

The Multimedia Press System - A Database System for Archiving, Electronic Press-Clipping and the Analysis of Print-, Online-, Radio- and TV-Media

By Wolfgang Schöhl¹

The *Multimedia Press System* is an integrated tool for electronic archiving and evaluating articles and documents from all types of media. It may be used by the following departments of a company, a government agency or other organisations in their work with the media and publications:

1. press and/or public relations
2. information and/or documentation
3. internal relations
4. investor relations
5. marketing and advertising
6. economics.

It may also be used in

7. independent agencies for public relations and advertising and in
8. media-archives of publishing houses and broadcasting stations.

The *Multimedia Press System* consists of up to 20 internal databases combined to a relationally interlinked system for facilitating the work of these departments with documents collected from the media and their own publications. The system is also designed for downloading from major online sources - with an emphasis on news agencies, newspapers, magazines and newsletters, such as:

- professional databases such as those of the host DATASTAR
- current news sources offered by online services such as COMPUSERVE or AMERICA ONLINE
- news sources and company information from the Internet, especially the World Wide Web.

The system also has a module for converting print articles through OCR (articles not available in electronic form or available only with delay). With this module you can rapidly prepare a daily bulletin with press clippings from newspapers and simultaneously incorporate the articles into the *press clipping database* which is an integral part of the *media database*.

Overview of the System

The complete *Multimedia Press System* consists of these 20 databases designed for integrated work with the media:

1. database for downloading all kinds of published (external) articles from online-sources - mostly from national and international agency, newspaper and magazine databases;
2. database for press articles clipped internally from print sources (mostly regional newspapers and smaller magazines) which are scanned and converted into ASCII format by optical character recognition (OCR);
3. database for clippings from external press clipping services;
4. database for a company's own publications such as press releases and newsletter articles.
The databases 1. to 4. together form the *press database*. They may also contain scanned images, tables and photographs of articles.

Next comes a group of databases with multimedia features: the

5. database for downloading and archiving internet pages; these can be indexed and thus be found easily by a fulltext index;

¹ Dr. Wolfgang Schöhl is a professor for Information Science and Media at the Department of Information and Documentation of the Fachhochschule Darmstadt, D-64295 Darmstadt, Germany and head of the research agency *Forschungsgesellschaft f? Medien- und Wirtschaftsinformation*, D-64283 Darmstadt, Germany

6. database for transcripts of radio broadcasts;
 7. database for transcripts of TV broadcasts;
- in these three databases the reports can be viewed in their original format as HTML-pages or as audio and video files in addition to a text transcript which is searchable via a fulltext index. The databases No. 6 and 7. together form the *broadcast database*; all databases 1. through 7. form the *media database*. They can be searched as one single database, as all databases together or in any combination of two or more databases. This facilitates media observation and *media resonance analysis* in more complex forms.

The sub-databases within the media database all may have relational links to other databases of the *Multimedia Press System*. One link that is often used is the one to an integrated

8. database of persons and their home addresses, which in itself is relationally linked to a
 9. database of addresses of firms and organisations.
- Both databases allow the user to associate all media documents with persons involved. These may be experts or people interested in an important article inside and outside a company or government agency; or journalists who are either authors or experts for the topic of an article. These two databases may also be used independently of the media database.

The *Multimedia Press System* further contains a

10. database for storing and retrieving a department's own images. It is mostly used for
 - photographs not pertaining to published articles (which would be stored within the *media database*), especially those used for illustrating press releases and newsletter articles
 - graphic images used in the publishing process
 - charts and organization chart used in publications or speeches.

Similarly the system has a

11. database for a company's own videos and films. These may be both films for education and vocational training and videoclips for advertising on TV. These videos differ from broadcast TV program clips included in the media database in that they need a different kind of description - for example, they must contain data about the copyrights of producers.

Press and marketing departments often distribute or lend photographs or videos to journalists and advertising agencies respectively. The system therefore has a

12. database for circulation of photographs or videos linked to the image and video databases. It serves to keep track of photographs lent or distributed and to avoid double distribution.

For departments who have their own specialized library the following two databases may be added to the *Multimedia-Press-System*: a

13. database for specialized literature from journals/magazines and for books and a
14. database for administering borrowed journals and books.

The database for literature often contains only abstracts or descriptors instead of full text. It does not reflect what has been written *about* a company but rather professional literature that is used *in* a company.

Any department which uses some or all of the above mentioned databases will have a lot of practical experience in the use of databases. It is therefore easy to use the *Multimedia Press System* in a similar way for finding details in all office communication. It thus has a

15. database for office communication, i.e. for internal and external letters and memos sent and received.

The system also has a

16. database for protocols and notes of conversations and agreements; it makes it easy to find details on face-to-face and telephone conversations after some time. This database is interlinked with a

17. database for the administration and monitoring of target and fixed dates in a company or its departments.

For better planning of various activities the *Multimedia Press System* finally offers a
18. database for planning and administration of projects in PR and marketing.

For the specific use in press and public relations departments the system offers a
19. database for addresses of German journalists from an electronic handbook which is updated monthly.

The whole *Multimedia Press System* would be somewhat difficult to understand if there was no electronic help system to find answers to the more important questions of how to use it. The system therefore may be completed by
20. an electronic handbook which consists of a fulltext database explaining the system. The user thus can rapidly find the relevant chapters und subchapters dealing with any problem he may have.

There are three optional modules that can be added to the *Multimedia-Press-System*:

- Import programs for automatic downloading from less frequented online-databases;
- Import programs for automatically transferring press releases written in Winword into the press database;
- a module for Electronic Press Clipping (scanning and layouting of a press clipping service and OCR).

The system has been programmed on the basis of the German database management system *LARS for Windows*, one of the very few systems which have both relational and fulltext retrieval properties. It may be used on any stand-alone PC under Windows 3.x and in networks under Windows NT and Novell Netware.

Media Observation and Analysis

On the following pages some simple uses of the media database shall be explained. Suppose we want to do a search for articles from online databases about the genetically altered tomatoes which were heavily discussed in the US in the early to mid nineties. We would want to make a media analysis for a European food company to answer the following questions:¹

- how many articles were published on this topic in which years?
- what were the main pro and con arguments?
- which parties were for or against the introduction of the gene-tomato?

The relevant articles were first searched and downloaded from many German, English and US-Newspaper-databases, amongst others the databases in the pool PAPERS of the host DIALOG. They were then imported by automatic import programs into the media-database for detailed media analysis.

For finding relevant articles from various online hosts, the retrieval screen of the media database then may be used for searching in the following fields: find the words flavr and savr or gene (truncated) and tomato in the fulltext field and limit the search by type of document either to published articles (PA) or search in all types of documents of the media-database:

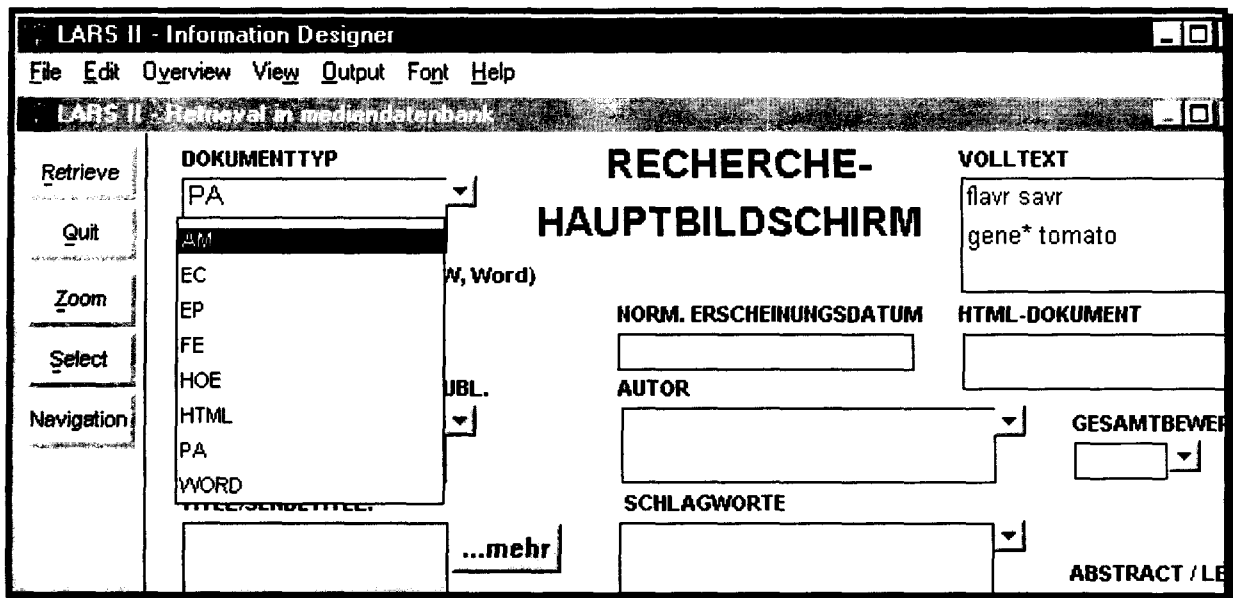


Figure 1: The retrieval by form-screen of the media-database

(The meaning of the German field names of the media-database is explained in English in the table in Annex 1.)

A subset of retrieved articles in German and English then appears on the general view of the database as follows:

LARS II - Query #2: 8 documents				
	ID	Titel	Dok	Normdatum
1	00039	Die*Gen-Tomate*kommt rot und teuer auf den amerik.anischen Markt.		12/07/1994
2	00041	In Amerika bald*Gen-Tomaten*zu kaufen.		
3	00042	Bald transgene*Tomaten*in Amerika?		
4	00057	USA: WE SAY TOMATO, THEY SAY*FLAVR SAVR.*		
5	00058	USA: FRANKENSTEIN'S TOMATO -*Calgene*Inc.	PA	07/11/1992
6	00059	UK: SCIENCE - FRIED GENE TOMATOES.		
7	00060	UK: GROWING ALARM AT GENE HARVEST - DR NIGEL POOLE - AGEN		
8	00061	USA: OUT OF AMERICA - AT LAST - A RIPE, GENETICALLY ENGINEERI		04/08/1993

Figure 2: A general view of titels of articles retrieved

Each article may then be briefly inspected in an introductory screen which shows the most important formal fields of the article.

LARS II - Modify Document mediendatenbank[58]

Pressedatenbank-Eingangsbildschirm zum Übersichtsbildschirm

DOKUMENTTYP	NORM. ERSCHEINUNGSDATUM	IMPORT. ERSCHEINUNGSDATUM	AUSGABE
PA	07/11/1992	921107.	Nov. 11, 1992
QUELLE/SENDUNG/EIGENPUBLIKATION	AUTOR	LÄNGE	
The-Daily-Telegraph, DT 7/11/	Edwar Miller	14,96	
TITEL/SENDETITEL			
USA: FRANKENSTEIN'S TOMATO -*Calgene*Inc. (DT).			...mehr
SCHLAGWORTE			
CALGENE-INC-USA, Horticulture, (I01002). UNITED-STATES, (USA	Industrial-gases-other-m CORPORATE, (CCAT);	SIGNATUR	
		BILD	<input checked="" type="checkbox"/> Ja <input type="checkbox"/> Nein
		SCAN-ARTIKEL	<input checked="" type="checkbox"/> Ja <input type="checkbox"/> Nein
WORD-DOK.			

Figure 3: The formal description of a relevant article

Each article can then be evaluated by the highlighted words and paragraphs under various aspects in the full text part of the documents.

LARS II - Modify Document mediendatenbank[58]

TITEL/SENDETITEL: *Copyright Forschungsges. für Medien- & Wirtschaftsinformation*

USA: FRANKENSTEIN'S TOMATO -*Calgene*Inc.

ABSTRACT/LEADPARAGRAPH:

VOLLTEXT:

In California, Charles Laurence tucks into a **genetically** engineered **tomato** and lives tell the tale. After a week or so, it was obvious that this was no ordinary **tomato**. It sat on shelf in my kitchen, round and ripe and red and juicy, and simply stayed that way. Its cousins came and went, from grocer to fridge to salad bowl, some going squishy in the steamy heat of an American summer before they could complete the cycle for which they were grown. But not CR3. Each morning it bore finger-prodding scrutiny as the coffee

RUBRIK: SEITE: 12 RESSORT: politics

Figure 4: The Fulltext- and Lead-Screen of the print-documents

The Multimedia Press System can also be used for detailed **Image and/or Content Analysis** by using a **database for evaluations-results** attached to all relevant articles. It contains the following fields:

		BEWERTUNG und KOMMENTARE							
Save	Quit	GEWICHTUNG der QUELLE:	<input type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	QUELLE:	The-Daily-Telegr
			A	B	C	D	E		
Select	Navigation	GEWICHTUNG des ARTIKELS/BEITRAGS:	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	DATUM:	921107.
			A	B	C	D	E		
		GESAMTBEWERTUNG PRODUKTBEWERTUNG PERSONENBEWERTUNG							
◀ ▶		BEWERTUNGEN:	<input type="text"/>	<input type="text"/>	<input type="text"/>				
Document	Mediendat	KOMMENTARE:	Copyright Forschungsg		1	en- & Wirtschaftsinformation Darmstadt			
#5 / 8	Field				2				
				3					
				4					
				5					
				KEINE					

Figure 5: The screen for the evaluation of media-documents

By means of the media database in combination with online databases one can thus make

- short media analyses
- comprehensive media analyses (image analyses)
- detailed content analyses.

They can refer to

- products and brand names
- leading personalities of companies or of governments or public life
- company names(one's own or competitors'), political parties, organisations.

While without databases these analyses take several weeks or even months, with databases they can be made in several days.

Annex 1: Fields in Screens *Pressdatenbank-Eingangsbildschirm (General Display-Screen)* and *Recherche-Hauptbildschirm (Main-Retrieval-Screen)*

German	English
Dokumenttyp	Type of Document = Sources of a document (PA=Press-article online, EP=own publication, EC=Own Press-Clipping, AM=Foreign Clipping HOE=Radio; FE=TV HTML=WWW WORD=Word with original format plus ASCII
Quelle/Sendung/Eigenpublikation	Source of Document: Name of newspaper/of broadcast/of own publication
Titel	Headline
Autor	name of author
Norm. Erscheinungsdatum	Standardized date of publication
Import. Erscheinungsdatum	Imported date of publication
Länge	length of article
Schlagworte	descriptors
Signatur	call number of print article/ of broadcast copy on audio/video cassette
Bild	picture available with article yes/no
Scan Artikel	scanned article (image of article) yes/no
Kontaktpersonen	Persons contacted on account of this document; transfer to Database of Persons/Adresses
WORD-DOK.	Word Document indexed and with link to original Word-Format
HTML-Dok. lokal	HTML document indexed and with link to original HTML-Format on local hard disk (Netscape-Browser opens)
HTML-Dok. lokal	HTML document indexed and with link to original HTML-Format on WWW (Netscape-Browser opens)
ID-Nr.	Identification Number
Dateiname	Filename (if document is a from word processing but not from winword)
Erfassungsdatum	Date of inclusion in media-database (or composition of a document)
Unternehmensbereich	department of firm for which a document is of concern or which has composed a document (in case of press releases)
Gesamtdokument	Total document (all fields in one); used for documents from WWW or Compuserve without or with varying field structure
Buttons leading to other display or retrieval screens;	these screens contain only fields typical to the specific type of document; fields for:
Text	Fulltext and/or abstract-documents
Bewertung	evaluation of article/report
Image/Bild	images as part of article
Hörfunk/Fernsehen	Radio/TV-reports
Eigenclipping	own clipping
Argusclipping	clipping by external clipping service
WWW-Dok.	WWW-documents (only titel and rest of document)
Formalangaben	headline plus fulltext plus contents of other (formal) fields in one field (e.g. author, URL, Date of issue)

ⁱ (this study was in fact made by the author for Nestlé Deutschland AG, a company that uses the *Multimedia Press System* jointly with all its affiliate companies). For further information about the issues raised in this article contact:

Prof. Dr. Wolfgang Sch?l, Forschungsgesellschaft f? Medien- und Wirtschaftsinformation (FMWI),
Wilhelminenstrasse 22, D-64283 Darmstadt, Germany

Phone: *49/6151/177889 Fax: *49/6151/177839 E-Mail:FMWI@compuserve.com