An evaluative study of the classification of the biochemical literature of Nigeria

F. Edu-Uwem Etim
University of Cross River State
Uyo

Abstract

This study examined the treatment of Biochemistry, a multidisciplinary subject, in the context of two traditional classification schemes, the Library of Congress and the National Library of Medicine schemes. In carrying out the study, the two schemes were surveyed for their treatment of Biochemistry. A test was also set up using the two schemes to classify a sample of biochemical literature of Nigeria. The results show that:

1.   The Library of Congress classification scheme is more suitable than the National Library of Medicine classification scheme for the classification of biochemical literature of Nigeria.

2.   For both schemes, 44% of the concepts are scattered into classes outside those specified for Biochemistry and these include very important areas of modern biochemical research like nutritional surveys, drug metabolism and cancer research.

3.   The two schemes are deficient in their organisation of the biochemical literature of Nigeria for easy retrieval. Consequently, there is need for a special faceted classification scheme for Biochemistry.

Introduction

The problem of the classification of knowledge is as old as philosophy and as open and unsettled as a philosophical system[1]. Vickery[2] recognises two opposite factors which counter balance each other as playing a dominant role in the sphere of classification. There is, on the one hand, the imperative desire of the intellect to understand the overwhelming richness of the phenomena with which it has to cope and, on the other, the varied, varying and seemingly endless multiplicity of data. It was this situation which urged Lord Rayleigh[3] to comment:

> In science, by a fiction as remarkable as any to be found in law, what has once been published, even though it be in the Russian language, is spoken of as known, and it is too often forgotten that the rediscovery in the library may be a more difficult and uncertain process than the first discovery in the laboratory.

These words have become far truer now than when they were written and, in 1948, the editors of a Royal Society Conference declared that "the task of keeping up with scientific literature is becoming an impossible one and is in turn leading to inefficiency and to a certain amount of frustration in scientific research and in the application of science"[4]. It is for this reason that the problem of information retrieval assumes such importance in science today.

The classification of subject matter may be carried out for all sorts of special purposes - to arrange books on shelves, to group inventions in patents, to classify the raw materials, intermediates and products of importance to a particular manufacturer. What is important is that the system used reveals the manifold inter-relations among the concepts of science. The complex techniques of information retrieval, now so widely discussed among science librarians and documentalitists, have been criticized by those who consider that the traditional techniques of the Bliss Classification, Library of Congress, the Universal Decimal Classification and so on, polished a little, will serve all purposes. These criticisms are probably based on lack of appreciation of what the modern information retrieval system aims to achieve - to design systems which allow access to specific subjects because the typical specific subject in science or technology is a multifacted one.

Biochemistry is the study of the substances that occur in the living organism, the process by which the substances enter into or are formed in the organisms and react with each other and the environment and the methods by which the substances and processes are identified, characterised and measured[5]. Biochemistry is thus an interdisciplinary subject which cuts across the subject areas of Biology, Chemistry, Physics, Medicine, Pharmacy, Nutrition, Technology and Agriculture. In such a heterogeneous discipline, subjects invariably overlap, their contents may be related to other subjects and this varies from time to time and place to place.

Biochemistry is a young science, having been known under that term only about 1900. By that time, most of the traditional classification schemes had already been formed. With the expansion of Nigerian biochemical literature, the librarian is faced with two main problems - the Nigerian literature of Science and Technology, past and present on one hand and, on the other, an enquirer with questions. How does the librarian select from vast masses of concepts, the few that are most closely relevant to an enquiry? It is this selection process that calls for an efficient and suitable system of classification.

The most important characteristic of documentary classification is that it is concerned with subjects - specific subjects, i.e. themes on which books, parts of books or articles are written. The specific subject of an article may not be simple concept but may consist of several concepts, i.e. a composite subject. In classifying a document, account is taken not only of the natural kinds of organisms or substances, but of their properties, behaviour, interactions and operations on them. Consequently, apart from hierarchies of the natural kinds, subject classification also provides for other categories which figure as elements in compound

subjects.   These various hierarchies that occur in the classification of a subject field have become known as facets of the subject.

In the classification of Biochemistry as a main subject in this study, its treatment by the schemes of the National Library of Medicine (NLM) and the Library of Congress (LC) has been examined in terms of enumeration of facets within the section allocated to Biochemistry the specificity in classification of items using these facets and the scattering of the literature into other classes in the scheme.

Research design

The objective of the study was to evaluate the adequacy of classification schemes for biochemical literature of Nigeria. In these schemes, the parts allocated to Biochemistry were examined and noted.   A test was also set up using the schemes.   A sample of materials was selected from monographs, journals and journal articles of Nigerian biochemical literature.   Each of the materials was classified using the LC and NLM classification schemes. It was noted whether the class mark given specified the subject exactly and whether the item was scattered outside the main area of biochemistry or placed with another subject.   The results were analysed for each scheme and these were used to evaluate the suitability of each scheme for the classification of biochemical literature.

Criteria for Selection of Schemes

1.   The NLM scheme[6] is the standard scheme proposed by the National Library of Medicine, U.S.A. and is used by E. Latunde Odeku Medical Library and many other medical libraries in Nigeria. Biochemistry is studied in Nigerian medical schools as a preclinical and para-medical subject.

2.   The LC scheme[7] is used by the University of Ibadan Library (Kenneth Dike Library) and many other university libraries in Nigeria.   It is also used by the National Library of Nigeria in Lagos.

Criteria for Selection of Materials

A sample of biochemical literature of Nigeria was randomly selected for classification.   These materials had to meet the following requirements:

(a)   Monographs and Research Papers

(i)   Materials written by a Nigerian anywhere in the world, about Nigeria or by a non-Nigerian working in Biochemistry in Nigeria.

(ii)   Materials in Biochemistry that can be grouped according to the National Library of Medicine Medical Subject Headings (MeSH).

(b)   <u>Journals</u>

Journals were randomly selected. The selection was based on a recent Bibliography of Biochemical Literature of Nigeria (1970 – 1985)[8] Analysis of this bibliography showed the ten journals most frequently used by Nigerian biochemists to publish their research papers. These ten journals were used for the study. They included interdisciplinary journals as the bibliography revealed that Nigerian biochemical literature is scattered in medical, agricultural, nutritional and pharmaceutical journals.

Sample   Composition

The sample distribution was based on the index accompanying the bibliography. This index revealed that over 70% of the scientific communication among Nigerian biochemists is done through journal articles, theses and dissertations. On the basis of this, the sample was distributed in this form:

Research Papers	–	60
Monographs	–	30
Journal Titles	–	110

The titles of research papers were selected from each area of biochemistry in the researcher's bibliography according to the NLM list of Medical Subject Headings (MeSH). The number of titles selected was proportional to the amount of work done in that area, as set out below.

Clinical Biochemistry and Immunology	–	10
Comparative Biochemistry and Metabolism	–	10
Nutritional Biochemistry	–	20
Toxicology and Cancer Research	–	<u>20</u>
		60

Discussion   and   Analysis   of   Results

A synopsis of the 4th edition of the NLM classification published in 1978 is as follows:-

Preclinical Sciences

QS   –   Gynab Anatomy

QT   –   Physiology

QU   -   Biochemistry

QV   -   Pharmacology

QW   -   Bacteriology and Immunology

QX   -   Parasitology

QY   -   Clinical Pathology

QZ   -   Pathology

Medicine and Related Subjects

Q   -   Medicine (General)

WA   -   WZ - Different Aspects of Medicine

The main class QU - Biochemistry, placed between QT - Physiology and QV - Pharmacology, two closely related subjects to Biochemistry which often overlap each other, enumerates only about 80 concepts of Biochemistry.

Table 1 illustrates the number of titles whose concepts could specifically or non-specifically be classified using the NLM scheme. For example, while a monograph title, recent Advances in Steroid Biochemistry, could be classified specifically under QU - 19 (steroids), another title, Prostaglandins and Reproduction, could not be classified specifically but broadly under QU-90 (fatty acids).

Out of a total sample of 100 titles, only 36 could be classified using specific concepts while 64 were classified in broad and related areas. Within class QU - (Biochemistry), 56 titles were classified but only 20 out of these were classified using specific concepts while 36 could only be classified using broad or related concepts. Similarly, out of 13 titles classified under QW - (Bacteriology/Immunology), only two could be classified using specific concepts while 11 were classified into related areas. The two titled classified under class WITO - (Toxicology) were both classified into related areas, not specific concepts.

Within class QU - Biochemistry, areas which could be specifically classified were:-

-       Body components and derivates, e.g. proteins, carbohydrates, fats, fatty acids, steroids, trace elements, etc.

-       Nutritional surveys.

-       Metabolism

-       Vitamins

Areas like

- Acids/Esters/Salts

- Postaglandings, glycosides

- Enzymes

- Antigens/antibodies

- Biochemical phenomenon, e.g. Transcription could only be classified broadly into related terms.

Table 1 also shows that 44 titles classified using the NLM scheme were scattered into classes other than QU - Biochemistry. Areas into which titles were generally scattered include Pharmacology - QW, Clinical pathology - QY, Pathology - QZ, Practice of Medicine - WB, Bacteriology/Immunology - QW, and Deficiency Disease - WD100. Journals were classified separately in W1 (Biomedical, preclinical) according to type of publication.

Table 1:   Scattering of concepts, specificity and non-specificity in classification of biochemical literature: NLM

S/N	Notation	Class name	Specific	Not specific	Total scatter
1.	QU	Biochemistry	20	36	56
2.	QV	Pharmacology	2	1	3
3.	QW	Bacteriology/ Immunology	2	11	13
4.	QY	Clinical Pathology	2	–	2
5.	QZ	Pathology	–	3	3
6.	WB	Practice of Medicine	2	–	2
7	WD100	Deficiency Disease	4	–	4
8.	WK	Endocrine System	1	1	2
9.	WN	Radiology	–	1	1

10.	WD105	Protein Calorie Malnutrition	1	–	1
11.	WP	Gynaecology and Obstetrics	–	2	2
12.	WI	Gastrointestinal System	1	–	1
13.	WINI	Nigerian Journal	–	2	2
14.	WIAE	West African Journal of Biology and Applied Science	–	1	1
15.	WIAF	African Medical Journal	–	1	1
16.	WING	Nutritional Reports	–	1	1
17.	WITO	Toxicology	–	2	2
18.	WIFE	Federation of European Biochemical Society	–	1	1
19.	WK770	Virilism	1	–	1
20.	WIBI	Biochemical Pharmacology	–	1	1
			36	64	100

The Library of Congress (LC) classification scheme, first published in 1901 but with a current and sixth edition in 1973, has Biochemistry under QP - Physiology.   The main purpose of the LC scheme was to provide for the Library of Congress.   The scheme covers all United States publications. However, priority is given to subjects of interest to the Library of Congress e.g. American History.   Class Q - Science, first published in 1905, has the following subclasses:

Q   —   Science General

QA   —   Mathematics

QB   —   Astronomy

QC   —   Physics

QD   —   Chemistry

QE   —   Geology

QH   —   Natural History

QK   —   Botany

QL   —   Zoology

QM   —   Human Anatomy

QP   —   Physiology

QR   —   Microbiology

Biochemistry is treated extensively under class QP - Physiology as Animal Biochemistry (QP501 - 801). There are adequate see and see also references to similar or allied classes as indicated in the index.   The scheme enumerates more than 1,000 facets, in both broad and specific terms, under general areas such as protein, amino acids, etc.   Within these broad areas, more facets are further enumerated alphabetically.

Examples are:

QP551   —   Protein, Amino acids, etc.

.C3   —   Carnosine

.C35   —   Casein, etc.

Table 2 illustrates the number of titles whose concepts could be specifically or non-specifically classified using the LC scheme. Results of the classification shows that 51 titles could be specifically classified while 49 classified into broad and related areas.

Within class Q - Physiology, areas which could be specifically classified were periodicals, vitamins, amino acids, body constituents, body components, enzymes components and biochemical techniques. However, areas like nutrition, toxins and hormones were only broadly classified.

Table 2 also illustrates the scattering of biochemical concepts into classes other than class QP - Physiology. While 56 titles could be classified within QP, 44 titles, as in the NLM scheme, were scattered into classes like QC - Physics, QD - Chemistry, QH - Biology, QK - Botany, and QR - Microbiology. Beside class QP, the class with the highest number of items was class RC - Internal Medicine, under which areas like protein deficiency and cancer research were classified. These are key areas in modern biochemical research.

Comparing the two schemes, it was found that while the LC scheme treats toxicology in the main class QP - Physiology, the NLM scheme treats it in class QW - Bacteriology and Immunology. On the other hand, while biochemical processes like conjugation, metabolism, protein synthesis and biochemical techniques are classed in the main class QU - Biochemistry in the NLM scheme, the LC scheme scatters them into QH - Biology and QD - Chemistry. In the LC scheme also, cancer research and protein calorie malnutrition studies are classed in RC - Internal Medicine, while the NLM scheme scatters them into classes QZ - Pathology and WD100 - Deficiency Diseases.

In conclusion, this study has shown that the two schemes are deficient in their specific treatment of biochemical literature and also both scatter aspects of modern biochemical research into other classes of sciences. The deficiencies of the schemes are aggravated in the case of Biochemistry by the subject's wide scope and complexity and the rapid expansion of its literature into fields hitherto regarded as distant to it.

Table 2:   Scattering of concepts, specificity and non-specificity
           in classification of biochemical literature: LC

S/N	Notation	Class name	Specific	Not specific	Total scatter
1.	QC	Physics	1	1	2
2.	QD	Chemistry	2	2	4
3.	QH	Biology	7	3	10
4.	QK	Botany	1	1	2
5.	QP	Physiology	25	31	56
6.	QR	Microbiology	3	–	3
7.	R	Medicine	–	2	2
8.	RA	Public Medicine	2	–	2
9.	RB	Pathology	–	1	1
10.	RC	Internal Medicine	6	6	12
11.	RG	Obstetrics and Gynaecology	1	–	1
12.	RM	Pharmacology	3	1	4
13.	SF	Animal Husbandry	–	1	1
			51	49	100

References:

1. Wojciechowski, J.A. (1971). The philosophical relevance of the problem of the classification of knowledge In: *Conceptual basis of the classification of knowledge*. Proceedings of the Ottawa Conference, 1971.

2. Vickery, B.C. (1975). *Classification and indexing in science*. 3rd ed. London: Butterworths.

3. *Ibid.*

4. *Ibid.*

5. *McGraw-Hill Encyclopaedia of Science and Technology*. Vol. 2. p. 301. New York, 1971.

6. *National Library of Medicine Classification; a scheme for the shelf arrangement of books in the field of Medicine and its related sciences.* 4th edition. Washington: National Library of Medicine, 1976.

7. *Library of Congress Classification Scheme Class Q: SCIENCE.* 6th edition. Washington: Library of Congress, 1973.

8. Etim, F.E. (1986). *Biochemical literature of Nigeria (1970 – 1985):* A bibliography. MLS term paper, University of Ibadan.