
UNIT 13 KEYWORD INDEXING

Structure

- 13.0 Objectives
- 13.1 Introduction
- 13.2 Keyword Indexing – Concept
- 13.3 Structure and Format of Keyword Indexing
- 13.4 Indexing Process
- 13.5 Variants of Keyword Indexing
 - 13.5.1 KWIC
 - 13.5.2 KWOC
 - 13.5.3 KWAC
 - 13.5.4 KWWC
 - 13.5.5 KEYTALPHA
 - 13.5.6 WADEX
 - 13.5.7 DKWIC
 - 13.5.8 KLIC
- 13.6 Advantages and Disadvantages of Keyword Indexing
- 13.7 Summary
- 13.8 Answers to Self Check Exercises
- 13.9 Keywords
- 13.10 References and Further Reading

13.0 OBJECTIVES

After reading this Unit, you will be able to:

- explain the meaning of keyword indexing;
- describe the structure of keyword indexing;
- demonstrate the preparation of entries in a keyword index; and
- discuss the various versions of keyword indexing with suitable examples.

13.1 INTRODUCTION

Specialised indexes to technical literature are an accepted means for directing scientists to sources of information pertinent to their interest. Indexes based on the titles of documents or authors of the documents are a poor substitute for a micro-document. Limitations of titles as indicators of document contents or absence of terminology control leads to various problems in indexing micro-documents. The establishment of index entries is a matter of judgment and experience and constitutes a considerable part of the intellectual effort involved in the manual compilation of indexes. The accelerated pace of scientific development along with demand for speedier communication have accentuated to establish an alternative method of subject indexing. It is argued that such demand can be satisfied by using machine in the form of a series of extractions each containing a significant, or key word as its nucleus.

13.2 KEYWORD INDEXING – CONCEPT

Keyword Indexing is a system of indexing technique which uses the natural language for indexing keywords or significant terms of a title. Significant words are the words which have relatively high correlation with the actual thought contents of the documents. The concept was first given by Andrea Crestadoro in 1864 in the name of Keyword-in-Title (KWIT). There was another term, ‘*catchword indexing*’, which was used to refer keyword indexing during 19th century. British Books in Print used catchword indexing for quite a long time and the well recognised journal ‘*Nature*’ also used this technique to derive keyword entries for their journal articles. In keyword indexing, the generation of keywords is done without use of any vocabulary control device like thesaurus. The word chosen may be a single word, multiple words or even phrases that convey the contents. The significance of such keywords could be determined only by referring to the statement from which the keyword had been chosen. The statement acts as a modifier pointing up the more specific sense in which a keyword has been applied. Several keywords may be selected for a title to provide access from different access point of users. This principle has been applied in present days in various indexing systems with a slight variation. There are eight known variants of keyword indexing: KWIC, KWOC, KWAC, KWVC, KEYTELPH, WADEX, DKWIC and KLIC.

13.3 STRUCTURE AND FORMAT OF KEYWORD INDEXING

An entry of a keyword index is in three parts:

- a) **Keywords:** Subject denoting words or significant words which can be approach terms
- b) **Context:** Remaining part of title serving as the context to the keyword.
- c) **Identification or Location Code:** An identification number (usually the serial numbers of the entries in the main part) to provide location of the document where the document will be available.

13.4 INDEXING PROCESS

The overall process of indexing involves the followings steps:

- a) **Choosing Keyword:** Keywords are chosen either from the title and /or abstract of the document. An indexer or an editor marks the significant terms and also marks words in the ‘stop list’. The stop-list is the list of words which are considered to have no significant value for indexing and no indexing entries need to be produced through those terms. These insignificant terms include articles (a, an, the), prepositions, conjunctions, pronouns, auxiliary verbs together with some general terms (like aspect, view-point, reference to, etc.). The larger this list, the fewer are expected index entries. The keywords thus selected serve as approach terms. For a document titled “Treatment of skin diseases by using Homeopathy”, the significant terms or keywords will be ‘treatment’, ‘skin’, ‘disease’, and ‘Homeopathy’ and the stop list will be ‘of’, ‘by’, ‘using’.
- b) **Entry Generation:** Index entries in a KWIC index or any of its versions are generated in association with all of the words in the batch of titles that are not stored in the stop-list. The title is so manipulated that the keyword comes in the beginning (or in the middle) followed by rest of the title. The word is printed and

displayed ‘in context’, that is, together with the remainder of the title in which it appears. Each significant word is normally written in either in bold face or in capital letters. In this way, a single line entry, which includes title and source reference of some type, is produced for significant/keyword word in the title. In the above example there will be four index entries for four significant words and each of them coming in the beginning by rotation. The last word and first word of the title is separated by a symbol ‘stroke’ (/) or other symbol like asterisk (*) or equal (=). An identification number or reference number is provided at the right end of each entry to link one entry with other.

TREATMENT of skin disease by using Homeopath	133
SKIN disease by using Homeopath/ Treatment of	133
DISEASE by using Homeopath/ Treatment of skin	133
HOMEOPATHY/ Treatment of skin disease using	133

c) **Filing:** All the index entries are arranged alphabetically.

DISEASE by using Homeopath/ Treatment of skin	133
HOMEOPATHY/ Treatment of skin disease using	133
SKIN disease by using Homeopath/ Treatment of	133
TREATMENT of skin disease by using Homeopath	133

In the above example, the first words (disease, homeopathy, skin, and treatment) are the ‘keywords’. The part of the title followed by each keyword is the ‘context’. The number attached to every entry on the extreme right is the ‘identification code’.

Self-Check Exercise

- Note:** i) Write your answers in the space given below.
 ii) Check your answers with the answers given at the end of this Unit.

- 1) What do you mean by keyword indexing?
- 2) What are the various elements of keyword indexing?

.....

.....

.....

.....

13.5 VARIANTS OF KEYWORD INDEXING

13.5.1 KWIC (Keyword –in- Context)

KWIC was developed by H.P. Luhn of IBM in the International Conference of Scientific Information held at Washington in 1958. This mechanised system is based on titles of documents indexed on the principle that title of a scientific document represents its contents. The significant words in the title indicate the subject of the document. The index is produced by rotating each significant term in the title at the beginning. The remaining part of the title also appears with each significant term to keep the context intact (in-context).

Keyword-in-context indexing may be carried out on various levels depending on the purpose that the index is to serve. The process may be applied to the title of an article, its abstract or its entire text. Keywords can be defined as those which characterise a subject more than others. To derive them, rules have to be established for differentiating between what is significant and non-significant. Since significant is difficult to predict, it is more practical to isolate it by rejecting all obviously non-significant or common words. Such words may include terms like 'report', 'analysis', 'theory' and the like, as well as conjunctions, prepositions, auxiliary verb etc. The remaining significant or 'key' words would be extracted from the text together with a certain number of words that precede and follow them. By making the keywords assume a fixed position within the extracted portions and by arranging these portions in alphabetic order of the keywords, the KWIC index is generated.

Let us take the title '*Prevention of diseases of wheat caused by insects*' to demonstrate the index entries generated through KWIC principle. In this title 'of', and 'by' have no significance and 'prevention', 'disease', 'wheat', 'and', 'insects' are the keywords. While generating entries through KWIC, we have to keep in mind that every keyword should come as approach term. The remaining terms should be written in such a manner so that context of the document must be maintained intact. To maintain the context, stroke (/) should be placed in proper position so that searcher of the document can understand what the starting or ending point of the title is. Although, capitalisation, or making bold is the job of the machine, indexer has to identify the significant terms associated with the document. It is the job of the editor to mark the keywords before the title is punched so that the key puncher is able to tag them and instruct the computer.

Diseases of wheat caused by insects/Prevention of	13243
Insects /Prevention of disease of wheat caused by	13243
Prevention of disease of wheat caused by insets/	13243
Wheat caused by insect/Prevention of disease of	13243

In the computer generated index, the keywords can be positioned at centre also.

Prevention of	DISEASE of wheat caused by insects/	13243
Disease of wheat	INSECTS /Prevention of	13243
caused by		
Wheat caused	PREVENTION of diseases of	13243
by insects/		
Of diseases of	WHEAT caused by insect/Prevention	13243

Let us take another example. We have a document entitled '*Monetary Policy of India*' with identification number 3642. The entry will be as,

INDIA/Monetary Policy of	3642
MONETARY Policy of India/	3642
POLICY of India/Monetary	3642

Or

Indexing – Sears List of Subject Heading

Policy of	INDIA/Monetary	3642
Of India/	MONETARY Policy	3642
Monetary	POLICY of India/	3642

We have another document entitled '*Indexing of articles in medical library*' with identification number 24513. The entry will be as,

ARTICLES in medical library/Indexing of	24513
INDEXING of articles in medical library/	24513
LIBRARY / Indexing of articles in medical	24513
MEDICAL library/Indexing of articles in	24513

Or

Indexing of	ARTICLES in medical library/	24513
Medical library/	INDEXING of articles in	24513
In medical	LIBRARY /Indexing of articles	24513
Articles in	MEDICAL library/Indexing of	24513

Let us take another example: '*Classroom behaviour of urban students in India*' with identification code 46135.

BEHAVIOUR of urban students in India/Classroom	46135
CLASSROOM behaviour of urban students in India/	46135
INDIA/Classroom behaviour of urban students in	46135
STUDENTS in India/Classroom behaviour of urban	46135
URBAN students in India/Classroom behaviour of	46135

Or

Classroom	BEHAVIOUR of urban students in India/	46135
In India/	CLASSROOM behaviour of urban students	46135
Students in	INDIA/Classroom behaviour of Urban	46135
Of urban	STUDENTS in India/Classroom behaviour	46135
Behaviour of	URBAN Students in India/Classroom	46135

In another example: '*Online information storage and retrieval*' with identification number 66213. The entry will be as,

INFORMATION storage and retrieval/Online	66213
ONLINE information storage and retrieval/	66213
RETRIEVAL/Online information storage and	66213

STORAGE and Retrieval/Online Information 66213

Keyword Indexing

Or

Online INFORMATION Storage and Retrieval/ 66213

Retrieval/ ONLINE Information Storage and 66213

Storage and RETRIEVAL/Online Information 66213

Information STORAGE and Retrieval/Online 66213

In the above examples the context of the document remains intact in spite of that fact that four different approach terms appeared one-by-one at the first.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

3) What are various variants of keyword indexing?

4) Prepare index entries for title: Collection Development Policy in Academic Libraries of India through KWIC.

.....
.....
.....
.....

13.5.2 KWOC (Key-Word Out-of-Context)

KWOC is another version of keyword indexing, in which each keyword is taken out and printed separately in the left hand margin with the complete title to provide complete context. The identification number of document is written on the right.

For example, entry for the document '*Prevention of diseases of wheat caused by chemicals*' under KWOC will be as:

DISEASES Prevention of diseases of wheat caused by chemicals 13243

CHEMICALS Prevention of diseases of wheat caused by chemicals 13243

PREVENTION Prevention of diseases of wheat caused by chemicals 13243

WHEAT Prevention of diseases of wheat caused by chemicals 13243

Sometime, keyword is printed as heading and the title is printed in the next line instead of the same line as follows.

DISEASES

Prevention of diseases of wheat caused by chemicals 13243

CHEMICALS

Prevention of diseases of wheat caused by chemicals 13243

PREVENTION

Prevention of diseases of wheat caused by chemicals 13243

WHEAT

Prevention of diseases of wheat caused by chemicals 13243

In earlier example: '*Indexing of articles in medical library*' with identification number 24513. The entry will be as,

ARTICLES	Indexing of articles in medical library	24513
INDEXING	Indexing of articles in medical library	24513
LIBRARY	Indexing of articles in medical library	24513
MEDICAL	Indexing of articles in medical library	24513

Or

ARTICLES

Indexing of articles in medical library 24513

INDEXING

Indexing of articles in medical library 24513

LIBRARY

Indexing of articles in medical library 24513

MEDICAL

Indexing of articles in medical library 24513

If we have a document entitled '*Monetary Policy of India*' with identification number 3642. The entry will be as,

INDIA	Monetary Policy of India	3642
MONETARY	Monetary Policy of India	3642
POLICY	Monetary Policy of India	3642

Or

INDIA

Monetary Policy of India 3642

MONETARY

Monetary Policy of India 3642

POLICY

Monetary Policy of India 3642

For title: '*Classroom Behaviour of urban students in India*' with identification code 46135, entries will be:

BEHAVIOUR	Classroom behaviour of urban students in India	46135
CLASSRO	Classroom behaviour of urban students in India	46135

OM INDIA	Classroom behaviour of urban students in India	46135
STUDENTS	Classroom behaviour of urban students in India	46135
URBAN	Classroom behaviour of urban students in India	46135

Keyword Indexing

Or

BEHAVIOUR

Classroom behaviour of urban students in India 46135

CLASSROOM

Classroom behaviour of urban students in India 46135

INDIA

Classroom behaviour of urban students in India 46135

STUDENT

Classroom behaviour of urban students in India 46135

URBAN

Classroom behavior of urban students in India 46135

For earlier example: '*Online information storage and retrieval*' with identification number 66213, the entry will be as,

INFORMATION	Online information storage and retrieval	66213
ONLINE	Online information storage and retrieval	66213
RETRIEVAL	Online information storage and retrieval	66213
STORAGE	Online information storage and retrieval	66213

Or

INFORMATION

Online information storage and retrieval 66213

ONLINE

Online information storage and retrieval 66213

RETRIEVAL

Online information storage and retrieval 66213

STORAGE

Online information storage and retrieval 66213

In a true KWOC an asterisk appears in the printed title in the position of and instead of the keyword that features as heading. Following is an example which explains how asterisk appears in title while index entries are generated.

ADULTS

A description study of Indian * in North Indian Region.

ADVANCED

The * evolution of globular cluster stars.

The KWOC format preserves the title verbatim. But it fails to provide the convenient adjacent context which the KWIC format provided.

Self-Check Exercise

- Note:** i) Write your answers in the space given below.
ii) Check your answers with the answers given at the end of this Unit.
- 5) Differentiate between KWIC and KWOC.

.....
.....
.....
.....

13.5.3 KWAC (Key-Word Augmented-in-Context)

KWAC also stands for ‘key-word-and-context’. It is the most improved version of KWIC. In many cases, title cannot represent the thought content of the document. The keyword of a fancy title (title which is not self expressive) can be a problem in the retrieval of relevant documents. In order to solve the problem of retrieving irrelevant documents, KWAC provides additional keywords, taken either from the abstract or from the original text of the document or its contents, and then inserted into the title to give further index entries. The meaning of ‘augmented’ here is ‘in addition to’. The CBAC (Chemical Biological Activities) of BIOSIS uses KWAC index where is enriched by another title-like phrase formulated by the indexer.

Let us understand the KWAC by an example: *African women in agricultural development*.

Here the concept ‘African women’, in reality, does not include all women of Africa but only the women of Sierra Leone. The main focus of the study is employment conditions of women in agriculture. After detailed study of the document, it came to light that ‘working hours’ is the central theme of the document. An indexer should augment these concepts in keywords for effective retrieval .

Sierra Leone	African women in agricultural development	6452
Employment conditions	African women in agricultural development	6452
Working hours	African women in agricultural development	6452

Similarly a title entitled: *The use of peer tutoring and programmed radio instruction* cannot be understandable unless the term *peer-tutoring* is explicitly indexed. Terms like *Educational radio, programmed instruction, non-formal instruction* are added to augment the actual concept of the document.

Let us take another example: ‘*Prevention of disease of wheat caused by insects*’ with identification number 13243. The entry will be as,

Streak mosaic	Prevention of disease of wheat caused by insets	13243
Arthropods	Prevention of disease of wheat caused by insets	13243

Here, streak mosaic is the name of disease, arthropods are the insects causing the disease and the wheat variety on which this investigation took place is Malvika. So, we have to augment all these concepts into the original title to make it searchable.

In another example: *Classroom behaviour of urban students in India* with identification code 46135.

Discipline	Classroom behaviour of urban students in India	46135
Boarding school	Classroom behaviour of urban students in India	46135
Hostellers	Classroom behavior of urban students in India	46135
Metro Cities	Classroom behavior of urban students in India	46135

Here the term behaviour means discipline and this study was conducted for boarding school hostellers of metro cities. Unless these concepts are added the actual content of the document cannot be understandable and retrieval is also affected.

Let us take another example. We have a document entitled '*Monetary Policy of India*' with identification number 3642. The entry will be as,

Reserve Bank of India	Monetary Policy of India	3642
Recession	Monetary Policy of India	3642
Interest Rate	Monetary Policy of India	3642

Here monetary policy of Reserve Bank of India is main concern of the document. Among monetary policy 'recession policy' and 'interest rate' are concern areas of the document. So these concepts need to add while indexing the documents.

In another example: '*Online information storage and retrieval*' with identification number 66213. The entry will be as,

Searching Techniques	Online information storage and retrieval	66213
Indexing	Online information storage and retrieval	66213
Digital Libraries	Online information storage and retrieval	66213

'Online information storage and retrieval' is a very vague term and can be used for various other contents too. Here, the prime emphasis of the document is digital library scenario. So these concepts need to be added while indexing.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

6) Why KWAC system is said to be an advanced version of KWIC?

.....

.....

.....

.....

13.5.4 KWWC (Key-Word-With-Context)

The key word with context is another version of KWIC, where only the part of the title (instead of full title) relevant to the keyword is considered as entry term. Sometimes, for clarity and delimitation of the scope of the subject, a secondary aspect of the scope of the subject is added to the primary aspect, separated by a punctuation mark. Keyword index of Indian Science Abstracts (ISA) used KWWC for its indexing. An example of KWWC is as follows for a title '*Bacterial contamination in Indian distilleries using waste molasses*'.

Alcohol Fermentation / Effect of Bacterial contamination	963
Bacterial contamination / Effect on Alcohol Fermentation	963
Fermentation, Alcoholic / Effect of Bacterial contamination	963
Molasses / Alcohol Fermentation ; Effect of Bacterial	963

Let us see another example: *Implacability of distillation theory in radio frequency breakdown in isomeric liquid*.

Distillation theory

Implacability in *radio frequency breakdown in isomeric liquid* 667

Isometric liquid

Implacability of distillation theory in radio frequency breakdown 667

Radio-frequency breakdown

Implacability of distillation theory breakdown in isomeric liquid 667

13.5.5 KEYTALPHA (Key-Term Alphabetical)

It is permuted subject index that lists only keywords assigned to each abstract. In KEYTALPHA, only keywords are taken from title instead of full title. Keywords are arranged side by side without forming a sentence. The entry generation of a document entitled '*Acquisition of journals in a university library in Japan*' will be as follows:

Acquisition, Journals, Japan, Library, University	335
Journals, Japan, Library, University, Acquisition	335
Japan, Library, University, Acquisition, Journals	335
Library, University, Acquisition, Journals, Japan	335
University, Acquisition, Journals, Japan, Library.	335

Considering with the earlier example of '*Prevention of diseases of wheat caused by insects*' index entries will be:

DISEASES , Insects, Prevention, Wheat	13243
INSECTS , Prevention, Wheat, Diseases	13243
PREVENTION , Wheat, Diseases, Insects	13243
WHEAT , Diseases, Insects	13243

In another example: '*Indexing of articles in a medical library*' with identification

number 24513. The entry will be as:

ARTICLES , Indexing, Library, Medical	24513
INDEXING , Library, Medical, Articles	24513
LIBRARY , Medical, Articles, Indexing	24513
MEDICAL , Articles, Indexing, Library	24513

If we have a document entitled '*Monetary Policy of India*' with identification number 3642. The entry will be as:

INDIA , Monetary, Policy	3642
MONETARY , Policy, India	3642
POLICY , India, Monetary	3642

In another example: '*Classroom behavior of urban students in India*' with identification code 46135.

BEHAVIOUR , Classroom, India, Students , Urban	46135
CLASSROOM , India, Students , Urban, Behaviour	46135
INDIA , Students, Urban, Behaviour, Classroom	46135
STUDENT , Urban, Behaviour, Classroom, India	46135
URBAN , Behaviour, Classroom, India, Students	46135

For example: '*Online information storage and retrieval*' with identification number 66213, the entry will be as:

INFORMATION , Online, Retrieval, Storage	66213
ONLINE , Retrieval, Storage, Information	66213
RETRIEVAL , Storage, Information, Online	66213
STORAGE , Information, Online, Retrieval	66213

Oceanic Abstracts Service is using this technique to derive indexing entries for their documents.

13.5.6 WADEX (Word and Author Index)

WADEX stands for Word and Author Index. It is an improved version of KWIC index where the names of authors are also treated as keyword along with significant subject terms and facilitates to satisfy author approach also. WADEX can be used for both retrospective and current search. Alphabetically sequenced entries are printed out of context. Words are single words or hyphenated word pairs or 'tagged words'. WADEX is arranged in two columns, both columns printed simultaneously with dictionary entries at top and pagination at bottom by computer. It is used in 'Applied Mechanics Review'. **AKWIC** (Author and keyword in context) index is another version of WADEX.

For example, a title entitled '*Selected Water Resources Index for Wisconsin*', can be index through various terms like Wisconsin University, Madison Water Resources Center, Water resources, Water rights, Water supplies, Water treatment. Here Wisconsin University, Madison Water Resources Center is the corporate author of the title and

Water resources, Water rights are the significant terms of the title. So the index entry will consist of:

Wisconsin University Water Resources for Wisconsin6636

Madison Water Resources Center Water Resources for Wisconsin6636

13.5.7 DKWIC (Double KWIC) Index

It is another improved version of KWIC index. The keyword is extracted from title and regarded as main index term. The remaining words are rotated using a wrap-around format. In DKWIC index, specific search for two words in combination is possible. DKWIC indexes have most of the features, facilities and failures of KWIC and KWOC indexes. Entry for the title ‘The computer analysis of weather forecast’ will be:

COMPUTER

Analysis of weather forecast - The* 36

Forecast –The* analysis of weather 36

Weather forecast – The* analysis 36

ANALYSIS

Computer* of weather forecast -The 36

Weather forecast -The computer* of 36

Forecast -The computer* of weather 36

Entry term is indicated by asterisk (*) and is not repeated in the rotated titles.

13.5.8 KLIC (Key-Letter-In-Context)

KLIC was developed by Kent. This system allows truncation of word, instead of complete word. The truncation of titles, either at the beginning or at the end, can be specified and the computer will pick up any term containing that fragment.

The Chemical Society (London) published a KLIC index as a guide to truncation. The KLIC index indicates which terms any particular word fragment will capture.

Self-Check Exercise

Note: i) Write your answers in the space given below.

ii) Check your answers with the answers given at the end of this Unit.

7) Prepare KWIC, KWOC and KWAC entries for title: “Classification and Cataloguing for Students”

8) Prepare a KEYTALPHA index entry for title “Role of Mass Media in eliminating illiteracy”

.....
.....
.....
.....

13.6 ADVANTAGES AND DISADVANTAGES OF KEYWORD INDEXING

Like every indexing system keyword indexing has its own advantages and disadvantages.

Advantages of Keyword Indexing

- 1) The main merit of keyword indexing is the speed and economy with which an index can be produced, owing to its machine origin. It has therefore, become feasible to issue keyword indexes at frequent intervals, perhaps monthly.
- 2) Absence of interpretation of contents leads to perfect consistency and predictability. If a word appears in the title of the document then it is certain that, unless the word is in the stop-wordlist, an entry will be generated under that word. There is no possibility of inconsistency in the allocation of terms in such systems as there is with human indexers. No scope for error or dispute exists.
- 3) In the knowledge-based economy, human intellect has become costlier day-by-day and difficult to find. It does not require intellectual effort or trained indexing staff as it is produced by computer. The computer interfiles the entries from a number of batches of input and prints the cumulative index.
- 4) Today most user search document by title rather than its author. Keyword index is very useful for such use.
- 5) Keyword-in-context permit the cross correlation of subjects to an extent not realisable by conventional procedures.
- 6) Now a days the authors, especially in the field science and technology use representative and meaningful titles which express the contents of documents. So use of keyword indexing leads to better results.
- 7) Keyword indexing satisfies the current approaches of users.
- 8) There is no need of vocabulary control devices to handle indexing.

Disadvantages of Keyword Indexing

- 1) It is good for science and technology but not effective in the field of humanities and social sciences, where no controlled vocabulary is used.
- 2) It depends on title of the document. If the title is not representative the system will become ineffective, particularly in humanities and social science subjects. For example what conclusions about subject content can you drawn from the title like: 'Black money white power?' or 'Integration: for and against'. There is some evidence to suggest that titles are becoming more informative, but titles will always be limited in the amount of information that they can convey.
- 3) KWIC is normally produced by computer and the format is of fixed field length. Due to this fixed field length, part of the title is truncated when the title is lengthy. This truncation leads to funny results far away from context.
- 4) Scattering of related terms occurs because of alphabetical arrangement. Search of a topic may have to be done under several keywords. It takes very high search time. Several pages of entries under one information for some documents to be rapidly rejected. Sub-arrangement at entry terms can break down long sequences of entries listed under the same keyword.

- 5) Searchers very often face to high recall and low precision.
- 6) It is ineffective for exhaustive approach for a large collection.

Self-Check Exercise

- Note:** i) Write your answers in the space given below.
ii) Check your answers with the answers given at the end of this Unit.
- 9) Identify any two advantages of keyword indexing?
 - 10) Name two Indian integrated library management Software that use keyword indexes while preparing index entries.

.....
.....
.....
.....

13.7 SUMMARY

Keyword index may offer various advantages including that of performing well for retrieval of information in computer environment. In fact, it is one of the most widely used indexing services in Online Public Access Catalogues. Various integrated library management software like LIBSYS and NewGenLib use this technique for their OPAC services. However, one has to keep in mind that machine products can never reach the level of perfection that humans are capable of and that there will always be residual effort left for humans.

Keyword indexing is based on significant words and arrangement of these keywords is done alphabetically. So, related and identical subjects are scattered throughout the index file. Because of no reference system is used for connecting two related terms, synonymous terms, singular and plural form of terms, alternative spellings, etc. the users should search under all possibilities to identify documents.

13.8 ANSWERS TO SELF CHECK EXERCISES

- 1) Keyword indexing is a process of indexing where indexer prepares entries by choosing significant words or keywords of a document. The significance of such keywords could be determined only by referring to the statement from which the keyword had been chosen. The statement act as A modifier pointing up the more specific sense in which a keyword has been applied. This indexing is mostly based on natural language and no human intellect is essential for preparing index entries. The concept was first launched by Andrea Crestadoro in 13856 in the name of Keyword-in-Title (KWIT).
- 2) The entries of Keyword indexes are one line three parts:
 - a) Keywords: Significant or subject denoting words which can be as approach terms
 - b) Context: Remaining part of title serve as the context to the Keyword.
 - c) Identification or Location Code: An identification number (usually the serial numbers of the entries in the main part) to provide location of the document where the document will be available.

- 3) There are eight known variants of keyword indexing has been developed: KWIC, KWOC, KWAC, KWWC, KEYTELPHA, WADEX, DKWIC & KLIC.
- 4) Collection Development Policy in Academic Libraries of India/ Academic Libraries of India/ Collection Development Policy in India/ Collection Development Policy in Academic Libraries of India.
- 5) In KWIC, index entries are prepared on the basis of keywords available in the document. Sometime key term of an title is not clearly reveals the subject of the document. For clarity and defining the scope of subject, a secondary aspect of the scope of the subject is added to the primary aspect. In KWAC, indexer chooses keyword by reading the title, abstract or contents of the *document*. *The flair of indexer is important in choosing terms in KWAC. Although, preparation of entries through KWAC bit difficult than KWIC, the precession of KWAC is much more higher than KWIC.*
- 6) The KWAC provides for the enrichment of the keywords of the title with additional significant words taken either from abstract of the document or its contents. In KWIC such additional significant words are not chosen instead it select keyword as it is available in title. So, making document more searchable with high satisfaction ration KWAC can be said to be as advanced version on KWIC.
- 7) KWIC:
 For Indian. Classification and
 Cataloguing Classification and Cataloguing for
 Indian Cataloguing for Indian. Classification and
 KWOC:
 Classification Classification and Cataloguing for
 Indian Cataloguing Classification and Cataloguing for
 Indian Indians Classification and Cataloguing for Indian
 KWAC:
 Classification and Cataloguing for Indian
 Cataloguing for Indian. Classification for
 Indians. Classification and Cataloguing for
- 8) Eliminating, Illiteracy, Role, Mass Media,
 Illiteracy, Role, Mass Media,
 Eliminating, Mass Media, Eliminating,
 Illiteracy, Role Role, Mass Media,
 Eliminating, Illiteracy
- 9) Simplicity and Cost are two very common advantage of keyword indexing.
- 10) Name of Indian Integrated Library Management Software that are using Keyword Index while preparing Index entries.

13.9 KEYWORDS

- Keywords** : Important Word or words having relative importance of the title.
- Context** : The parts of something written or spoken that immediately precede and follow a word or passage and clarify its meaning.

13.10 REFERENCES AND FURTHER READING

- Prasher, R.G. *Index and Indexing System*. New Delhi: Medallion Press, 2005. Print
- Satpathi, J.N. Keyword indexing system. *Indexing System*/ Rajan, T.N. Kolkata: IASLIC, 1983. Print
- Chakraborty, A.R. Chakrabarty, B. *Indexing: Principles, Process and Products*. Calcutta: World Press, 1984. Print
- Ramakrishnan, M.N. *Indexing System and Techniques*. Trivandrum: Laxmishree Publ., 1986. Print
- Lancaster, F.W. *Indexing and Abstracting in Theory and Practice*. London: Facet, 2003. Print
- J.E., *Abstracting and Indexing*, London: Clive Bingley, 1988. Print
- Chakrabarty, A. *Documentation*. Calcutta: Academic Press, 1989. Print