and communities at large. Lesser data and experiences exist on how and in what form information is extracted from the data bases. Therefore this paper describes patterns and trends emerging from a data-mining process using the CesrIKi medicinal plant database.

3 DATA MINING IN MEDICINAL ENVIRONMENTS

Data mining is the analysis of data to discover previously unknown relationships that provide useful information. Another definition of Data Mining by [10] is that it is the extraction of interesting (non-trivial, implicit, previously unknown and potentially useful) patterns or knowledge from huge amount of data. Data mining is applied in different fields like: Pharmaceuticals, Healthcare, Banking, Telemarketing and direct marketing, Airlines, Manufacturers, Insurance companies, Telecommunications, Distribution and retailing, Remotely sensed data etc. Given databases of sufficient size and quality, data mining technology can generate new business opportunities by providing these capabilities:

a. Automated prediction of trends and behavithes
b. Automated discovery of previously unknown patterns

Han, Kamber and Pei [11] Suggest that data mining techniques fall into the following major categories:

3.1 Classification

Given a set of data elements classification maps each data element to one of a set of pre-determined classes based on the difference among data elements belonging to different classes. The goal is to discover rules that define whether an item belongs to a particular subset or class of data.

3.2 Clustering

A cluster is a collection of data objects. Clustering is unsupervised classification. A good clustering method produces high quality clusters with high intraclass similarity and low interclass similarity

3.3 Sequencing

Sequence algorithm finds sequential patterns of certain events, e.g. Web page hits, and groups similar patterns together, e.g. all visitors to a Web site exhibiting similar browsing behavithe will form a separate group. For instance, the Microsoft Sequencing algorithm finds the most common sequences by grouping identical sequences together [12].

3.4 Association

Association techniques search all transactions from a system for patterns of associated occurrences. That is, this technique attempts to find groups of items that are commonly found together. A common method is *market basket analysis*, in which the set of products purchased by thousands of consumers are examined. It discovers what items are usually purchased with others, predicting the frequency with which certain items are purchased at the same time. Results can be portrayed as percentages; for example, "30% of the people that buy steaks also buy charcoal"

4 Data Mining in Medical Environments

Data mining techniques have been used in medical research for many years and have been known to be effective in extracting information from medicinal data [13]. A Hybrid Data Mining Method has helped, for example, with a medical classification of chest pain. A hybrid methodology that combined data mining techniques (such as association rules and classification trees) was used in order to solve problems faced by emergency departments such as long-waiting time, congestion, and delayed patient care. The methodology has also been applied to analyze emergency data collected from a hospital. The results are expected to help physicians make more accurate classification of chest pain diseases. Abdullah, Ahmad and Ahmed [14] demonstrate that association techniques can be used to yield association rules in medical billing data. Their motivation in this research work is to uncover patterns of resemblance between medical bill and purchase bills. The Association technique in particular, has been applied through the apriori algorithm, in numerous medical projects. A research by [15] studies how data mining techniques are used for the data analysis and knowledge discovery in medical sciences using the apriori algorithm and a self developed algorithm. The author claims to have used realistic values from a medical database which makes their output reliable, efficient and precise for decision making. Specifically results may be used for analyzing and improving performance by the medical staff. In another research [16], association rules are used in a healthcare system to conduct an intelligent diagnosis and extract information and build important knowledge bases quickly and automatically. The authors are looking specifically at trends and patterns of diabetic data repositories by applying the apriori algorithm to a database containing records of diabetic patients. In the paper it is noted the results are good enough to conclude that the methodology followed is of good value to the diagnostic procedure, especially when large data volumes are used. Hence this paper applies the Apriori Algorithm approach to ethno-survey data of medicinal plants from 6 administrative districts in Botswana

5 Conducting The Research

The methodology used is a quantitative method of research which required the researchers to do surveys that involved consultative workshops to introduce the project to communities, followed by interviews for generating data. The research required collection of data in large quantities as it can only be feasible with large populations of data. So far 6 regions have been visited, North East Botswana, Central Botswana, Kgalagadi, Mmankgodi, Ghanzi and Maun. In all these regions 80 practitioners have been consulted and interviewed which makes it around 20 practitioners and community representatives involved per region on average. They collectively gave data on 1897 medicinal plants. Some of the data includes 828 records about preparatory methods and 322 records about symptoms. Quantitative research helps in identifying cause and effect relationships, testing the hypotheses and examining different relationships. The diagram below shows statistics about the research data.

Table 1: Data Statistics

Attribute	No. of Records
Regions	6
Practitioners	80
Medicinal Plants	1897
Medicinal Uses	1670
Symptoms	322
Preparatory methods	828

6 Proposed Solution

The research proposes to implement manual searches as part of the data mining process followed by an application of the Apriori Algorithm whose purpose is to augment manual data mining. The starting point is to look at the structure of the database itself which has been nicknamed 'Dingaka Database'. The database is set out to collect data on plants, medicinal uses, symptoms, researcher, ngaka/traditional doctors, sample, dosage, symptom preparatory methods, side effects etc, as shown in Fig.8 below.

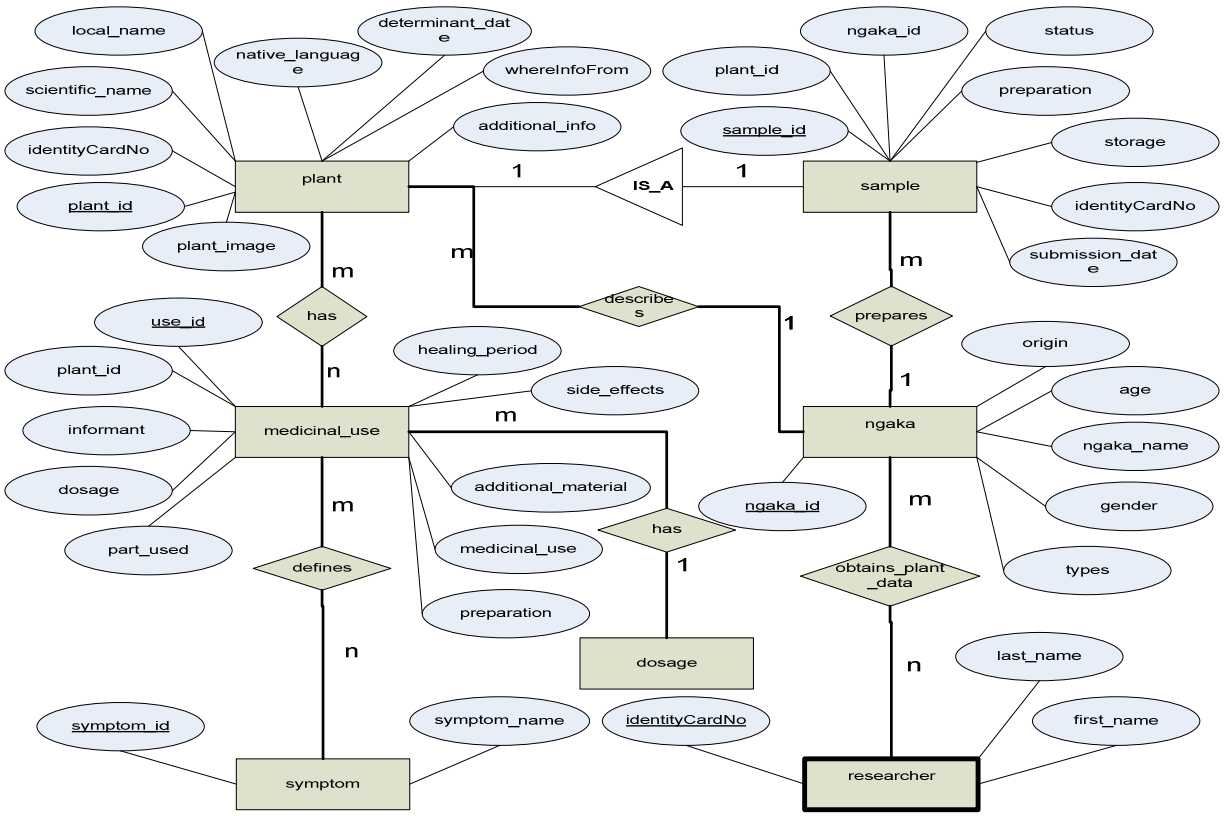

Fig.1. Technical Representation of ethno medicinal data: Entity Relational Diagram

6.1 Application of Manual Searches

Various search combinations of key features of the data objects within the database were constructed manually in order to create meaningful output that would communicate a pattern or a trend form the data. Hence on the left hand side of Table 2 you will see these combinations and on the right hand side you will see the reasons.

Table 2. Sample Application Of Manual Searches

Search Combination	Possible Reason
Plant, Plant Part, Medicinal Use, Symptom, Preparatory Method/Dosage, Practitioner	To get individual statistics of the entities and to derive the most popular occurrences among the attributes.
Plant against medicinal use.	To see why it is that some plants are used for several usages? What medicinal plants are associated with each other?
Medicinal use against plants.	What diseases are associated with each other, and what associates them? How do medicinal uses associate with plants used?
Plant against Preparatory method	Why is it that a group of plants are prepared in a similar way?
Plant against Plant parts against medicinal use	Which is the most used medicinal plant parts and for what medicinal use?
Medicinal Use against Plant against Practitioner	To see the relationship of the plants for the same medicinal use by different practitioners.
Plant against Medicinal Use against Symptom	What medicinal use relates to a particular Symptom

To implement this process, a structured PHP script was used to run through the large data sets. The script helped to manually search and get plant name against medicinal use and medicinal use against plant name results from database. The search was designed to extract both plant names and medicinal uses and then take a plant name against medicinal uses, and give the list of medicinal uses for a given plant. Then furthermore extract medicinal uses against plant names and give a list of medicinal plants for a given disease. For instance Mophane medicinal plant shows its medicinal uses as high blood pressure, setlhabi, heart diseases, sores in the stomach to the throat and diarrhea while a medicinal use like Heart Diseases shows plants as moretologa, motlalemetsi, motsididi, mophane and xaoji. (For sample results refer to Table 3 and 4).

Table 3. Manual Search; Plant Name Against Medicinal Use

Plant Name	Disease
1.pelobotlhoko	miscarriage
2.sefaphabadimo	Popelo, cancer
3.setlhare sa kgwedi	lack of menstrual cycle
4. mosokelatsebeng + motoroko	sugar diabetes
5.tlhotlhamadi	hydration & kidney failure
6. moraro	persistant cough , Tuberculosis, gonorrhoea, bladder and kidney problems, diphate
7.mophane	high blood pressure, setlhabi, heart diseases, sores in the stomach to the throat, diarhoea
8.monepenepe	high blood pressure, setlhabi, snake bite, witchcraft, swollen legs

+ Means combined with

Table 4. Manual Search; Medicinal Use Against Plant Name

Disease	Plant Name
1.Tubercolosis, gonorrhoea, bladder and kidney problems, diphate	moraro
2. sugar diabetes	mosokelatsebeng + motoroko
3.Blood Purification, Stroke	mosukujane, legala-la-tshwene
4. Asthma	modumela, mosetlha, mosotlho
5.Mental Illness	motsididi, mphaphama, mokunyo, tokitsone
6. liver problems	sebetjana, sekaname, morolwana, monnamontsho, makanyana, qae, mogatana
7. heart diseases	moretologa, motlalemetsi, motsididi, mophane , xaoji
8.STDs	Makgonatsotlhe, thothamadi, sekaname , mositsana , monamani , morolwana , masigomabe , motalemetsi , borumolane , motsitsane , kgalemela, moarasupe, mmatlakgomo, phekolola, mabeleapodi

Manual Search proved to be a tedious process but thorough hence the importance of objective 2, Data Mining through computing algorithms to technically analyze the database. Below are a few manual statistics which as few as they are proves how demanding is manually going through the database.

Table 5. Sample Manual Statistics

Attribute	Attribute Item	No. of records
1.Plant Name	Mabeleapodi	2
	Moarasupe	6
2.Medicinal use/Disease	Nutrition	85
	Snake Bites	4
3.Dosage	Boil	400
	Mix	100
	Boil Roots	75

6.2 Application of Association Apriori to Sample Ethno Database

The Apriori Algorithm performs the searches just like the above through establishing associated classes of data sets based on similarities or commonalities. . It uses “prior” knowledge of the dataset to generate the candidate set. It employs two approaches; Support shows the frequency of the patterns in the rule, it is the percentage of transactions or item sets that contain both A and B.Support = (# of transactions involving A and B) / (total number of transactions). Confidence is the strength of implication of a rule, it is the percentage of transactions that contain B if they contain A. Confidence = (# of transactions involving A and B) / (total number of transactions that have A). It first calculates the support of items in the item sets, then the confidence. A minimum support, and minimum confidence threshold were set to qualify item sets for credibility as illustrated in Fig.2 below.

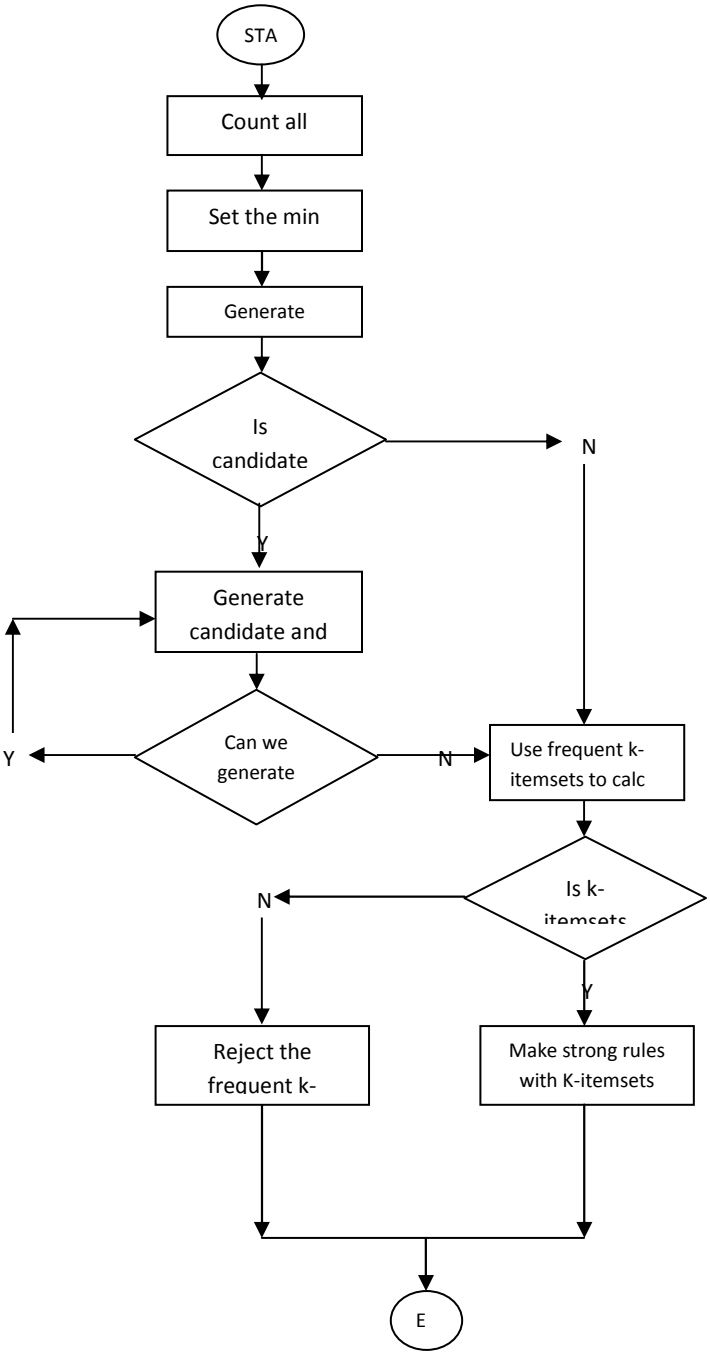

Fig.2. The Apriori process chart

The Apriori algorithm is implemented through a technological application called Weka which is an open source data mining software. Weka supports machine learning algorithms for data mining techniques like classification, regression, clustering, and association rules [17]. The whole suite is written in java, so it can be run on most platforms. Data is loaded from a Weka compatible Attribute relational file but it can also be extracted from a CSV file format or extracted from a database using an SQL query. The following figure shows a typical Weka window with data mining techniques.

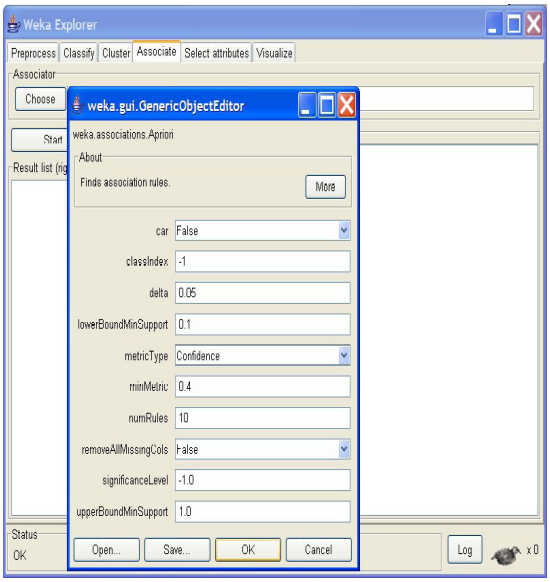

Fig.3. Typical Weka window

The result produced 622 rules/items sets Apriori was set to 0.1 minimum support which is 10% then the Confidence was set to 0.9(90%) threshold respectively. **Table 5** shows sample results of 10 randomly picked data sets through association using Apriori.

Table 6. Association (Apriori) Best Rules

Rule NO.	RESULTING RULE/622	RULE INTERPRETATION	Confi-dence	CONCLUSION
5	plant part=roots additional material=blood of kgage signs/symptoms=lump around anus TD/CM=TD ==> dosage&prep cont=3 times a day	To test the relationship of plant part AND additional material AND signs/symptoms AND TD/CM TO dosage&prep	1.0	This rule qualified for both the support minimum and confidence threshold and EVERYTIME the plant part is roots AND additional material is blood of kgage AND data source is a Traditional doctor **THEN** dosage&prep cont is 3 times a day
17	healing period=< 1 week TD/CM=CM ==> additional material=blood of kgage	To test the relationship of healing period AND TD/CM TO additional material	1.0	This rule qualified and it says EVERYTIME the healing period=< 1 week AND data source is a Traditional doctor **THEN** additional material is blood of kgage
164	dosage & prep=boil and drink half a cup 3 times a day ==> dosage&prep cont=3 times a day	To test the relationship of dosage & prep TO dosage&prep count	1.0	This rule qualified and it says EVERYTIME when the dosage & prep is boil and drink half a cup 3 times a day **THEN** the dosage&prep cont is 3 times a day.
190	plant part=roots dosage&prep cont=3 times a day side effects=none healing period=1-2 weeks ==> TD/CM=TD	To test the relationship of plant part AND dosage&prep count AND side effects AND healing period TO TD/CM	1.0	This means rule 196 qualified and EVERYTIME there was NO side effects and the healing period=< 1 week there was NO additional info and NO storage&prep.
204	name of practitioner=Sentsima Gaofenngwe ==> side effects=none	To test the relationship of name of practitioner TO side effects	1.0	The rule qualified and EVERYTIME the part used was roots there was NO additional info.
621	plant part=roots healing period=< 1 week TD/CM=TD ==> side effects=none	To test the relationship of plant part AND period AND TD/CM TO side effects	0.9	This means rule 621 qualified for both the support minimum and confidence threshold and most of the time the plant part used was roots and healing period was =< a week and data source was a Traditional doctor there were no side effects.

7 Discussion of Results

The first objective which is Manual data mining using manual searches was met and it helped in analysing data by going thoroughly through the data manually and several trends were visible but not enough to make concrete decisions. The second objective, Data Mining through Apriori was met and it helped to improve on the first objective, as it managed to handle the large amount of data within the database and some of the human limitations. Patterns and trends were found, some inconsistencies in the data collected came into the open. This allowed for technically studying and analyzing the data furthermore. The study will contribute to the general knowledge of ethno survey data and hence to Indigenous knowledge by showing extensively relations in medicinal plants and other attributes.

7.1 Rising Questions on Validity of Data

For example, rule 204 in Table 5 above, says EVERYTIME the name of practitioner is Sentsima Gaofenngwe there are NO side effects. One can ask themselves how is it the mentioned practitioner provides/practices with only those medicinal plants without side effects. Does it necessarily mean he is very good at what he does, or he can be wrong in some cases or he only provided a few medicinal plant data which happen not to have side effects? This can be a debate to be discussed by other practitioners and stakeholders about one rule while contributing more to general Indigenous knowledge.

7.2 Contribution To Further Development of The Database-Default Data

After Manual data mining and data mining using Apriori, the discoveries show an influence in the future development of the database by informing researchers, the expert system and refining the development of the database. Looking at Rule 5 which says EVERYTIME the plant part is roots AND additional material is blood of kgage AND data source is a Traditional doctor THEN dosage & prep cont is 3. When populating more data from other regions into the database, and for a particular plant the practitioner somehow does not include the dosage and all the variables are as follows: Plant part is roots AND additional material is blood of kgage AND data source is a Traditional doctor then one can safely fill in the space for dosage as 3 times. The same can be done with all other attributes that have been qualified by the Association apriori algorithm. All these additions from data collectors and practitioners will help in further development for the database developer to create the attributes that are very vital in cases where other attributes do not have enough data that can allow for decision making. This would make it more sensible to apply data mining using apriori to the ethno medicinal data.

7.3 Frequent Occurrences /Associations

A discovery of relationships and trends of data elements as they occur frequently could be observed. An example is seen through Sengaparile, that most times a patient is having womb problems or delivery problems Sengaparile is one of the options.

7.4 Missing Data

Some attribute data is yet to be collected in terms of data population. These will continue to remind the developers of the gaps waiting to be filled. Further research questions can be thrown back to the practitioners to give more input in another data collection cycle. For instance attributes like Additional material and combined with have very insignificant values (around 5% filled fields each) while there are 1500 records. This means these attributes affect the outcomes of the results while they are not significant enough for decision making.

7.5 Global Conclusions

The global view of practitioners on diseases and plants can be influenced positively and active further research can be recommended on these indigenous plants and their attributes. For instance, scientific tests can be done to examine what contents Sengaparile has that can help in womb problems or delivery problems. Another example, when capturing data into the database, the Apriori algorithm will interpret the following plant names as different: Motsitsan, Motsitsane and Motsitasne. Henceforth this prompts for standardization of terminology and representation in order to make it understandable to other stakeholders.

7.6 Conclusion and Further Work

This research has proven successful in achieving the main objectives and giving way to future developments. It has contributed to the discussions on documenting Indigenous Knowledge. The question is how experiential and orally-transmitted knowledge can be integrated into a formal traceable structure using technical Indigenous knowledge Systems (IKS). *Dingaka* database serves as the documentation of ethno medicinal data while Data Mining techniques sought to thoroughly analyse patterns emerging from the data. While this is good research, going forward accurate information is very important as the study deals with very vital and sensitive data which if not handled with care can contribute to health problems and even instability within communities involved. In addition the data mining process awarded discoveries on anomalies in the data ranging from unfilled fields to errors in pronunciation of attribute names. The concept of data mining has proved to be a real research problem, especially on ethno medicinal IK collected in Botswana for the first time. It is a ground breaking project that shows a sense of recognition and appreciation for *Dingaka* (traditional doctors) community and even cultural pride for Batswana. In addition the project will promote networking or collaborations between traditional practitioners, researchers, modern medicine practitioners and other concerned relevant stakeholders as they share their knowledge and experiences.

8 REFERENCES

[1]. J. Hays, “Learning indigenous knowledge systems”, University of Tromsø, Retrieved June 11, 2011, Available: http://rwl5.uwc.ac.za/usrfiles/users/99062813/documents/Hays_Jennifer_367.doc, 2004

[2]. S.K. Sukula, “Developing indigenous knowledge databases in India”, [Online journal], Vol 24, No. 1, pg. 83-93. Retrieved September 02, 2012, Available: http://www.emeraldinsight.com. Emerald Group Publishing Limited. India, 2006

[3]. I. Hedberg and F. Staugård, “Traditional Medicine in Botswana”, Ipeleng Publishers, Botswana, 1989

[4]. M. Ishtiaq, W. Hanif, M.A. Khan, M. Ashraf, and A.M. Butt, “An ethno medicinal survey and documentation of important medicinal folklore food phytonims of flora of Samahni valley”, Laboratory of Ethno botany and Plant Taxonomy, Quaid-e-Azam University, Islamabad, 2007

[5]. A. Ribeiro, M. Romeiras, J. Tavares and M. Faria, “Ethno botanical survey in Canhane village, District of Massingir, Mozambique”, medicinal plants and traditional knowledge, Availbable:

http://www.ethnobiomed.com/content/6/1/33, 03/12/2010

[6]. P. Pillay, V.J. Maharaj and P.J. Smith, "Investigating South African plants as a source of new antimalarial drugs", 2008 Oct 28; 119(3):438-54, Epub, 2008

[7]. H. Yineger, D. Yewhalaw and D. Teketay, "Ethno medicinal plant knowledge and practice of the Oromo ethnic group in southwestern Ethiopia", J Ethnobiol Ethno med, Ethopia, 2008

[8]. M. Baerts and J. Lehmann, "Prelude Medicinal Plants Database", Université catholique de Louvain, Belgium, 1996

[9]. T. Leslie, "The healing power of rain forests plants", Retrieved April 15, 2011. Available: http://www.amazon.co.uk/Healing-Power-Rainforest-Herbs-Understanding, 2005

[10]. P. Swaroop and B. Golden, "Data Mining Introduction and a Health Care Application", Robert H Smith School of Business, University of Maryland, 2009

[11]. J. Han, M. Kamber and J. Pei, "Data Mining Concepts and Techniques (3rd Ed.) Chapter 8", University of Illinois, Illinois, 2010

[12]. Z. Tang and J. MacLennan, "Data Mining with SQL Server 2005", Wiley Publishing Inc. Indianapolis, USA, 2005

[13]. S. Ha and S. Joo, "A Hybrid Data Mining Method for the Medical Classification of Chest Pain", World Academy of Science Engineering and Technology, 2010

[14]. U. Abdullah, J. Ahmad and A. Ahmed, "Analysis of effectiveness of apriori algorithm in medical billing data mining", Emerging Technologies, Islamabad, 2008

[15]. I. Ullah, "Data Mining Algorithms And Medical Sciences", International Journal of Computer Science and Information Technology (IJCSIT), Vol 2, No 6, Dec. 2010

[16]. S. Stilou, P. Bamidis, N. Maglaveras and C. Pappas, "Mining association rules from clinical databases, an intelligent diagnostic process in healthcare", The Medical School, Aristotelian University, Greece, 2001

[17]. E. Frank, M. Hall, G. Holmes, R. Kirkby, B. Pfahringer, I. Witten and L. Trigg, "A Machine Learning Workbench for Data Mining", Data Mining and Knowledge Discovery Handbook, Springer US, USA, 2010