

In Quest of Sustainable Information Databases, Database Analysis, and Database Research - Exploration of a Research Field

Helmut M. Artus

GESIS-IZ, Bonn, Germany

Prefatory remark

To make things clear at the very beginning: My paper is a plea *for* database research, *not against* it, although it sometimes may seem so. But supporting something cannot and must not mean: resigning to any criticism or to sceptical questions. So my criticism is supposed to be a positive and constructive one.

Introduction

Whether or not we believe in the *obsolescence* of scientific literature, it is a fact that the actual *usage* of literature decreases in the course of time (except so-called *citation classics*). As a logical consequence, we must assume a *correlative 'obsolescence' of information contained in databases*.

Databases, however, are much more than just registers for bibliographic information, but are also systematic and well structured storehouses of information on science history: on innovations and changes, on developments and discourses as well as on social structures, on socio-cognitive networks of scientists and researchers and so on. So using such information for purpose of analysis may contribute to the *sustainability* of databases.

In order to get valid data, however, simple statistical processing is not sufficient by far. Database research, as I see it, is a rather complex and above all an interdisciplinary enterprise. As far as I see, there are five different research areas which I summarize under the notion of 'database research'. The requirements and opportunities are described in this paper.

1 What is Database Research? A short overview

As mentioned before, database research consists of five different parts or research areas (see figure 1):

- The **first** part refers to the *methodology for database construction and management*. This, I think, is the most trivial part of all in so far as it concerns the every-day business of each person who and each institute that has to do with databases. So it does not require further explication.
- The **second** part is a traditional field connected with public databases, and although it is generally called user research it also comprises the research of database *usability* and *us-*

age and sometimes even that of their *usefulness*. Again, this is rather trivial. It should be mentioned, however, that the quality of such should be somewhat higher than it uses to be. A closer co-operation with empirical Sociology could be useful.

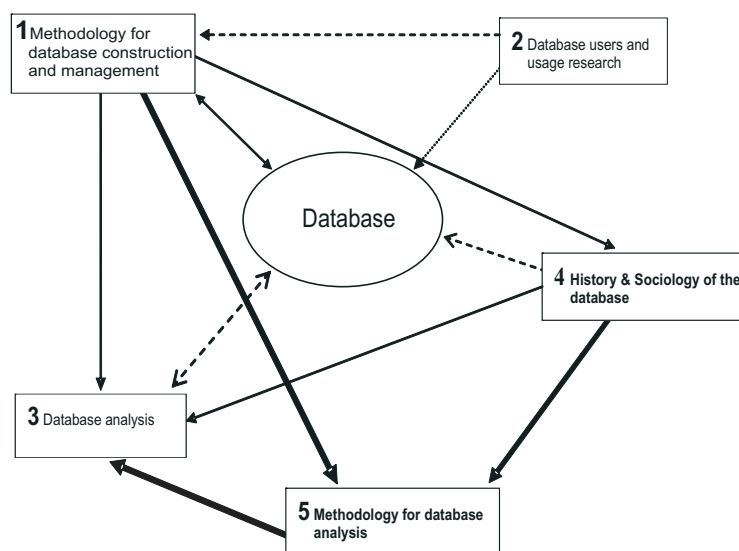


Figure 1: The elements of database research

The **third** and perhaps most familiar area is so-called *database analysis*, which means that bibliographical¹ data are taken as data base for quantitative analysis of a scientific field in order to find out its cognitive and social structure, be it at a time given, be it in the course of time. A few examples may show the range of questions that can be examined by database analysis:

- cognitive structure of sub-disciplines;²
- research networks³;
- temporal development of scientific concepts⁴;

1 In this paper, I only refer to bibliographic databases. Most issues and arguments, however, can be applied to research project databases as well; to a smaller extent there may also be an applicability to certain other types of databases, e. g. concerning scientific and research institutes. – For certain reasons, my paper is completely focussed on the GESIS databases SOLIS, SOFIS and ZEITSCHRIFTEN and of cases of using them as bases for quantitative analyses. That may explain that most examples are rather old (but by no means obsolete), since database research in the GESIS-IZ has been suspended.

SOLIS = Social Sciences literature from German language countries; SOFIS (formerly FORIS) = Social Sciences research projects from German language countries; ZEITSCHRIFTEN = German Social Sciences journals.

2 See e. g. Binder/Herfurth and Binder/Stahl (1993) and Best/R. Only in *References 1* as well as Sahner in *References 2* at the end of this paper.

3 See e. g. Güdler; Mutschke, *ibid.*

4 See e. g. Renner, *ibid.*

- changes in scientific publication behaviour as shown by a comparative analysis between printed and online journals⁵;
 - combination of data sets from an empirical survey by questionnaire with data sets extracted from databases⁶;
 - thematic structure of university teaching⁷;
 - knowledge transfer from research to university teaching by matching different databases⁸
- etc.

The **fourth** part which I put into the center of my paper consists of two areas: the *history of databases* and all their indexing instruments and rules, external influences and the like, and the *sociology of databases* dealing with rules and the way they are carried out, it includes social structures and social action, social change and social conflicts, social conformism and deviance and much more.

And the **fifth** and last part is dedicated to the *methodology of database analysis* which – to a certain degree – has to draw the methodological consequences from the results of the history and sociology of the database and combine them with the knowledge of the *methodology for database construction and management*.

In this paper, I have to concentrate on parts 4 and 5, i. e. History and Sociology as well as Methodology.

2 What is a database? Two different perspectives

If we ask for a definition of the concept of ‘database’, we are usually given an explication like this: “A computer database is a structured collection of records or data that is stored in a computer system.”⁹

This is undoubtedly true, but undoubtedly it is only one side of the coin. We could give a completely different definition which approximately reads like this:

“A database is a structured collection of data usually produced by a lot of more or less normal human beings under different conditions of work which in many cases are far from being ideal, adequate or only normal ...” etc.

Actually, no one - except me – gives such a definition. But whenever you talk with someone who is involved in database construction, you will hear lots and lots of stories about this human respectively social side of the database, lots and lots of complaints, lots and lots of critical remarks.

So there is a deep gap between the view of theorists and practitioners of the database. In the following table, I try to give a rough and surely not exhausting overview over these two diverging perspectives on the database.

5 See e. g. Artus (2005), *ibid.*

6 See e. g. Artus (1988; 1992); *ibid.*

7 See e. g. Artus (1996), *ibid.*

8 See e. g. Binder/Stahl (1994; 1996), *ibid.*

9 <http://en.wikipedia.org/wiki/Database>, Article “Database” (Wikipedia)

There are two completely different ways to look at databases:

- The first one (which I call the structuralist view) treats databases as formal structures filled with a content created according to strict rules. In this view, it seems to be self-evident, that the data are a perfect and undoubtable basis for quantitative statistical analyses.
- The second one (which I call the sociological view) has an eye on the people carrying out these rules and actually creating the database as well as on the social processes of this database creation as well as on the conditions on which all this happens and these people have to work. In this view, a database is a work of man and as such imperfect, so that it is not a self-evident basis for a quantitative analysis.

The double nature of databases

Structuralist view	Sociological view
target (i. e.: what should be)	performance (i. e.: what really exists)
ideal situation of construction of indexing rules and instruments	real (imperfect) situation; maybe lack of competence, at least in certain disciplines, sub-disciplines or topical areas
unhistorical, beyond change or development	historical ('philosophies' or 'ideologies' of the database, political influences, development and changes of scope, indexing rules and instruments such as manuals, thesauri, classifications and other descriptors as well as of the rules of their application and application control; manpower and fluctuation etc.)
deterministic rules	application of rules as a highly social, psychic and cognitive process of understanding (rules, but also the document in question), interpreting (rules and document), selecting and finding (descriptors of any kind) and decision (Is my description pertinent? And is it sufficient? Maybe too circumstantial?), that means: indexing as a highly complex, accident sensitive and to a certain degree even hazardous process
unambiguousness by formal structures	real ambiguity despite any formalisation
social norms of action	real action, including deviation, conflicts, rivalry etc.
ideal work situation	real work situation with time pressure, competing priorities etc.
ideal agent presupposed	a) real (i. e. imperfect) agent b) not one single, but lots of agents with different minds and qualifications, different interests and motivations, different time budgets, different loyalties (concerning to their different institutes), c) not only lots of agents with all their differences, but many of them not only working at the same time but interacting (and interfering) with each other: competition and rivalry, conflicts of any kind and so on d) loyalties against a lot of different institutions with special interests, policies, clientele, priorities and the like (in the case of SOLIS and SOFIS there have been about twenty different cooperating institutes in the course of time, usually about ten to twelve at the same time) etc.

As I have mentioned before, this list is not exhaustive by far.

This simple juxtaposition shows clearly, that we have to draw consequences – at least three of them. That is why before using databases as the basis for quantitative analyses we

should make them objects of research themselves unless we would risk severe biases and mistakes.

- **Consequence number one** is that we need a ‘*History of the database*’ (which has always to be read as ‘the history of the special database in question’). For each database that we use as the basis of a quantitative or statistical analysis, we have to write a rather detailed history of its development, of all considerable changes, be it in scope, be it in the classification codes or methodological descriptors, be it the history of the thesaurus, be it the rules of the manual.
- **Consequence number two** is that we need a ‘*Sociology of the database*’ in order to understand the influences which at the same time form and deform the social system that is the basis of database-construction and the social action of all people who have to do with the creation of the database and the rules according to them it is built.
This ‘Sociology of the database’ should be a general theory as well its application to a special database.
- But there is still a **third consequence** concerning *methodology* or *epistemology*, notably the so-called *problem of the empirical basis* (“Basisproblem”). The question is whether data that have been constructed and edited for a completely different purpose can without further justification be used as the basis for an allegedly empirical analysis.

3 The History of the database

As I have said above, we need a rather detailed history of the development and of all considerable changes of the database, be it in scope, rules of indexing or descriptors of any kind.

Let us take the disciplinary and bibliographical scope of the database SOLIS as an example.

SOLIS started in early 1980 with the official and political¹⁰ appointment to cover all sociological literature in the German speaking countries, i. e. the two Germanies which still existed at that time, Austria, and Switzerland. About two years later, a retrospective documentation took place going back as far as 1972 – but it did not include books and grey literature which SOLIS had been documenting from its very beginning, but only journal articles – and not from all journals but only from the most important ones. Nearly at the same time, further bibliographic data going back as far as 1945 were added – but without abstracts and only with a very rudimentary indexing which is a difference that spoils any quantitative analysis.¹¹

So within three years, there were not less than three important changes of the scope in bibliographic respect:

- All sociological literature – i. e. journal articles, articles in anthologies, books and grey literature (i. e. unpublished research reports) – with abstracts and detailed description starting with the year 1980;

¹⁰ In order to understand this chapter, we must keep in mind that all this happened under the terms of public financing. All institutes in question were subject to the same conditions – except our partners in Austria and Switzerland.

¹¹ The latter bibliographic data came from a printed bibliography that was based on a survey (!): All German sociologists had been asked to report their publications.

- only articles of the sociological core journals with abstracts and detailed description for the period from 1972 till 1979;
- and books, journal articles and articles in anthologies without abstracts and with a poor indexing covering the years from 1945 to 1971.

It is self-evident, that such changes would have an enormous influence on any quantitative analysis of the database. But another important change was still to come:

- In 1983, the scope was dramatically extended. Instead of one single discipline – *Sociology* – SOLIS was now obliged to cover the literature of all *Social Sciences* in the German speaking countries.

So, this was the fourth important change of the scope within only three years. But this is only a small part of the ‘scope history’.

When the IZ took charge to cover the Social Sciences as a whole, one of the political conditions was that the IZ should not document all disciplines by itself but should close cooperation contracts with specialized institutes. That is why SOLIS at certain times was a cooperative product of not less than twelve institutes at a time.

But actually there should have been *thirteen* instead. The last one, however, preferred its independence and refused cooperation. Nevertheless, the IZ was strictly forbidden to document the concerning discipline – which was *Political Science* - itself. For about nine years there was a state of war which ended up in a series of lawsuits and reached its climax with the menace of a complaint of unconstitutionality which fortunately was not into practice. At a certain point of this process, the IZ took over about 3,600 documents - which some months later (in May 1992) had to be removed from the database by judicial order. One lawsuit and about three years later, they could be incorporated into SOLIS again. For the intervening period, the IZ had been documenting only a small amount of Political Science literature published in general social scientific journals but not in core journals.

So at least for a period of twelve or more years, the number and share of Political Science publications in SOLIS was by no means an indicator for the *scientific output* of this discipline (as would have been assumed by database analysis) but actually an indicator for the *network of institutions* cooperating – or not cooperating - in the project ‘database SOLIS’. Years later, the IZ closed another contract of cooperation with an institute that is very reliable but has very strong own interests which sometimes rub against the documentalist ‘philosophy’ of SOLIS.

The case of Political Science was the most striking conflict but other systematical quantitative biases and distortions occurred within the system of cooperation which has a non-hierarchical structure. So the IZ was in a jam: It is neither allowed to document scientific discipline or their parts that have already been documented by other institutes, nor has it the right to enforce the general rules holding for the commonly built database.

As a necessary consequence of the development and the changes of scope, a good deal of the instruments of indexing had to be completed or to be differentiated and specified, e. g. the thesaurus with all its controlled terms, the comprehensive list of classification codes and the methodological descriptors.

This was only one segment of the history of our database. As the examples have shown, the written history of the database must as well be a history of political influence, of institutional cooperations and conflicts and of the institute which is responsible for the database as

a whole and for the coordination of diverging institutes and interests. Among others, this is a history of institutional and database ‘philosophy’, of social, political and financial influences, of number, qualification and specialisation of database personnel as well as of its fluctuation, of technical equipment etc. Last but not least, it is as well a history of *technology*.

A quantitative database without all such ‘historical’ knowledge could not estimate the risks, the possible mistakes or even the degree of deformations and could easily lead to research artefacts.¹²

4 The Sociology of the database

Such database *histories* are necessary but not sufficient. They have to be completed by a *sociology* of the database.¹³

A completely unexploited research field is the analysis of databases not only as formal structures, created according to strict manuals, thesauri, lists of classification, of methods etc. but as the results of the action and interaction of a multitude of people counting by hundreds and sometimes the co-operation of dozens of institutes with different interests or ‘philosophies’ which sometimes leads to more or less serious conflicts, as we have seen in the preceding chapter.

The reconstruction and empirical analysis of such complex socio-cognitive systems should be the object of a *sociology of the database*, which – in reverse – would unveil the *social conditionality of data* derived from databases. We should keep in mind that this is the crucial point of any database analysis: the social conditionality of the data to be analysed.¹⁴

Influences responsible for the formation and deformation of the social system are, e. g., co-operations with other institutes contributing special parts covering scientific disciplines or sub-disciplines to the database and in reverse make claims which need not be compatible with the general ‘philosophy’ and orientation of the database project, as we have seen above.

The perhaps most interesting and important topic of such a Sociology of the database, however, is the systematic empirical analysis and theoretical explication of social action within the institutes creating a certain database as well as between them.

Though as an empirical science, Sociology has to do with the rules of action, its essential topic is action itself – notably in comparison with and contrast to the rules and norms this action should be obliged to. That is why Sociology asks for the normative conformity or deviance of action as well as for the social conditions of deviance or even anomia or normlessness.

This question for the relationship between norms and rules on one side and their real application on the other is not restricted to the area of ‘normal’ social action but holds for the *documentalist* action as well. And the same holds for the conditions, determinants, and pos-

12 See also Artus (1992) in *References 2* at the end of this paper.

13 Just to mention it, the History of the database should refer to both, the individual database in question as well as the database as such and in general, whereas the Sociology of the database could be confined to the general aspect. Both, however, must be strictly empirical enterprises – or they are useless.

14 For an analysis of the process of repression of this social conditionality see Artus (1996) in *References 2* at the end of this paper.

sibilities of social and documentalist action as well as for the mechanisms of control and/including sanctioning and the manner and degree of their institutionalisation.

This is the place for topics like the following ones:

- Social structure of departments responsible for database construction including creation, development, control and updating of manuals as well as indexing rules and instruments.
- Quality control, feedback, communication and interaction; degree, mode and quality of cooperation between colleagues.
- Social control, sanctions, conflicts and conflict management – be it social conflicts, be it (cognitive) conflicts concerning the interpretation and application of rules, norms and instruments. In case that there are external colleagues, the questions of control, criticism and on-the-job training are even more difficult than in the case of inhouse control.
- Work situation including the question whether there are enough colleagues who are not only qualified in a general way but also have the necessary specialisation for exactly the subject areas they are responsible for. This is also the place for the questions for recruitment and fluctuation of such specialists.

But even if an institute has accomplished a high degree of qualification and specialisation, of control and cognitive consonance and has managed to institutionalise all this, one important question remains: In how far can action – social action as well as documentalist action – be *determined* by norms and rules? In most cases, we should be happy if the real action is *compatible* with the rules.

To explain this, let us take a quite simple example:

- “*In case the red light flashes up, press button A!*” This is a determinist rule combining an unambiguous event (“red light on”) in an unambiguous way with a required and likewise unambiguous reaction “push button A!”.
- “*Describe the most important topics of any publication!*”, however, is a non-determinist statement. It is rather a formulation of the guiding idea of the indexing process than an advice for concrete action. One can understand the meaning of this principle, but it is impossible to deduce concrete or even unambiguous consequences for action resp. indexing.

If, however, determinist action is just an illusion, we cannot expect a unity of indexing but should be satisfied if the predominant part of all cases is at least within the ‘horizon of consent’ of the indexing team resp. institute. That means that database analysts as well cannot expect data generated to clear determinist rules but can only expect ‘compatible’ data at best. So the inescapable question is that for the variance and spreading of such data.

Eventually, there is a problem which could be called *relation of uncertainty*. Science aims at precision, but precision is hard to get in database analysis. The more precise and differentiated indexing rules and instruments are, the more difficult and complicated is the application, the more qualification and specialisation is required, and at last the more time for documentation is needed. In other words – to be honest (or realistic, which is the same) -, the higher the precision, the higher the probability of mistakes and in the end: of *research artefacts*.

All these facts & factors that have to do with the history and sociology of the database, its indexing rules and instruments and their application have consequences for any database

analysis. Some of these consequences appear immediately, some of them show up only in the context of special analyses. Nevertheless, we have to take account of them all. Wherever people collaborate or institutes cooperate, such phenomena and problems occur necessarily. So their consideration in database analysis and in the *methodology* of database analysis is inevitable.

5 The Epistemological Question

This desirable and necessary research field deserves much more attention and interest than it has been receiving as yet. If at all, researchers completely concentrate on methods and methodology of data analysis, i. e.: merely on an application of the tools of statistics.

The really interesting (and important and complicated) questions, however, are those related to the generation of such data, to their validity, to the methodological prerequisites of validation, and their comparison with social reality.

As I have said above, *the question is whether data that have been constructed and edited for a completely different purpose can without further justification be used as the basis for an allegedly empirical analysis.*

The problem is that database analysis only pretends (or *believes*) to analyse science but actually refers to a database which as such is an *artefact* representing science under specific aspects and for specific purposes. The data contained in the database have undergone processes of selection and transformation according to the database manual and the indexing rules. Thus the database is not a true image of its object – in this case of science – but an artefact. Or more precisely: a *first degree artefact*.

In the course of database analysis, this first degree artefact is subject to further processes of selection, reduction and transformation. No matter whether the database data are qualitative or quantitative – they are transformed into a set of structured quantitative or at least computable data, thus fabricating a new, a *second degree artefact*. (See figure 2!)

It is a moot point to argue whether this transformation is a *methodological or an epistemological problem* – if only the consequences are borne. The problem is not that we have to do with an artefact but that at the same time it is a *research artefact*. The transformation of database data to data for database analysis is by no means just a formal transformation of data without changing or anyhow affecting their content and meaning. Actually, by this transformation ‘weak’ data become hard facts – which unfortunately turn out to be artefacts.

To understand this we should have a look at the process of database searches which is the usage databases are made for. After having entered descriptors and an appropriate search logic we receive a number of documents fulfilling our search criteria. But this is by no means already the end of the search. What follows is the intellectual browsing and selection of the really pertinent hits which means that only in these cases the descriptors we have used really describe the central topic of a publication.

When executing a database search, we are still aware that there is (or at least can be) an important difference of relevance between the descriptors used in a document. If, however, the same descriptors are used in database analysis, suddenly all of them have/receive the same importance, which - at least in relation to a single document - causes a total bias. A perfunctory topic has the same importance as the topical core.

This is not the only source of such biases. Similar effects show up with classification and methodological descriptors, but also with descriptors which do not refer to the contents but to formal categories like sources (e. g. journals) or document types (e. g. article, book, research report etc.).

So the problem is not *that* data from databases are taken as the basis of a quantitative analysis but *how* this is done.¹⁵ A database search without intellectual reoperation (control, selection) is very problematic and does not meet the *state of the art*. A database analysis renouncing to intellectual reoperation runs the risk of failing scientific reality by far and of producing an artefact.

Or in other words: this is the short way *from weak facts to hard artefacts*.

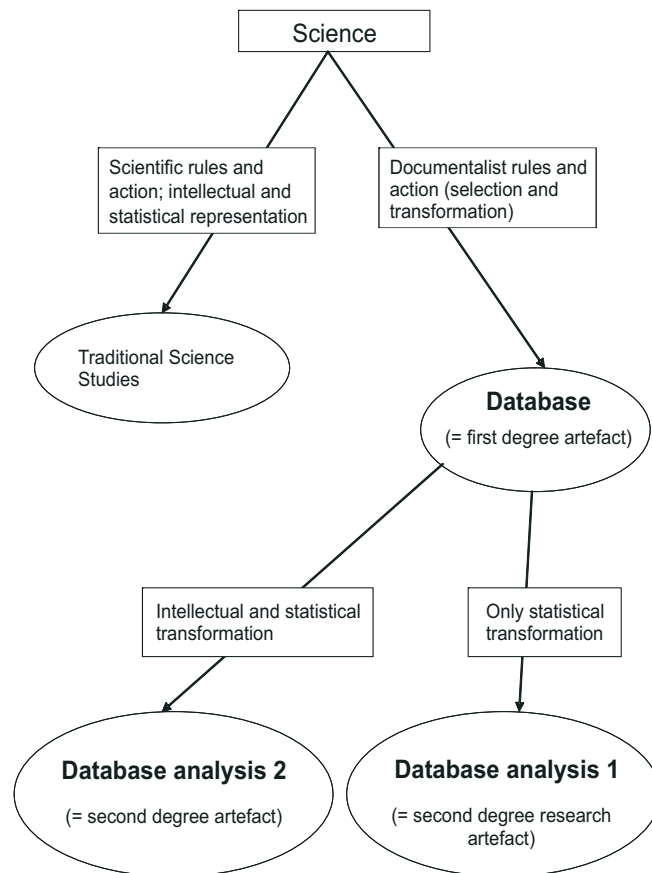


Figure 2: Facts, Artefacts and Research Artefacts

15 The discussion about so-called process produced data in the 1970s and 1980s – usually concerning data from public administration or from historical sources – is on one side a proof for the usability of certain data for certain purposes under certain conditions, but it was on the other much too shallow for an application to the problems tackled in this paper.

6 The first ‘Grand Project’ of Database Research

My argumentation is based on a lot of scientific as well as documentalist experience, on theoretical (even meta-theoretical) as well as on empirical as well as on practical experience. So I am convinced that it is logically untouchable.

Whether, however, the consequences are likewise untouchable and undoubtable, is another question, i. e. a question that cannot be decided theoretically or by logical argumentation but only *empirically*. Maybe all the effects that I have mentioned cancel each other out or are so marginal that any discussion were much ado about nothing. The most probable result, however, could be that in parts the quality of formal transformation is acceptable whereas in other parts the data need more or less intellectual rework[ing]. But as I said: all this is a question that can only be decided empirically.

So the first grand project - or work order – for database research is clearing this question – and I mean *properly* clearing it.

- That means - *firstly* – that we need a separate, independent and valid empirical survey covering exactly the same segment of science covered by database analyses; and this survey must be carried out according to the theoretical as well as methodological state of the art of the empirical Social Sciences; and the same holds for the subsequent comparison between empirical and database analytical data. Above all, any circularity must be strictly avoided.
- And it means *secondly* that a single survey and comparison is not enough. As we know from documentalist database searches, the pertinence of descriptors is highly correlated with special subject matters and sub-disciplines. The empirical surveys must take account of these differences, if their results should be valid, i. e. they should be the justification of any further procedure of database research.

Clearing this question will surely be hard work, but it is the only alternative to the procedure I have called “database analysis 2” (in figure 2).

7 Qualification, scientific interaction & transdisciplinary cooperation

The explications given above should have shown that database research can only be successful if conceptualized as an interdisciplinary enterprise. As shown above (figure 1), the sociology of the database has strong reactions on database research, on methods and methodology of database construction as well as of database analysis. The close relationships between a sociology of databases and the research in users, usage, usability and usefulness are obvious. This holds likewise for questions, methods and theoretical approaches.

Researchers working in the area of database research, I conclude, should have a double qualification in their ‘home discipline’ (for instance sociology, psychology, statistics, informatics etc.) and in database construction. The latter qualification does at least require several years of professional involvement in the field of database construction and management. Only a close relationship between the different sectors and a profound understanding of each others perspective can guarantee a sound and valid research based on databases and thus contribute to the sustainability of information contained in databases,

References 1

Selected publications referring to one or more GESIS-IZ databases and using their data for quantitative analyses

The databases used are:

<i>SOLIS</i>	Social Science literature from the German speaking countries
<i>SOFIS (formerly FORIS)</i>	Social Science research projects from the German speaking countries
<i>LEHRE</i>	University teaching in German Social Sciences (database abandoned meanwhile)
<i>ZEITSCHRIFTEN</i>	Social Science journals from the German speaking countries

Artus, Helmut M.: Graue Literatur: Zum informellen Kommunikationssystem der Sozialwissenschaften: Abschlußbericht. Bonn 1992, 342 pp, 101 tables

Artus, Helmut M.: The transition from 'grey' to 'white' literature: a study in the communication and publication behaviour of social scientists. pp. 126-144. In: Kühnhold, W. W.; W. P. Kirchner (eds.): Publications as an integral part of scientific research. Proceedings of the 5th International Conference of Scientific Editors, Hamburg, 1987 June 14 -19, Hamburg 1988 (Microfiche edition)

Examples for the combination of survey data (concerning communication and publication processes) in combination with database data concerning the same research project under formal aspects (number of researchers, duration of project, commissioned research; number of publications etc.)

Artus, Helmut M.: Soziologielehre: eine räumliche und thematische Analyse. pp. 59-86. In: Artus, Helmut M.; Matthias Herfurth (eds.): Soziologielehre in Deutschland: Lehre, Studium, beruflicher Verbleib, Lehrangebot, Studien- und Prüfungsordnungen. Opladen: Leske + Budrich 1996

Example for a traditional quantitative analysis of the database LEHRE (= University teaching; database abandoned)

Artus, Helmut M.: Old WWWine in New Bottles? Developments in electronic information and communication: structural change and functional inertia. pp. 9-16. In: The Grey Journal: an International Journal on Grey Literature 1, No. 1, 2005

Example for a traditional quantitative analysis of the database ZEITSCHRIFTEN (= scientific journals)

Best, Heinrich; Renate Ohly: Entwicklungstendenzen der deutschsprachigen Soziologie im Spiegel ihrer führenden Fachzeitschriften - Ergebnisse einer Korrespondenzanalyse. pp. 575-592. In: Best, Heinrich; Brigitte Endres-Niggemeyer; Matthias Herfurth; H. Peter Ohly (eds.): Informations- und Wissensverarbeitung in den Sozialwissenschaften: Beiträge zur Umsetzung neuer Informationstechnologien. Opladen: Westdeutscher Verlag 1994, 623 pp.

Best, Heinrich; Renate Ohly: From paradigms to eclecticism: thematic profiles of German language core sociology journals 1984 -1991. pp. 95-113. In: Bulletin de Méthodologie Sociologique (BMS) März (1994) 42

Examples for traditional quantitative analyses of databases (SOLIS) referring to/concerning thematic profiles and developments in Sociology

Binder, Gisbert; Matthias Herfurth: Quantitative analysis of bibliographic databases. pp. 33-43. In: Saelen, Kirsti T.; Arnaud Marks, György Rozsa; Karl A. Stroetmann; Vladimir Vinogradov; Jiri Zahradil; Manfred Krause; Harald Koch; Liparit Kiuzadjan (eds.): The role of social science infor-

mation in knowledge creation: integrative aspects in information, communication and knowledge; proceedings of the Vth ECSSID General Conference, Berlin, GDR, January 22-24 1989; Vol. 2. Wien 1990, 206 Seiten

Binder, Gisbert; Matthias Stahl: Sozialforschung 1990 in den fünf neuen Bundesländern: Ergebnisse einer scientometrischen Analyse der Projektdatenbank FORIS. pp. 561-570. In: Neubauer, Wolfgang; Karl-Heinz Meier (eds.): Deutscher Dokumentartag 1992: Technik und Information. Markt, Medien und Methoden. Frankfurt a.M.: Deutsche Gesellschaft für Dokumentation 1993. (DGD-Schrift (DOK-5) 1/93)

Example for traditional quantitative database analysis concerning thematic as well as formal aspects

Binder, Gisbert; Matthias Stahl: Der thematische Zusammenhang von Forschung und Lehre: eine scientometrische Analyse. pp. 593-612. In: Best, Heinrich; Brigitte Endres-Niggemeyer; Matthias Herfurth; H. Peter Ohly (eds.): Informations- und Wissensverarbeitung in den Sozialwissenschaften: Beiträge zur Umsetzung neuer Informationstechnologien. Opladen: Westdeutscher Verlag 1994, 623 pp.

Binder, Gisbert; Matthias Stahl: Der Forschungsbezug von Lehrveranstaltungen in den Sozialwissenschaften. pp. 87-106. In: Artus, Helmut M.; Matthias Herfurth (eds.): Soziologielehre in Deutschland: Lehre, Studium, beruflicher Verbleib, Lehrangebot, Studien- und Prüfungsordnungen. Opladen: Leske + Budrich 1996

Examples for the combination of quantitative analysis and thematic comparison of data derived from different databases, in these cases SOLIS (Social Science literature) and LEHRE (= University teaching; database abandoned)

Güdler, Jürgen: Kooperationsnetzwerke in der Forschung. Entstehung, Struktur und Wirkung am Beispiel der Soziologie. Bonn: Informationszentrum Sozialwissenschaften 2003, 238 pp. (Forschungsberichte, Band 5)

Güdler, Jürgen: Dynamik der Medienforschung. Eine szientometrische Analyse auf der Grundlage sozialwissenschaftlicher Fachdatenbanken. Bonn: Informationszentrum Sozialwissenschaften 1996, 136 pp. (Forschungsberichte, Band 1)

Examples for network analysis based on database data, in this case cooperative structures and research networks.

Mutschke, Peter: Autorennetzwerke: Verfahren der Netzwerkanalyse als Mehrwertdienste für Informationssysteme. April 2004, 49 pp. (IZ-Arbeitsbericht Nr. 32)

Example for network analysis based on database data and its usability for upgraded information services in the Social Sciences.

Renner, Ilona: Soziale Kohärenz und Innovativität. Struktureffekte zur Akzeptanz neuer Themen in sozialwissenschaftlichen Forschungsfeldern. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 49, 1997, pp. 74-97.

“Social Coherence and Innovativeness. Structural Effects on the Acceptance of New Subjects in Social Science Research Fields”. Example for

References 2

Other published and unpublished literature concerning the subject matter of this paper.

Artus, Helmut M.: Zu einer Soziologie der Datenbank (Diskussionspapier), IZ Sozialwissenschaften; Bonn, Sept. 1992, 21 pp. (unpublished)

Artus, Helmut M.: Science Indicators Derived from Databases: the Case of the Social Sciences. pp. 297-311. In: *Scientometrics* Vol. 37, 1996, No. 2

Sahner, Heinz: Theorie und Forschung. Zur paradigmatischen Struktur der westdeutschen Soziologie und zu ihrem Einfluß auf die Forschung. (= Beiträge zur sozialwissenschaftlichen Forschung; Bd. 34), Opladen 1982: Westdeutscher Verlag, 338 pp.

Example for the completely unusual case of a bibliographical database especially created for a single, personal research project of the author.

Contact Information

Dr. Helmut M. Artus

GESIS - Informationszentrum Sozialwissenschaften

Lennéstr. 30

53113 Bonn

Germany

E-Mail: helmut.artus@gesis.org