

# **TRANSFORMING DATA DUST TO DATA GOLD**

*An inaugural lecture delivered  
at the University of Ibadan*

*on Thursday, 25 August, 2011*

*By*

**ADENIKE OYINLOLA OSOFISAN**

*Professor of Computer Science*

*Faculty of Science*

*University of Ibadan*

*Ibadan, Nigeria.*

**UNIVERSITY OF IBADAN**

Ibadan University Press  
Publishing House  
University of Ibadan  
Ibadan, Nigeria.

© University of Ibadan, 2011  
Ibadan, Nigeria

*First Published 2011*

*All Rights Reserved*

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, photocopying, recording or otherwise, without permission in writing from the publisher.

**ISBN: 978 - 978 - 8414 - 60 - 5**

*Printed by: Ibadan University Printery*

*The Vice-Chancellor, Deputy Vice-Chancellor (Administration), Deputy Vice-Chancellor (Academic), Registrar, Librarian, Provost of the College of Medicine, Dean of the Faculty of Science, Dean of the Postgraduate School, Deans of other Faculties, and of Students, Distinguished Ladies and Gentlemen.*

I feel highly honoured to deliver this inaugural lecture, the second from the Computer Science Department since its inception in 1975. The first was given by Prof. Olu Longe in the 1983/84 series with the title "Ifa Divination and Computer Science". The gap in time (1983/84 to 2010/11) is an indication of the state of affairs of the department over time. But I will return to this issue later.

Mr. Vice-Chancellor Sir, today is of historical importance because this is the very first inaugural lecture to be delivered in Nigeria by a female professor in the field of Computer Science. The title of my lecture is "Transforming Data Dust to Data Gold". I chose this topic advisedly, first, because of the significance of data to the "Knowledge Economy" of the present century, especially in our country; then secondly, because of the many opportunities that we have lost as a result of not adequately collecting and storing historical data over time; and thirdly, because of the vast benefits of this untapped knowledge that we have thereby missed. And finally, I also chose this topic because it encapsulates much of the thrust of my research activities in Data Communication and Knowledge Discovery in Databases (KDD) as well as my administrative and professional activities within and outside the university environment. Hopefully therefore, this inaugural lecture will reveal how data that have hitherto been gathering dust in our various walks of life may be turned into valuable powerful decision-making tools that will in turn create national wealth and add value to our economy and our lives.

## **Definitions**

First let me begin with definitions.

**Dust** in the sense that it is used in this lecture refers to the wasted material, like the particles of dust that settle especially on household surfaces; something that is worthless. **Gold** on the other hand, is something precious, like the soft, yellow and heavy element that bears the name; something quite rare but very valuable. The weights of these two definitions rest on those two adjectives—**worthless** and **valuable**.

**Data** are facts or observations about physical phenomena or business transactions. Specifically, data are objective measurements of the attributes of entities such as people, places, things and events. Data can be numbers, words, measurements, observations or even just descriptions of things. Data in Computer Science are often viewed as the lowest levels of abstraction from which information and then knowledge are derived. This process of deriving Knowledge from Data is what has led to a new research area in Data Bases, referred to as “Knowledge Discovery in Data Bases” or KDD (Frawley, Piatetsky-Shapiro & Matheus 1992; Date 2001).

The mechanics of effectively storing, organizing, processing and communicating data lead to its transformation from data into **information**. Information is therefore the data that have been interpreted and understood by the recipient of a message (Lucey 1997; Mannino 2001). Consequently, for data to become information, it has to be communicated, and both the user and the sender must be involved in such transformation of data. It is only when these data are communicated and understood by the recipient that value is added to the transformed data, that is, information. And as we know, information is needed to improve decision making, increase knowledge, reduce uncertainty and add value.

Now, there is hardly a transaction that does not generate data somewhere, internally or externally, for a particular environment. These data accumulate daily in our environment as hard copy reports, memos, students’ results, letters, graphs, videos, audios, animations, etc. All these data that we carelessly allow to gather dust in our various work places and homes hold valuable information, such as trends and patterns,

which could be used to improve the business of governments, schools, or private enterprises, or even home decisions, and therefore optimize our operational success. This business of communicating data through computer systems has led to several areas of research in Data Communications, one of which is Computer Network in Computer Science. The main thrust of this evening's lecture, Mr. Vice-Chancellor, will therefore be on Computer Networks and Knowledge Discovery in Data Bases which are my own areas of Research and specialization, and I will be talking about my contributions to them.

### **Computer Networks**

The origin of computer networks has been ascribed to a series of memos written by one J.C.R. Licklider of MIT in August 1962 discussing his "Galactic Network" concept. Licklider envisioned a globally interconnected set of computers through which everyone could quickly access data and programs from any site. In spirit, the concept was very much like the Internet of today (Wikipedia 2011).

As Licklider envisaged it, a computer network would be a collection of autonomous computers called **hosts** through which users can quickly access resources using **protocols** to communicate over transmission media. These media of communication may be copper wire, lasers, microwaves, cables (coaxial, optical) and earth satellites. The distances between the nodes may range from a few meters to several millions of kilometers. This computer network would differ from ordinary computer channels in that it would provide:

- A high degree of insulation between processing nodes;
- Greater variability in communication protocols;
- Less emphasis on very low latency; and
- The possibility of spanning longer distances.

One criterion for classifying networks is their scale as shown in table 1.

**Table 1: Classification of Interconnected Processors by Scale**

<b>Inter processor distance</b>	<b>Processor located in same place</b>	<b>Example</b>
0.1 meter	Circuit board	Data Flow Machine
1.0 meter	System	Multicomputer
10 meters	Room	Local Area Network (LAN)
100 meters	Building	
1 kilometer	Campus	
10 kilometers	City	Metropolitan Area Network (MAN)
100 kilometers	Country	Wide Area Network (WAN)
1,000 kilometers	Continent	
10,000 kilometers	Planet	The Internet.

Needless to say, Licklider’s dream has become a reality today. Let us proceed to other definitions and clarifications.

The unit of information transferred on a network is called a **message** in complete form or in shorter packets. In any network, there exist collections of hosts intended for running users’ programs. The hosts are connected by a communication subnet. The work of the subnet is to carry messages from host to host, and it consists of two basic components:

- Switching elements, and
- Transmission lines

Let me briefly try to explain how the system works. The switching elements are generally specialized computers often called:

- Communication computer or Frontend,
- Packet switch,
- Node, and
- Data switching exchange.

Transmission lines are often called channels or circuits. Broadly, there are two types of designs for the subnet:

- Broadcast channels, and
- Point-to-point channels.

In broadcast channel design, there is a single communication channel shared by all nodes. Packets sent by any node are received by all other nodes. A portion of control information in the packet specifies to whom it is intended, so that if a node receives a packet not intended for it, such a node just ignores the packet.

Broadcast channel may further be subdivided into static and dynamic, depending on how the channel is allocated. Static allocation could divide time up into discrete intervals and run the channel in a round robin fashion allowing each node to broadcast only when its time slot comes up. Static allocation has the problem of wasting channel capacity when a node has nothing to send during its time allocation. Dynamic allocation method is either centralized or decentralized. In centralized channel allocation method, there is a single entity which determines which node goes next. But in the decentralized allocation method, there is no central entity, each node must decide whether to transmit or not. There are algorithms designed to bring order to any chaos that might arise.

In point-to-point channels, on the other hand, the network contains cables or leased telephone lines, each one connecting a pair of nodes. If two nodes do not share a cable and they wish to communicate, they must do so indirectly via another node. When a packet is sent from one node to another via one or more intermediate nodes, this packet is received at each intermediate node, stored there until the required output line is free and then forwarded. A subnet using this principle is called a point-to-point store and forward subnet. An important design issue is what the node interconnection **topology** should look like.

Topology is the virtual shape or structure of a network. This virtual design does not correspond to the actual or the physical shape of the computer network. The transfer of data in a network is referred to as **Logical Topology** (the basic network) while the **Physical Topology** (the core network) accounts for the physical structure of the network that carries

devices, cable installations and locations (Horak 1999; Bartee 1985; Schwartz 1977; Tanenbaum 1996).

Network topologies are categorized into five basic types as shown in figures 1-5. They are,

- Star
- Bus/Line
- Tree
- Mesh/Fully Connected
- Ring

Let us quickly look at these topologies one by one, to understand how they work (Wikipedia 2011).

First, a Star Network features a central connection point called a "hub" that may be a switch or router. Devices typically connect to the hub with Unshielded Twisted Pair (UTP) Ethernet. Compared to the bus topology, a star network generally requires more cable, but a failure in any star network cable will only take down one computer's network access and not the entire LAN. If the hub fails, however, the entire network also fails. Many home networks use the star topology.

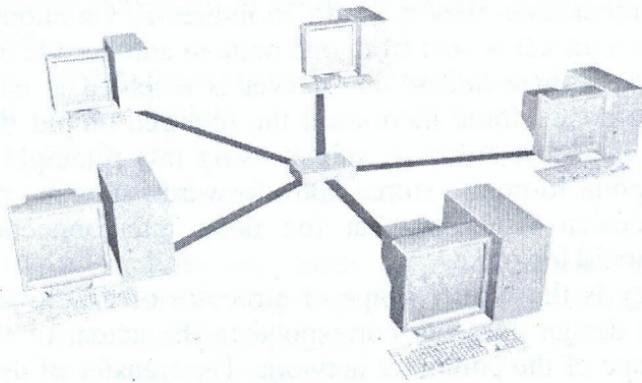


Fig. 1. Star Topology

Bus Networks use a common backbone to connect all devices. A single cable that is the backbone functions as a shared communication medium that devices are attached to or tapped into with an interface connector. A device wanting to communicate with another device on the network sends a broadcast message onto the wire that all other devices see, but only the intended recipient actually accepts and processes the message. Ethernet bus topologies are relatively easy to install and do not require much cabling compared to the alternatives. 10Base-2 ("ThinNet") and 10Base-5 ("ThickNet") were both popular Ethernet cabling options many years ago for bus topologies. However, bus networks work best with a limited number of devices. If more than a few dozen computers are added to a network bus, performance problems will likely result. In addition, if the backbone cable fails, the entire network effectively becomes unusable.

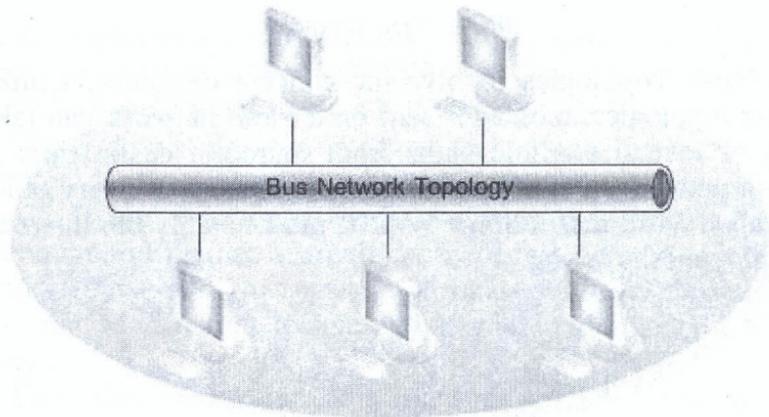


Fig. 2. Bus Topology

Tree Topologies integrate multiple Star topologies together onto a Bus. In its simplest form, only hub devices connect directly to the tree bus and each hub functions as the "root" of a tree of devices. This bus/star hybrid approach supports future expandability of the network much better than

a bus (limited in the number of devices due to the broadcast traffic it generates) or a star (limited by the number of hub connection points) alone.

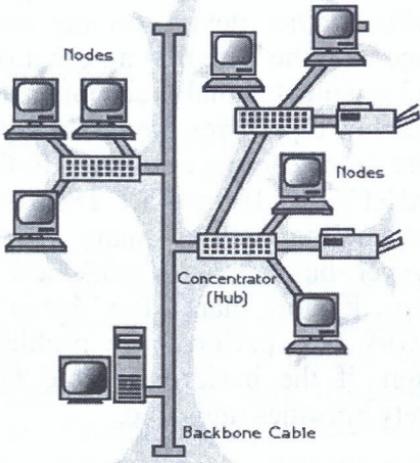


Fig. 3. Tree Topology

Mesh Topologies involve the concept of routes. Unlike other topologies, messages sent on a mesh network can take any of several possible paths from source to destination. A mesh network in which every device connects to every other is called a full mesh. Some WANs, most notably the Internet, employ mesh routing.

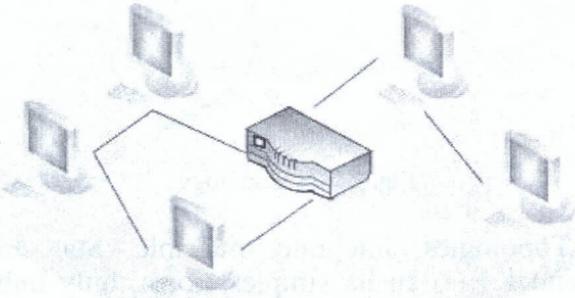


Fig. 4. Mesh Topology

In a Ring Network, every device has exactly two neighbours for communication purposes. All messages travel through a ring in the same direction either "clockwise" or "counter clockwise". A failure in any cable or device breaks the loop and can take down the entire network.

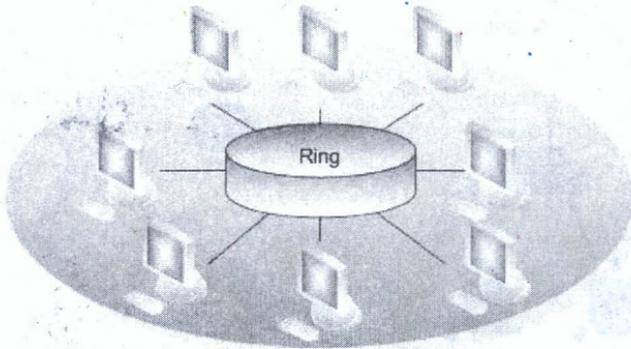


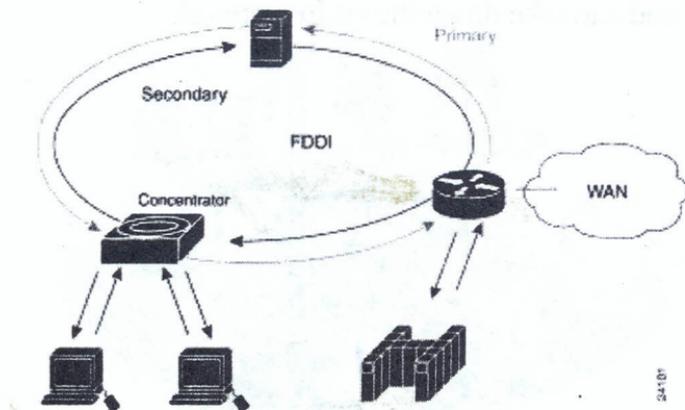
Fig. 5. Ring Topology

To implement a ring network, one typically uses Token Ring, or FDDI (Fiber Distributed Data Interface), or SONET (Synchronous Optical Network).

The Token Ring is a data link technology for LANs. It operates at layer 2 of the Open System Interconnection Model. It was developed by IBM during the 1980s as an alternative to Ethernet. It significantly decreased in popularity in the 1990s and was gradually phased out of business networks as Ethernet technology began to dominate LAN designs.

The Fiber Distributed Data Interface (FDDI), in its own case, specifies a 100-Mbps token-passing, dual-ring LAN using fiber-optic cable. FDDI is frequently used as high-speed backbone technology because of its support for high bandwidth and greater distances than copper. Furthermore, FDDI uses dual-ring architecture with traffic on each ring flowing in opposite directions (called *counter-rotating*). The dual-rings consist of a primary and a secondary ring. During normal operation, the primary ring is used for data

transmission, and the secondary ring remains idle. The primary purpose of the dual rings is to provide superior reliability and robustness. Figure 6 shows the counter-rotating primary and secondary FDDI rings.



**Fig. 6.** FDDI uses counter-rotating primary and secondary rings.  
 ([www.pulsewan.com/data101/fddi\\_basics.htm](http://www.pulsewan.com/data101/fddi_basics.htm)) (July 24, 2011)

SONET on the other hand, is a physical layer network technology designed to carry large volumes of traffic over relatively long distances on fiber optic cabling. SONET was originally designed by the American National Standards Institute (ANSI) for the USA public telephone network in the mid-1980s. SONET possesses several characteristics that make it appealing on the Internet today. For instance,

- It defines clear interoperability standards between different vendors' products;
- It carries nearly any higher-level protocol (including IP); and
- It includes built-in support for ease of management and maintenance.

One could arrange the Home Network in a circle but it does not replicate Ring Topology. Graph Theory is used for studying network topology. Nodes' distance and inter-connectivity, and the rate of transmission and types of signals

of two networks might vary, but their topologies could be identical. Complex and efficient networks can be built as hybrids of two or more of the above basic topologies.

It is therefore clear that Network Topology is not something one can just configure by a rule of thumb or by favouritism of one node (department or faculty) over the other. The University of Ibadan has just learnt the bitter lesson of this. Our UI network topology was not configured using computer networks engineering and scientific principles as was advised by my humble self when it was being done. Now many years afterwards it has taken an expert from the University of Oregon to tell UI ICT team that there is the need to address some technical aspects of the Network design so that when the bandwidth is increased, it will be well utilized across all nodes. Also, the optical fibre links should be circular, such that any damage to any route should not cut off the nodes on that route, that is, there should be alternative signal routes. Technically, there should be only one Network Address Translation (NAT) instead of the present two that UI has, and the issue of authentication at every node should be limited to a central authentication server.

The simple truth, Mr. Vice-Chancellor Sir, is that node locations, interconnectivity and rates of transmission cannot be determined except through hard core data communication principles. Data dust will be transformed to data gold only when one uses computer communication theories and not the rule of thumb, favouritism and/or ignorance.

### **Knowledge Discovery in Databases (KDD)**

Mr. Vice-Chancellor Sir, this is where I come to the subject of Knowledge Discovery and, specifically, of Data Mining, which is my own special field. KDD (Knowledge Discovery in Databases) is a field of computer science that includes the tools and theories to help us extract useful and previously unknown information and knowledge from large collections of digitized data. KDD consists of several steps, and Data Mining is one of them. Data Mining is the application of a specific algorithm in order to extract patterns from data. It is a

field of Computer Science that deals with the mapping of low-level data into other forms that are more compact, abstract and useful. KDD achieves this by creating short reports, modelling the process of generating data and developing predictive models for future cases.

Due to the exponential growth of data, especially in areas such as business and technology in our modern times, KDD has become a very important tool to convert this large wealth of data into business intelligence. The manual extraction of patterns has become increasingly outmoded and inadequate in the past few decades. As a result, KDD is now the preferred choice for various applications such as social network analysis, fraud detection, scientific research, investment, manufacturing, telecommunications, data cleaning, sports, information retrieval, marketing, and several others.

The KDD process has several steps as shown in figure 7. Starting with the development of the application domain and the goal, it then proceeds to the creation of the target dataset. This is followed by cleaning, preprocessing, reduction and projection of data. The next step is using Data Mining algorithms (explained below) to identify patterns. Finally, discovered knowledge is consolidated by visualizing and/or interpreting these patterns.

There are two major Data Mining goals as defined by the specific application, and they are, namely, verification, or discovery. Verification is checking and confirming the user's hypothesis about data, while discovery is automatically finding interesting patterns among the data. There are four major data mining tasks:

- Clustering, which is identifying similar groups from unstructured data;
- Classification, which is learning rules that can be applied to new data;
- Regression, which is finding functions with minimal error to model data; and
- Association (Summarization) which is looking for relationships between variables.

After this, the specific data mining algorithm needs to be selected. Depending on the goal, the algorithm could be one or a combination of the following:

- Market Basket Analysis
- Memory-Based Reasoning
- Automatic Cluster Detection
- Link Analysis
- Decision Trees/Rule Induction
- Artificial Neural Networks
- Genetic Algorithm

The first, Market Basket Analysis (MBA), is a form of clustering used for finding groups of items that tend to occur together in a transaction. As a clustering technique, MBA is useful for problems where one wants to know what items occur together or in a particular sequence. The resulting information is often used for purposes such as planning store layouts and bundling of products.

The second, Memory-Based Reasoning (MBR), is a directed data mining technique that uses known instances in a model to make predictions about the unknown. It is a technique directly borrowed from the field of Artificial Intelligence (AI). MBR looks for the nearest neighbours in the known instances and combines their values to assign classifications or prediction values. MBR may be used on more complex data types, such as texts and images, when distance functions are available in these domains.

The third, Clustering Detection (CD) is the building of models that find data records that are similar to each other. There are many methods for finding clusters – geometric methods, statistical methods and neural methods. CD is a very good way to start any analysis on data. Self-similar clusters can provide a starting point for knowing what is in the data and for figuring how best to make use of them.

The fourth, Link Analysis (LA), follows relationships between records to develop models based on patterns of relationships. Link analysis is useful in the law enforcement

profession, where clues about crimes are linked together to help solve them.

The fifth, Decision Trees (DTs), are a powerful model produced by a class of techniques that include Classification and Regression Trees (CART) and Chi-squared Automatic Induction (CHAID). DTs are used for directed data mining, particularly classification. The decision tree model is explainable because it takes the form of explicit rules. Therefore, the results can be evaluated and the key attributes in the process can be identified. The rules can be expressed easily as logic statements in a language such as SQL, so that they can be applied directly to new records. Along with Link Analysis, Decision Trees arose from graph theory and its application to Data Structures in Computer Science.

The sixth, Artificial Neural Networks (ANN), are simple models of neural interconnections in brains adapted for use on digital computers. ANN detects patterns in data in a manner analogous to human thinking. It can be used for both directed and undirected Data Mining (DM). In the simplest form, ANN learns from a training set, and then generalizes patterns inside it for classification and prediction. It is used for undirected DM in the form of self-organizing maps and related structures and time series predictions. Neural networks are of particular interest because they offer a means of efficiently modelling large and complex problems in which there may be hundreds of predictor variables that have many interactions. Neural nets may be used in classification problems (where the output is a categorical variable) or for regressions (where the output variable is continuous).

Finally, the seventh, Genetic Algorithms (GA) are not used to find patterns per se, but rather to guide the learning process of data mining algorithms such as neural nets. Essentially, genetic algorithms act as a method for performing a guided search for good models in the solution space. They are called genetic algorithms because they loosely follow the pattern of biological evolution in which the members of one generation (of models) compete to pass on

their characteristics to the next generation (of models), until the best (model) is found. The information to be passed on is contained in “chromosomes,” which contain the parameters for building the model. For example, in building a neural net, genetic algorithms can replace back propagation as a way to adjust the weights. The chromosome in this case would contain the weights. Alternatively, genetic algorithms might be used to find the best architecture, and the chromosomes would contain the number of hidden layers and the number of nodes in each layer. As you can see, Genetic Algorithms and Neural Network came from attempts to create computing models using biological processes. Genetic Algorithms are also used to enhance MBR and neural networks (Fayyad et al. 1996, 2002; Fausset 1994, Berry & Linoff 1997; Golberg 1989; Haykin 1994; Speckt 1991; Witten & Frank 2000; Ferruci & Lally 2004).

The two terms KDD and Data Mining (DM) are often used interchangeably, but actually they refer to two related yet slightly different concepts. While KDD is the overall process of extracting knowledge from data, DM is a step inside the KDD process, which deals with identifying patterns in data. In other words, DM is only the application of a specific algorithm based on the overall goal of the KDD process as shown in figure 7.

Mr. Vice-Chancellor Sir, my research work in KDD has focused mainly on the use of ANN algorithms combined with Genetic Algorithm and/or Decision trees. With your permission I shall proceed now to provide highlights of my humble contributions in academics to this process of turning data dust to data gold.

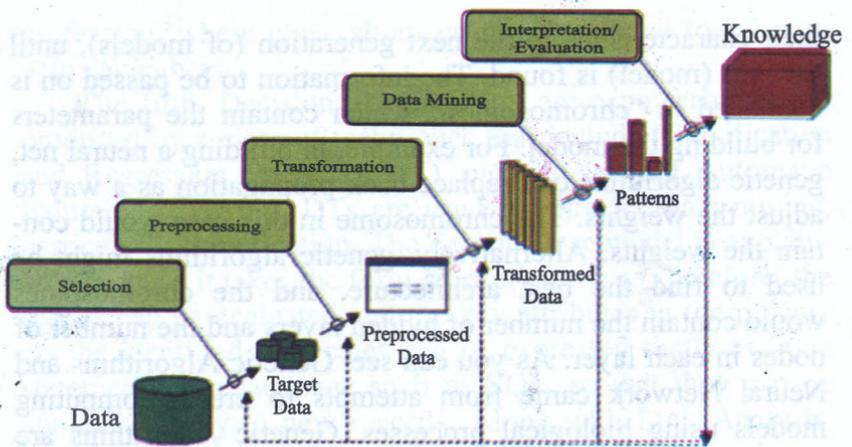


Fig. 7. KDD Process (Adapted from Fayyad et al. 1996).

### My Contributions

My contributions have been mainly in Data Communications, in Computer Networks in particular, and Knowledge Discovery in Data Bases.

### NINET

My first contact with research in Data Communications was during my PhD work at the Obafemi Awolowo University, Ile-Ife. I had initially wanted to do my research in Data Bases but there was no one senior enough in the area to supervise the work. So the then Ag. HOD, Dr. A.D. Akinde (now Professor Akinde and Bishop of Lagos Mainland Anglican Communion) asked me if I could switch to Data Communication, which I gladly did. He agreed to supervise me. Little did I know then that I would also be making history at the end of the research work that would lead to the design of an efficient national computer network, especially for data handling purposes. We came up with a 3-level hierarchical (Tree Topology) computer communication network, which we named the N<sup>I</sup>gerian NETwork (NINET), as well as a prototype distributed Operating System (NINOS) (Akinde & Osofisan 1994). Four design questions were posed and answered in the research work as follows:

1. *Topological Design.* That is, given the geographical location of computers and other message sources, as well as their expected traffic characteristics, where should network nodes be located? How should they be connected? What form should the network take? How many connections between nodes would be needed?
2. *Line Capacity Allocation.* What size trunks, in units of data capacity bits/sec, should be used throughout the network?
3. *Routing Procedures.* How did different routing algorithms compare? Should there be local or centralized routing control? How often should routing information and/or routing control be updated? How should this be done? What should be the significant parameters in determining routing?
4. *Flow Control Procedures.* How should one ensure smooth traffic flow throughout the network? How are bottlenecks and deadlocks to be prevented? How should one stop any one user, data source, or node from tying up the network?

In answering these questions the performance criteria we used included:

- (a) Specified response time,
- (b) Specified cost, and
- (c) Specified reliability.

Load capacity and response time were worked out between every pair of nodes. The simulation result showed that the maximum line rate possible then was 56,000kps due to the Nigerian Telecommunication environment. Our NINET was therefore assigned to bring about resource sharing, cost advantages and performance improvement of computer distributed systems to large communities of computer users in Nigeria. This, Mr. Vice-Chancellor Sir, was the first major research work in Nigeria along this line.

## **NINOS**

After this, the work then went ahead to propose a prototype model of a heterogeneous distributed operating system using clustering, which we called NINOS (Adagunodo, Akinde & Osofisan 1999).

Mr. Vice-Chancellor, it is very unfortunate that up till today, Nigerian Universities are still grappling with the issue of getting connected together after the dismal failure of two projects called NUMIS and NUNET, which are similar in nature to NINET. And it is more disheartening to find that the reason for this is that personal interest has been allowed to take precedence over collective interest. Needless to say, Mr. Vice-Chancellor Sir, personal interest over collective interest will never transform our data dust to data gold.

## **NUMAPS**

Still in the search for Nigeria to use the Internet as an excellent mechanism for the distribution of educational resources to improve teaching and learning, in the absence of effective Internet presence in Nigeria, I then again got involved with a project funded by UNESCO situated at the International Institute of Theoretical and Applied Physics (IITAP) of Iowa State University (ISU), Ames, during the summers of 1995 and 1996. There we successfully developed a simulated model of the Internet for use in an educational environment. This was christened NUMAPS (Network for Uses in Mathematics and Applied Physical Sciences). Three inter-woven projects were successfully implemented in NUMAPS. They were:

- Netshell,
- Netpack, and
- IntraNET (Osofisan & Fils 1996; Osofisan 1996).

These three projects provide educators, teachers and students the opportunity to acquire knowledge about how the Internet works, explore the electronic frontier, create World Wide Web (WWW) pages, and use Uniform Resource Locators (URLs) to locate resources from other computers even without full access to the Internet.

Mr. Vice-Chancellor Sir, I am glad to say that the University of Ibadan was the first to benefit from this project in Nigeria, because as soon as I moved from the Polytechnic, Ibadan, to UI in March 1997, I started supervising final year projects in the area of Internet Technology. The under-listed projects were designed and implemented at the University of Ibadan and our students became the very first to be introduced to Web Technology in Nigeria. The testimony is in the following successful projects:

- Development of a Website—Case Study: Faculty of Science (by Alejo H.S., Durodola P.O. & Ebroo R., in February 1998).
- Design of a Website for University of Ibadan, College of Medicine (by Amore B. O., Ajayi F.T., Olusanya O.A. & Egwaikhide J.O., in Feb., 1999).
- Design of a Website for the Offices of the Vice-Chancellor, Registrar and the Registry (by Adeyemo A., Ebitigha A.K.A. & Ezeh V.N., in Feb. 1999).

These projects were demonstrated to the management of the University at the time, with the hope that Computer Science Students would be used to develop the University of Ibadan Website. The idea was well received by the then Vice-Chancellor, but surprisingly not by those who were supposed to make it happen. Hence up till now, since 1998, UI is still struggling with getting a decent website of world-class standard.

However, Mr. Vice-Chancellor Sir, I wish to seize this opportunity to congratulate you on UI's recent outing on the Webometric ranking as announced last month. As a result of your fresh initiatives recently, UI is now first in Nigeria. This is undoubtedly the beginning of good times for us in Information Technology, although there is still more work to be done of course. All of us now live in the world of Knowledge Economy and UI must begin earnestly to transform its data dust to data gold by using sound principles and theories in Information Technology, instead of the rule of thumb, favouritism and/or ignorance as has hitherto been the practice.

## **E-Commerce**

In 1998 we extended the idea of NUMAPS to commerce and came up with a design of an e-commerce site where customers and sellers may negotiate on price, participate in an auction and come to an agreement about prices at which the merchandise would move. The design can also handle personalized bundled goods and customized interactions with negotiations over value rather than prize. Thus, our design gives a practical example of the ability of e-commerce to negotiate both price and quality in addition to mass customization (Osofisan 2000b).

Again, Mr. Vice-Chancellor, this knowledge was transferred to our final year students, some of whom carried out their final projects in the area of website design for various industries in Nigeria. An example of such a project is:

- Designing a website for advertisement of products and services: Case study of Lever Brothers Nig. PLC, Dockyard Road Apapa, Lagos (by Oso, O.O., Orioye, B.D. & Akinwale, O. N., in May 2000).

I can assure you Sir, that Lever Brothers, eagerly took advantage of this work. There are many of such projects produced by students of Computer Science under our supervision, which have been of Nigerian industries.

## **NETBILL**

In 2001, Osunade, Adeyemo and Osofisan (2001) came up with a business model called NETBILL, that is, a set of protocols and software implementation that allows customers to pay owners and retailers of information. This is an e-commerce solution for businesses dealing in the sale of information such as universities, research institutes, etc. NETBILL uses a single protocol that supports charging in a wide range of service interactions. Mr. Vice-Chancellor Sir, there is no reason why the University of Ibadan cannot engage in the sale of information and also make money through foreign exchange. Companies such as Google, Yahoo, and Amazon, to mention just a few, have made

fortunes out of such sales of information. These companies have transformed data dust to data gold. We have the knowledge and the expertise to follow their example.

### The UI Network

In 2000, in the bid to get the University of Ibadan network correctly configured, we embarked on a research on cooperative caching in web servers. Historical trace data were collected from two ISP's in Nigeria. With our traces, we evaluated quantitatively the performance improvement potential in inter-proxy cooperation between 500 to 1000 clients of one ISP operating within a small geographical location, and over 5000 clients nationwide. We used existing models to demonstrate that cooperating caching has performance benefits within a particular population boundary to confirm earlier findings. The results are displayed in figures 8 and 9.

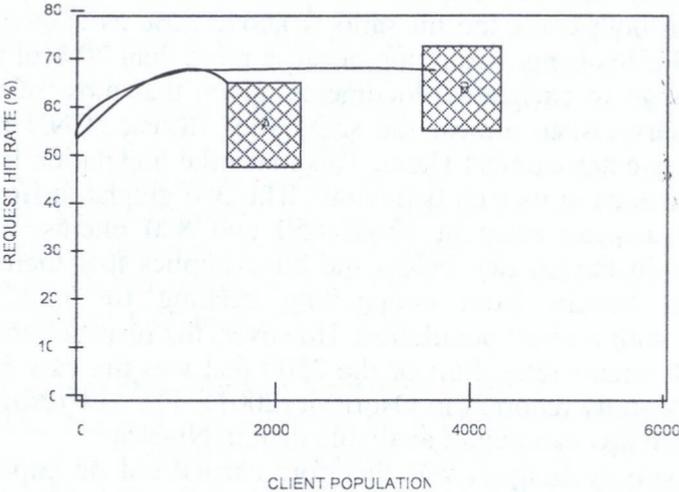
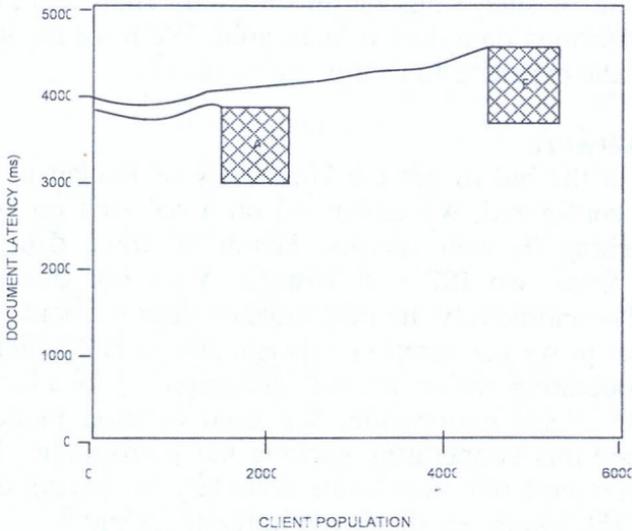


Fig. 8. Cacheable curves for ISP A and B.



**Fig. 9.** Mean latency as a function of client population.

The cacheable curves of the two ISPs almost overlie each other. In both cases the hit ratio is above 50% as shown in figure 8. This is not surprising because more than 70% of the requests go to cacheable documents, given that most of the Web users visited almost the same sites (Music, CNN and some few entertainment sites). This made the population look homogeneous in its web behaviour. The two graphs in figure 8 have a sharp knee at about 650 and 800 clients. The increase in the hit rate below the knee implies that there is potential benefit from cooperating caching for multiple proxies with a small population. However, the result of about 650-800 clients falls short of the 2500 that was the case in a previous study reported in Osofisan (2001). The low result is due to the low bandwidth available then in Nigeria.

Sensitivity analysis was therefore carried out on populations that are strictly heterogeneous in nature, such as a university environment. The graph of mean latency as a function of client population is shown in figure 9. This figure shows that caches have little impact on the mean latency

beyond a particular population. Therefore, unless the bandwidth offered by ISPs in Nigeria increases, no ISP will be able to support the needs of the university, because the university audience cannot exhibit the homogeneous behaviour that the population we studied exhibited. We immediately shared this information at that time so that some solution would be proffered and we would get value out of our network. We were eager to transform our data dust to data gold.

To our surprise however, rather than praise, what we got was a severe rebuke for our efforts. Well almost immediately afterwards, we were proven right, because when more people in UI came on board the network, the hit ratio of web access fell drastically and the university Internet users became disenchanted. Hardly could any user get a mail across from the Faculties except from the ICT unit. Going to the ICT unit to send and receive mails and download files became the norm. As we all know, University of Ibadan users are completely heterogeneous in nature; even a single department in the university cannot possibly exhibit homogeneity in its web behaviour because there are various research interests and pastimes within it. That is why we have been encountering the kind of problems we are having and will continue to have, until the situation is rectified. Fortunately as we indicated earlier, the new Administration is fully aware of the problem and is taking steps already to solve it. For this, I thank you Mr. Vice-Chancellor Sir, and also pledge that we on our part will not relent in our efforts to make the University of Ibadan network functional as in the other world universities. May the Almighty God help us all. Amen.

### **Content Distribution Networks (CDNs)**

The Internet has experienced a phenomenal growth worldwide as a result of increasing demands for contents, content distributions and other services. Present development trends in content networking facilities or resources provisioning has stirred up interest in interconnecting CDNs. In order to achieve a cost effective content delivery and better overall

service, distinct CDN providers seek ways to cooperate and coordinate their services. Customers' interests or preferences are forming a very important part in the provisioning of CDN services while taking into account some specific QoS requirements. Analytical modelling is a very good and effective tool that can be used to solve the resource sharing and management problems among autonomous CDNs in order to justify the overall system goals.

CDNs have since evolved as cooperative and collaborative groups of networks over the Internet where contents are replicated over the surrogate servers for efficient delivery performance to the clients and improved service cost to the CDN providers (Osofisan & Idowu 2009a). However, a CDN is limited in terms of Point of Presence (PoP) and capability. Therefore, in 2011 we became concerned about content object replication among peering CDNs and then explored the issue of Web content delivery among them (peering CDNs) (Osofisan & Idowu 2009b, 2011). We precisely considered content object replication. We then provided an analytical model for the replication problem in terms of a constrained optimization problem subject to a mix of QoS requirements (bandwidth and delay).

Our objective was to minimize the service cost which consists of both the storage and consistency management costs. In order to ensure the consistency of content objects in replica placement, different values of reading and writing rates were considered, assuming a flat update delivery. We developed a greedy algorithm to obtain a near-optimal solution using AMPL/CLEX. The computational results obtained are shown in tables 2 and 3.

Table 2: Content Objects Replication Costs

	No of CDN			X>=0	X=0 or 1			X>=0	X=0 or 1			X>=0	X=0 or 1
	CDN-1	CDN-2	CDN-3			CDN-4	CDN-5			CDN-6	CDN-7		
2	Obj1	0.741	0	0									
	Obj2	0	0.302	0.603									
				<b>16.25</b>	<b>29.27</b>								
	Obj1	0.819	0	0			0.273	0					
	Obj2	0	0.333	0.667			0	0.222					
								<b>35.35</b>	<b>94.11</b>				
	Obj1	0.969	0	0			0.242	0			0.162	0	
	Obj2	0	0.395	0.789			0	0.263			0	0.056	
											<b>53.02</b>	<b>193.74</b>	
3	Obj1	1.403	0	0									
	Obj2	0	0.571	0									
	Obj3	0	0	0.281									
				<b>19.99</b>	<b>55.05</b>								
	Obj1	0.771	0	0			0.444	0					
	Obj2	0.229	0.543	1.086			0	0.362					
	Obj3	0	0	0			0	0					
								<b>57.88</b>	<b>119.89</b>				
	Obj1	0.551	0	0			0.414	0			0.276	0	
	Obj2	0	0.673	0			0	0			0.276	0	
	Obj3	0	0	0.905			0	0.453			0	0.071	
											<b>72.63</b>	<b>157.44</b>	
4	Obj1	1.900	0	0									
	Obj2	0.099	0.873	0									
	Obj3	0	0	0.430									
	Obj4	0	0	0									
				<b>30.57</b>	<b>55.05</b>								



**Table 3: Content Object Replication Cost Summary Results**

No of objects	No of peering CDN	Total replication cost (trc)	
		X>=0	X=0 or 1
2	3	16.35	29.27
2	5	35.35	94.11
2	7	53.02	193.74
3	3	19.99	55.05
3	5	57.88	119.89
3	7	72.63	157.44
4	3	30.57	55.05
4	5	77.00	128.03
4	7	101.51	266.35

This is still an ongoing research. We believe that further work on modelling in Content Distribution Internetworking (CDI) in particular with respect to clients' requests, re-direction and content replications can still be explored. This can be done by focusing on two main aspects;

1. New model realization to include new situations or parameters to help facilitate the solution of problems being represented.
2. Algorithm development for more complex situations with more QoS requirements.

These two issues are being addressed in the doctoral work of S.A. Idowu, whom I am supervising.

### **Cyber Crime**

The use of the Internet in Africa has grown so rapidly with the explosion in the number of Internet Service Providers (ISPs), Internet cyber cafés and access points within the last decade. This has had several positive impacts on the social, economic and educational sectors on the continent. Unfortunately, the image of nations such as Nigeria, Ghana and Cameroon has also suffered as a result of the nefarious activities of some users who, instead of exploring the Internet for constructive purposes, turn it into a ready channel for the perpetration of criminal activities. We are all familiar, for

instance, with the 'Advance Fee Fraud' (AFF), known more familiarly as "419".

Nigeria, in particular, has gained recognition as a source of fraudulent Spam mails characterized by bogus business proposals and fraudulent joint ventures. If these mails do not entirely emanate from the suspected nations, then spammers have scored another point by allowing efforts to be directed and focused solely on a few nations in Africa that only contribute a small percentage to the advance fee fraud problem. Since most of these nations do not even have reliable databases for apprehending cyber criminals, these criminals can simply change location, move to other nations even in the Western world and continue to perpetuate their heinous crimes.

The increase in the volume of cyber criminal activities noticed in the United States, Canada and parts of Europe may be a consequence of this scenario. Are spammers migrating to the Western world from resource-poor regions of the world; from sub-Saharan Africa in particular? The implications for cyber security and information systems research is that measures must be adopted that correctly identify causation and locate criminals in order to direct preventive efforts in the right direction and provide appropriate solutions.

Thus the correct identification of the origins of advance fee fraud mail is a major concern. Although previous research (Cuckier et al. 2007; Gbenga 2007; Igwe 2007; Profgame 2007) suggest that these mails originate mainly from Nigeria and other West African Countries, we decided in 2010 to carry out research using available tracking tools to validate previously held notions about the issue of advance fee e-fraud mails (Longe & Osofisan 2011).

We harvested in real-time aggregated advance fraud e-mails over a two-year period using the sinkhole aggregation methodology as proposed by Abhinav et al. (2008). Using freeware e-mail and Internet protocol address tracers, we obtained results (table 4) that deviate from the generally held belief about the origins of advance fee fraud e-mails. In table 4, the amount lost by selected fraud type is reported as Monetary Loss.

**Table 4: Amount Lost by Selected Fraud Type for Reported Monetary Loss**

<b>Complaint Type</b>	<b>% of Reported total dollar loss</b>	<b>Average (median) \$ loss per complaint</b>
Nigerian letter fraud	1.7%	\$5,100.00
Check fraud	11.1%	\$3,744.00
Investment fraud	4.0%	\$2,694.99
Confidence fraud	4.5%	\$2400.00
Auction fraud	33.0%	\$602.50
Non-delivery	28.1%	\$585.00
Credit/debit card fraud	3.6%	\$427.50

Source: Internet Crime Complaint Centre Report (2006-2008), <http://www.ic3.gov/media/annualreports.aspx>

As shown in table 5, the ICC report for the period 2006-2008 reflects that it is the United States that tops the list of nations that perpetuate cyber crime (62% - 66.1% in three years). This is followed by the United Kingdom where cyber crime activities by percentage dropped from 15.9% to 10.5% between 2006 and 2008. Nigeria is only third on the list, with a marginal increase of 1.6% between 2006 and 2008. Two other African countries, Ghana and South Africa are listed in the midst of other European and Asian nations.

**Table 5: Top Ten Countries—Perpetrators of Cyber Crime**

Year 2006		Year 2007		Year 2008	
United States	60.9%	United States	63.2%	United States	66.1%
United Kingdom	15.9%	United Kingdom	15.3%	United Kingdom	10.5%
Nigeria	5.9%	Nigeria	5.7%	Nigeria	7.5%
Canada	5.6%	Canada	5.6%	Canada	3.1%
Romania	1.6%	Romania	1.5%	China	1.6%
Italy	1.2%	Italy	1.3%	South Africa	0.7%
Netherlands	1.2%	Spain	0.9%	Ghana	0.6%
Russia	1.1%	South Africa	0.9%	Spain	0.6%
Germany	0.7%	Russia	0.8%	Italy	0.5%
South Africa	0.6%	Ghana	0.7%	Romania	0.5%

Source: Internet Crime Complaint Centre Report (2006-2008) – <http://www.ic3.gov/media/annualreports.aspx>

What is not reflected in table 5 however, is the typology of these cyber crimes. For instance, even though the United States is consistently ahead of other nations in terms of cyber criminal activities as reflected in the table, the diversity of cyber criminal activities responsible for the figure includes advertisement spam, cyber pornography, and product marketing, bulk mails, and so on. However, a focus on cyber fraud in table 4 shows that monetary loss due to advance fee fraud tops the list of reported dollar loss to cyber criminal activities. This unfortunate scenario begs the question, Are these losses traceable to mails that emanate from Nigeria and other West African Countries?

We answered this question in Longe and Osofisan (2011) by using a modified version of the sinkhole methodology in real-time as proposed by Abhinav et al. (2008). Our e-mail accounts were activated and used from locations outside Nigeria, specifically in the United States of America, United Kingdom, and Canada to ascertain the source of these mails. Subscription to all forms of online promotion and marketing were avoided in order to avoid other forms of spam. We received over 36,000 fraud spam mails over a two-year period, and out of these we made a random selection of 400 for analysis. To find the actual locations from which the e-mails originated, we picked the "Received From" IP that is at the bottom of the list on the header view. We ran the e-mails through specific open-source e-mail and IP address tracers such as IPSLocator, E-Mail Trace, and IPGeocator. The results we obtained are shown in tables 6 and 7 as well as figures 10-13.

**Table 6: Different Forms of E-mail Harvested from Our Experimental E-mail Accounts**

Type	Spamming technique employed	Total no of mails (%)
Bayesian poisoning	E-mail with links to other websites with an e-mail body filled with meaningless sentences to confuse the Bayesian filters using tokens manipulations i.e AgeiNg numbe3rs etc.	35
Dating spam	E-mail in response to dating site activities asking for funds transfer, pictures or photographs of the recipient etc.	22%
Spoofing/phishing	Fictitious sender addresses usually direct users to respond to an entirely different address(es) from the sender address. Usually, the two addresses look similar. i.e. <i>akinadekolapo444@yahoo.com</i> <i>kolapoakinade444@yahoo.com</i>	17%
Academic spam	For conference invitation – variations in the form of e-mail address and the required e-mail address for communication	13%
Attachment only e-mail	These e-mails contain just an attachment or two with no contents at all	9%
Individually targeted e-mail	Electronic mails targeted at individuals with concrete correct information purportedly asking for funds transfers to a friend(s) in trouble or stranded somewhere far abroad	4%

**Table 7: Advance Fee Fraud E-mail Distribution by Continents**

Country	Percentage of 400 mails
AMERICA AND CANADA	28.5
EUROPE	23.2
AFRICA	20.4
SOUTH AMERICA	15.2
ASIA	9.5
AUSTRALIA	4.2

From table 6, one very disturbing trend in the spam war is the emergence of individually targeted e-mails and academic conference (invitation) spam e-mails. Quite a number of individuals have fallen victims to this genre of spam. The fraudsters use key-loggers and access to privileged information on individuals such as travel itineraries, to target unsuspecting victims.

Findings from our experiment as shown in table 7, also showed that advance fee fraud e-mails do not only emanate from Nigeria and some West African nations alone, but also from the Western world (America and Europe), the Middle East, and Asia. This result is very significant for us because it reflects the fact that the usual focus on Africa, and West Africa in particular, as the major source of fraudulent mailing and cyber activities is misleading. It remains to be investigated however if the increase noticeable in the volume of fraudulent spamming activities emanating from these other parts of the world correlates with the number or volume of Africans, Asians and other immigrants' moving into the Western nations, as some have alleged.

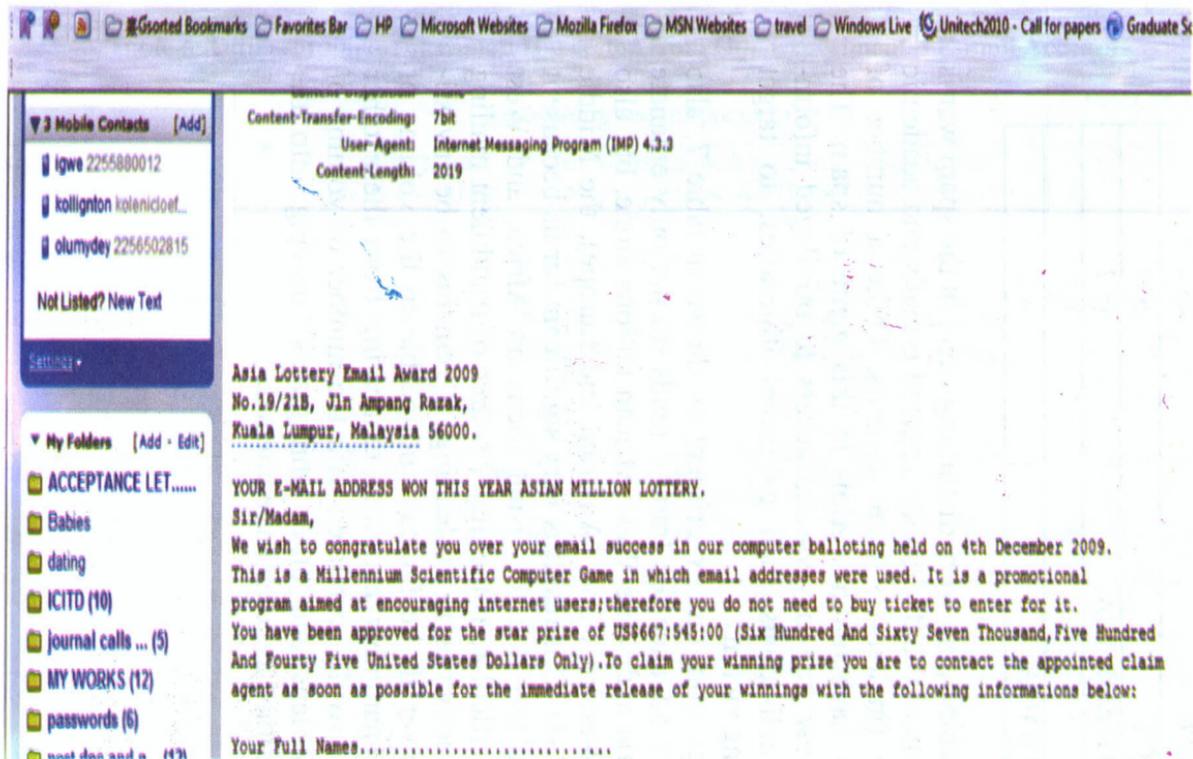


Fig. 10. Screenshot of e-mail purportedly from Kuala Lumpur.

**Email Tracing successful!**

**The email sender**

used at Fri, 04 Dec 2009 00:48:21 -0200 the IP address: **200.164.80.186** from **Brazil**.

**Email sender**

This Email was sent at Fri, 04 Dec 2009 00:48:21 -0200 from this IP address: **200.164.80.186**  
IP Location: in Brazil.

[Click here for a big satellite image.](#)

POWERED BY **Google** 50 m | 100 Mbps, data © 2009 MapLab/Tele Atlas - Terms of Use

**Email Sender IP address location & IP address info:**

IP address [?]:	<b>200.164.80.186</b> [Copy][Whois]
IP address country:	<b>Brazil</b>
IP address state:	n/a
IP address city:	n/a
IP address latitude:	-10.0000
IP address longitude:	-55.0000
ISP of this IP [?]:	Tele Norte Leste Participações S.A.
Organization:	Telemar Norte Leste S.A.
Host of this IP [?]:	ntsrvt003.etlce.ce.gov.br [Whois]

**Applicant Tracking Software**  
Applicant Tracking Software Info. Access 10 Search Engines At Once.  
[www.info.com](http://www.info.com)

**Instant Email Trace**  
1) Type In An Email Address 2) Find Owner's Name & Info  
[EmailFinder.com](http://EmailFinder.com)

**Easy Produce Traceability**  
The leading food traceability system from field to fork. See it work.  
[www.HarvestMark.com](http://www.HarvestMark.com)

Ads by Google

SHARE

Fig. 11. E-Mail trace result for the email showing it originates from Brazil.



Mail Search

[Previous](#) | [Next](#) | [Back to Messages](#)

Delete

Reply ▾

Forward

Spam

Move... ▾

**Access To Your ATM Card Has Been Suspended.**

From Interswitch Nigeria Limited &lt;A9&gt; Sat Oct 17 12:07:14 2009

**X-Apparently-To:** longeelumide@yahoo.com via 67.195.15.210; Sat, 17 Oct 2009 05:07:40 -t**Return-Path:** <onlineservices-email@interswitchng.com>**X-YahooFilteredBulki:** 160.36.0.84**X-YMailISG:** Rg2vGYeWLDsK8XHm9pWPFvGGGALn7YftqCuHyYujscpkj12Ee5HR2N51yj0Ck**X-Originating-IP:** [160.36.0.84]**Authentication-Results:** mta295.mail.re4.yahoo.com from=, domainkeys=neutral (no sig); from=; e**Received:** from 160.36.0.84 (EHLO jaguar.mail.utk.edu) (160.36.0.84) by mta295.ma**Received:** from User (adsl-99-154-19-102.dsl.ipvnc.a.abcglob.net [99.154.19.102]) (i**Message-Id:** <200910171207.n9HC788H008656@jaguar.mail.utk.edu>**From:** "Interswitch Nigeria Limited <A9>" <onlineservices-email@interswitchng.com>**Subject:** Access To Your ATM Card Has Been Suspended.**Date:** Sat, 17 Oct 2009 05:07:14 -0700**MIME-Version:** 1.0

Fig. 12. Letter purportedly from Interswitch Nigeria Limited.

IP Address	Country (Short)	Country (Full)	Flag	Region	City	ISP	Map
160.36.0.84	US	UNITED STATES		TENNESSEE	KNOXVILLE	UNIVERSITY OF TENNESSEE	
160.36.0.84	US	UNITED STATES		TENNESSEE	KNOXVILLE	UNIVERSITY OF TENNESSEE	
160.36.0.84	US	UNITED STATES		TENNESSEE	KNOXVILLE	UNIVERSITY OF TENNESSEE	

**Fig. 13.** Result from IP2Location on the validity of the IP address.

What is certain anyway is that as the Internet expands, opportunities for unethical use will continue to increase if nothing pragmatic is done in terms of policies, technology and legislation to protect users against online criminals. This is reflected in the routine activity theory which posits that crime can be motivated by opportunities provided in routine activities. Incidentally, the use of the web falls perfectly into the domain of routine activities, howbeit on a global scale (Cohen et al. 1979). But correlating poverty with cyber crime and using this as a major factor to point to Africa and some resource-poor environments as the sources of advance fee fraud mails does not capture the entire picture and could be totally misleading. Future research will explore the contributions of other subtle but extraneous factors, such as the economic meltdown and immigration policies, to the increase in fraudulent cyber activities now noticeable in the Western world.

### **Data Mining in the Service of UI**

In 2002, we obtained a 3-month fellowship from the UNDP to Swinburne University of Technology in Australia, and it was while there that we researched into a project in KDD using the University of Ibadan as case study. We in particular, set out to find a solution to the Vision goals that the University Governing Council set in 2001. The Vision Committee had set out to find answers to the following questions:

- Where were we before?
- Where are we now and how did we get there?
- Where do we want to be and how do we get there?

Interestingly enough, the University over the years had accumulated valuable data in various formats (text, video, audio, etc) which could be used to answer their questions. But unfortunately, the data were mostly gathering dust in several shelves, files, and storerooms, scattered and unharnessed together, all over the place. Within these data lies valuable information from which knowledge can be derived to find solutions to the three questions outlined by the Vision Committee. We came up with a KDD framework that we successfully demonstrated in a seminar presented at the Swinburne University of Technology for solving the three questions (Osofisan 2002).

Unfortunately, Mr. Vice-Chancellor Sir, the difficulties encountered in obtaining relevant data from appropriate units of the university has stalled the conclusion of this work. The framework for the Decision Support Tool is ready. It is a matter of urgency that UI must develop an enterprise Data Base for the University, and all I can say is that given the right cooperation, the Department of Computer Science will be willing to assist in getting it done. As a matter of fact, we had already proposed, as far back as ten years ago, a workable Data Warehouse Architecture for transactions in an academic environment (Osofisan 2000a). The Architecture is a hybrid of data marts that feed information into the central data warehouse (Osofisan 2000b). These data marts were designed with regard to cross-functional information requirements. The architecture is appropriate for any environment, such as ours, that consists of an inconsistent legacy system of islands of data. The proposed architecture combines the advantages of locality of reference, lower entry cost, faster implementation, and centralized system management.

## Students' Records

In 2009, in a bid to see how we can get around the problem of frequently missing or confused data, Olamiti A.O. and Osofisan A.O. embarked on another research of how to make sense out of some students' historical data available in the department. We investigated two treatment methods for dealing with missing values in students' data in the department, and found out that the mean error rate of Embedded-C4.5 method was 3.2% less than that of Pre-processing-CD method (Olamiti & Osofisan 2009a). The statistical t-test indicated that there is no significant difference between the two methods at 5% level of test. However, the two methods showed that there were some misclassifications in the final grades awarded to some students. Some fields that were expected not to have null values, such as the SSCE results in English, were found to be null. These are interesting facts that need to be further investigated. We would like to state that the historical data mostly fell within the period we classified as the period of decline in our previous study (Osofisan 2002). Further work was carried out on the same set of data (Olamiti & Osofisan 2009b), and the following results were derived:

1. Only students with grade A1, A2, or A3 in SSCE Mathematics graduated with First Class honours.
2. Students with grades of A1, A2, or A3 in SSCE Mathematics but who spent extra years before completing the programme graduated with at least Second Class Lower only if they took further Mathematics at SSCE.
3. All students who graduated with First Class honours had A1 in SSCE Chemistry.

Knowledge 1 is to be interpreted as meaning that if a student graduated with First Class, then the Student's grade in SSCE Mathematics is either A1, A2, or A3. The converse of this knowledge is not true, that is, all students with A1, A2, or A3 in SSCE Mathematics graduated with First Class.

Knowledge 2 supports this converse interpretation, but also gives another interesting knowledge—that any student who took Further Mathematics in SSCE would graduate with at least Second Class Lower degree in Computer Science. Knowledge 1 and 2 therefore underscore the importance of mathematics to Computer Science.

Knowledge 3 is very interesting too. It shows that all students who graduated with First Class degree had A1 in SSCE Chemistry. Now combining knowledge 1 and 3 means that we can make the following proposition: ***“Any student who comes into the Computer Science department with at least A3 in SSCE Mathematics and A1 in SSCE Chemistry is a potential First Class Student”***.

We will continue to investigate this proposition for validity with historical data from other universities in Nigeria. But if it proves to be true, then it may be necessary to find out the reasons behind the proposition for decision making in respect of science education and the production of highly skilled software engineers. Students who score any grade of A in SSCE Mathematics and A1 in SSCE Chemistry may therefore be targeted for special training in Software Engineering. This crop of Software Engineers may bring fortune to Nigeria in terms of production of world-class marketable software as India is currently doing. This might lead to wealth creation for Nigeria.

### **Contributions to the Outside Community**

Mr. Vice-Chancellor Sir, my contributions have not been limited to the university alone. In all modesty, I have also tried to use my knowledge in Computer Science to analyse existing problems in the outside society and create solutions to them. For the rest of this lecture, I will try to give a few instances of these.

### **The School System**

The first concerns schools and the educational system. In 2003, we decided to focus our research effort in knowledge discovery on this area by using historical data of JSS examination results in one state to determine the success or

failure of the 6-3-3-4 Educational system, most specially, the 3-3 segment. We created a data warehouse consisting of student enrollment, performance, local government, school ownership, core subjects, pre-vocational and elective (non-vocational) subjects as specified in the National Policy on Education Revised (1982). We found that overall performance P can be any of the following eight values:

- All Rounder (good in core subjects, pre-vocational and elective) – A
- Science/Humanities – S/H
- Science/Vocational – S/V
- Humanities/Vocational – H/V
- Science only – S
- Humanities only – H
- Vocational only – V
- No classification – N

The Data warehouse relation, School (S), has three components:

- Facts
- Dimension
- Attributes

Facts attributes (LGA-ID, Sch-ID, Std-ID, Time-ID, Name, Sex, Age, Results in courses registered for, P)

Dimensions (School, LGA, and Time)

Attributes (Ownership, Location, School Type)

We then used Classification and Clustering technique to find data groupings, data dependencies, relationships, as well as classifications.

The result was that we found a cluster representing about 25.9% that could not proceed to Senior Secondary School, and another cluster of about 3% that did not fit into any academic classification even though they passed six subjects. These two groups should be candidates for trade centers as proposed by the National Policy on Education. This means that the state government should make provision for 28.9% of JSS students to enter various trade centers in the state.

However, the statistics of the number of trade centers in the state showed that only 14.5% of these students could be absorbed by the available trade centers. This means that, for government to make a success of the 3-3 segment of the educational system in the state, trade centers must be increased by about 620% (Osofisan et al. 2003). The state therefore needed to set up many more trade centres for various artisans in areas like plumbing, carpentry, electrical fittings, and so on, that is, artisans in the building industry, who are almost going into extinction nowadays.

## **Banks**

In 2004, because of the ailing nature of the Bank Industry in Nigeria, we investigated the Portfolio Management of one of the Banks, also using the KDD procedure (Osofisan, Inanga, & Adeyemo 2004). We found out that the Central Bank of Nigeria Policy as well as the internal management of the bank affected the bank's portfolio management. Precisely, we found out that no bank can adhere strictly to all the CBN guidelines, because some of the CBN policy guidelines are conflicting. For example the credit expansion goal and capital adequacy goals should complement each other, but analyses show that they conflict. The analyses also show that bank managements tend to set unrealistic standards. The basic result of goal achievement is shown in table 8 while the results for variants 1 to 5 of the basic model are illustrated in tables 9-11.  $P_k=0$  implies that the goal at the  $k$  priority level is achieved, while  $P_k=1$  means that the goal at the  $k$  priority level is not achieved. None of the variant models gave the same level of goal achievement as the basic model solution, and none of them was identical to another (see tables 9-11). The results of these various combinations explain the effect of changes in the internal policy on the performance of the bank. Also in all the variants as well as in the basic model, at least one of CBN statutory requirements was routinely violated.

**Table 8: Degree of Goal Attainment of the Basic Model**

Priority level	Constraints	Goal achievement				
		T1	T2	T3	T4	T5
0	Artificial	Y	Y	Y	Y	Y
1	Fund allocation	Y	Y	Y	Y	Y
2	Sector credit allocation	Y	Y	Y	Y	Y
3	Other statutory goals	N	N	N	N	N
4	Customer deposit mix/profit	Y	Y	Y	Y	Y
5	Principal ratios	N	N	N	N	N
6	Profit/deposit growth	Y	Y	Y	Y	Y
7	Low bound constraints	Y	Y	N	Y	N
8	Balance sheet	Y	Y	Y	Y	Y

**Table 9: Sensitivity Analysis Results for Time Periods for Variants 1 & 2**

S/N	Constraints	Revised model combination									
		Variant 1					Variant 2				
		1	2	3	4	5	1	2	3	4	5
1	Fund availability	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$
2	Sector credit allocation	$P_3=1$	$P_4=1$	$P_5=1$	$P_2=0$	$P_2=0$	$P_3=1$	$P_4=1$	$P_5=1$	$P_2=0$	$P_2=0$
3	Other statutory allocation	$P_2=0$	$P_3=1$	$P_3=0$	$P_3=1$	$P_3=1$	$P_2=0$	$P_3=1$	$P_3=1$	$P_3=1$	$P_3=1$
4	Customer deposit mix/profit	$P_5=1$	$P_5=1$	$P_6=1$	$P_6=1$	$P_6=0$	$P_5=1$	$P_5=0$	$P_6=1$	$P_6=1$	$P_6=1$
5	Principal ratios goals	$P_4=1$	$P_2=1$	$P_4=1$	$P_4=0$	$P_5=0$	$P_4=1$	$P_2=1$	$P_4=1$	$P_4=1$	$P_5=1$
6	Profit and deposit growth goals	$P_6=1$	$P_6=1$	$P_7=0$	$P_7=0$	$P_7=0$	$P_6=1$	$P_6=0$	$P_7=1$	$P_7=0$	$P_7=0$
7	Lower bound constraints	$P_7=1$	$P_7=1$	$P_2=0$	$P_5=1$	$P_4=1$	$P_7=1$	$P_7=1$	$P_2=0$	$P_5=1$	$P_4=1$
8	Balance sheet identity	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$

**Table 10: Sensitivity Analysis Results for Time Periods for Variants 3 & 4**

S/N	Constraints	Revised model combination									
		Variant 1					Variant 2				
		1	2	3	4	5	1	2	3	4	5
1	Fund availability	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$	$P_1=0$
2	Sector credit allocation	$P_3=1$	$P_4=1$	$P_5=1$	$P_2=0$	$P_2=0$	$P_3=1$	$P_4=1$	$P_5=1$	$P_2=0$	$P_2=0$
3	Other statutory allocation	$P_2=0$	$P_3=1$	$P_3=0$	$P_3=1$	$P_3=1$	$P_2=0$	$P_3=1$	$P_3=1$	$P_3=1$	$P_3=1$
4	Customer deposit mix/profit	$P_5=1$	$P_5=1$	$P_6=1$	$P_6=0$	$P_6=0$	$P_5=1$	$P_5=0$	$P_6=1$	$P_6=1$	$P_6=1$
5	Principal ratios goals	$P_4=1$	$P_2=0$	$P_4=1$	$P_4=0$	$P_5=0$	$P_4=1$	$P_2=1$	$P_4=1$	$P_4=1$	$P_5=1$
6	Profit and deposit growth goals	$P_6=1$	$P_6=1$	$P_7=0$	$P_7=0$	$P_7=0$	$P_6=1$	$P_6=0$	$P_7=1$	$P_7=0$	$P_7=1$
7	Lower bound constraints	$P_7=1$	$P_7=1$	$P_2=0$	$P_5=1$	$P_4=1$	$P_7=1$	$P_7=1$	$P_2=0$	$P_5=1$	$P_4=1$
8	Balance sheet identity	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$	$P_8=0$

**Table 11: Sensitivity Analysis Results for Time Periods for Variant 5**

S/N	Constraints	Revised model combination				
		Variant 5				
		1	2	3	4	5
1	Fund availability	P <sub>1</sub> =0	P <sub>1</sub> =0	P <sub>1</sub> =0	P <sub>1</sub> =0	P <sub>1</sub> =0
2	Sector credit allocation	P <sub>3</sub> =1	P <sub>4</sub> =1	P <sub>5</sub> =1	P <sub>2</sub> =0	P <sub>2</sub> =0
3	Other statutory allocation	P <sub>2</sub> =0	P <sub>3</sub> =1	P <sub>3</sub> =0	P <sub>3</sub> =1	P <sub>3</sub> =1
4	Customer deposit mix/profit	P <sub>5</sub> =1	P <sub>5</sub> =1	P <sub>6</sub> =1	P <sub>6</sub> =0	P <sub>6</sub> =0
5	Principal ratios goals	P <sub>4</sub> =1	P <sub>2</sub> =1	P <sub>4</sub> =1	P <sub>4</sub> =0	P <sub>5</sub> =0
6	Profit and deposit growth goals	P <sub>6</sub> =1	P <sub>6</sub> =1	P <sub>7</sub> =0	P <sub>7</sub> =0	P <sub>7</sub> =1
7	Lower bound constraints	P <sub>7</sub> =1	P <sub>7</sub> =1	P <sub>2</sub> =0	P <sub>5</sub> =1	P <sub>4</sub> =1
8	Balance sheet identity	P <sub>8</sub> =0	P <sub>8</sub> =0	P <sub>8</sub> =0	P <sub>8</sub> =0	P <sub>8</sub> =0

All these show therefore that, for the banking system to function properly, the CBN Monetary Policy and bank managements policies must be scientifically determined, based on historical facts.

### **The Lagos Ibadan Express Way**

In 2005, we turned our attention to the issue of accidents on the Lagos-Ibadan Express Road. We wanted to look for interesting patterns in the nature of the accidents that had occurred in the past. Historical data about road accidents were therefore collected from the Road Safety Commission for the first 40 kilometers from Ibadan to Lagos. We employed Neural Network using Multilayer Perception as well as Multidimensional Data Analysis method and discovered that the dark spot on the first 40 kilometers of the Ibadan-Lagos road is between kilometers 10-20 where 60.2% of all accidents took place. We found three interesting clusters at 0-9 kilometers, 10-20 kilometers, and 21-40 kilometers. All the accidents due to wrong overtaking took place in cluster 2, while no accident caused by robbery took place in cluster 1. One third of the accidents that occurred in cluster 1 were caused by loss of control, followed by over-speeding and dangerous/careless driving. Therefore, the accidents that occurred in cluster 1 can be attributed to the drivers, and may therefore be drastically reduced if drivers have the right attitude to driving.

Cluster 2 has a fair share of all causes of accidents, with tyre burst (71.3%) being the highest, followed by over-speeding and loss of control. Going by the fact that in cluster 1, only 2.1% of the accidents occurred as a result of tyre burst, with 71.3% in cluster 2, and 26.6% in cluster 3, then tyre burst in cluster 2 may be a result of other extraneous causes. It may therefore be necessary to further investigate cluster 2 in order to identify these extraneous causes in order to reduce accidents in this cluster.

In all the clusters, most of the accidents occurred in the morning (41.4%), and in the afternoon (42.1%). Only 16.5% of the accidents occurred at night. Furthermore, there were more accidents in the dry season than in the wet season. Then, no accident due to robbery occurred at night, and big vehicles were more involved in accidents than small vehicles in the ratio of 59.9:40.1. This is useful information for the authorities concerned. The results show the measures to be taken to reduce accidents and make the roads safe (Osofisan, Komolafe, & Akinola 2005).

### **University Administration**

Mr. Vice-Chancellor Sir, I wish to turn now very briefly to our modest contributions in the areas of administration. When in September 1999, I became the Acting Head of Computer Science Department, there were only 5 lecturers, and I was the only one among them with a PhD degree. Hard decisions had to be taken therefore on how to move the department forward. The first decision was to start a postgraduate programme, but how? After a lot of reflection and consultation, we decided that our best option was to start with MPhil/PhD conversion for all staff in the department with an M.Sc. degree. The next question to be answered then was: With me as the only supervisor, would all of them now be constrained to do their research in my area and so end up with making the department a monolithic research department? The answer of course had to be in the negative, since that would not be good for the system and I was determined to build a robust department. Then God, Who sees the heart in His infinite mercy, sent help to us.

This help came through the following: Prof. A. David of the University of Nancy 2 in France along with his colleagues, Prof. Odile Thiéry, and Dr. G Duffing; Dr. V. H. Dang of IIST, United Nations University, Macau; Prof. A. B. Sofoluwe (Current Vice-Chancellor of Unilag), Prof. O. Abass; Prof. J. O. A. Ayeni and Prof. C. O. Uwadia, all of University of Lagos; Drs. E. R. Adagunodo and G.A. Aderounmu, now both professors of Obafemi Awolowo University, Ile-Ife. With the assistance of this selfless lady and the gentlemen, we were able to diversify our research base in the department. It also meant my shifting slightly away from my original research areas in Data Communication and Data Base to other areas of research in Computing since I had to co-supervise PhD work in these other areas with colleagues within and outside Nigeria.

I take pride in the fact that, within the time of my tenure, seven of the members of staff listed below have graduated with PhD degrees within the department.

- OLUWADE, Bamidele Ayodeji in Code Theory (“Design and Analysis of Computer Coded Character set”, 2004)
- AKINKUNMI, Babatunde Opeoluwa in Knowledge Representation (“Temporal Properties and Application of Recurrent Entities”, 2005)
- OSUNADE Oluwaseyitanfunmi in Data Communication (“Data Migration Patterns for JAVA-Based Mobile Agent Systems”, 2007)
- OKIKE, Ezekiel Uzor in Software Engineering (“Measuring Class Cohesion in Object-Oriented Systems Using Chidamber and Kemerer Metrics: Java as Case Study”, 2007)
- AKINOLA, Solomon Olalekan in Software Engineering (“A Comparative Study of the Effectiveness of Checklist Based and Ad-Hoc Software Inspection Reading Technique in Paper and Tool-Based Environments”, 2010)

- ONIFADE, Olufade Williams in Intelligent Systems (“A Model for Information Risk Management in Economic Intelligence Systems”, 2010)
- OLADEJO, Fausat Bolanle in Knowledge Capitalization (“User-Centred Capitalization of Knowledge in the Context of Economic Intelligence System”, 2010)

I am also happy to say that, apart from the 6<sup>th</sup> and 7<sup>th</sup> candidates who were co-supervised by me with Professors A. David and Odile Thiéry respectively, all the remaining five PhD candidates were solely supervised by my humble self.

Furthermore, the first PhD student, B. A. Oluwade has since joined the services of Salem University in Kogi State as a Professor. He is here with us in the audience today. Three other members of staff also got their PhD degrees as listed below:

- ROBERTS, Charles Abiodun in Systems Modelling (‘L’annotation Pour La Recherché D’information Dans Le Context D’intelligence Économigne’. University of Loria 2 France, 2007).
- ADEYEMO, Barnabas Adesesan in Data Mining (“Development of Data Mining System for Oil Well Lithology and Fluid Contents Analysis”. Federal University of Technology, Akure, 2008).
- LONGE Olumide in Cyber Crime (“SPAMAng – A Domain Specific Outbound Antispam system for Filtering Fraudulent Electronic Mail”. University of Benin, 2010).

The result, Mr. Vice-Chancellor Sir, is that today, our department can boast of a staff body of 21, 9 of whom have PhDs. Sir, there is no other Computer Science Department in Nigeria to my knowledge that can boast of a similar strength.

### **Conclusion**

Mr. Vice-Chancellor Sir, I wish to emphasize that, talking about these projects and challenges, my aim is to showcase

the almost limitless possibilities that exist within the field of Computer Science, and especially of KDD, to grapple with, and offer solutions to the many problems that beset us in this nation. Sir, the creation of Enterprise Data Bases and of Data Warehouse is a must in every sector of the Nigerian enterprise in this era of knowledge economy.

As many nations have discovered, the organization of life and progress in the modern world is almost impossible without the systematic gathering and organization of information. That is what is meant when people say we live in the information age. Unfortunately, we in this part of the world tend to be careless about the collection of data, and extremely nonchalant towards the management of information. This is principally why we remain backward in the 21<sup>st</sup> century and why we continue to suffer the terrible nightmare of underdevelopment. But we have the means of changing our lives for the better and joining the globalization train not just as passengers, but as drivers and pilots. One of the weapons in our hands, Mr. Vice-Chancellor, is Information Technology (IT), and by IT, I mean the Convergence of the Telecommunication and Computing Industries. I hope, in the brief scope of this lecture, I have been able to demonstrate sufficiently how IT can surely transform our Data Dust to Data Gold!

### **Acknowledgements**

I now wish to make a few acknowledgements. First and foremost, I acknowledge the Almighty God, my Father and my Maker Who has moved in many mysterious ways in my life for His wonders to perform.

I thank my father, Late Hon. Chief Josiah Orisabinu Adedipe, the Elemo of Akure and Mayegun of Osogbo, (*aka Yagbo Yaju*, Osun Osogbo) who gave me not only sound education, but also moral uprightness. Papa, you taught your children to respect all and fear only God. You taught us dignity of labour, and the pursuit of excellence in all that we do. Above all you gave equal opportunities to all your children whatever their gender, and as a result Papa, you have

three female professors (daughters) to your credit. May I here humbly mention and thank my two sisters as follows:

- Prof. Adefunke Oyemade (Medicine). First Female Medical Doctor in the old Ondo Province of the defunct Western State;
- Prof. Adeola Abaelu (Bio-Chemistry), one of the first two students to be awarded the PhD degree of the University of Ife, now the Obafemi Awolowo University, Ile-Ife; and
- I am, of course, the third professor after them, and the very First Female Professor of Computer Science, in all of sub-Saharan Africa.

Papa, no legacy is as good and precious as a successful child. We can never thank you enough.

As for you, my mother, Late Chief Mrs. Felicia Fehintola Adedipe (Nee Ogundare, and the former Akuwajo of St. Paul's Anglican Church, Emure Ekiti), words alone cannot express my gratitude to you. You were selfless and big-hearted. You took care of all our children while they were growing up. You were the "Ark of God" in our household.

Your Royal Highness, Oba Adebisi Adesida, the Deji of Akureland, I feel highly honoured by your presence. I remain truly your obedient subject. To all the members of the Adedipe Dynasty of Akure land here present, led by our Patriarch, High Chief Hon. Bola Adedipe, the Elemo of Akure land and former Minister of the Federal Republic of Nigeria, I say a big thank you, to you all. Our family is a special one in Akure, whose enduring hallmark is love for one another. We are always there for each other through thick and thin. I thank God for creating me to be an Adedipe, and now an Osofisan.

I thank my brother in-law, Chief Adefemi Adekanye and my sister, Chief Mrs. Adefunmilayo Adekanye, the Erelu of Akure land for their love and particularly for looking after Wale and Yemi while I was away studying at the Georgia Institute of Technology, Atlanta, Georgia.

I thank all my relations from Emure-Ekiti here tonight. I remember with nostalgia, my maternal grandmother, Late Mrs. Alice Siyanbola Ogundare, who taught me lots of words of wisdom. I can never forget her telling me that '*Inu mimo ja ju ogun lo*'. This statement has been proven right in my life many times over.

I sincerely appreciate the entire Osofisan family, and its extensions, especially those here present. You accepted me the way I am as a daughter and not as a wife. These include Prof. Philip Babatunde Osofisan, my husband's older brother, and his other siblings—Mr. Olusola Osofisan, now in the USA, and Mrs. Dupe Ilori, a retired civil servant. They also include their spouses and numerous relations all over the world. I have indeed blossomed among you. I also appreciate my in-laws – the Akinrindes; the Nwabuezes; and the Usidames. Thank you for giving me three more lovely daughters—Jumoke, Uzo and Ehi. I appreciate my brother Engr. Leke Adedipe, retired staff of Shell Petroleum Development Company, who accompanied me to my new marital home in 1974; my cousin, Hon. Mrs. Bunmi Adelugba, a Commissioner in Ekiti State, and my brother, Hon. Adegboyega Adedipe, the present Ondo State Secretary of the Action Congress of Nigeria—who were the first “two children” I raised. I am proud of you. Words are not enough to express my deep love for late Funmi Jeyifous *aka* “Funmi sharp” my late adopted daughter, may your soul rest in perfect peace. I appreciate the dedication of my wards, Elijah and Anu as well as my nephew Akinyemi who are currently staying with me.

I thank all the teachers who have taught me right from primary school to the university. I wish to appreciate most especially, Miss Rosaline Jane Pelly, my principal at Fiwasaiye Girls' Grammar School Akure, for teaching her students time management. I thank Dr. Olumide Kuti, the then Comprehensive High School Aiyetoro Guidance Counsellor, for asking me to change my A' Level subjects from Physics, Chemistry and Biology to Physics, Chemistry

and Mathematics so that I could study Computer Science, even when Computer Science was not studied in Nigeria then in 1968. I appreciate Dr. Isaac Odeyemi for making my studying Computer Science possible. I remember vividly that day at the University of Ife Computing Centre in 1972, when I visited you and you requested me to change my course from Physics honours to Computer Science/Economics that was just starting then in September 1972. I remember the argument that took place between you and Prof. Abiodun Oluwole (NNOM) who would rather have me study Computer Science/Mathematics. My innate spirit chose Economics even though I knew nothing about the subject. That was how I started a journey into the world of two unknowns—one that controlled world economy then and the other that would move the world economy into that of knowledge. It is amazing!

My gratitude goes to my PhD Supervisor, Prof. Bishop Adebayo Dada Akinde, for facilitating my goal of having a PhD degree in Computer Science. As I said earlier, little did we know that history was in the making that day in your office in 1984, that I would become the first female to have a PhD degree in the Faculty of Technology, Obafemi Awolowo University Ile-Ife, as well as the first female to have a PhD degree in Computer Science in Nigeria. God's hands were obviously working behind the scene.

I thank those who in one way or the other contributed to the success of my PhD work, among whom I must mention Prof. Biodun Jeyifo, my husband's bosom friend, and my cousin, Mr. Tubosun Adedipe living in USA, especially for paying for my subscriptions to various international journals, conference proceedings and books. I also thank the Joint Komputer Kompany (JKK), DEBIS Nig Ltd., and the then DATAMATICS owned by Chief Dapo Sarumi, former Minister of the Federal Republic of Nigeria, for sponsoring my attendance at the 10<sup>th</sup> World Computer Congress in Dublin in 1986; and the International Council for Computer Communications (ICCC) for sponsoring my attendance at all ICCC conferences between 1986 and 1989.

I thank all my classmates at the University of Ife, in particular, the foundation set of Computer Science Education in Nigeria – Yemisi Osiname (née Odedeyi), Saka Fawole, Femi Aladesulu, Funso Akinniyi, Gbadebo Adeleke, Charles Ikemere, Kola Ogunlana, Gbenga Okusaga, and Taiwo Yakubu, for your comradeship. I thank every member of Fiwasaiye Old Girls' Association (FOGA), most especially, Wura Ajibola for always being there for me.

My special appreciation goes to Prof. Akinbo Adesomoju for being instrumental to my moving from the Polytechnic Ibadan to the University of Ibadan in 1997, and for believing in my ability. Prof. Funso Olorunsogo, my technical classmate at All Saints' Primary School, Osogbo, thank you for encouraging me to stay on in UI when the road was very rough and the environment extremely hostile. Your godly statement—"You are not going back to Egypt"—struck a chord within me and I resolved to cross the Red Sea, fight all the 'Canaan's kings' and pull down the walls of Jericho with God by my side and He, in His infinite mercy, saw me through. May God continue to use you both as vehicles of grace. To Professors O. Osonubi, L. A. Hussain, A.T. Hassan, R.A. Oderinde and T. Odiaka, I say thank you all for the roles you have played in my life. And to all the 'Canaan's kings', I also say thank you.

To every member of staff and student of Computer Science Department, past and present since I joined the services of UI, I say God bless you all, especially my postgraduate students and my colleagues collaborating with me in research, who have been sources of inspiration to me. You will all reach the top of your profession in Jesus' name (Amen).

Many thanks to my professional colleagues in the Computer Professional Registration Council of Nigeria (CPN), the Nigeria Computer Society (NCS) and the Nigerian Institute of Management (NIM) (chartered). I thank most especially, Prof. Senator Iya Abubakar and Senator Joy Emordi for the role they played while I was the President and Chairman of CPN. During my tenure, the Council was transformed from a

one-office affair in Lagos, with just five members of staff, to a nationwide organization with nine zonal offices in Lagos, Abuja, Owerri, Jos, Ibadan, Port Harcourt, Gombe, Kano and Yola, and a total staff membership of 78! The Council was also able to acquire 5.404.03sq.mtr of land in Abuja, at least one vehicle for each zonal office, and a modern e-library, all within the 4 years of my stay in office. To God be the Glory. Great things He has done.

My appreciation goes to my brothers and sisters in Christ, at The Chapel of the Resurrection, University of Ibadan, most especially members of the Resurrection Lilies; to the entire congregations of the Church of Transfiguration, Anglican Communion at Ikolaba Ibadan and St. David's Cathedral Akure, as well as every member of the Bible Study Fellowship, Ibadan EW and EM.

To my Children – Wale (“Oshofs1”), Yemi (“Shoffybrown”), Akin (“Oshofs3”) and Tomi (“SisOshofs”), MamOshofs says you are precious and dearly loved and my little Oshofs –Toni, Mayo, Rayoke, Todun and Lase—you are bundles of joy. My darling husband, Femi “the Elereko”, you are the love of my life, my pillar of strength and the husband of my youth. You have supported me all through my entire academic and professional career. So I say to you:

*Like an apple tree among the trees of the forest  
is my lover among the young men.*

*I delight to sit in his shade,  
And his fruit is sweet to my taste.  
He has taken me to the banquet hall,  
And his banner over me is love*

—Song of Songs 2:3-4

Now to God Immortal, Invisible, the only Wise, be all honour and glory now and for evermore. Mr. Vice-Chancellor Sir, ladies and gentlemen, I thank you for listening.

## References

- Adagunodo, E.R. Akinde, A.D. & Osofisan, A.O. (1999) Protocols specification & verification: A graph – Theoretic model approach. *The Journal of Computer Science and its Application*. 7(1) June, pp. 55 – 64.
- Akinde, A.D. & Osofisan, A.O. (1994) A conceptual design and specification of a data handling computer network. *Ife Journal of Technology* 4(1) pp. 49 – 53.
- Bartee, T.C. (1985) *Data communications, networks and systems*. New York: Howard W. Sams & Co.
- Berry, J.A. & Linoff, G. (1997) *Data mining techniques*. New York: John Wiley & Sons, Inc.
- Cohen, Lawrence, and Marcus Felson (1979) Social change and crime rate trends. *American Sociological Review* 44: 588. doi:10.2307/2094589.
- Cukier Wendy, Eva J. Nesselroth, Susan Cody (2007) “Genre, Narrative and the Nigerian Letter in Electronic Mail”. *Proceedings of the 40th Annual Hawaii International Conference on System Sciences HICSS 2007*: 70.
- Date, C.J. (2001) *An introduction to database systems*. Seventh edition. New Delhi: Addison-Wesley.
- Fausset, L. (1994) *Fundamentals of neural networks*. Upper Saddle River, NJ: Prentice Hall.
- Fayyad, U., Piatetsky-Shapiro, G. and Smyth, P. (1996) From data mining to knowledge discovery in databases AAAI 1 pp. 37-54.
- Fayyad, U. Wiese, A. and Grinstein, G.G. (2002) *Knowledge discovery and information systems*. San Fransico: Morgan Kaufmann Publishers.
- Federal Govt, Nigeria (1982) *National Policy on Education Revised*.
- Ferrucci, D. & Lally, A. (2004) “UIMA: An architectural approach to unstructured information processing in the corporate research environment” *Natural Language Engineering* 10(3/4), pp 327-348.
- Frawley, W., Piatetsky-Shapiro, G., & Matheus, C. (1992) Knowledge discovery in databases: An overview. *AI Magazine*, Fall 1992, pgs 213-228.
- Gbenga, S. (2007) “The growing Menace of Cyber Crime in Nigeria-Causes, Effects & Solutions”  
<http://events.tigweb.org/12657>

- Goldberg, D.E. (1989) *Genetic Algorithms*. Reading M.A.: Addison Wesley.
- Haykin, S. (1994) *Neural networks: A comprehensive foundation*. New York: Macmillan.
- Hofstetter, F.T. (2005) *Internet technologies at work*. Boston: McGraw Hill Technology Education.
- Horak, R. (1999) *Communications systems and networks*, 2<sup>nd</sup> ed. Chicago: M&T Books.
- Igwe, C. (2007) "Taking Back Nigeria from 419: What to do about the Worldwide E-mail Scam - Advance Fee Fraud". *Bloomington: iUniverse*.
- Longe, O. & Osofisan, A. (2011) On the origins of advance fee fraud electronic mails: A technical investigation using internet. *Information Systems* 3(1), pp 17-26.
- Lucey, T. (1997) *Management information systems*. London: DP Publications.
- Mannino, M.V. (2001) *Database application development & design*. Boston: McGraw Hill.
- Nisbert, R., Elder, J. & Miner, G. (2009) *Handbook of statistical analysis & data mining applications*. Elsevier Inc. Amsterdam.
- O'Brien, J.A. (2003) *Introduction to information systems*. Irwin: Boston: McGraw Hill.
- Olamiti, A.O. & Osofisan, A.O. (2009a) "Academic background of students and performance in a computer science programme in a Nigerian university". *European Journal of Social Science* 9(4), 2009.
- \_\_\_\_\_ (2009b) Experimental comparison of missing value treatment methods in students' enrolment data. *European Journal of Scientific Research* 33(4).
- Osofisan, A.O. (1996) Designing and writing Web documents: A case study of NUMAPS. *INFOTECH* (no. 2) pp 184 – 198.
- \_\_\_\_\_ (2000a) A virtual data warehouse architecture for academic environment. *Journal of Science Research*. 6(2) pp 122 – 125.
- \_\_\_\_\_ (2000b) "Creating Synergy on the Net" In *Computer Association of Nigeria Conference Proceedings on Deployment of Telematics Systems: Trends, Techniques and Tools*, Vol. 11, Eds. C.O. Uwadia, H.O.D. Longe, and A.D. Akinde, pp. 255 – 261.

- Osofisan, A.O. (2001) Performance trace analysis of abstract co-operative caching in web servers. *The Journal of Computer Science and its Application*. 8 (1) June pp. 13 – 16.
- \_\_\_\_\_ (2002) “Application of Data Warehousing Technology to Economic Intelligence: A case Study of University of Ibadan” In *Information and Communication Technology Applied to Economic Intelligence ICTEI’2002* Institut National De Recherche en Informatique et en Automatique, INRIA, Lorraine. pp 21– 31.
- Osofisan, A.O. & Fils, D. (1996) “Network for Use in Mathematics and Applied Physical Science”. *Computer Association of Nigeria Conference Proceedings on Trends in Computing: Hardware, Software and Networking*. Vol. 7. Eds. E.R. Adagunodo, A.D. Akinde, L.O. Kehinde and T. Onabanjo, pp. 85 – 115.
- Osofisan, A.O. & Idowu, S.A. (2009a) “Content and Service Delivery Techniques on the Internet: Trend of Development and Characterization” *IJAS Conference for Academic Disciplines* held on the Suffolk University campus in downtown Boston (June 22-25, 2009).
- \_\_\_\_\_ (2009b) QoS-based request-routing among peering content distribution networks (CDNs) in a Virtual Organization (VO)-Based model *IJCSIS* 1(1), May.
- \_\_\_\_\_ (2011) QoS-based objects replica placement among peering Content Distribution Networks (CDNs). *IJCSI* 8(2), March. ISSN (Online): 1694-0814 [www.IJCSI.org](http://www.IJCSI.org) pp 490-500.
- Osofisan, A.O., Inanga, E.C. and Adeyemo, O.O. (2004) Bank asset portfolio management model using goal programming techniques. *ICMCS*. 1, pg. 171-195.
- Osofisan, A.O., Komolafe, O.P. & Akinola, S.O. (2005) Discovering knowledge in road accident database. *ICMCS* 2, pp 31-43, 2005.
- Osofisan, A., Ajayi, F., Adeoye, A., Akinola, O., & Olamiti, A. (2003) “Data Warehouse Application in Education Information System: A Case Study of Junior Secondary Educational System (Oyo State, Nigeria)” In *Intelligence Economique: Recherches & Applications Formation and Communication IERA’2003 a l’INIST-CNRS*, Nancy, France, Institut National De Recherche en Informatique et en Automatique, INRIA, Lorraine. pp 79–99.

- Osunade, O., Adeyemo, A. & Osofisan, A.O. (2001) "NETBILL: An Ecommerce Solution for Sale of Information" In *Computer Association of Nigeria Conference Proceedings on Impact of e-Commerce on National Economy and Development (E-Commerce-ned)*. Eds. S. Juwe, H.O.D. Longe and A.D. Akinde, Vol. 12 pp. 165 – 180.
- Profgame Blogspot (2007) "Smile out of Poverty – Cyber Crime in Nigeria – A Sociological Analysis:"  
<http://profgame.blogspot.com/2008/08/cyber-crime-in-nigeria-sociological.html>
- Schwartz, M. (1977) *Computer-communication network design & analysis*. Prentice-Hall Inc.
- Speckt, D.F. (1991) A generalized regression neural network. *IEEE Transactions on Neural Networks* 2(6), pp 568-576.
- Tanenbaum, A.S. (1996) *Computer networks*. New Jersey: Prentice-Hall International.
- Wiederhold, G. (1981) *Database design*. Johannesburg: McGraw-Hill International Book Company.
- Wikipedia (2011) *Computer network topologies*.  
<http://www.en.wikipedia.org> 6/8/2011.
- Witten, I. H. & Frank, E. (2000) *Data mining: Practical machine learning tools and techniques*. New York: Morgan Kaufmann.

## **BIODATA OF PROFESSOR ADENIKE OYINLOLA OSOFISAN**

Adenike Oyinlola OSOFISAN was born on 11<sup>th</sup> March 1950 in Osogbo, to the late Chief Honourable Josiah Orisabinu Adedipe, the Elemo of Akure and Mayegun of Osogbo and late Chief (Mrs.) Felicia Fehintola Adedipe (nee Ogundare), the Akuwajo of St. Paul Anglican Church Emure Ekiti.

She started school at All Saints Primary School, Osogbo (1956-59); and then transferred to Saint Steven's Primary School, Akure (1960-61). Then, for her secondary education, she went first to Fiwasaiye Girls' Grammar School in Akure (1962-67) where she finished in First Division and as the first student of the school to have a grade of distinction in Mathematics. Then for her Higher School Certificate education, she attended the Aiyetoro Comprehensive High School, spending an extra year (1968-70) in order to convert from Arts to Science.

Adenike was a pioneer student of the department of Computer Science of University of Ile-Ife (now Obafemi Awolowo University), and graduated with a 2<sup>nd</sup> Class Upper Division Honours Degree in Computer Science/Economics. She also has a Masters degree in Computer Science in 1979 from Georgia Institute of Technology in Atlanta and PhD Computer Science in OAU to become the first Nigerian woman to hold a PhD degree in Computer Science and the first female to have a PhD degree from the Faculty of Technology of Obafemi Awolowo University Ile-Ife. She also has the MBA (Finance & Accounts) of University of Ibadan in 1993, where she graduated with unprecedented nine distinctions.

Prof. Osofisan worked at the University of Ibadan as a Programmer between 1976 and 1978. She started her teaching career as Lecturer II at The Polytechnic, Ibadan in 1979. She was promoted Lecturer I in 1981; Senior Lecturer in 1984; and was appointed, Head Department of Computer Science in 1986, a position she held till 1995. While in the saddle, she

was promoted to the position of Principal Lecturer in 1990. Through dint of hard work, Prof. Osofisan rose to the highest position of Senior Principal Lecturer at the Polytechnic, Ibadan in 1993. She was head of Computer Studies for nine years (1986-1995) and Dean of Faculty of science for two years (1995-97) before she transferred her services to UI as a Senior Lecturer. She was promoted Reader in 2003 and Professor in 2006 to become the first black African Female Professor of Computer Science in Sub-Saharan Africa.

Professor Osofisan was the Pioneer President, Nigeria Women in Information Technology (IT) and first female President and Chairman of Council, Computer Professionals Registration Council of Nigeria (CPN). Prof. Osofisan was the longest serving female member on the National Executive Council of Nigeria Computer Society.

She was a Council Member of the Nigeria Institute of Management (NIM) for two years, Branch Chairman of NIM, Ibadan chapter. Currently, she is a member of Governing Council, National Mathematical Centre, Abuja; Member of the Governing Council, Achievers University Owo, Ondo State; Treasurer Nigeria Association of Women in Science, Mathematics, Engineering and Technology (NAWSTEM), Oyo State Branch and Patroness of the Nigerian Association of Computer Science Society (NACOSS). She belongs to several professional bodies such as Board of Trustees of Nigeria Internet Registration Association (NIRA). She is a Board Member OMATEC Computers; Coordinator – Four Nigeria Institutions of Higher Learning/Nancy 2 University, France Collaboration Programmes. She is also a Director of Computerize Nigeria Project.

Prof. Osofisan has been the recipient of many national and international prizes, scholarships and fellowship awards such as, the International Council for Computer Communication Fellowship (ICCC) to New Delhi and Bombay (now Mumbai) India, University Development Linkage programme (UDLP), International Women in Science and Technology (IWISE) Iowa USA, UNESCO, South Africa, United Nations

University (UNIST), and University of Technology Melbourne, Australia among others. Prof. Adenike Osofisan is a Fellow of the Nigeria Institute of Management (NIM) chartered; Fellow Nigeria Computer Society (NCS); Life Member, Nigerian Economic Society, Member, Computer Society of IEEE; and Member Association of Computing Machinery (ACM).

She has received numerous awards for the development of Information Technology and Computer Science Education in Nigeria, such as Meritorious Award (NACOSS) 1995, Women of Achievement Award, (NACOSS), 1996; Digne Ambassadeur, Computer Science Students Association Obafemi Awolowo University, Ile-Ife 1997; Builder of Community Award, Akure 1998; Branch Merit Award NIM, 2000; Award of Excellence, The Compronians, 2001; The Woman of Merit Gold Award, 2002; Role Model Award, NAPATAN, Ondo State Chapter, 2006; Woman of Distinction Merit Award, 2006; Outstanding Dedication to the Growth and Development of NACOSS, 2006; NCS Award towards advancement of IT in Nigeria, 2007; the 2007 IT Woman Personality of the Year; NASU University of Ibadan Award towards the progress and development of Humanity, 2007; ICT Female Legend, 2008; Distinguished National Award of Excellence 2008; TITAN of Technology, Nigeria Award 2008; Information Technology Champion, ITAN 2009, Fiwasaiye Old Girls Association (FOGA) Golden Jubilee Merit Award for being the first Nigerian Professor of Science and meritorious service to FOGA and Humanity February, 2010 and ICT for Africa 2010 Lifetime Achievement Award in Recognition for Excellence in the Research and Practice of ICT in Africa, March 27, 2010.

Professor Osofisan has attended many conferences, seminars both at national and international level presenting papers of her various research works. She has over 70 publications authored in national and international journals and conference proceedings.

Prof. Adenike Osofisan is a devout Christian, serving the Christian Community in various positions such as Teaching Leader, Ibadan Evening Women Bible Study Fellowship International; Member Parish Church Council, Church of the Transfiguration, Anglican Communion, Ikolaba Ibadan; Bishop Nominee for the 2011-2013 Synod of The Ibadan Diocese, the Church of Nigeria Anglican Communion. She was the former Secretary and the former President of The Resurrection Lilies Society, Chapel of The Resurrection, University of Ibadan.

Prof. Adenike Osofisan is married to a literary giant, Prof. Femi Osofisan of the Department of Theatres Arts, University of Ibadan and former General Manager, National Theatre Iganmu Lagos. The marriage is blessed with children, and grandchildren.