

# Integrated Use of Knowledge Organizing Tools: Case Studies with Classic Sacred Works

K.N. Prasad\*

Sarada Ranganathan Endowment for Library Science, 702, 42<sup>nd</sup> Cross, III Block, Rajajinagar, Bangalore 560 010, Karnataka State, India

**Abstract:** Presents case studies of using integrated bilingual and multilingual thesauri linked to lexicons in the domain of sacred works of the Hindus. The hyper-linking is shown to facilitate data mining and knowledge discovery. The paper emphasizes the importance of collaboration between subject specialists / scholars and information specialists in developing integrated knowledge organizing tools.

**Keywords:** Knowledge organization, tools for knowledge organization, classics, sacred texts, multi-lingual thesauri, hyperlinking.

## 1. OBJECTIVE

In all disciplines there are core subjects of study and research interacting with and drawing from, several cognate and supporting fields. The case studies presented here demonstrate such interfacing and mutual synergy between studies in the subject domain on the one hand and knowledge organization tools (KOTs) on the other. These studies resulted from our work of designing and developing (1) a trilingual (English, Farsi and Sanskrit) thesaurus for 'mysticism and allied subjects', (2) a Tamil-English bilingual information retrieval thesaurus, and (3) an online database of extracts from spiritual and religious sources.

Specifically, based on our work on KOTs, this paper briefly describes and gives examples of

- Collaborative development of KOTs scholars and information professionals;
- Value addition to information retrieval by integrated use of multiple KOTs in the religious and spiritual domains;
- Providing support to data mining;
- Creating the possibility of knowledge discovery through KOTs ; and
- Enhancing the value of KOTs beyond their use in vocabulary control.

The discussion and examples in this paper relate mostly to classics and sacred works of the Hindus. However, the methodology can equally be applied to other domains.

## 2. INTRODUCTION

The sources of basic information on the ancient established religions such as Jainism, Buddhism, Hinduism, Con-

fucianism, Judaism, Islam, Christianity and their many sects / denominations are embodied in textual materials, collectively called sacred texts or works. Such collections are usually organized on library shelves or their surrogates in catalogues and bibliographies using one or other of the schemes for subject classification – the general ones, such as, the Dewey Decimal Classification, the Library of Congress Classification, the Colon Classification, etc., or other special schemes for these domains. With the developments in information and communication technologies (ICTs) large volumes of information sources are accessible in digital form on CDs, DVDs, etc. and on the Internet/WWW. Disseminating, exchanging and sharing of information and experiences across global distances on any subject have become easier than ever before through blogs, twitter, UTube, FaceBook etc on the Internet. Wireless technologies such as the mobile phone have extended the reach and coverage of communications. By creating virtual scenarios on the personal computer or other hand-held e-devices one can 'enter a temple or church and the like, pray in your own way and using your own language, call your favourite priest to do the rituals (e.g. *puja*) for you to your favourite deity', without leaving your home or workplace. Debates, discussions, discourses, and reality shows on spiritual matters in different languages are becoming popular thanks to the television.

Knowledge organization tools (KOTs) include ontologies, taxonomies, schemes for subject classification, thesauri, WordNets, semantic nets, self-organizing systems, etc. These are useful in vocabulary management, formulating search expressions and retrieval from databases (including online web-enabled sources) and also in organizing information resources (e.g. documents, non-documentary materials) and their surrogates (e.g. entries in catalogues and bibliographies) in a helpful sequence. Some of the KOTs, especially classification schemes and faceted thesauri, for specific subject fields can provide a good mapping of the domain and its concepts and their interrelationships [1]. Ontologies "firstly, help sustain a communications framework around the domains of interest between people, organizations, and systems

\*Address correspondence to this author at the Sarada Ranganathan Endowment for Library Science, 702, 42<sup>nd</sup> Cross, III Block, Rajajinagar, Bangalore 560 010, Karnataka State, India; Tel: (080) 23305109; E-mail: prasadkn@vsnl.net

by providing a shared and common understanding of the domain. Secondly, they enable knowledge re-use and sharing since other researchers can adopt or integrate an ontology for their own purposes..... They facilitate inter-operability among systems by specifying and translating different concepts and languages in a domain (or across several domains). A well-developed ontology produces cost-time benefits by eliminating or reducing the cost of re-inventing a knowledge-base system for each use. Furthermore, ontologies help users learn domain knowledge, and, in addition, browse and search for information by providing structured knowledge representation” [2].

However, the integrated use of two or more KOTs in the religious and spiritual domains appear to be quite limited, if at all. The work reported in this paper demonstrates the value of such integrated use. In other words, a Knowledge Organizing System (KOS) in which two or more types of KOTs that are interlinked or integrated so as to serve vocabulary management functions in information indexing, searching, retrieval and presentation, and *additionally to support knowledge discovery, data mining and enhance the scope for research in the subject domain and also in the field of KOS theory and practice.* The development of such integrated KOS very much depends on the collaborative contributions of information specialists and subject specialists.

Our work on information support for religious and spiritual studies began about a decade ago with the design and development of an online database of extracts from spiritual, religious texts and scholarly philosophical works. This OM02 database currently holds over 18000 extracts from works on the major religions of the world. The search can be index-based or direct free-text, or a combination of the two. Boolean searches are provided for. A detailed description of the main database of extracts (OM02) and the associated databases of lists of sources used (OMBIB), the biographical database (OMBIO) etc. is given in [3]. The database has been distributed to several interested persons and institutions. As the extracts were derived from such sources as the Vedas and Upanishads, Sufi literature, besides sources in English users felt that a glossary of the Farsi (Persian) and Sanskrit (Vedic) terms would be useful. Such a glossary (GLOS) has been developed. This glossary of over 5000 terms has subsequently been used in the development of thesauri described below. Some of the source materials were suggested and /or provided by scholars.

### 3. WORKING DEFINITIONS

*Classic* -- “a creation of the highest excellence ... of recognized authority” [4] (WordNet 2.1).

“Work usually having embodiments in several versions, adaptations, and translations, attracting other works on itself, and getting copied out and/or brought out in print (or electronic form) even long after its origin.” [5, FA7] There are classics in all major disciplines and subjects. Many of the sacred works are examples.

*Quasi-Class* – “Work which is treated as a class in a scheme for classification, or whose title is used as a subject-heading in cataloguing practice.” [5, FA8]. A classic is an example.

Data mining: The process of extracting hidden information from large databases to explore patterns of relationship among entities to enable forecasting of trends in the field concerned.

*Typology of expositions*– According to Vedic tradition standards of exposition in discourse / composition are of four types:

1. Seminal work (Prabhu Sammita (= Jnana Sammita));
2. Near-seminar work (Suhrit Sammita);
3. Popular or non-specialist work (Kantha Sammita); and
4. Childrens’ work (Sisu Sammita).

*Seminal work*: Such work embodies ideas of a fundamental nature, mainly arising spontaneously from intuition of great minds. Such a work expresses depth of thought and implies layers of implications. It helps in the fertilization of new ideas in other minds. In its original nascent form the ideas embodied may be fully comprehended only by a few scholars. The ideas would reach others through restatement and translation in a variety of ways and at different standards. There are many works of this nature in the fine arts and literature, religion, spirituality and philosophy.

*Near-seminar work*: Such work embodies near-seminar ideas arising from researches carried out by intellectuals. They may be guided by the fundamental ideas (seminal works) - commentaries, evaluations, analytical studies and even taking new directions. Such intellectual work, largely embodied in scholarly communications, *Suhrit Sammita*, make up the major source of information and knowledge in the domains they deal with.

*Non-specialist (popular) work*: The ideas embodied in seminal and more particular in near-seminar works have implications to and applications in a many situations and disciplines. These, in turn, may have social implications and, therefore, need to be widely disseminated. This may be achieved through translations in different languages and the exposition should be such as to make the seminal and near seminal ideas and their implications and applications intelligible and understandable to non-specialists [6].

There may be and should be expositions of selected ideas to suit the learning needs of children.

The above categorization is helpful in document selection in terms of their level of exposition and their potential user groups, and in preparing annotations to such works in surrogates (e.g. catalogues and bibliographies).

Other types of information sources include commentary, critique, review, state-of-the-art and trend reports, abridgement, adaptation, summary, digest for executive / technician etc. Some of these cut-across the boundaries of the main categories mentioned above.

Commentary and evaluation (critique) are of particular interest to scholars researching classics. In our work on Tamil classics the following guidelines on commenting and evaluation were noted [7]:

*urayiRkOTal*: An accepted canon of exegesis which permits of additions being made to the subject matter of a text

by the commentator in the course of his exposition. As per this canon:

- *urayirkoL*: To make additional statements in a commentary about a subject mentioned in the text;

- *eTuttOttu*: Detailed comment, exposition; expositive statement;

- *caTutti*: Examination, inspection, review, search; and

- *uyttuNaravaippu*: Use in an exposition of such expressions as would stimulate thought or further enquiry.

These guidelines are helpful not only to the reviewer and commentator on texts but also to the cataloguer in preparing annotations, and in deep indexing of the texts which can assist comparative studies and data mining.

With a similar purpose in view and to facilitate deep classification and indexing an extensive analytical study of subforms of Tamil poetry was done by Ranganathan and Thilainayagam [8]. Such analysis is helpful in responding to queries of the type: What subforms of poetry or hymnals are more frequently devoted to a particular deity (e.g. Siva, Ganapati, Murugan, Saraswati, Lakshmi, etc.) and during which type of *puja* or offerings. Also useful to the beginner / learner for priesthood, and for a scholar's study of Tamil poetry.

#### 4. REQUIREMENT OF SCHOLARS

In designing and developing the Colon Classification (CC), the author S.R. Ranganathan had frequent interactions with the users and potential users of information sources in various subject fields. In respect of domains such as, fine arts, literature, linguistics, religion, mysticism, philosophy, etc. he understood the expressed need of scholars and researchers to bring together the classics in the different domains along with their respective authors, commentaries, evaluations, translations, etc. in a helpful sequence on the shelves and in the catalogue entries, and also to provide appropriate access to these materials through the index entries. An example from CC, ed. 6 (page 3.23) [9]:

[Note: x1 is the notation for the Classic Device in CC

Q Religion

Q2 Hinduism (post-Vedic)

Q21 Saivism, Smartaism

Q22 Vaishnavism

In the extract from the schedule, given below, the name of the author / source is given in parenthesis

Q21 Smartaism / Saivism

(*selections from Vedas*)

Q21,21x1,1 purusha-sUktam

Q21,22 purANam

Q21,221 brahma purANam

Q21,222 mArkantEya purANam

Q21,223 bhavishya purANam

.....

Q22 Vaishnavism

Q22,417 Hymn

Q22,417x1 nAlAyira-divya-pirapantam

Q22,417x1,1 First thousand

Q22,417x1,11 tirup-pallANTu (PeriyAzhvar)

Q22,417x1,12 tiru-moLi (PeriyAzhvar)

Q22,417x1,13 tiru-pAvai (ANTAL)

Q22,417x1,14 nAcciyAr-tiru-moLi (Andal)

.....

Q22,417x1,2 Second thousand

Q22,417x1,21 periya-tiru-moLi (Tirumangkai Azhvar)

Q22,417x1,22 tiruk-kurun-tANTakam(Tirumangkai

Azhvar)

.....

Q22,417x1,3 Third thousand

Q22,417x1,31 mutal-tiru-vantAti (Poikai AahvAr)

Q22,417x1,32 iranTam-tiru-vantAti

(PutattAahvAr)

.....

Q22,417x1,4 Fourth thousand

Q22,417x1,41 tiruvAy-moLi (NammAzhvar)

Q22,417x1,42 iRmAnucA nURRanTAti

Q22,4173x1 vishNu-sahasra-nAmam

Q22,4173x1,1 Sankara's commentary

.....

It may be noted here that many of the early classic sacred works were presented in the form of verses and hymns. In such a case the work may be assigned to the Literature class. Some of them have verses dealing with spiritual matters and/or eulogizing a particular deity. Therefore, the KOT used should provide for analytical indexing and point to the spiritual content and interlinking among concepts facilitating data mining. The input and collaboration of scholars is obviously necessary here.

Here is an example of an anthology classified under Tamil literature (poem) that contain verses treating religious, moral and social conduct, human values and virtues, causes of pain and suffering

[Link from bi-lingual thesaurus term [patin-en-kil-k-kanakku](#) to Tamil Term in the *Tamil Lexicon* [7]

O,31,1C [O = Literature, 0,31 Sangam (period) literature]

[பதினெண்கீழ்க்கணக்கு](#) (p. 2476) [[patinenkizkkanakku](#)] n [patin-en-kil-k-kanakku](#) . < id. +.

The group of 18 short classics, mostly in venpā metre, on one or more of the themes aram, porul, inipam, namely, nāl-ati vār, nānmani-k-katikai, innā-nārpātu, iniyavai nārpātu, kār-nārpātu,

kalavali-nāṛpatu, ain tinai-aimpatu, aintinai-elupatu, tinaimoli-y-aimpatu, tinaimālai-nūrāimpatu, tirukkural, tirikatukam, ācāra-k-kōvai, palamoli, ciṇu pañca-mūlam, mutumoli-k-kāñci, ēlāti, kainilai; அடிநிமிர்வில்லாது பெரும்பான்மை வெண்பா வான் அமைந்து அறம், பொருள், இன்பங்களுள் ஒன்றை யேனும் பலவற்றையேனும் கூறும் நாலடியார், நான்மணிக்கடிகை, இன்னாநாற்பது, இனியவைநாற்பது, கார்நாற்பது, களவழிநாற்பது, ஐந்திணையம்பது, ஐந்திணையெழுபது, திணைமொழியைம்பது, திணைமாலை நாற்றைம்பது, திருக்குறள், திரிகடுகம், ஆசாரக்கோவை, பழமொழி, சிறுபஞ்சமூலம், முதுமொழிக்காஞ்சி, ஏலாதி, கைந்நிலை என்ற பதினெட்டு நூல்கள்.

ஆசாரக்கோவை (p. 0210) [ ācāra-k-kōvai ] n ācāra-k-kōvai . < id. +. A classic work treating of religious, social and moral conduct in 1 stanza by Peruvāyīn-mulliyār, one of patinen-kīl-k-kanakku; பதினெண்கீழ்க்கணக்கு ளொன்று.

இன்னாநாற்பது (p. 0369) [ innānāṛpatu ] n innā-nāṛpatu . < id. +. An ancient didactic work by Kapilar consisting of 4 stanzas in each of which are mentioned certain outstanding causes of pain and suffering in general, one of patinen-kīl-k-kanakku; பதினெண்கீழ்க்கணக்கு ளொன்று.

இனியவைநாற்பது (p. 0371) [ iniyavaināṛpatu ] n iniyavai-nāṛpatu . < id. +. An ancient didactic work by Pūtañ cētanār, consisting of forty stanzas, and describing various matters considered pleasurable, delightful and good, one of patinen-kīl-k-kanakku; பதினெண்கீழ்க்கணக்கு ளொன்று.

ஏலாதி (p. 0566) [ ēlāti ] n ēlāti . < elā + ādi. Lit., 'Cardamom and other ingredients', name of an ancient didactic work by kaṇi-mēṭāyiyār, containing 1 stanzas and mentioning six virtues in each stanza, one of patin-en-kīl-k-kanakku.; பதினெண்கீழ்க்கணக்குக்களு ளொன்று.

In another paper [10] it is shown that the CC class number can be used as a code to switch between languages.

## 5. THESAURI

### 5.1. Bi-lingual Thesaurus

We present here experiences from designing a bilingual thesaurus and a trii-lingual thesaurus We have used Unesco's WINISIS DBMS application software for the design and development of thesauri and to link the terms to the corresponding related terms in the different schedules within the thesaurus and also to online glossary database, text files in non-English languages (e.g. Tamil) and the remotely located online lexicon.

The Centre of Excellence for Classical Tamil was established by the Government of India, about five years ago, with the view to promoting studies in classical Tamil. The Centre started in Mysore at the Central Institute of Indian Languages has been moved to Chennai (former Madras), South India. Included among several programmes initiated by the Centre one relates to the development of a digital library for supporting classical Tamil studies . A component

of this programme is the design and development of a bi-lingual (English-Tamil) thesaurus for information management. Work on the thesaurus was started a couple of years ago and currently contains about 100,000 descriptors (transliterated terms) together with an index. The English transliterated terms (descriptor, BT, NT, RT, USE, UF terms) are hyperlinked among themselves and to terms in the online *Tamil lexicon* [7]. A more detailed description of the design and development and use of these thesauri are given in other papers [3,4]. For building such tools for a subject-field, documentary sources, existing relevant thesauri, glossaries and lexicons in machine-readable format (or online) plus interactions with subject experts and their inputs constitute the main base on which information specialists design and develop.

At the earlier stages of the development of the bi-lingual thesaurus a couple of demonstrations were presented to Tamil scholars. Among other things, a desire was expressed that the thesaurus should include not only concept terms especially from the 41 Tamil classics of the Sangam period (300 BCE – 200 CE ) but also the names of the classics enumerated in a helpful sequence together with the names of

the respective authors, compilers, commentaries, commentators, etc. The latter will normally be a function of a catalogue or subject bibliography. Nevertheless, there are also advantages in including the names of the classics, their authors, commentaries, names of authors of commentaries, etc. in the thesaurus or a scheme for classification. It amounts to treating a classic as if it were a 'subject' by itself. That is as a quasi-class (see Section 3).

Initially the 41 classics were listed in a MSword file and linked to the term *cangkanUI* (Tamil classics) in the thesaurus. The list will be displayed when clicked on the term *cangkanNul* (tamil classics) in the thesaurus. Each of the 41 classics gets an entry in the thesaurus with indication of related works, authors' names, etc. and linked to related entries in the online *Tamil Lexicon* [7]. Each of these entries can be clicked to surf to the broader anthology name and / or subsections, commentary(ies) etc. Subsequently at the suggestion of the scholars (users) the MSword file was abandoned and the titles of the 41 classics, the related commentaries, names of authors, etc. were entered in one record. Each line can be clicked to get a link to the entry for the commentary, section of the classic, etc. Please see Annex 1.

Working on the bilingual thesaurus led to a discovery of information about the Tamils of the Sangam period. In early times the Tamils, like communities in most parts of the world congregated in particular types of land environment which provided them the means of living and enjoyment. The Tamil communities gave names for the different land types (tracts) and the varieties of flora (plants, climbers / creepers, trees and shrubs) that identified each of the tracts. For example:

*kuRinci* ீே௧்யு௧ = Conehead varieties; *marutam* ீௌ௩/௪ = Paddy; *mullai* ீௌ = Jasmine varieties; *neytal* ீௌ/௪ = Water lily; *pAlai* ீ௩ = Desert vegetation.

The inhabitants lived in harmony with nature and were obviously influenced by nature's offerings; they identified the different regions by the type of principal flora and gave the name of the flora to the respective tracts, lands or regions. Thus:

*kuRinci* ீே௧்யு௧ = Hilly land / tract; *mullai* ீௌ = Pastoral or sylvan land / tract; *marutam* ீௌ/௪ = Agricultural plain / land / tract; *neytal* ீௌ/௪ = Coastal land / Maritime tract; *pAlai* ீ௩ = Desert land / tract.

Subsequently the particular way of life in the different tracts / lands (*nilam*), rules of conduct (*tinai*), the poetic love conventions (*akattinai* = love, as a mental experience of lovers, of seven forms, namely, *kuRinci* ீே௧்யு௧, *mullai* ீௌ, *marutam* ீௌ/௪, *neytal* ீௌ/௪, *pAlai* ீ௩, *kaikkilai* ீ௩, *peRuntinai* ீௌ/௪), and melodies prevalent in the different tracts, came to be identified by the respective names of the tracts. Thus, to quote from *tolKappiam*:

**"akattiNai:** (ௌ/௪) A distinctive idealized love career, treated in the *akam* literature of the *Cankam* poetic corpus, falling into two phases, viz., *kaLavu* (ௌ) (pre-marital clandestine course) and *karpu* (ௌ) (wedded course), and manifesting in seven kinds of emotional behaviour, namely *kaikkiLai* (ௌ) (unrequited sexual passion), *kuRinci* ீே௧்யு௧ (union), *mullai* ீௌ (forbearance), *pAlai* ீ௩ (separation), *marutam* ீௌ/௪ (sulking), *neytal* ீௌ/௪ (lamentation) and *peruntinai* ீௌ/௪ (abnormal lustful passion) set against the backdrop of *kuRinci* ீே௧்யு௧ (mountainous), *mullai* ீௌ (sylvan), *pAlai* ீ௩ (desert-like condition of mountainous and pastoral regions in abnormally hot season), *marutam* ீௌ/௪ (cultivated plains) and *neytal* ீௌ/௪ (coastal) regions" (Murugan, 2000).

Similar parallel practices among indigenous peoples of North America was also reported. [11]

Hyper-linking of concepts of the thesaurus to corresponding term(s) in the online *Tamil Lexicon* and generating a KWIC index therein, indicated several types of relationships among concepts (See Fig. 1 to 4). Another example: the concept / term Lotus (and its Tamil equivalents *Tamarai* and *Kamalam*) occurs as representing the lotus flower, as a pre-

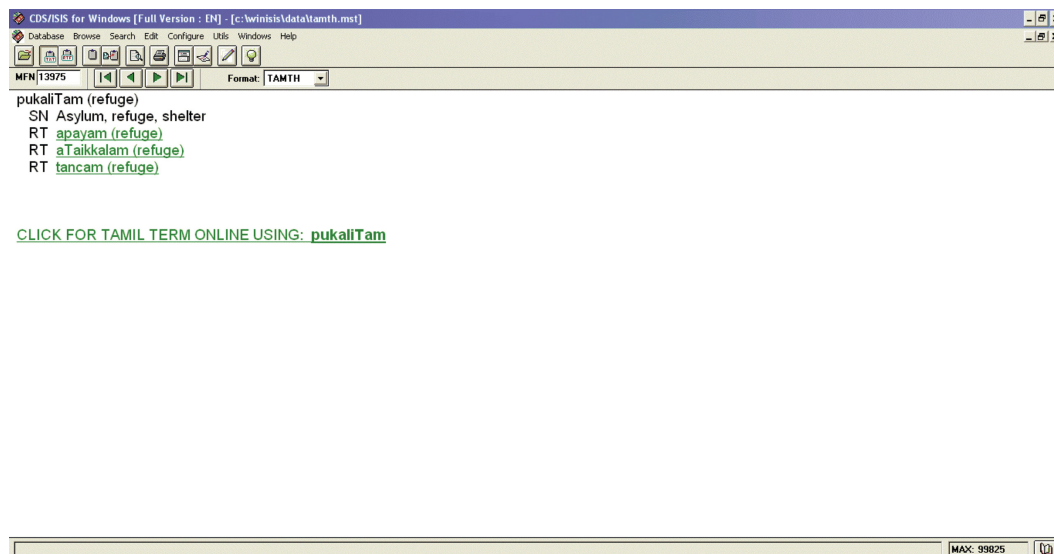


Fig. (1). Screenshot Extract from English – Tamil Thesaurus.

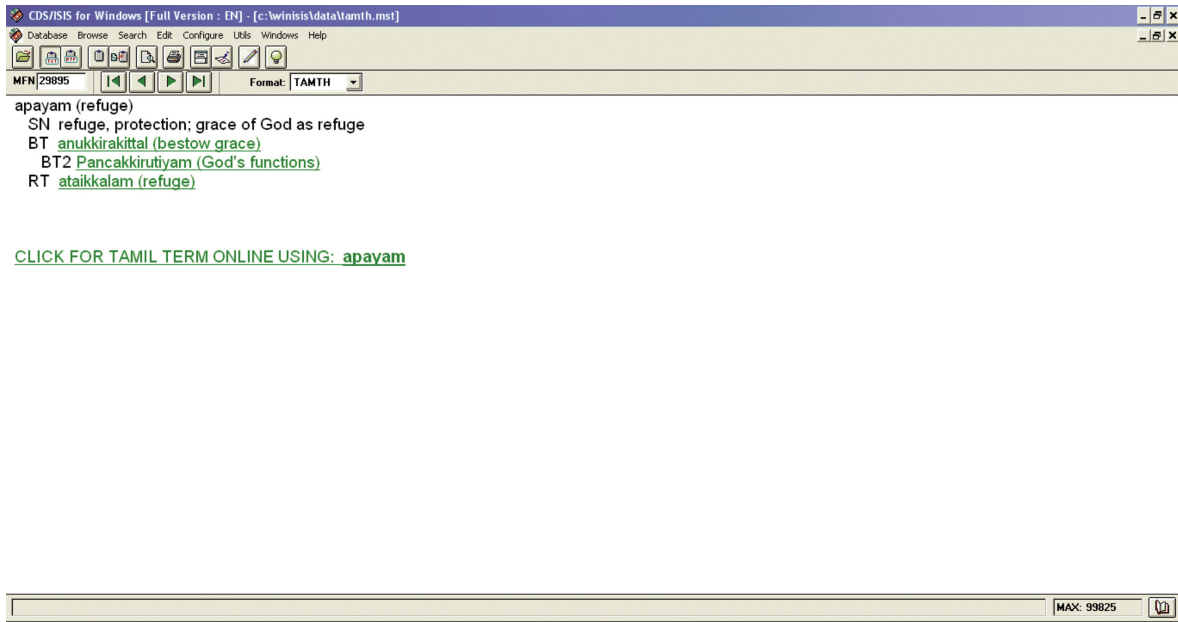


Fig. (2). Screenshot Thesaurus Record for 'apayam (refuge)'.

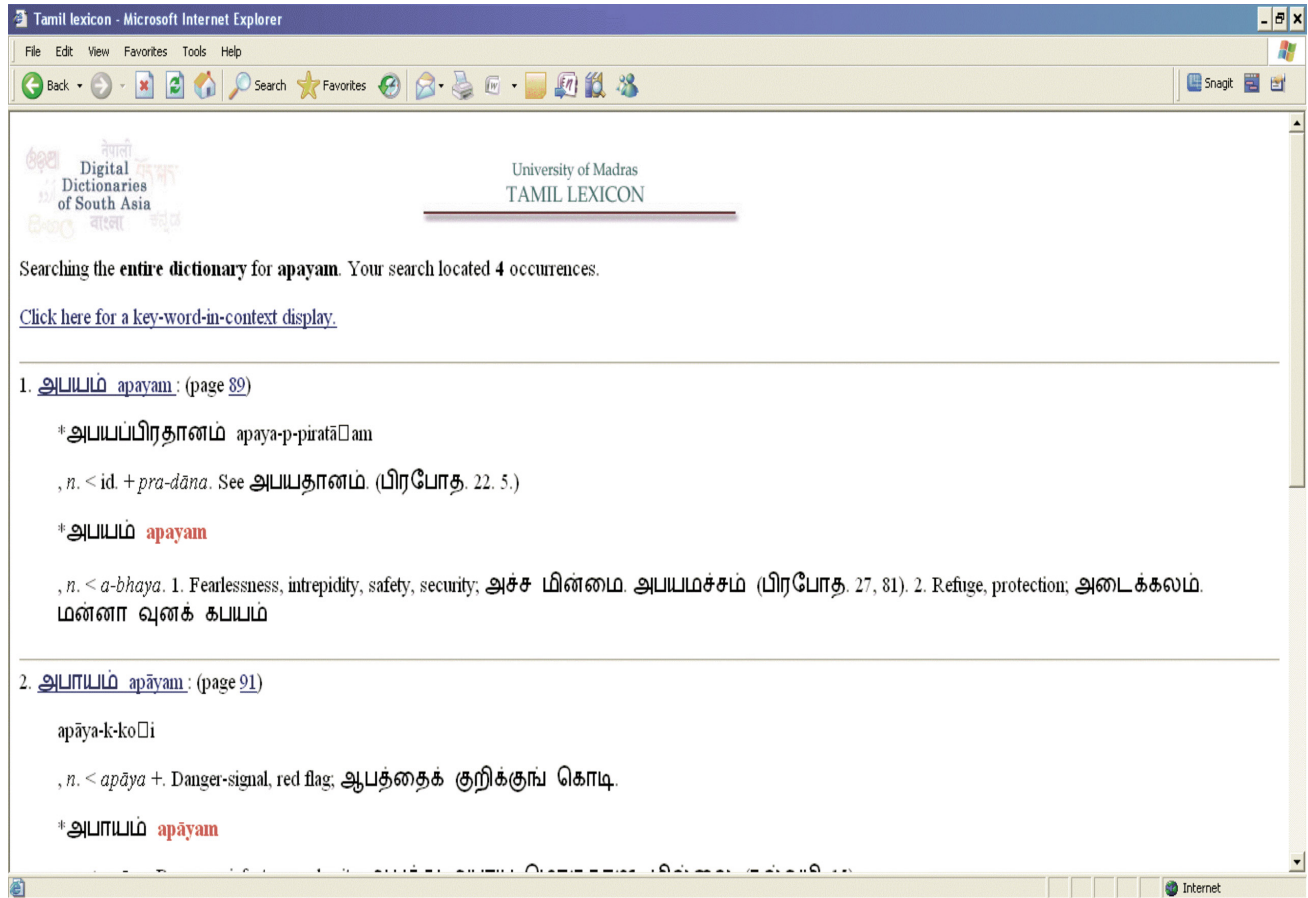


Fig. (3). Screenshot Partial Set of Records on 'apayam' from online Tamil Lexicon (Univ. Chicago).

ferred flower in worship (in temple or home), as a qualifier to the goddess Lakshmi (as one seated on lotus), as lotus pose in Indian classical dance *mudras* (poses) (see below), as lotus shape and patterns in temple architectural elements, as decoration on vessels or other utensils used in worship rituals, lotus-like property (soft as lotus-petal), as one pos-

sessing lotus petal shaped eye (e.g. kamalanayanan, kamala kaNNan, kamalAksi, names of deities), and so on. The KOT used should disclose / indicate these occurrences to facilitate data mining [10].

In India and in many other countries particular plants, flowers, and products made with or out of them are associ-



Fig. (4). Screenshot KWIC Index for Records in Fig. (3).

ated with particular deities, and sacred rituals, in addition to their being used, for instance, in medicinal preparations. Examples: Lotus, Basil (Tulasi), Sandal wood etc. Even when they are used for their medicinal value, such applications may be accompanied by chanting of certain hymns and / or *mantras*. Tulasi leaves and consecrated water presented as offering to the idol in a Vishnu temple or at home is distributed to worshippers or members of the congregation. Similar practice is observed with sandal wood and sandal paste not only in worship in Hinduism but also in Buddhism. Chinese, Japanese and Zoroastrian worship. Use of sandal wood as an incense is a 4000 year old practice. Such practices are considered to facilitate the praying and communication with the deity. The KOT should display such relationships among concepts.

## 5.2. Tri-lingual Thesaurus

A description of the design and development of a tri-lingual (English, Farsi (Persian), Vedic Sanskrit/ Pali) thesaurus (F-THES) for mysticism and allied fields is given in a paper by Neelameghan and Raghavan [12]. The terms in Farsi and Sanskrit are transliterated into Roman script. The index of terms (of the thesaurus) is a single alphabetical file. A user can select a term in any one of the languages (English, Farsi or Sanskrit) to formulate the search expression. The suffix to the index term indicates the language of the concept term. The output will be the matching descriptor (in the language of the search term) with the BT, NT, and RT

also in the same language, followed by the equivalent term in the other two languages. Example:

Let us say that a user look up the term 'bhakti' in the index to F-THES and he / she selects BHAKTI (V) where the suffix (V) indicates Vedic Sanskrit. Similarly the suffix (S) indicates Sufi (Farsi) term and (E) English. Clicking the term enters it in the search box. Executing the search and display options displays the following record (Fig. 5):

In the above screen-shot (Fig. 5) the first set of NT and RT terms, suffixed with (V) relate to the descriptor bhakti (V), then the equivalent / near-equivalent terms in English and Farsi (S). User may select and click any of the underlined terms. For example, if he/she clicks on munib (S), the following thesaurus record in which munib is the descriptor will be displayed (Fig. 6).

User may select and click on any of the terms in this schedule and surf the thesaurus.

The design of F-THES is such that three separate thesauri for Mysticism – of English, Farsi and Sanskrit terms respectively – have been derived. Each term in F-THES can be hyperlinked to relevant glossary or lexicon in machine readable form either stored as a database in the computer, in a network or accessed in remote locations using the URL.

The linking of related parts of the thesaurus schedules for a given concept term different chains / networks – terms in the same language or corresponding terms in different lan-

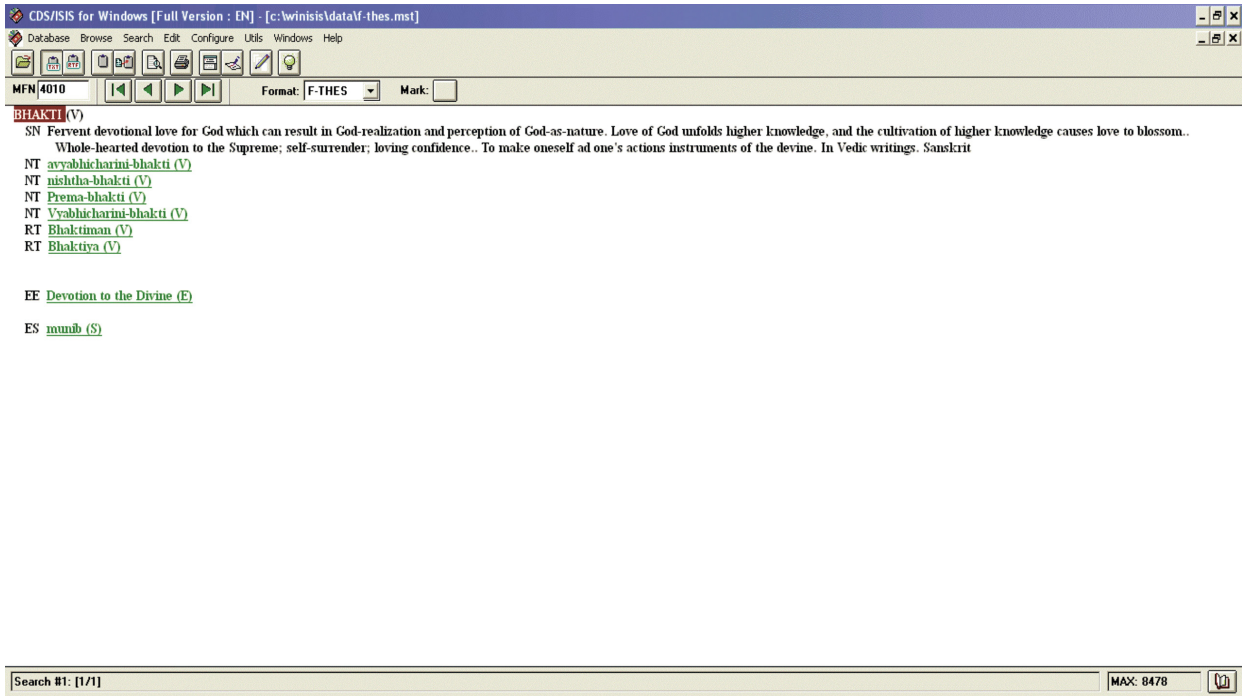


Fig. (5). Screenshot Tri-lingual Thesaurus Schedule for 'bhakti'.

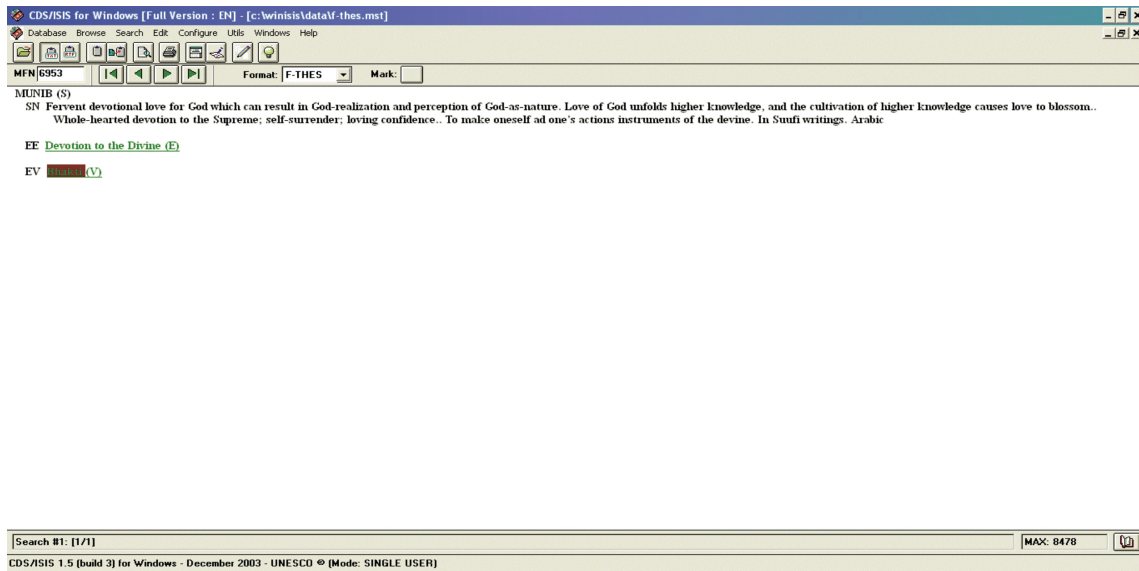


Fig. (6). Screenshot Linked Record for 'MUNIB(S)'.

languages (e.g. English, Farsi, Sanskrit) - are formed. It provides the basis for comparative study of the hierarchic and non-hierarchic associative term links and corresponding relationships in the different languages. The difference in the linkages and relationships may arise from the particular characteristics of the languages and / or from the cultural differences of the native speakers of the concerned languages. Such studies could be collaborative work between language/ linguistics specialists and KOT specialists.

**6. CONCLUDING REMARKS**

Case studies of integrated applications of classification schemes, thesauri hyperlinked to lexicons, etc. have been presented to demonstrate value addition to information retrieval service, creating and facilitating data mining and

knowledge discovery. The importance of collaborative contributions of scholars and information specialists in the development of such integrated systems has been emphasized.

Compared to subjects in science and technology, terminology in the humanities is fuzzier and tends to be interpreted differently in different contexts. It means that in the humanities problems are encountered in managing vocabularies and also because of the occurrence of large numbers of homographs and homonyms for many descriptors; further, conceptual relationships are often culture-specific and reflected in the language of cultural groups or native speakers of the language and exact equivalents may not be found in the languages of other groups. These issues are further discussed in the case studies.



## ANNEX 1.

CDS/ISIS for Windows [Full Version : EN] - [c:\winisis\data\tamth.mst]

Database Browse Search Edit Configure Utils Windows Help

MFN 14029 Format: TAMTH Mark:

**cangkanUl** (Tamil classics)

SN The 41 Tamil classics of the Sangam period

NT [tolkAppiyam](#) (Tamil grammar)

NT2 [poruLatikAram](#) (division of TolKAppiyam)

NT [patineNmElkaNakku](#) (upper classics)

NT2 [ETTuttokai](#) (anthology)

NT3 [naRRinai](#) (anthology)

NT3 [kuruntokai](#) (anthology)

NT3 [aingurunURu](#) (anthology)

NT3 [patiRRuppattu](#) (anthology)

NT3 [paripAtal](#) (anthology)

NT3 [kaLittokai](#) (anthology)

NT3 [akanAnUru](#) (anthology)

NT3 [puranAnUru](#) (anthology)

NT2 [pattuppATTu](#) (anthology)

NT3 [tirumurukARRuppaTai](#) (poem)

NT3 [porunArARRuppaTai](#) (poem)

NT3 [cirupANaRRuppaTai](#) (poem)

NT3 [perumpANaRRuppaTai](#) (poem)

NT3 [mullaippATTu](#) (poem)

NT3 [maturaikkANci](#) (poem)

NT3 [neTunAlvATai](#) (poem)

NT3 [kurincippATTu](#) (poem)

NT3 [paTTinappAlai](#) (poem)

NT3 [malaipaTukaTAm](#) (poem)

---

NT [patiNeNkilKkaNakku](#) (lower classics)

NT2 [nAlaTiyAr](#) (Tamil classic)

NT2 [nAnmaNikkaTikai](#) (didactic work)

NT2 [nAnArpatu](#) (poem)

NT2 [innAnArpatu](#) (didactic work)

NT2 [inivavainArpatu](#) (poem)

NT2 [kAmArpatu](#) (love poem)

NT2 [kalavaLinArpatu](#) (poem)

NT2 [aintiNaiaimpatu](#) (poem)

NT2 [aintiNaielUpatu](#) (love poem)

NT2 [tiNaimoLiaimpatu](#) (love poem)

NT2 [tiNaimAlainURRaimpatu](#) (love poem)

NT2 [tirukkuraL](#) (classic on virtue)

NT3 [ArattuppAl](#) (poem)

NT3 [porupAl](#) (poem)

NT3 [kAmattuppAl](#) (poem)

NT2 [tirukatukam](#) (didactic work)

NT2 [AcArakkOvai](#) (didactic poem)

NT2 [paLamoLi](#) (didactic poem)

NT2 [cirupancamUlam](#) (didactic poem)

NT2 [mutumolikkAnci](#) (didactic poem)

NT2 [EiAti](#) (didactic poem)

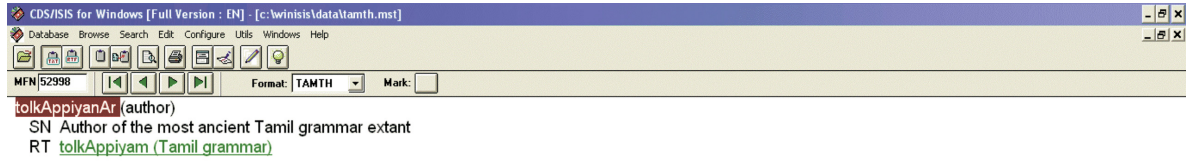
NT2 [kainnilai](#) (didactic poem)

NT [muttoLLAyiram](#) (poem)

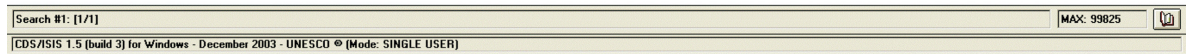
NT [kalaviyal](#) (poem)

CLICK FOR TAMIL TERM ONLINE USING: [cangkanUl](#)

Search #1: [1/1] MAX: 99825



CLICK FOR TAMIL TERM ONLINE USING: [talkAppiyam](#)



Partial Set of the Records from online Tamil Lexicon (Univ. Chicago)

**REFERENCES**

- [1] Neelameghan A. Itself an education. *Inf Stud* 2008; 14(3): 133-4.
- [2] Park OK. Opening ontology design: a study of the implications of knowledge organization for ontology design. *Knowl Org* 2008; 35(4): 209-21.
- [3] Rajashekar TB, Ravi SA, Neelameghan A. Designing a multimedia information service for the Internet and CD-ROM. *Inf Stud* 1998; 4: 125-42.
- [4] Word Net 2.1 (<http://www.wordnet-online.com>)
- [5] Ranganathan SR. Classified catalogue code with additional rules for dictionary catalogue code. Bangalore, India: Sarada Ranganathan Endowment for Library Science; 1964; p. 119, Sections FA7 and FA8.
- [6] Ranganathan SR, Ed. Documentation and its facets. Chapter B, Section 4, 1963.
- [7] Tamil lexicon. Published under the authority of the University of Madras. Madras, India University of Madras; 1924-1939. Reprint 1982. v.I-V + Supplement VI. [Online version, Dictionaries of South Asian Languages, Chicago, US: University of Chicago. <http://dsal.uchicago.edu/dictionaries/tamil-lex/>]
- [8] Ranganathan SR, Thillainayagam V. Subforms of Tamil poetry and their classification. *Ann Lib Sci* 1963; 10: 14-37, 175-86.
- [9] Ranganathan SR. Colon classification. 6<sup>th</sup> ed. Bangalore, India: Sarada Ranganathan Endowment for Library Science; Part 3: Schedules of classics and sacred books with special names. 1960; 27<sup>th</sup> reprint 2006.
- [10] Neelameghan A. Value addition by online integration of knowledge organization tools: a case study of a bilingual thesaurus + Colon classification + Lexicon. International Conference on Future of Knowledge Organization in the Networked Environment, 3-5 September 2007; In: Raghavan KS, Ed. Bangalore, India. IKONE 2007. 2007; p. 115-130
- [11] Neelameghan A, Raghavan KS. A knowledge organizing system for humanistic disciplines with enhanced capabilities: Case studies. *Inf Stud* 2009; 15(2): 75-94.
- [12] Neelameghan A, Raghavan KS. An online multi-lingual, multi-faith thesaurus: A progress report. *Webology* 2005; 2(4): article 14, (<http://webology.ir/2005/v2m4/a14html>).

---

Received: March 30, 2010

Revised: May 07, 2010

Accepted: May 07, 2010

© K.N. Prasad; Licensee *Bentham Open*.This is an open access article licensed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted, non-commercial use, distribution and reproduction in any medium, provided the work is properly cited.