Working Experiences with a Cumulative Index on CAD: "CUMINCAD"

Bob Martens and Ziga Turk

To researchers in every discipline, Internet is quickly becoming the dominating environment to do literature studies. Commercial bibliographic databases tend to be too general, are not up-to-date and require special skills and effort to be searched. On the other hand researchers also publish on the Web and collaboratively that can create indexes of relevant publications. CUMINCAD is a bibliographic index that compiles papers related to computer aided architectural design. Implemented with a database, it allows searching and browsing in the ways usual on the Web. It provides a "historical evolution" to learn from previous efforts and draws attention to older original works that could have been ignored because they could not be found on the Web otherwise. The authors believe that CUMINCAD will help focus future CAAD research and improve the education. CUMINCAD work started in 1998 and is available at <u>www.fagg.uni-lj.si/cumincad/</u>. This paper focusses on the design and development of the database and presents some ideas concerning its advanced use in the analysis of research efforts.

Keywords: CAAD-related Publications, Web-based Bibliographic Database, Searchable Index, Retrospective CAAD Research

Introduction and related work

CUMINCAD - an acronym for "<u>Cum</u>ulative <u>Index</u> on <u>CAD</u>" - provides a valuable resource, allowing access, via a searchable index, to the conference papers for both eCAADe and other related organizations with special attention to publications on computer-aided design in architecture, implemented in a web-based environment. Exposure to the Web has changed the ways in which scientists are doing literature studies and retrieve copies of papers. In the past, several information services have been established to assist them, however, most of them are commercial, not available on Internet, or both (i.e. ICONDA, RWSB, URBADISC). Each uses a proprietary search technique, requires a special software and may only be available in the local library where one is supposed to walk to physically. They are not updated as promptly as new contents is being authored. Getting full texts is even more awkward - paper copies are ordered and may take weeks to arrive. The Web also changed the ways in which information is being published. Many scientists publish their papers or drafts on the Internet, sometimes even before a paper copy is available. But finding publications stored on remote servers at some faculty is guite difficult. The use of general Internet search engines such as Yahoo or AltaVista usually does not end up with useful results. One possibility are specialised search engines, based on proprietary harvesting [1] of a sub-sub set of the Internet (i.e. all departments of architecture in Europe), and data-mining the results set [2], but the relevance of such searches is limited as well, compared to the approach taken by the authors. The development of the Web service and the data entry started in 1998 [3]. Results from this first working phase are presented here and in part statistically evaluated. This sets the stage for further evaluations and a second working phase.

CUMINCAD Database

CUMINCAD is a rather informally structured bibliographic database. It contains bibliographic information including abstracts and e-coordinates of the author, for over 2000 papers related to CAAD that have been published since 1970 (three exceptions: 1954, 1961 and 1961!). Fig. 1 shows the diagram of the number of papers vs. publication year.

CUMINCAD is an Internet database. It is not only used, but also managed and maintained over the Internet. In the background is a database system called WODA [4]. WODA is a CGI application written in Perl. Its design goal was to create a smart and simple tool which would allow very rapid creation of small to medium size database applications, which could be used and managed using Web tools. WODA is tightly integrated with Web technology, supports multimedia contents (such as full text articles) file uploads, full text searches and includes software agent technology. WODA can talk to the user in English, German, French, Spanish and Slovene. It has been used locally for the Faculty information system, nationally to support the Building Center of Slovenia and in several international cooperation projects such as Esprit-SCENIC, Copernicus ATEM, EASY earthquake engineering slide database, CONNET databases etc.

WODA supports several search and browse approaches similar to common navigation techniques of most popular Web sites. Simple search is similar to AltaVista's simple syntax and the user simply types in a few words he/she is looking for. The composition of advanced searches is simplified by a user-friendly search form.

Number of Records

The first working stage focussed on the more or less complete entry of contributions from all existing Conference Proceedings, produced by the different worldwide acting CAAD-Associations. It is guite a challenge to figure out "what is already in". Although there is still a "missing stock" the overall situation is quite satisfying with regard to the degree of completeness. The 1999 proceedings and later ones will we be entered in the course of the corresponding year. CUMINCAD offers the feature "Browse contents" which allows the user to find the actual number of records which is related to the different associations resp. series (see fig. 1, which is self-explanatory). Most contributions have been created during the last two centuries. The significant growth after 1996 is a result of the CAADRIA-Conferences.

CUMINCAD in fact contains two databases, one of the publications in English and the others in Spanish - SiGraDi. The main reason to do this was the fact that nearly all contributions are in Spanish. However, in this "CUMINCADES"-database 148 records are available (see http://www.fagg.uni-Ij.si/cumincad/ cumincades.cgi)

The CADLINE-Database [5] developed under Yehuda Kalay's coordination has been imported into



Figure 1a (left). CUMINCAD papers by source.

Figure 1b (right). CUMINCAD papers by publication year.



Papers by publication year

CUMINCAD as well. The focus on the category "journal papers" will be a topic for the second working phase. Also considerations towards new categories such as "Doctoral Dissertations" or "Research Reports" have to be made. However, CUMINCAD will soon include thousands of records, too much to browse through. We would like to offer "guest-editors" a tool that would let them define annotated subsets of CUMINCAD - selections of best papers with their comments. This is a topic which has to be discussed intensively with the worldwide acting CAAD-Associations, canalyzing proposals from individuals. In this respect criteria will have to be developed ("guidelines"), such as the relationship to CAAD, publication medium, online-availability, etc.

Rating of Key- and Titlewords

First of all an overview on key- resp. titlewords was retrieved independently from each other by using the corresponding features. CUMINCAD produces a listing, which was transferred into an Excelspreadsheet and then "cleaned" manually. This cleaning work was done in a small part of the listing, focussing on topics which were mentioned more than 20 times, f.i. accumulating together "architecture" with "architectural" resp. CAD and CAAD or harmonizing singular resp. plural expressions. There were f.i. 867 different keywords with a total of 7.711 mentions resp. 3.248 titlewords with 13.718 mentions. Fig. 2a shows the seven most frequently mentioned terms. The result as such is not really surprising. It is in line with the aims of the associations. However, the number of mentions for "education" is lower than expected. Even "worse" seems to be the number of hits for "practice" resp. "research" (Fig. 2b).

Main authors in CUMINCAD

CUMINCAD generates a listing with no differentiation between names and surnames. A limit to the manual cleaning work was placed at the minimum of 10 mentions. According to fig. 3 "first prize winner" is John Gero. Examining this specific result (+series:"CADline" + Gero) shows that 74 hits are allocated to CADline. In the same way 31 hits from Yehuda Kalay are related to CADline.

It has to be stated, that this is just the actual account for the moment and that "full completeness" for the whole field of CAAD-related publications is not given.

Index	Number #K	Number #T
1 Design	402	753
2 Architecture/Architectural	297	270
3 CAD/CAAD	242	222
4 Systems	241	290
5 Computer/Computational	224	293
6 Algorithms	180	31
7 Model (ing)	168	234
	1754	2093

Conference Themes

Taking into account that two decades of conference activities have taken place, the question arises which were the corresponding themes in the course of the time.

Index	#K	#T	Index	#K	#T	IL TECEIVE	u a uispii	oportional a
AI	84		Knowledge (-Based)		119	followed b	v a disapp	ointment. an
Aided		53	Knowledge	165			, ,,	
Analysis	60	56	Language	60	36			Index
Application(s)	76	27	Layout	24				96 dero
Approach		81	Learning	27	41	6 420/	140	44 kolou
Artificial		23	Logic	25		6,43%	140	44 Kalay
Automation	33		Management	40	24			35 bentl
B-Rep	23		Media		32			32 oxma
Based		48	Methods	47	24			21 0000
Building	79	146	Multimedia		30			ST COVIR
CAM	21	223	Networks	27				26 rasdo
Case		33	Operations	28				26 shavi
Collaborative		25	Optimization	28				24 brow
Communication		20	Pearls		27			24 0101
Computer-Aided		57	Performance	39	22			22 eastr
Conceptual	20	20	Planning	32	33			22 mitch
Constraints	28	20	Plans	28	22	10 98%	239	21 radfo
Construction	116	28	Practice Process (Processing	110	53	10,0070	200	20 make
Data	110	60	Process/Processing	124	72			20 mane
Database	62	28	Prototypes	24	12			19 johns
Development	02	50	Reasoning	38	21			19 roser
Digital		42	Relational	21	2.			1.9 wood
Drafting	21		Representation	148	71			10 0000
Dynamic		23	Research	23				17 carra
Education	86	104	Shape	36	30			17 mave
Electronic		24	Simulation	45	39			17 cacar
Energy	22		Software	42	26			10
Engineering	116	48	Solid	48	28			16 mart
Environment		44	Space/Spatial	21	75			15 asand
Evaluation	69	43	Standards	23				15 flemr
Expert	109	79	Structure(s)	73	77			15 0000
Form		25	Studio		64			15 gross
Future		26	Support		32			15 schm
Generation		32	Surface(s)	82	45			13 balac
Geometry/Geometric	129	36	Synthesis	52				12 forme
Grammars	27		Teaching		63			13 ICINC
Graphics	148	52	Techniques	53	12121			13 nover
Graphs	23		Technology		33			12 brida
Hardware	23		100I(S)		92			11 chenr
Information	37	55	Urban	70	36			11
Integration	50	20	Virtual	73	62			11 galle
Intempence		20	Virtual	25	20	13,19%	287	11 kouta
Interface	79	27	Visualization	60	29	69.39%	1510	<10 Rest
internation international inte	19	61	Trandilizacion	1607	1547	1000	2170	

In the initial stage the conferences had more of a "workshop character", thus creating a place of encounter for teachers of CAD and to enhancing possibilities of exchange, paying less attention to the presentation form of the Conference Proceedings. Remarkably, is f.i. that explicit conference themes have not always been available. In the course of time a growing "professionalization" can be registrated ("comprehensive descriptions").

Brandon and Betts [6] claimed that research topics in construction information technology follow an evolution of five stages (Fig. 6). First a topic is recognized. Then it is enthusiastically accepted and it received a disproportional amount of attention, nd critique which, later

0	36				Index of words in field(s) au
4				96	ger o
7	41	6,43%	140	44	kalay
0	24			35	bentley
	32			32	oxman
7	24			21	coupe
-	30			20	coyne
R				20	rasdorf
в				26	shaviv
	27			24	brown
9	22			22	eastman
2	33			22	mitchell
1	33	10,98%	239	21	radford
9	54	0	- 10	20	maher
4	72			19	johnson
9 B	21			19	rosenman
1	2.			18	woodbury
в	71			17	carrara
3	20			17	mayor
5	30			17	maver
2	26			17	sasada
в	28			16	martens
1	75			15	asanowicz
3				15	flemming
5	64			15	gross
	32			15	schmitt
2	45			13	balachandran
Z				13	ferwee
2	63			12	novembri
5	33			10	heidere
	92			12	bridges
	36			11	cheng
3				11	galle
5	03	13,19%	287	11	koutamanis
0	26	69,39%	1510	<10	Rest

Figure 2a (far left top). Top seven of Keywords (#K) and Titlewords (#T).

Figure 2b (far left bottom). Main entries of Keywords (#K) and Titlewords (#T).

Figure 3 (left bottom). Table of Author Index

Figure 4 (right). Chronological Overview on Conferences resp. Conference Themes (Search for +title:proceedings).

Record key	Year	Association	Location	Theme
4b04	1983	eCAADe	Brussels	
c57c	1984	eCAADe	Helsinki	
e34b	1985	ACADIA	Tempe	
e115	1985	CAAD Futures	Delft	
	1986	ACADIA		Practice in the Next Decade
	1986	eCAADe	Rome	Teaching and Research Experience with CAAD
01c2	1987	ACADIA	Raleigh	Integrating Computers into the Architectural Curriculum
6abc	1987	eCAADe	Zurich	Architectural Education and the Information Explosion
651b	1987	CAAD Futures	Eindhoven	
c7f4	1988	ACADIA	Ann Ar bor	Computing in Design Education
	1989	ACADIA		Computing in Design - Directions for the 1990's
45e6	1989	eCAADe	Arhus	CAAD: Education - Research and Practice
8bf3	1989	CAAD Futures	Cambridge	The Electronic Design Studio:
				Architectural Knowledge and Media in the Computer Era
	1990	ACADIA		New Ideas and Practice
	1991	ACADIA		From Research to ?
61c0	1991	eCAADe	Munich	Experiences with CAAD in Education and Practice
c5a8	1991	CAAD Futures	Zurich	
6d34	1992	ACADIA	Los Angeles	Mission - Method - Madness
4857	1992	eCAADe	Barcelona	CAAD Instruction: The New Teaching of an Architect?
a01a	1993	ACADIA	Texas	Education and Practice: The Critical Interface
edcb	1993	eCAADe	Eindhoven	
a9bb	1993	CAAD Futures	Pittsburgh	
26b4	1994	ACADIA	Saint Louis	Reconnecting
807e	1994	eCAADe	Glasgow	The Virtual Studio
802c	1995	ACADIA	Seattle	Computing in Design - Enabling, Capturing and Sharing Ideas
44f6	1995	eCAADe	Paler mo	Multimedia and Architectural Disciplines
	1995	CAAD Fut ur es	Singapor e	The Global Design Studio
e678	1996	ACADIA	Tucson	Design Computation: Collabor ation, Reasoning, Pedagogy
4b55	1996	eCAADe	Lund	Education for Practice
2b9a	1996	CAADRIA	Hong Kong	The Introduction of Technology
aca6	1997	ACADIA	Cincinatti	Design and Representation
29cf	1997	eCAADe	Vienna	Challenges of the Future
e8aa	1997	CAADRIA	Taiwan	
8504	1997	CAAD Fut ur es	Munich	
634d	1998	ACADIA	Quèbec	Digital Design Studios: Do Computers Make a Difference?
4a59	1998	eCAADe	Paris	Computer ised Cr aftsmanship
a5cd	1998	CAADRIA	Osaka	

results in a better understanding of the problem and realistic value of the approach can be appraised. Is this true for CAAD research as well?

A comprehensive database, such as CUMINCAD can be used to quantify these claims. Fig. 7 shows the relative frequencies of occurrences of some popular topics in titles, abstracts and keyword sections of CUMINCAD data. The database was searched for all approximations of a listed keyword (e.g. integration, integrated ...) and the count of matching papers was compiled and normalized for the total number of papers published in that year. Because some conferences seemed to be happening on a two-year frequency, the curves were smoothed

	Buzzword (multiple)
8	Computer\$ / Computing
8	Design
8	Education/Teaching
6	Practice
5	Architectural
4	CAAD
3	Research
3	Studio
2	Experience(s)
2	Ideas

using an averaging function over three years.

This diagram shows e.g. that in the early years, there has been a lot of attention to the development of various algorithms, presumably related to rendering and visualization. After these were implemented in Mac, PC and workstation CAAD packages (mid 1980s), there has been a sharp decline of publications in this area. "Intelligence", followed by "expert" and "knowledge based systems" followed the explosion of artificial intelligence research in computer science (mid seventies), but it appears that by the mid 1980s researchers became much more cautions of using the word "intelligence" and moved on to "expert system". The hope in AI seems to have culminated in the mid 1980s but declined sharply in the early nineties. Since there are few commercial applications of AI the topic seems to be in the "negative reaction" state. Two important topics are on the rise - there is a solid upward trend in the concern for education since the 1980s when faculties of architecture started education in CAAD. The other rising topic is multimedia, gaining a lot of popularity in the mid 1990s. This discussion is a preliminary result of a more extensive study that will be carried out in the next year.

Conclusion and Outlook

It can be said that the current state of CUMINCAD presents a useful working tool for the CAAD community. By means of collecting facts towards aspects like the numbers of records, authors, keyand titlewords related to the year of publishing resp. conference themes a rough overview could be created in the framework of this contribution. This way of stocktaking is, however, necessary for the next upcoming working phase, which will focus on the broad entry of journal papers (including dissertations and research reports). The former CADline team was strongly supported by an international editorial board. A "revival" of such a board should be considered seriously. With regard to the fact this work was not continued after 1991, means that the nineties will have to be covered. On the other hand also the fifties and sixties should be examined (3 hits at the moment). The number of people involved in the second phase of CUMINCAD will be much larger, taking into account that a certain amount of "fuzzyness" is acceptable. This and further extensions have to be discussed with the CAAD-associations resp. individual users.

eCAADe's mission statement contains the idea to "make specific proposals towards the creation of closer relations between education, research and practice". The relation between education and research is probably closer than towards "practice". One of the reasons might be, that both education and practice are accommodated in an "academic" environment, producing output in a specific "format" such as conference papers. The "output" of CAADpractice is not necessarily made in the same "format".

Buzzword (ond	ce)	
Capturing	Era	Method
Challenges	Explosion	Mission
Collaboration	Future	Multimedia
Craftsmanship	Global	Pedagogy
Critical	Information	Reasoning
Curriculum	Instruction	Reconnecting
Decade	Integrating	Representation
Difference	Interface	Sharing
Digital	Introduction	Technology
Directions	Knowledge	Virtual
Disciplines	Madness	
Electronic	Media	

Figure 5a (top left). Multiple buzzwords in Conference Themes (2-8 times).

Figure 5b (bottom right). Buzzwords in Conference Themes (only one mention). Figure 6 (right). Interest as expressed by investment in new technology (Brandon and Betts, 1997).

Figure 7 (far right). Evolution of some topics in CAAD.



Stars of this scene like Greg Lynn, Kas Oosterhuis and Ben van Berkel - propagating "newest" tendencies - use different presentation fora. Something like a collection of Weblinks could therefore be useful. This could be regarded as a rough idea for a third working phase.

In order to make CUMINCAD available as much as possible - the portion to a CD-ROM in the framework of the eCAADe '99 Proceedings has good chances to be realized and may come in handy regarding off-line utilization. To keep its use as simple as possible and make sure that the CD can be ported to Macs, PCs and UNIX systems the CD will include standard Web data only and anyone with a CD drive (of course) and a Web browser will be able to use it. The search will be implemented in the JavaScript language. The straightforward layout, in which WODA keeps its files, simplifies the creation of mirror sites world-wide. At the moment there are running negotiations to have a mirror site in the US because the Atlantic is a well-known bottleneck of the Internet.





References

 C. Mic. Bowman et. al., Harvest: A Scalable, Customizable Discovery and Access System. Technical Report CU-CS-732-94, Department of Computer Science, University of Colorado, Boulder, August 26th 1994.



- S.R. Lockley and R. Amor, The Construction Information Gateway, in: The second European Conference on Product and Process Modelling. London, 19-21 October 1998, pp. 337-348.
- B. Martens and Z. Turk, The Creation of a Cumulative Index on CAD: "CUMINCAD", presented at eCAADe '98, Paris [publication forthcoming].
- 4. Z. Turk,WODA A Web Oriented Database, <u>http://www.fagg.uni-Ij.si/woda/</u>
- Y.E. Kalay, CADLINE: A Bibliographic Database on CAAD-Related Publications. Buffalo, State University of New York, 1991.
- P. Brandon and M. Betts, Veni Vidi Vici, in: Pre-Proceedings of the International Meeting "Global Construction IT Futures" 1997. University of Salford, Construct IT, pp. 86-89.

Bob Martens and Ziga Turk Vienna University of Technology and University of Ljubljana <u>bmartens@email.tuwien.ac.at</u> <u>ziga.turk@ikpir.fagg.uni-lj.si</u>