

Application of WINISIS/GENISIS Software in Newspapers Clippings

Lavji N. Zala & Niraj R. Patel

Abstract:

The paper describes how Knowledge Exchange and Information Centre (KEIC) of Mudra Institute of Communications, Ahmedabad makes the newspaper clippings accessible for its target audiences (users). The bibliographic information (title, author, newspaper's name, date, product category, brand name, and company name, etc.) as well as full text (down-loading from website of concerned newspaper) of the article is recorded in the newspaper module. The whole procedure is done in WINISIS software and then the data is converted/up-loaded in GENISIS software to make accessible the web-catalogue of newspaper clippings for its end-users. Over the period of time using CDS/ISIS software (DOS, windows, GENISIS versions) which type problems were faced and resolved, the merits and demerits of every version, etc., have been discussed in this paper. In the age of information technology, the IT professionals and other organisations are designing effective but very costly software. The small libraries don't have capacity to carry too much financial burden by purchasing these types of software. However they wish to get their libraries automated and make the resources available to their users by applying very cost-effective software. This paper describes how the CDS/ISIS software will become a boon for these types of information providers (libraries/information centres) and information receiver (users/clients). This paper focuses on every stage of evolution of CDS/ISIS and its successful implication in KEIC.

1. INTRODUCTION

Library denotes itself as a well-organised information system, which enables the users to find out the required information in minimum time. The printed reading materials like manuscripts, books, etc., were useful source of information in the past. The users demanded for specific titled books. But with the passage of time the explosion of literature resulted in users are demanding information exhaustively and pin-pointedly. Information today exists in many forms than just as printed material this led to libraries into digital era.

Libraries started the use of computers and other technology for housekeeping operations

and information services. Digital libraries break the barriers of physical boundaries and allow to users to access the digital resources across in house or varied domains of the world. Along with journals the newspapers are playing important role to disseminate latest information in all spheres.

The Knowledge Exchange and Information Centre (KEIC) run by Mudra Institute of Communications, Ahmedabad (MICA) has taken initiative to build up a newspaper module for its more than 75,500 (including 21,878 articles with full text) newspapers clippings using WINISIS/GENISIS software to help its end-users to find out information as per requirement.

2. KNOWLEDGE EXCHANGE AND INFORMATION CENTRE (KEIC)

MICA is an academic institution that has been set up in 1991 to develop professional communication skills in India. The institute primarily aims at imparting relevant education to young men and women with a view to provide them with the expertise to become effective communications professionals.

Further, the institute also aims to conduct research in communication as applied to the needs of the communications industry. Through such research, the intent is to create and share new knowledge in this fast growing and increasingly important field. As part of its agenda, MICA also means to collaborate with other institutions, both Indian and foreign, to jointly work on expansion of the knowledge base in the field of communication and its related areas. To fulfill the goal of MICA and its human capital including students, faculty, research associates, alumni and staff, the KEIC was established in 1994 with the basic objective of harnessing the huge body of information available throughout the world on all aspects of people in their role as consumers and on all forms of communication.

□ Collection Development

Table 1 shows the number of records of collection of different types of reading materials in databases.

3. DIGITIZATION OF NEWSPAPER CLIPPINGS

Digitization is a process of converting any item, it being in printed format, manuscript, image, or sound, film and video recording- from one format (print or analogue) into a digital. Basically it involves capturing images of the physical object using various types of scanners and digital cameras- and converting to digital format that can be stored electronically and accessed via computer.

4. EVOLUTION OF NEWSPAPER CLIPPING PROCESS

The articles associated with prime subjects of MICA are stored in well organised manner in three stages using WINISIS/GENISIS software. The evolution of digitisation of newspaper database passed through following phases:

4.1 Phase I (1994-1999)

KEIC has adopted the automation since its inception using DOS version of ISIS. Along with many modules, the newspaper module was also created using this software at the beginning. The newspaper module covered following fields or access elements in the field definition table (FDT) to store bibliographic information of selected news articles from selected major Indian newspapers. The FDT provides information on the contents of the master records in a given database. Each field has its own tag so that software knows only the tag of concern field and shows a

Table 1. Database of reading materials

Sr. No.	Name of database	Number of records	Mode of database
1.	AAA	1478	BI
2.	Book	5121	BI
3.	Case Study	947	BI
4.	Dissertations and Summer Report	1233	BI(728), FT(505)
5.	Newspaper Clippings	75500	BI(53622), FT (21878)
6.	Periodical	12200	BI
7.	Television Commercial	16915	Movies
8.	Print Ads	1539	Images

BI- Bibliographic Information FT- Full Text

result of a query accordingly. Tag is a unique number of a field/an access element in a given database. For example tag no. 10 has been allotted for author, tag no. 20 for title, tag no. 30 for newspaper name and so on.

- ❑ **Author(s)**
Name of author who is responsible for the express of concept, ideas and writings of the article
- ❑ **Title**
Title of news article
- ❑ **Newspaper's Name**
Name of the newspaper in which the news article is appeared
- ❑ **Date**
Date of publication of the newspaper. e.g. yyymmdd (20050605)
- ❑ **Edition**
Edition of newspaper e.g. Times of India (Ahmedabad, Mumbai), Telegraph (Calcutta)
- ❑ **Page No.**
This is the page number in MBS (monthly brief service) where the news article is appeared.
- ❑ **Subject**
Subject of news article. e.g. advertising, service marketing, branding, etc.
- ❑ **Product**
If particular product is discussed in news article it's name is mentioned in this field. e.g., soaps, biscuits, apparels, footwears
- ❑ **Company's Name**
Name of company that may be manufacture, consultant, service provider, marketer of a product whose name appeared in the article. E.g. Hindustan liver limited (HL)
- ❑ **Keywords**
Important words of news article. e.g. ad spend, turn over
- ❑ **Agency Name**
Suppose the news article is related to advertising campaign, the ad agency's

name is mentioned in this field which is taken responsibility to create print ad for specific brand. E.g. Mudra advertising agency, O & M

4.1.1 Procedure of Article Selection

The news articles are selected on the basis of institution's curriculum and users' requirements so that the users (students and faculty of the institute) may get information about happenings in the various fields like advertising, marketing, media, branding, management, retailing, business etc. Industry point of view is also considered at the time of selecting articles from concern newspapers.

- ❑ **Cutting/Pasting**
Once article was selected, it was cut and pasted on A-4 size paper.
- ❑ **Numbering**
All loose pages (A-4 sized papers) containing news article were arranged by dd/mm/yy. Thus every page went through the numbering process.
- ❑ **Indexing**
Indexing follows the numbering process. Each article was indexed as per the fields (author, title, newspaper name, brand name, product category, note, date, edition, page number, key words, etc). In this stage only bibliographic information was entered into the newspaper module.
- ❑ **Photocopying**
After completion of numbering according to date the process of photocopying begins.
- ❑ **Binding**
All these articles were compiled date-wise and were bound monthly. This is called the monthly briefing service (MBS). This process is done every month.

4.1.2 Procedure of Article Retrieval

All MBS (printed form) are arranged according to months of appropriate year to make accessible of them to users while searching through newspaper database, the users have to take reference of article like the date and the page no. and refer to MBS.

It means only bibliographic information was accessed through the newspaper database in the first stage.

❑ Demerits

- Users could get the access of only bibliographic information.
- Users had to refer the MBS for printed articles.
- All articles from selected Indian business newspapers were compiled and prepared in the form of MBS. So the procedure of cutting, pasting, numbering, indexing (in all fields), photocopying, binding and retrieval etc., effected to the cost effectiveness factor.

4.2 Phase II (1999-2003)

In the second phase selected articles went through the entire process of first stage as mentioned above but one more step was added, which is scanning of the original articles. Scanned articles (JPG format) were linked to a newly created field called 'image'. By clicking on the link, the scanned page was displayed on the screen. The users could find both options of their choice of getting articles in scanned form and bibliographic form. The scanned articles are also available as print copies. The user could also use 'expert search' in case of exhaustive search and 'guided search' in case of specific search

(e.g. specific brand, product, company, agency, subject)

The second stage was a little hope for the user to view the scanned article(s) of his interest. The digitisation process started from this stage. Figure 1 & 2 show bibliographic details and the scanned article.

❑ Merits

- Users could get the access of bibliographic information as well as full image of (JPG Format) of the concerned article.
- It saved the time of information seekers from referring to hard copies for articles.

❑ Demerits

- Search was not possible within JPG form
- Storage of JPG files needed much storage capacity in server.

4.3 Phase III (2003 onwards)

The third stage of evolution of digitisation has changed the outlook of services. Another field called full text (the entire text of article except charts, illustration, images, and tables) has been introduced in the newspaper database. Selected articles have directly been downloaded from the concerned newspaper's website and they have been pasted in the field of 'full text'. User enables to see the full text of required articles from any computer of MICA premises because the database is

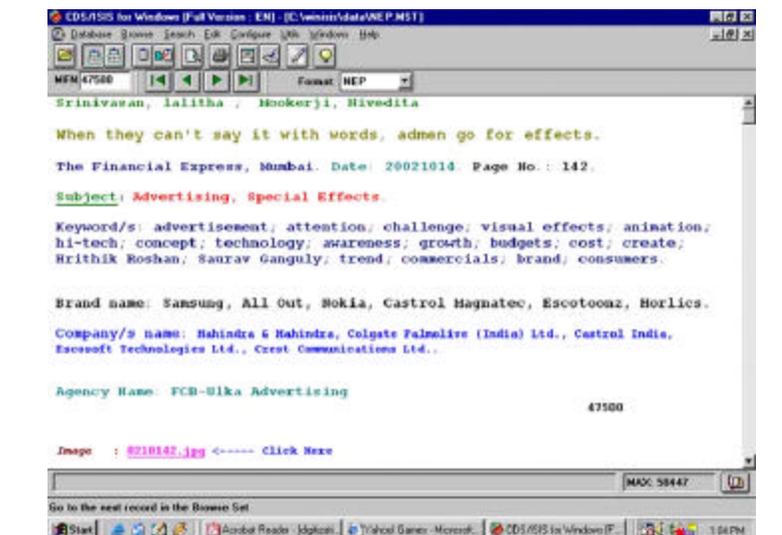


Fig 1. View of bibliographic details

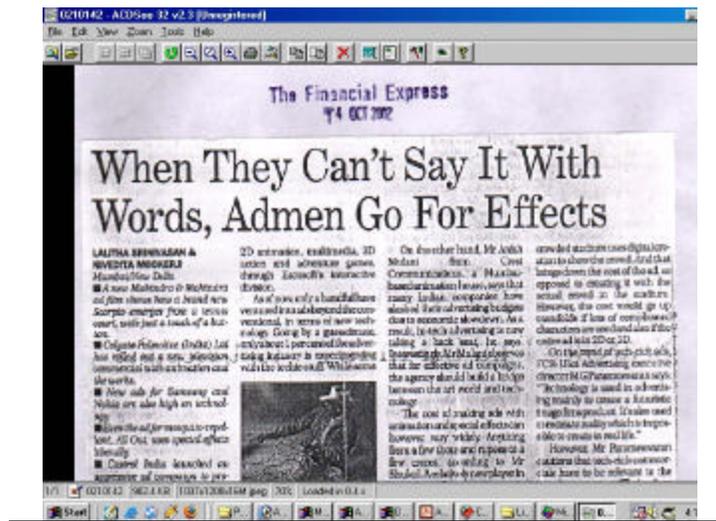


Fig 2. View of scanned article

placed on intranet. Figure 3 shows the full text view of article.

Those articles with images or tables, illustrations, charts (indexer feels these physical components are very important for information seekers) and some of articles are not found on website, they are gone through the process of cutting, pasting, numbering, indexing and binding (which was seen in first stage) in the form of hard copy.

□ Merits

- Those articles which were indexed in first and second stage of evolution of this database can be accessed through each of the fields or access elements.
- Those articles which covered pure text as well as some other physical components like illustration, tables, charts and images (these components are not supported by this software), and indexer feels the all components are very important for end-users to understand a higher level of authenticity, the impression of the whole page attraction, and covered important tabulated data and the pi-chart, these articles were indexed as whole (along with full text) and compiled the printed form of the concerned article. So that user can get both option of viewing full text as well as other physical components e.g., images, charts, illustration, table, etc., within the article(s) in hard copy.

- If users wish to search article with full text, the search through 'guided search' selecting 'full text' field can be retrieved as well search within full text is also possible.
- If users wish to search any information in whole database that is possible through 'expert search'.

□ Demerits

- The WINISIS does not support to physical components like images, charts, tables, illustration, etc. so these components are not covered in 'full text' field.
- The JPG files were removed in this stage so the users could not see the scanned articles (JPG Files). In that case if users need to go through the particular article, they have to get reference form database and has to refer hard copy of clippings.
- The important fields like brand name, product category, company name, agency name, etc., can not be accessed through 'guided search' because no one field covers any information or data in it. However users can search or access the information in same article through 'expert search'.

4.4 Phase IV

■ Web Cataloguing/GENISIS

Since end of 2003, KEIC has been using GENISIS version 3.0.0. Selected fields which

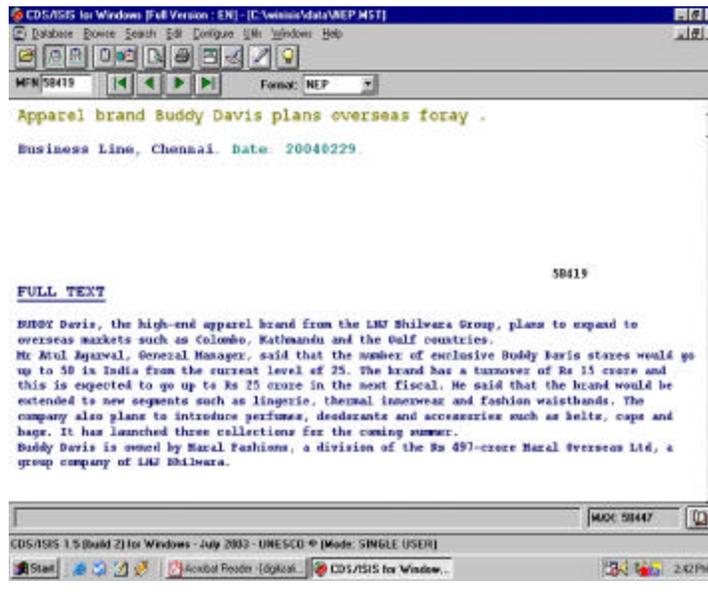


Fig 3. View of full text article

have been used in previous stages of implementation of CDS/ISIS software in newspaper module have been used in newspaper module in GENISIS version. Converting/up-loading of all data is taken place from WINISIS to GENISIS in the newspaper database. In this phase search is possible by all fields like author, title, newspaper's name, date, subject, product, company name, brand name, keywords, note, agency name and full text, etc. Figure 4 shows the search database.

Figure 5 mentions search result which shows that the word 'advertising' is appeared in 3707 full text articles in newspaper database.

Merits

- Web query form (with field selection and index access)
- Display of query results
- Display of particular record's results
- Speedily search and accessible anywhere because web-based technology used
- Searching 'full text' and within 'full text' is possible

Demerits

- Search is not possible within PDF files

- Physical components like images, charts, illustration, tables are not supported within articles

5. RETRIEVAL INTERFACE IN WINISIS/GENISIS

The important feature of any database in any library is fast and accurate retrieval of information. If the database has not facility to retrieve the data or information of users' query, it is considered no value of the same database.

The WINISIS has facility to extract one or more elements from master file record using field select table (FST). Depending on the context in which a FST is being used, these elements may then be used to create inverted file entries for the record from which they were extracted, for sorting records in the desired sequence before producing a printed report, or to reformat records during an import or export operation.

An element can be generally defined as a fragment of a record resulting from particular process. Although in many cases elements will be actual data elements, i.e., a field or subfield, in other cases they may be words, phrases, or any other piece of data which has a particular meaning to a specific application.

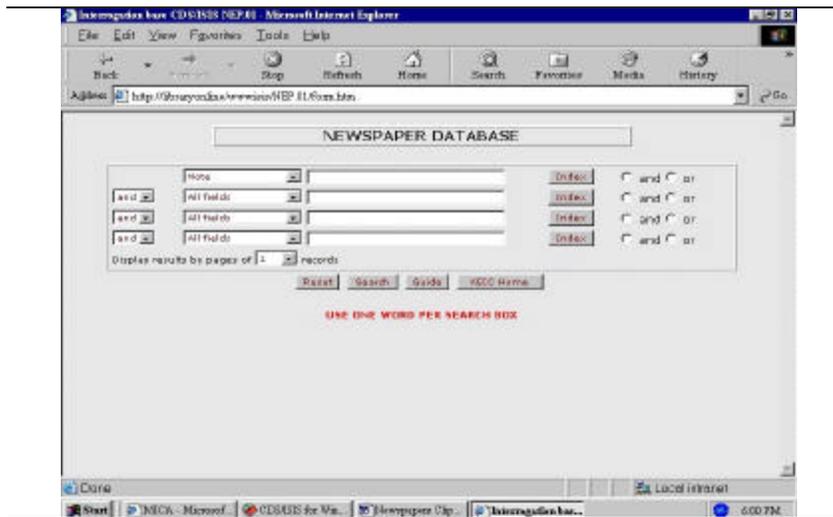


Fig 4. View of newspaper database in Genisis

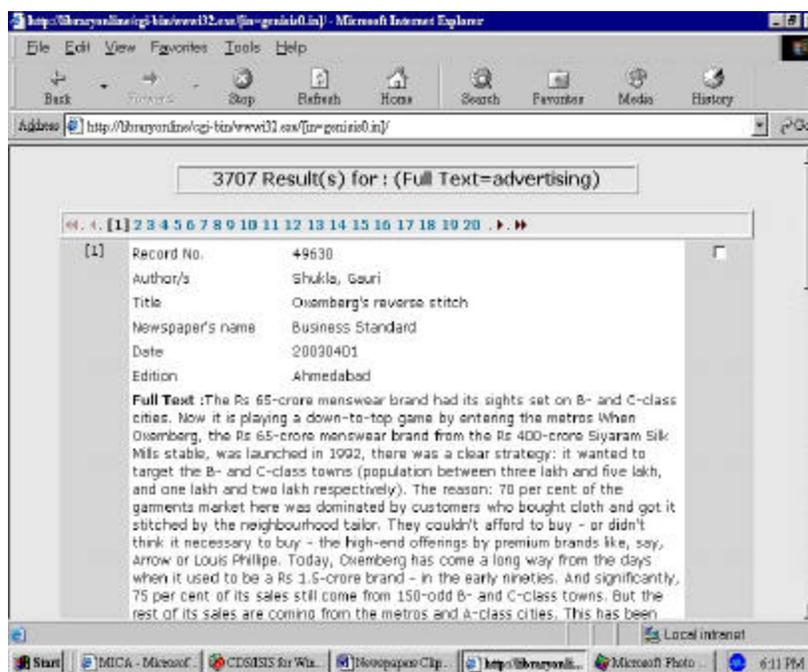


Fig 5. View of search result

Any field or access element can be made available for retrieval by selecting tag (field identifier) and techniques (indexing technique) option given in the FST. Techniques means how the access elements can be retrieved by accepting code numbers.

E.g. 0-by line, 1-by subfield of line, 2-by <term>, 3-by term, 4-by word, 5-by (prefix)subf/line, 6-by (prefix)<term>, 7-by (prefix)/term/, 8-by (prefix)word.

6. IMPLEMENTATION OF CDS/ISIS PACKAGE IN KEIC

6.1 CDS/ISIS and WINISIS

In 1985, UNESCO found a solution to the software problem by developing the Computerised Development System/Integrated Set of Information Systems (CDS/ISIS). Designed for small libraries and corporate libraries, as well as sections of

large academic libraries, the CDS/ISIS is non-numerical bibliographic information storage and retrieval software package. Most of the support for and the development of this package was provided by Del Bigio who adapted it from the mainframe version, developed in late 1960s.

In late 1985, version 1.0 was released. This version took into account the fact that many personal computers (PCs) had no hard disk. The data and programme were put on the same diskette so that they could run on such machines. The main focus of this version was the exchange of bibliographic data while its major shortcoming was that it could contain only 32000 records.

While version 2.0 was released in 1987, it was not widely circulated till march 1989. The next public release was version 2.3 which included improvements in the speed of the indexing and in the space used by the indexes. This was achieved in part by setting up two indexes, one for short and one for long terms. In May 1992, version 3.0 was released.

The main feature of this version is the support of local area network (LAN). Although it was possible to run a network by specifying certain parameters of with version of 2.3, there was a danger of file corruption if two users tried to update or make changes on a particular record at the same time. The addition of record locking and database looking feature has resolved this problem. The 3.07 version was released soon after, with some improvements.

The 1.0 version for Windows was released in 1998, it is designed to run in Windows environment.

6.2 GENISIS

GENISIS is an authoring software (for Win32) for visually producing web forms to query CDS/ISIS databases. There are two versions of the tool: GenisisWeb, for web publishing and GENISISCD for developing CD—ROM interfaces for CDS/ISIS database. GENISIS was originally developed by the former INISCUS Association (France) for UNESCO and it is now available free of

charge on the UNESCO FTP site. The software is written in Microsoft Visual Basic.

For more details see:

<http://www.unesco.org/isis/files/winisis/genisis/web/genisisweb.exe>

All the software designed by UNESCO are very cost effective. MICA, KEIC has used different versions of ISIS software in three phases.

7. HARDWARE AND SOFTWARE REQUIREMENTS

The minimum and recommended hardware and software requirements for running CDS/ISIS and GENISIS are following:

- CPU: 486 processor at 40 MHZ (Pentium at 100MHz or Higher recommended)
- RAM: 8Mb (16 Mb or more recommended)
- Floppy or CD-ROM unit - 1
- Hard Disk (with at least 4Mb free) - 1
- VGA 640*480 colour screen (super VGA 800*600 or higher recommended) - 1
- Printer - 1
- Scanner - 1
- Server (Unix, linux) - 1

Note: Although CDS/ISIS is a Windows 3.1 based program, it runs under Window95, 98, NT4, 2000 and XP without specific problems.

8. CONCLUSION

The theme of the article is low cost digitisation of newspapers clippings and used most of latest version of CDS/ISIS like DOS (ISIS), windows (Winisis) and web (Genisis). Though some drawbacks (not supported to image, table, charts, reading PDF, word file etc.) are found in this software, it is very cost effective. The KEIC has been using different functions of CDS/ISIS like inter-database link, image link, movie link, web link, full text of article, link of various files like word file, HTML file, PDF file, etc., in different modules and databases. The world is growing fast in all spheres. The library is not remained aloof from the changing environment. The revolution of information technology changed

the scenario of library services. Users can access this database from hostels, computer lab., class rooms, library through MICA Intranet and the faculty can access the same database from off-campus using login and password of their own.

REFERENCES

1. Buxton, Andrew & Hopkinson, Alan. The CDS/ISIS for windows handbook. UNESCO/CI, Paris, 2001.
2. Parekh, Harsha. Digitization: An overview of issues. *In CALIBER-2001: Creation and Management of Digital resources, edited by M.K.R. Naidu; Ashok Jambhekar; S.M. Salgar; & T.S. Kumbar.* Paper presented at the Eighth National Convention for Automation of Libraries in Education and Research Institutes, 15-16 March 2001, Pune. Information and Library Network Centre, (INFLIBNET), Ahmedabad, 2001. pp. 1-11.
3. Talagala, Deepali. Web interface for CDS/ISIS GENISIS web version 3.0.0 (Manual). Sri Lanka Library Association, Colombo. 2004.
4. Yagnik, Shailesh; Thomas, ABY K; & Patel, Niraj. Digitization of the MICA Library. *In CALIBER-2001: Creation and Management of Digital Resources, edited by M.K.R. Naidu; Ashok Jambhekar; S.M. Salgar; & T.S. Kumbar.* Paper presented at the Eighth National Convention for Automation of Libraries in Education and Research Institutes, 15-16 March 2001, Pune. Information and Library Network Centre, (INFLIBNET), Ahmedabad, 2001. pp. 91-99.

<p>Contributors: Ms Lavji N. Zala, Library Officer, Knowledge Exchange & Information Centre (KEIC), Mudra Institute of Communications, Ahmedabad (MICA), Shela Village, Ahmedabad - 380058, Gujarat.</p> <p>Sh. Niraj R. Patel, Dy. Librarian, Knowledge Exchange & Information Centre (KEIC), Mudra Institute of Communications, Ahmedabad (MICA), Shela Village, Ahmedabad - 380058, Gujarat.</p>
--