

Classification of Keywords Extracted from Research Articles Published in Science Journals

Bidyarthi Dutta
Librarian
St. Xavier's College
30, Park Street
Kolkata 700016
Krishnapada Majumder
Head
Department of Library & information Science
Jadavpur University
B.K.Sen
80, Shivalik Apartments
Alaknanda
New Delhi 110019

Abstract: This paper is based on an analytical study of 335 keywords extracted from titles and abstracts of 70 research articles, taking ten from each year starting from 2000 to 2006, in decreasing order of relevance, on the subject Fermi Liquid, which is a specific subject under the broad area of Condensed Matter Physics. The research articles have been collected from the bibliographic database of INSPEC. The keywords are indexed to critically examine its physical structure that is composed of three fundamental kernels, viz. keyphrase, modulator and qualifier. The keyphrase reflects the central concept, which is usually post-coordinated by the modulator to amend the central concept in accordance with the relevant context. The qualifier comes after the modulator to describe the particular state of the central concept and/or amended concept. The keywords are further classified in 16 classes on the basis of the four parameters, viz. Associativeness, chronological appearance, frequency of occurrence and category. The taxonomy of keywords will enable to analyze research-trend of a subject and to identify potential research-areas of a subject.

Keyword: Research-trend-analysis, Subject-analysis, Keyword-cluster-analysis, Keyword taxonomy, INSPEC, Fermi Liquid, Condensed Matter physics

Introduction

Most of the research articles begin with a title followed by an abstract and some keywords. These three features describe an article's contents. The title gives the snapshot of the content in a concise way. The abstract summarizes the content. The keywords indicate core concepts and central fields of concern. The keywords are the building blocks of the 'Descriptors' or 'Subject headings', because subject headings comprise several keywords. The four essential parameters of an article, viz, title, author, abstract and keywords are required in any bibliographic database designed to aid electronic information retrieval. The earliest use of keywords was found in 1975 in the *Journal of Applied Behaviour Analysis*¹. The keywords are extracted from titles and abstracts of journal articles. The relevance of titles as source of keywords is discussed by Bottle², Hansen³, Kraft⁴, Lancaster⁵, Olive⁶ and Ruhl⁷. The relative merits of using title, subject heading and abstract as sources of keywords is discussed by Byrne⁸. The comparative layout between title keywords and subject descriptors is discussed by Voorbij⁹. The subject- descriptors comprise controlled terms which are required for subject indexing, while title keywords are available directly without any such intellectual activities. Dubois¹⁰ and Taylor¹¹ summarized the advantages and disadvantages of both approaches. Studies on indexing show significant variation in the use of keywords selected by different indexers to represent the same topic or document¹². Surraud et al.¹³ observed the non-existence of well-defined keywords in newly-emerging subject areas, which makes bibliographic searches difficult. Bates et al.¹⁴ discussed about development in the structures of thesauri and in the designs of the online information systems. Hurt¹⁵ emphasized on renewal and expansion of indexing and classification systems. Soergel et al.¹⁶ also pointed out that existing classification schemes and thesauri lack well-defined semantics and structural consistency. Juvan et al.¹⁷ executed

keyword analysis to identify narrower research fields within the broader scientific field. They proposed a bibliometric methodology that was based on keyword analysis and the structuring of data into hierarchical tree system and could be used for the assessment of bibliographic databases and the identification of research trends.

Keywords allow readers to decide whether or not an article contains material relevant to their interest. Keywords provide readers with suitable terms to use in web-based searches to locate other materials on similar topics. Keywords also enable indexers and editors to retrieve related materials. Any subject, whether very specific or broad one, always undergoes through incessant changes which are very picturesquely reflected from the keywords pertaining to the concerned field. Keywords indicate growth of a subject in different orientations but the notable feature is that selection of keywords is a vital aspect. If the keywords are not selected in logical way but scattered over irrelevant peripheral areas then they will mislead both the information professional and clientele. Keywords occur within titles, abstracts, series-names, content-notes, subject-headings and index-terms supplied by the authors. The keywords may also be used as effective subject access points. It is to be noted that every word in titles, abstract or content-notes are not keywords. Articles, prepositions, conjunctions and general words are too common to serve as keywords.

Objectives

The main objectives of this work are summarized below:

- To develop a taxonomy of keywords that covers all possible parameters of keywords occurring in articles of various science journal
- To analyze the physical structure of keywords
- To present possible ways of analysis of research trend of a subject and the potential areas of research from this taxonomy

Scope and Methodology

The titles and abstracts of 70 research articles on the subject Fermi Liquid were collected from the bibliographic database of INSPEC during the span of seven years, i.e. 2000-2006. The first ten articles have been selected in the decreasing order of relevance from each year's database. The method of systematic sampling is thus followed here. The keywords have been culled out from the titles and abstracts of those 70 articles. The notable feature is that only research articles have been taken for study out of the entirely available published literature. The other forms of outcomes like monographs, conference-proceedings, short communications, reviews, letters, reports etc. have been excluded from the considered domain of the present study as the largest contribution to the full set of published literature comes only from the research articles. All keywords are not collected from titles and abstracts for study, but some keywords, which are belonging to the following categories, have been rejected.

- 1) Too lengthy keyword (e.g. Low-temperature specific heat coefficient $C/\text{sub } V//T$, this keyword is selected after cutting off the last part)

- 2) Too common keyword (e.g. Physics)
- 3) Acronym (e.g. MFT) (Acronym is considered after expansion; e.g. Magnetic Field Tuning for MFT)
- 4) Too specific jargon (e.g. $1/[T \ln/\sup 4/(T/\sub K//T)]$ divergence)
- 5) Keywords not directly related with central or allied theme of the subject concerned as manifested by the abstract (e.g. Local moments)
- 6) Symbol
- 7) Formula (e.g. $B \ln/\sup 2/(T/\sub K//B)$)
- 8) Numerical figure

Keyword Analysis

The collection of entire keywords was thoroughly studied. The keywords have been segmented in atmost three kernels, which comprise: 1) Keyphrase, 2) Modulator and 3) Qualifier. The keyphrase tells the central theme underlying behind the concept, the modulator amends the central theme in accordance with the relevant context. The modulator modulates the manifestation by the total spectrum of the central theme. The modulator polarizes the all-pervaded manifestation of the keyphrase in a specific orientation. The qualifier comes after the modulator to describe the particular state of the central concept and/or amended concept without disturbing the conceptual wholeness.

The notable point is that, there are differences between “Keyphrase” and “Keyword”. For instance, the “Keyphrase” is just a part of a “Keyword”, but not the entire keyword. A complete “Keyword” consists of all three above-mentioned components, i.e. Keyphrase, modulator and qualifier. There may be lot of keywords containing Keyphrase only, as evident from Table 5, but that is accidental. It is to be noted that all keywords studied here don’t consist of all these three parts. The keyphrase is mandatory component of a keyword, i.e. each and every keyword must have a keyphrase, whereas other two components are optional, i.e. a keyword may or may not contain either a modulator or a qualifier or both. In this study, some keywords contain only keyphrase, some other contains keyphrase and modulator, and a number of keywords contain keyphrase and qualifier; whereas the remaining keywords contain all these three kernels, viz. keyphrase, modulator and qualifier. The distribution of kernels over the keywords is shown in Table 1 below. The sequence of outcome of these three components is: Keyphrase ----- Modulator ----- Qualifier. The physical structure of a keyword thus takes the form as given below:

$$\text{Keyword} = (\text{Keyphrase} + \text{Modulator}), \text{Qualifier}$$

Mandatory Optional Optional

This is, actually the generalized structure of a keyword. In this study, a considerable number of keywords contain only the keyphrase part, and some keywords contain keyphrase and modulator or qualifier. The remaining keywords contain all the three components. The relative distribution of kernels over 335 keywords is presented in Table 1 and the diagrammatic presentation is given in Figure 1. The keywords containing all

three components may be indicated as “Complete keyword”. It is clear from Table 1 and Figure 1 below, that only 10% keywords have been found “Complete” from the whole sample of keywords. Only 15% keywords contain keyphrase and modulator; 28% keywords contain keyphrase and qualifier; and 47% keywords contain only keyphrase. The keyphrase is a subject-specific term/terms, the modulator may either be a subject-specific term/terms or a general word/words and the qualifier is a general word for majority of keywords. The qualifier hardly becomes a subject-specific term. The keywords have thus been undergone through the following four phases after collection.

- 1) The three kernels of each keyword were identified at first
- 2) The kernels were arranged in the following order:
Keyphrase Modulator Qualifier
- 3) The keywords were reorganized according to this order
- 4) After reorganization the keywords were arranged alphabetically

Table 1: Relative distribution of kernels over the keywords

Kernel	Only Keyphrase (K)	Keyphrase + Modulator (K+M)	Keyphrase + Qualifier (K+Q)	Keyphrase + Modulator + Qualifier (K+M+Q)
No. of keywords	156	50	95	34
Percentage of keywords	46.6	14.9	28.4	10.1

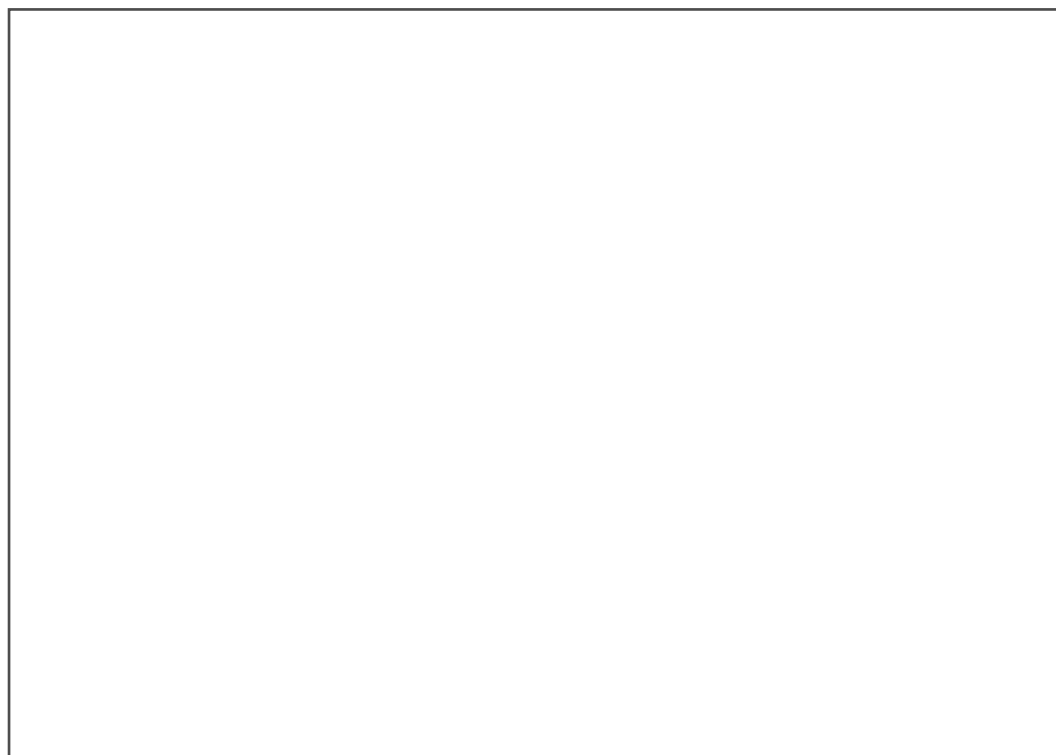


Figure 1

Keyword Taxonomy

Four fundamental properties of a keyword have been taken under consideration in this study for development of a new taxonomy. These four properties are listed below:

- 1) Mode of association or Associativeness with the subject-content,
- 2) Chronological appearance,
- 3) Frequency of occurrence and
- 4) Category.

The clarifications of these four properties are given here.

1) Mode of association or Associativeness with the subject-content: The keywords are collected over the said span of time from the journal articles and then organized in an ordered way. The organized keyword-set revealed some features very clearly. After organizing it was noticed that the majority of keywords occur in some flocks with a common keyphrase. Such keywords have been categorized as 'Clustered keywords', as

they form clusters of keywords having a common keyphrase. Some keywords appear in binary with a common keyphrase that have been termed as “Twin keyword”. A number of keywords don’t form any cluster, but appear in isolation, which are categorized as ‘Single keywords’. The keyphrase is common factor within a set of clustered keywords, though the modulator and qualifier vary, and thus the name of each cluster may be given in the name of the corresponding keyphrase. The Keyword clusters represent major keyphrases involved in the subject concerned. Keyword Cluster Analysis (KCA) is a very effective methodology to reckon potential keyphrases of a subject. The examples are cited below from the present study:

- i) Clustered keyword (K): Fermion; Fermion compound, heavy; Fermion metal, heavy;
(The common keyphrase here is “Fermion”)
- ii) Twin keyword (W): Fullerene-compound; Fullerene-compound, intercalation;
(The common keyphrase here is “Fullerene-compound” and it forms only two keywords)
- iii) Single keyword (S): Helium-3, superfluid; High-temperature-superconductor; Hubbard-like-coupling
(There is no common keyphrase among these keywords)

2) Chronological appearance: The classification with respect to the chronological appearance of the keywords divides the keyword in two groups, i.e. Ephemeral keyword and Non-ephemeral keyword; while the keywords falling under latter group may further be sub-divided in three sub-groups, i.e. *New* keyword, *Steady* keyword and *Obsolete* keyword. The explanation of this mode of categorization is given in the Table 2 below.

Table 2: Ephemeral and non-ephemeral keywords (Definition)

Year		Y-2	Y-1	Y (Concerned Year)	Y+1	Y+2	Total
Keyword-type							
Non-Ephemeral	New			A, B, C	A, C	A, B	w
	Steady	D, E, F	D, F	D, E, F	E, F	D, E	l
	Obsolete	G, H, I	H, I	G, H, I			b
Ephemeral				J, K, L			e
Total							n

The alphabets A, B, C etc. represent keywords. The year under consideration is represented by Y. The ephemeral keywords for the year Y occur only in the year Y, whereas the non-ephemeral keywords occur in other years also. If any keyword occurs all other years except the year Y, then it will not be considered as the keyword of the year Y. The keywords A, C are *New* with respect to the year Y, because they appear for the first time in the year ‘Y’; whereas *Steady* with respect to the year Y+1, because they appear in both preceding and following years of the year ‘Y+1’. Similarly, the keywords H, I are *Steady* with respect to the year Y-1, but *Obsolete* with respect to the year Y, because they appear for the last time in the year ‘Y’. The classifications *New*,

Obsolete and *Steady* are thus temporal classifications, i.e. classifications based on time and hence change from year to year. If a keyword starts to occur from the year 2004 and continued till 2007, i.e. it stopped occurring after 2007, then the keyword would be considered as “New” in 2004; “Steady” in the year 2005 and 2006 and “Obsolete” in 2007. The examples are cited below from the present study:

- i) Non-ephemeral keyword (F): The keyword “Band model, magnetism” is “New” in the year 2002, “Steady” in the year 2003 & “Obsolete” in the year 2005. Also, the keyword “Fermi level” is “New” in the year 2003 & “Obsolete” in the year 2004 and the keyword “Helium” is “New” in the year 2000 & “Obsolete” in the year 2004. These keywords became extinct before becoming “Steady”.
- ii) Ephemeral keyword (E): The keywords “Helium film” and “Hund’s coupling” are “Ephemeral” as they occurred once only during the span of seven years, i.e. from 2000 to 2006.

3) Frequency of occurrence: The keywords are divided in three groups with respect to its frequency of occurrence over the span of seven years, i.e. 1) Mono-frequent keyword (keywords having frequency of occurrence one during 2000-2006, i.e. the keyword appeared only once in one journal-article from 2000 to 2006. It is also clear that all mono-frequent keywords are ephemeral keywords also); 2) Di-frequent keyword (keywords having frequency two, i.e. the keyword appeared either once in two journal-articles, or two times in one journal-article at each instant) and 3) Multi-frequent keyword (keywords having frequencies larger than two). The examples are cited below from the present study:

- i) Mono-frequent keyword (I): “1/N-expansion”; “Absorption, collisionless” and “Acoustic-wave propagation” are mono-frequent keywords.
- ii) Di-frequent keyword (D): “Anderson-model”; “Antiferromagnetic-material” and “Antiferromagnetism” are di-frequent keywords.
- iii) Multi-frequent keyword (U): “Fermi-level”; “Fermi-liquid” and “Fermi-liquid fixed-point” are multi-frequent keywords.

4) Category: Lastly, the keywords are divided in six groups with respect to its categories, i.e. Action, Theory, Entity, Material, Property and Method. The elementary category *Action* indicates some sort of work or function to express the concept of “Doing”. The category *Theory* reflects the micro and macro conceptual organization of an idea in the concerned subject. The keywords expressing some model or parameter are also included in the scope of this category. The category *Entity* include manifestations having perceptual correlates, or having only conceptual existence, as contrasted with their properties and actions performed by them or on them. The scope of this category includes energy, light, plant, animal, place, time etc. The category *Material* represents the focal idea of the subject involved. The focal idea spots the central theme of the subject. The category *Property* indicates the concept of “Attribute” or “Quality”. The scope of the category Property includes effect, form, capability, utility etc. The category *Method* indicates some ways of any execution or performance, any sort of procedural concept etc.

The categories “Material”, “Property” and “Method” accord with Ranganathan’s “Matter-

Material”, “Matter-Property” and “Matter-Method”; the three sub-categories under one of the five fundamental categories “Matter”¹⁸. The category “Action” accords with Ranganathan’s fundamental category “Energy”; and lastly, the categories “Theory” and “Entity” are in accordance with the fundamental category “Personality”. In this study, no keywords from the fundamental categories “Space” and “Time” has been found.

The examples are cited below from the present study:

- i) Action (G): “Contact-interaction”; “Coulomb-interaction, long-range”; and “Coulomb-repulsion”.
- ii) Theory (T): “Fermi-liquid theory”; “Hall-effect” and “Kondo-model”.
- iii) Entity (X): “Entropy”; “Fermi-energy” and “Fermi-level”.
- iv) Material (M): “Fullerene-compound”; “Gallium-arsenide” and “Gluon”.
- v) Property (P): “Infrared-singularity”; “Logarithmic-temperature-dependence” and “Magnetic-property”.
- vi) Method (H): “Green’s-function-method”; “One-loop-calculation; and “Optical-study”.

The classification of keywords with respect to its “Associativeness with the subject-content” helps in providing content-description and in executing content-analysis of a Subject. The “Chronological appearance” of keywords reveals about obsolescence and recency of research on a subject. The “Frequency of occurrence” of keywords reflects the intensity of research on a topic. The impact of a subject with universe of knowledge is well understood from different “Categories” of keywords. The new taxonomy of keywords that has been proposed in this paper is presented in Table 3:

Table 3: Keyword Taxonomy

Major criteria for classification of keywords	No. of classified keyword clusters	Name of each cluster	Representative Notation
Mode of Association or Associativeness with the Subject-Content (ASC)	3	Clustered	K
		Twin	W
		Single	S
Chronological Appearance (CAP)	4	Ephemeral	E
		Non-Ephemeral (F)	J
		Stable	Y
Frequency of Occurrence (FOC)	3	Obsolete	O
		Mono-frequent	I
		Di-frequent	D
		Multi-frequent	U
Category (CAT)	6	Action	G
		Theory	T
		Entity	X
		Material	M
		Property	P
		Method	H

Characteristic features of keywords are indicators of research trend and the latest state-of-the-art of the concerned subject. After assembling the keywords it was noticed that a fraction of total collection of keywords occurred in different clusters, whereas all keywords assembled in each cluster contain some common keyphrase with different modulators and qualifiers. These features are clearly reflected from the alphabetical arrangement of keywords in Table 5. The statistics of occurrence of the non-ephemeral keywords (new, steady and obsolete) during the span of five years (2001-2005) is shown in Table 4. The starting and ending years concerned (2000 and 2006) are not shown as all keywords in the starting year 2000 are “New” and in the ending year 2006 are “Obsolete”. The non-ephemeral keywords have been classified according to the taxonomy presented in Table 3, i.e. new, obsolete and steady. The extreme right column of Table 3 gives the representative notations of each class of keywords. The relative percentage distribution of keywords belonging to each class is shown in figures (2), (3), (4) and (5). It is clear from Figure 2 that Ephemeral keywords are highly dominating (86%) over Non-ephemeral keywords (14%). The Stable keywords are rare in this subject. A particular keyword hardly sustains here. The keywords are continuously

waving on the vast ocean of the subject “Fermi liquid”. Such momentary occurrence of a large fraction of keywords indicates the dynamism of this subject, i.e. new research works are continuously executed in this area. It is obvious that only new and newer research projects can give birth to new keywords. The number of research projects is an indicator of dynamism of a subject. The Figure 3 depicts the dominance of Mono-frequent keywords (81%) over Di-frequent and Multi-frequent keywords. A keyword with frequency of occurrence one indicates the appearance of that keyword in one and only one journal article. The keywords are thus hardly repeated in more than once journal article in this subject area. Out of 335 keywords, 50% appeared in “Single” mode; 20% in “Twin” mode and 30% in “Clustered” mode, which is also clear from Figure 4. The relative distribution of six “Categories” over all keywords is presented in Figure 5. The keywords belonging to the category “Entity” is highest (54%) in occurrence.

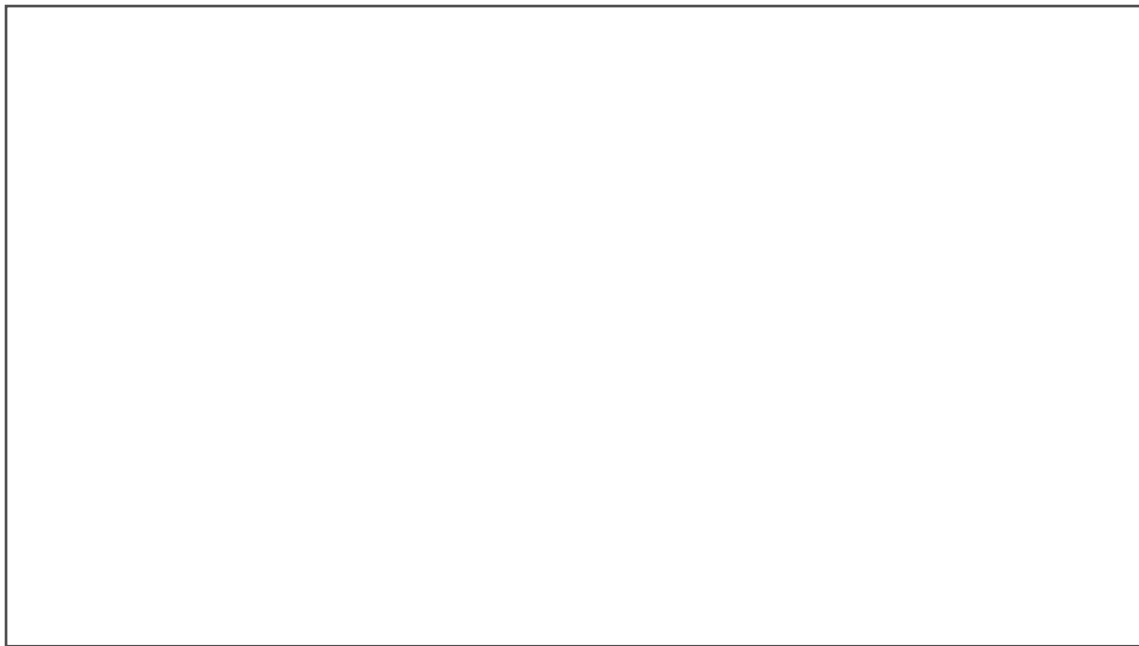


Figure 2: Chronological Appearance (CAP)

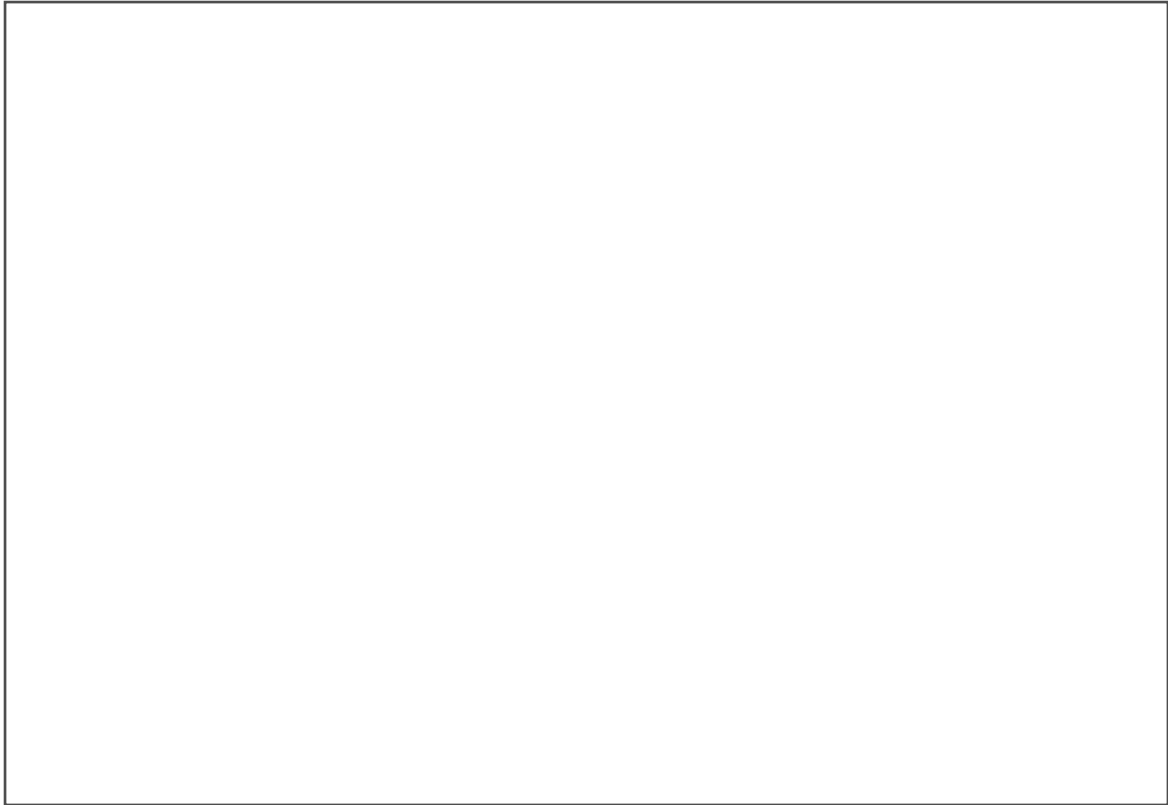


Figure 3: Frequency of Occurrence (FOC)

Figure 4: Associativeness with the Subject Content (ASC)

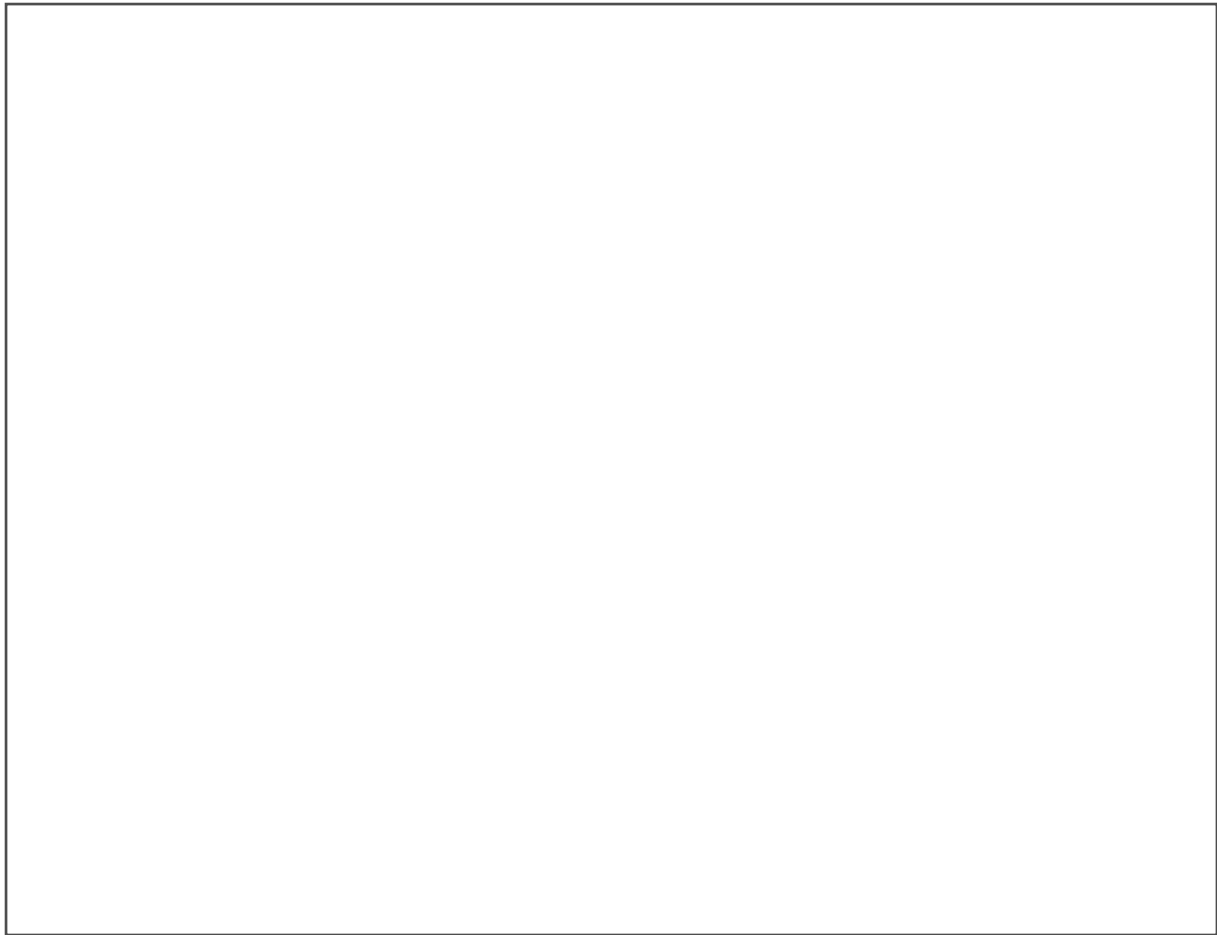


Figure 5: Category (CAT)

Table 4: Statistics of non-ephemeral keywords
from 2001 to 2005

	2001	2002	2003	2004	2005
New	7	9	7	3	0
Obsolete	1	4	8	11	15
Steady	8	9	13	4	2
All	16	22	28	18	17

The list of selected 335 keywords is arranged alphabetically in the Table 5. These keywords have been classified according to the taxonomy presented in Table 3. The frequency of occurrence in each year is shown in appropriate cells. Different kernels of each keyword are indicated by three brackets. The 1st bracket stands for “Keyphrase; the 2nd bracket stands for “Modulator” and the 3rd bracket indicate “Qualifier”. Keywords belonging to different classes are also indicated here by the representative symbols corresponding to each class as given in Table 3.

Table 5: List of classified keywords

KEYWORD = (KEYPHRASE + MODULATOR + QUALIFIER)

KEYPHRASE > ()

MODULATOR > { }

QUALIFIER > []

(The symbol > stands for 'Is indicated by') |ASC|CAP|FOC|CAT|2000|2001|2002|2003|2004|2005|2006|Total| |(1/N-expansion) |S|E| |H|1| | | | | | |1| |(Absorption), [collisionless] |S|E| |G| | | | | |1| |1| |(Acoustic-wave) {propagation} |S|E| |G| | | | | |1| | | | | |1| |(Adsorption) |S|E| |G| | | | | |1| | | | |1| |(Aluminium) {compound} |S|E| |M| |1| | | | | |1| |(Amplitude) {mode} |S|E| |X| | | | | |1| | | | |1| |(Analytic-continuation) |S|E| |H| |1| | | | | |1| |(Anderson-model) |S|F|D|T|1| | |1| | | | |2| |(Angular-momentum), [orbital] |S|E| |X| |1| | | | | |1| |(Anisotropic-large-dimension-limit) |S|E| |X| |1| | | | | |1| |(Anomalous-distribution-function) |S|E| |T|1| | | | | |1| |(Anomalous-exponent) |S|E| |X| |1| | | | | |1| |(Anomalous-property) |S|E| |P| |1| | | | | |1| |(Antiferromagnet), [metallic] |S|E| |M| |1| | | | | |1| |(Antiferromagnetic-material) |S|F|D|M|1| | | | | |1| |2| |(Antiferromagnetism) |W|F|D|X| | | | | |1| | | |1| |2| |

KEYWORD = (KEYPHRASE + MODULATOR + QUALIFIER)

KEYPHRASE > ()

MODULATOR > { }

QUALIFIER > []

(The symbol > stands for 'Is indicated by') |ASC|CAP|FOC|CAT|2000|2001|2002|2003|2004|2005|2006|Total| |(Antiferromagnetism), [itinerant] |W|E| |X| | | | | |1| | | | |1| |(Anvil-apparatus), [cubic] |S|E| |M| | | | | |1| |1| |(Atomic-system) |S|E| |X| | | | | |1| |1| |(Band) {filling}, [conduction] |K|E| |G| | | | | |1| | |1| |(Band) {model}, [magnetism] |K|F|U|T| | |2| |1| |1| |4| |(Band) {structure} |K|F|U|X|2| | | | | |1| |3| |(Band), [highly-1d] [half-filled] |K|E| |X| | | | | |1| |1| |(Band), [parabolic] |K|E| |X| |1| | | | | |1| |(Band), [valence] |K|E|D|X| | | | |2| | |2| |(Bethe-ansatz-solution) |S|E| |H| | | | | |1| | | | |1| |(Bias-voltage) |S|E| |X| | | | | |1| | |1| |(Binding-energy) |S|E| |X| | | | | |1| | |1| |(Bogolubov-method) |S|E| |H| | | | | |1| |1| |(Boson), [Schwinger] |S|E| |X| | | | | |1| | | | |1| |(Carbon-nanotube) |K|F|U|X| |1| |1| |1| | |3| |(Carbon-nanotube), [multi-walled] |K|E| |X| | |1| | | | | |1| |(Carbon-nanotube), [single-wall] |K|E| |X| | | | | |1| | | |1| |(Charge) {carrier mobility} |K|E| |X| | |1| | | | | |1| |(Charge) {transfer} |K|E| |G| | | | | |1| |1| |(Charge) {transfer complex}, [organic] |K|E| |G| | | | | |1| |1| |(Charge-density-wave) |S|E| |X| | | | | |1| | | |1| |(Chemical-potential) |S|E|D|X|2| | | | | |2| |(Collective-mode) |S|E| |X| | | | | |1| | | |1| |(Conducting-material) |S|E| |M| | | | | |1| |1| |(Conductivity) |K|E| |X| | | | | |1| | | |1| |(Conductivity), [1d] |K|E| |X| | |1| | | | | |1| |(Conductivity), [electrical] |K|E| |X| | | | | |1| | | | |1| |(Contact-interaction) |S|E| |G| |1| | | | | |1| |(Core-level) |S|E|D|X| | | | |2| | |2| |(Correlation-function), [density-density] |S|E| |T|1| | | | | |1| |(Coulomb-interaction), [long-range] |S|E| |G| |1| | | | | |1| |(Coulomb-repulsion) |S|E| |G| | | | | |1| |1| |(Coupling-constant) |W|E| |X| | | | | |1| | |1| |(Coupling-constant), [elementary-particle] |W|E| |X| | | | | |1| | |1| |(Critical-fluctuation) |S|E| |G| | | | | |1| | | | |1| |(Crossover-temperature) |S|E| |X| | | | | |1| | | |1|

KEYPHRASE > ()
MODULATOR > { }
QUALIFIER > []

(The symbol > stands for 'Is indicated by') |**ASC |CAP |FOC |CAT |2000 |2001 |2002 |2003 |2004 |2005 |2006 |Total** | |(Paramagnetic-material) |S |E | |M |1 | | | | |1 | |(Paramagnetism) |S |E | |X | | | | |1 | | | | |1 | |(Particle) {density} |S |E | |X | | | | |1 | | | | |1 | |(Particle-hole) {excitation}, [long-wavelength] |S |E | |G |1 | | | | |1 | |(Perturbation) {calculation} |W |E | |H | |1 | | | | |1 | |(Perturbation) {theory} |W |E | |U |T |3 | | | | |3 | |(Phase-relaxation) {time} |S |E | |X | | | | |1 | | | | |1 | |(Phase-transition) |S |E | |G | | | | |1 | | | |1 | |(Photoelectron) {spectra} |S |E | |X | | | | |1 | | | | |1 | |(Photoemission) {spectra}, [high-resolution] |S |E | |X | | | | |1 | | | | |1 | |(Photon) |S |E | |X | |1 | | | | |1 | |(Potassium) {compound} |S |E | |M | | | | |1 | | | | |1 | |(Power-law) {behavior}, [effective] |W |E | |P | |1 | | | | |1 | |(Power-law) {temperature-dependence} |W |E | |P | |1 | | | | |1 | |(Pressure-tensor) |S |E | |X | | | | |1 | | | | |1 | |(Quantum-chromodynamics) |W |F | |U |X | |2 | | | |1 | | | |3 | |(Quantum-chromodynamics), [high-density] |W |E | |X | | | | |1 | | | | |1 | |(Quantum-critical-point) |S |F | |U |X | | |1 | |2 | | | |3 | |(Quantum-electrodynamics) |S |E | |D |X | |2 | | | | |2 | | | |2 | |(Quantum-wire) |S |E | |D |X | | | | |2 | | | |2 | |(Quark-matter) |W |E | |M | | | | |1 | | | | |1 | |(Quark-matter), [dense] |W |E | |M | | | | |1 | | | | |1 | |(Quasi-1d-system) |S |E | |X | |1 | | | | |1 | |(Quasiparticle) |K |F | |U |X | |2 | |1 | |3 | |1 | |2 | |10 | |(Quasiparticle) {lifetime}, [inelastic] |K |E | |X | | | | |1 | | | | |1 | |(Quasiparticle) {mapping} |K |E | |G | | | | |1 | | | | |1 | |(Quasiparticle) {peak}, [sharp] |K |E | |X | |1 | | | | |1 | |(Recursion-method) |S |E | |H |1 | | | | |1 | |(Recursive-equation), [asymptotically-exact] |S |E | |T |1 | | | | |1 | |(Reflection-spectrum) |S |E | |X | | | | |1 | | | |1 | |(Renormalization) |W |F | |U |H |1 | |1 | |1 | |1 | |5 | |(Renormalization), [wave-function] |W |E | |H | |1 | | | | |1 | |(Renormalization-group) {flow} |K |E | |G | | | | |1 | | | | |1 | |(Renormalization-group) {technique}, [N-chain] |K |E | |H | | | | |1 | |1 | |(Renormalization-group), [dynamical] |K |E | |H | |1 | | | | |1 | |(Renormalization-group), [Euclidean] |K |E | |H | |1 | | | | |1 | |(Renormalized-vertex) |S |E | |X |1 | | | | |1 | |(Repulsive-potential-interaction) |S |E | |G | | | | |1 | | | | |1 | |(Resistivity) |W |F | |U |X |1 | | | |1 | |3 | |(Resistivity), [electrical] |W |F | |U |X |1 | | | |1 | |3 | |

KEYWORD = (KEYPHRASE + MODULATOR + QUALIFIER)

KEYPHRASE > ()
MODULATOR > { }
QUALIFIER > []

(The symbol > stands for 'Is indicated by') |**ASC |CAP |FOC |CAT |2000 |2001 |2002 |2003 |2004 |2005 |2006 |Total** | |(Resonance) {energy} |K |E | |X | | | | |1 | | | | |1 | |(Resonance), [giant] |K |E | |X | | | | |1 | | | | |1 | |(Resonance), [giant-multipole] |K |E | |X | | | | |1 | | | | |1 | |(Ripplon) |S |E | |D |X | | | | |2 | | | |2 | |(Scattering), [impurity-related] |S |E | |G |1 | | | | |1 | |(Scattering-amplitude), [divergent] |S |E | |X | | | | |1 | | | |1 | |(Scattering-vertex) |S |E | |X | |1 | | | | |1 | |(Self-consistent-medium) |S |E | |X |1 | | | | |1 | |(Self-energy) |W |E | |X |1 | | | | |1 | |(Self-energy), [off-shell] |W |E | |X |1 | | | | |1 | |(Semiconducting-tube) |S |E | |M | | | | |1 | | | | |1 | |(s-f-exchange-model) |S |E | |T |1 | | | | |1 | |(Short-ranged-interaction) |S |E | |G | |1 | | | | |1 | |(Short-wave) {limit} |S |E | |X | | | | |1 | | | |1 | |(Single-channel-case) |S |E | |X |1 | | | | |1 | |(Single-particle) {excitation} |K |E | |G | | | | |1 | | | | |1 | |(Single-particle) {level} |K |E | |X | | | | |1 | | | |1 | |(Single-particle) {self-energy} |K |E | |X |1 | | | | |1 | |(Skeleton-expansion) |S |E | |H |1 | | | | |1 | |(Soft-mode) |S |E | |X | | | | |1 | | | | |1 | |(Solid-state-system) |S |E | |X | | | | |1 | | | | |1 | |(Sound-wave) {propagation}, [longitudinal] |W |E | |G | | | | |1 | | | | |1 | |(Sound-wave) {damping}, [strong-anisotropic] |W |E | |G | | | | |1 | | | | |1 | |(Specific-heat) |S |F | |U |X |1 |4 | |1 | |1 | |7 | |(Spectral-density) |S |F | |D |X |1 |1 | | | | |2 | |(Spectral-function) |W |E | |X |1 | | | | |1 | |(Spectral-function), [single-particle] |W |E | |X |1 | | | | |1 | |(Spectral-shape) |S |E | |X | | | | |1 | | | | |1 | |(Spin) {component} |K |E | |X | |1 | | | | |1 | |(Spin) {fluctuation} |K |F | |D |G | | | | |1 | | | | |2 | |(Spin) {fluctuation}, [antiferromagnetic] |K |E | |G | | | | |1 | | | |1 | |(Spin) {Hamiltonian} |K |E | |X | | | | |1 | | | | |1 | |(Spin-polarized) {transport} |K |E | |G | | | | |1 | | | | |1 | |(Spin) {susceptibility} |K |F | |D |X | | | | |1 | | | | |2 | |(Spin), [local] |K |E | |X | | | | |1 | | | | |1 | |(Spin-density-wave) |S |F | |D |X |1 | | | | |1 | | | | |2 | |(Spin-orbital-degeneracy) {limit}, [large] |S |E | |X |1 | | | | |1 | |(Spin-polarized-transport) |S |E | |G | | | | |1 | | | | |1 | |(Spontaneously-broken-symmetry), [rotational] |S |E | |X | | | | |1 | | | | |1 | |(Strong-coupling-regime) |S |F | |D |X | | | | |1 | | | | |2 | |

KEYWORD = (KEYPHRASE + MODULATOR + QUALIFIER)

KEYPHRASE > ()

MODULATOR > { }

QUALIFIER > []

(The symbol > stands for 'Is indicated by') |ASC |CAP |FOC |CAT |2000 |2001 |2002 |2003 |2004 |2005 |2006 |Total | |(Strongly-correlated-system), [2d] |S |E | |X | | | | | | | | |(Superconducting-compound), [Copper-oxide-based] |S |E | |M | | | | | | | | |(Superconducting-material) |W |E |D |M | | | | | | | |2 |2 | |(Superconducting-material), [organic] |W |E | |M | | | | | | | | | | |(Superconducting-phase) |S |E | |X | | | | | | | | | | |(Superconducting-transition) {temperature} |S |E | |X | | | | | | | | | | |(Superconductivity) |W |E |D |X | | | | | | | |2 |2 | |(Superconductivity), [penetration-depth] |W |E | |X | | | | | | | | | | |(Superconductor), [ambient-pressure] |K |E | |M | | | | | | | | | | |(Superconductor), [doped-type] [organic] |K |E | |M | | | | | | | | | | |(Superconductor), [quasi-1d] |K |E | |M | | | | | | | | | | |(Surface-mediated-indirect-interaction) |S |E | |G | | | | | | | | | | |(Temperature-dependence) |W |F |D |X | | | | | | | | | |2 | |(Temperature), [finite] |W |E | |X | | | | | | | | | | |(Thermal-effect) |S |E | |X | | | | | | | | | | |(Thermodynamics) |S |E | |X | | | | | | | | | | |(Thermoelectric-power) |S |E |D |X | | | | | | | | | |2 | |(T-J-model) |S |E | |T | | | | | | | | | | |(Translational-invariance) |S |E | |X | | | | | | | | | | | | | | |(Transport) {property} |W |F |U |P | | | | | | | | | |3 | |(Transport) {theory} |W |E | |T | | | | | | | | | | |(Triplet-state) |S |E | |X | | | | | | | | | | |(Tunnel-junction) |S |E | |X | | | | | | | | | | |(Tunneling) |S |F |U |X | | | | | | | | | |3 | |(Unitary-transformation), [continuous] |S |E | |H | | | | | | | | | | |(Valence-compound), [mixed] |S |E | |M | | | | | | | | | | |(Velocity) {field} |S |E | |X | | | | | | | | | | |(Vertex-correction) |S |E | |H | | | | | | | | | | |(Ward-identity) |S |E | |X | | | | | | | | | | |(Wave-vector) {orientation} |S |E | |G | | | | | | | | | | |(Weak-coupling-limit) |S |E |D |X | | | | | | | | | |2 | |2 | |(Weak-interchain-hopping) |S |E | |G | | | | | | | | | | | | | | |(Weak-localization) |W |E | |G | | | | | | | | | | |(Weak-localization) {anomaly} |W |E | |X | | | | | | | | | | |(Weakly-damping-collective-mode) |S |E | |X | | | | | | | | | | |(Weakly-interacting-particle) |S |E | |M | | | | | | | | | | |(Zeeman-splitting) |S |E | |G | | | | | | | | | | |(Zero-bias-anomaly) |S |E | |X | | | | | | | | | | |(Zero-temperature-limit) |S |E | |X | | | | | | | | | |

Conclusion

A method of classification of keywords selected from articles published in science journals has been proposed here. This study has been executed on the subject "Fermi liquid". The viability of this classification scheme for the keywords from other subject areas will be studied later on. The keywords have been classified from four different criteria. In all, 335 keywords have been classified here. The keyword-collection portrays the core and allied contents of a subject. The mode of classification of keywords is an indicator of research trend of a subject. The variation of content in a subject due to various research projects from time to time will be reflected from the classified keywords over a stipulated time span.

References:

1. Hartley J and Kostoff R N, How useful are 'keywords' in scientific journals? *Journal of Information Science*, 29 (5) (2003) 433–438.
2. Bottle R T, The information content of titles in engineering literature, *IEEE Transactions on Engineering Writing*, EWS-13 (1970) 41-45.
3. Hansen I, Evaluation of the database CA condensates compared with chemical titles, *Journal of Chemical Documentation*, 12 (1972) 101-109.
4. Kraft D H, Comparison of keyword-in-context indexing of titles with a subject-heading classification system, *American Documentation*, 15 (1964) 48-52.
5. Lancaster F W, *Vocabulary Control for Information Retrieval*, (Information Resources Press; Washington DC), 1972.

6. Olive G, Terry J E and Datta S, Studies to compare retrieval using titles with that using index terms, *Journal of Documentation*, 29 (1973) 169-191.
7. Ruhl M J, Chemical documents and their titles: Human concept indexing vs. KWIC machine indexing, *American Documentation*, 15 (1964) 136-141.
8. Byrne J R, Relative effectiveness of titles, abstracts and subject headings for machine retrieval from the COMPENDEX services, *Journal of the American Society for Information Science*, 26 (4) (1975) 223-229.
9. Voorbij H J, Title keywords and subject descriptors: A comparison of subject search entries of books in the humanities and social sciences, *Journal of Documentation*, 54 (4) (1998) 466-476.
10. Dubois C P R, Free text vs. controlled vocabulary; a reassessment, *Online Review*, 11 (1987) 243-253.
11. Taylor A G, On the subject of subjects, *Journal of Academic Librarianship*, 21 (1995) 484-491.
12. Bertrand A and Cellier J M, Psychological approach to indexing effects of the operator's expertise upon indexing behaviour, *Journal of Information Science*, 21 (6) (1995) 459-472.
13. Suraud M G et al, On the significance of databases keywords for a large-scale bibliometric investigation in fundamental physics, *Scientometrics*, 33 (1) (1995) 41-63.
14. Bates M J, Wilde D N and Siegfried S, An analysis of search terminology used by humanities scholars: the Getty Online Searching Project Report, No.1, *Library Quarterly*, 63 (1) (1993) 1-39.
15. Hurt C D, Classification and subject analysis: looking to the future at a distance, *Cataloging and Classification Quarterly*, 24 (1-2) (1997) 97-112.
16. Soergel D et al, Re-engineering thesauri for new applications: the AGROVAC example, *Journal of Digital Information*, 4 (4) (2004). (Also available at: <http://jodi.ecs.soton.ac.uk/Articles/v04/i04/Soergel/>)
17. Juvan S, Bartol T and Boh B, Data structuring and classification in newly-emerging scientific fields, *Online Information Review*, 29 (5) (2005) 483-498. (Also available at: www.emeraldinsight.com/1468-4527.htm)
18. Ranganathan S R, *Prolegomena to Library Classification*, (Asia Publishing House; London), 1967.