

Priller, Eckhard

**Article — Digitized Version**

## Zur Anwendung und Nutzung von Datenbanken in der soziologischen Forschung - Möglichkeiten und Grenzen

Soziologie und Sozialpolitik: Beiträge aus der Forschung

**Provided in Cooperation with:**

WZB Berlin Social Science Center

*Suggested Citation:* Priller, Eckhard (1987) : Zur Anwendung und Nutzung von Datenbanken in der soziologischen Forschung - Möglichkeiten und Grenzen, Soziologie und Sozialpolitik: Beiträge aus der Forschung, ISSN 0232-9417, Akademie der Wissenschaften der DDR, Institut für Soziologie und Sozialpolitik, Berlin, Vol. 6, Iss. 4, pp. 76-99

This Version is available at:

<https://hdl.handle.net/10419/112459>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



## WZB-Open Access Digitalisate

## WZB-Open Access digital copies

---

Das nachfolgende Dokument wurde zum Zweck der kostenfreien Onlinebereitstellung digitalisiert am Wissenschaftszentrum Berlin für Sozialforschung gGmbH (WZB). Das WZB verfügt über die entsprechenden Nutzungsrechte. Sollten Sie sich durch die Onlineveröffentlichung des Dokuments wider Erwarten dennoch in Ihren Rechten verletzt sehen, kontaktieren Sie bitte das WZB postalisch oder per E-Mail:

Wissenschaftszentrum Berlin für Sozialforschung gGmbH  
Bibliothek und wissenschaftliche Information  
Reichpietschufer 50  
D-10785 Berlin  
E-Mail: [bibliothek@wzb.eu](mailto:bibliothek@wzb.eu)

The following document was digitized at the Berlin Social Science Center (WZB) in order to make it publicly available online.

The WZB has the corresponding rights of use. If, against all possibility, you consider your rights to be violated by the online publication of this document, please contact the WZB by sending a letter or an e-mail to:

Berlin Social Science Center (WZB)  
Library and Scientific Information  
Reichpietschufer 50  
D-10785 Berlin  
e-mail: [bibliothek@wzb.eu](mailto:bibliothek@wzb.eu)

---

Digitalisierung und Bereitstellung dieser Publikation erfolgten im Rahmen des Retrodigitalisierungsprojektes **OA 1000+**. Weitere Informationen zum Projekt und eine Liste der ca. 1 500 digitalisierten Texte sind unter <http://www.wzb.eu/de/bibliothek/serviceangebote/open-access/oa-1000> verfügbar.

This text was digitizing and published online as part of the digitizing-project **OA 1000+**. More about the project as well as a list of all the digitized documents (ca. 1 500) can be found at <http://www.wzb.eu/en/library/services/open-access/oa-1000>.

Eckhard Priller

Zur Anwendung und Nutzung von Datenbanken in der soziologischen  
Forschung - Möglichkeiten und Grenzen

Der Beitrag der Soziologie zur Ausarbeitung wissenschaftlicher Grundlagen für die Leitung, Planung und Prognostizierung sozialer Prozesse in der sozialistischen Gesellschaft ist eng mit einer niveauvollen empirischen soziologischen Forschung verbunden. Empirische Forschung reduziert sich dabei nicht auf die Durchführung und Verwertung speziell ausgerichteter empirisch-soziologischer Erhebungen, sondern umfaßt gleichzeitig die Einbeziehung weiterer für die Forschung verwertbarer empirischer Informationen. Diese können aus verschiedenen Quellen stammen und in unterschiedlicher Form vorliegen, als bereits durchgeführte empirisch-soziologische Untersuchungen zur gleichen Thematik oder mit dieser in Verbindung stehend, als empirisches Material angrenzender Wissenschaftsdisziplinen und als statistische Materialien, die von speziellen Kennziffernsystemen bis zu Daten der Berichterstattung gesellschaftlicher Einrichtungen (Staat, wirtschaftliche Einheiten, Organisationen, Institutionen usw.) reichen.

Der soziologischen Forschung steht somit eine umfangreiche empirische Materialbasis zur Verfügung, deren effektive Nutzung eine grundlegende Aufgabenstellung darstellt. Ihre erfolgreiche Bearbeitung erfordert die Anwendung moderner Informationstechnologien und die Entwicklung entsprechender Instrumentarien. Durch den Einsatz der Computertechnik und die Praktizierung computergestützten Forschens entstehen in dieser Hinsicht sowohl weitreichende Möglichkeiten wie gleichzeitig veränderte Anforderungen an das kooperative und arbeitsteilige Zusammenwirken in der soziologischen Forschung.

#### 1. Einsatzgebiete und Einsatz Tendenzen der Rechentechnik für die soziologische Informationsgewinnung

Die Bedeutung der elektronischen Datenverarbeitung (EDV) für die soziologische Forschung wird seit längerem allgemein anerkannt und findet eine breite Wertschätzung. Der größte Teil empirisch-soziologischer Untersuchungen, die in der DDR durchgeführt werden,

ist ohne den Einsatz elektronischer Datenverarbeitung nicht denkbar. Durch den Einsatz von EDV-Anlagen (EDVA) sind die Bearbeitungsschritte, um empirische Daten aufzubereiten und statistisch auszuwerten, in relativ kurzer Zeit zu bewältigen und die Anwendung rechenaufwendiger mathematisch-statistischer Verfahren erst in größerem Umfang möglich geworden.

Charakteristisch für den bisherigen Einsatz von EDV in der Soziologie der DDR sind folgende Faktoren:

1. Die Rechentechnik fand vor allem zur Aufbereitung und statistischen Auswertung der von Soziologen durchgeführten empirischen Primärerhebungen Verwendung.<sup>1</sup> Insbesondere in den letzten Jahren entwickelten sich unter der Nutzung moderner internationaler Standardsoftware gute Grundlagen für eine rationelle Aufbereitung und Auswertung empirischer Untersuchungen.

2. Der Einsatz von Rechentechnik konzentrierte sich hauptsächlich auf die Nutzung von Großrechnern in entsprechend ausgestatteten Rechenzentren der Industrie, von Akademieeinrichtungen, Hochschulen und Universitäten.

3. Unter forschungsorganisatorischen Gesichtspunkten wurde für die Konzipierung von Untersuchungen eingeräumt, daß neben den spezifischen Bedingungen im Untersuchungsbereich und solchen Faktoren wie Forschungskollektiv, straffe Organisation des Forschungsverlaufs, Zeitpunkt und Zeitspanne sowie zur Verfügung stehende finanzielle und technische Hilfsmittel auch entsprechende Zugangsmöglichkeiten der Datenaufbereitung und Datenverarbeitung als eingrenzende bzw. die Spezifik einer Untersuchung bestimmende Faktoren Berücksichtigung finden müssen.<sup>2</sup> Konzeptionierung, Vorbereitung und Durchführung sowie die Auswertung der Untersuchungen mittels EDV bleiben dabei relativ eigenständige Etappen, die zwar wechselseitige Prämissen setzen, sich aber nur in begrenztem Maße gegenseitig durchdrängen. Die damit einhergehende arbeitsteilige Realisierung rechentechnischer Aufgaben durch Mathematiker und EDV-Spezialisten beschränkte sich weitgehend auf eine Charakterisierung als Dienstleistungsfunktion. Die potentiellen Möglichkeiten des Einsatzes moderner Rechentechnik bereits bei der Konzipierung empirischer Forschung wirksam werden zu lassen und sie als integrativen Bestandteil der empirischen Forschung einzusetzen (zum Beispiel im Dialog mit dem Rechner zu

arbeiten), wurden dabei nur partiell ausgeschöpft.

Im Ergebnis der Erhöhung der Leistungsfähigkeit der modernen Rechentechnik, insbesondere der Mikrorechentechnik, sowie deren zunehmender gerätetechnischer Verfügbarkeit erweitert sich gegenwärtig der Einsatz moderner Rechentechnik in der soziologischen Forschung bedeutend. Ausgehend von Schlußfolgerungen und Erfahrungen, wie sie allgemein aus der wissenschaftlichen Arbeit vorliegen, lassen sich zwei generelle Tendenzen reflektieren, die für die soziologische Forschungspraxis zunehmende Bedeutung gewinnen:

1. Der Einsatz von Rechentechnik bleibt in der soziologischen Forschung nicht auf die Aufbereitung und Auswertung von Daten beschränkt, sondern wird für andere Etappen und Bereiche erschlossen. Diese Entwicklung stimmt mit der anderer Wissenschaftsdisziplinen überein. In dem von einem Autorenkollektiv verfaßten Band "Technik für die Wissenschaft"<sup>3</sup> werden Einflüsse, Merkmale und Tendenzen, wie sie mit dem Vordringen moderner Rechen- und Nachrichtentechnik verbunden sind, als Hauptanwendungsgebiete der Forschung in folgenden Richtungen angesehen:

- Erfassung, Verwaltung, Bereitstellung und Austausch wissenschaftlicher Informationen;
- Realisierung wissenschaftlich-technischer Rechnungen;
- Entwicklung und Konstruktion;
- Forschungsleitung und -planung.

Im Rahmen soziologischer Informationsgewinnung gewinnt vor allem der Aspekt der Erfassung, Verwaltung, Bereitstellung und Austausch von Informationen an Bedeutung. Auf einige der damit verbundenen Fragestellungen zur Datenbankproblematik ist deshalb in diesem Beitrag einzugehen.

2. So, wie allgemein aus den Erfahrungen und Erkenntnissen der Wissenschaft geschlußfolgert wird, daß neue technische Mittel, speziell der Informationsverarbeitung und Kommunikationstechnik, neue Methoden zur Nutzung dieser Mittel hervorbringen, sich bestimmte Methoden effektiver oder überhaupt erst praktisch anwenden lassen bzw. neue Methoden auf der Spezifik der neuen Technik basierend entwickelt werden,<sup>4</sup> vollziehen sich entsprechende Entwicklungen in der soziologischen Forschungsmethodik besonders zu Aspekten soziologischer Informationsgewinnung. Damit kommt zu-

gleich das wechselseitige Durchdringen beider Momente zum Ausdruck.

Demonstrativ und beispielhaft sind im folgenden einige Anwendungsmöglichkeiten der Rechentechnik auf jene Seiten soziologischer Informationsgewinnung vorzustellen. Dabei kann in den meisten Fällen noch nicht auf umfangreiche praktische Erfahrungen und Ergebnisse verwiesen werden. Hingegen sind in Auswertung internationaler und nationaler Erfahrungen sowie erster eigener Ergebnisse Schlußfolgerungen für die weitere Arbeit zur Anwendung moderner Rechentechnik für die Informationsgewinnung zu entwickeln. Schwerpunkt der Darstellung ist die Behandlung der Datenbankproblematik, da sie unter den gegenwärtigen Aufgabenstellungen soziologischer Forschung und dem verstärkten Drang nach sekundäranalytischer Gewinnung und Nutzung von Daten von besonderer Relevanz ist. Andere Anwendungsgebiete, die neben dem umfassenden Einsatz moderner Rechentechnik bei der Auswertung empirischer Untersuchungen sich vor allem auch direkt auf die Erhebung beziehen, wie sie bereits in entsprechenden Ansätzen zur Durchführung computergestützter Befragungen<sup>5</sup>, computergestützter Telefoninterviews<sup>6</sup> oder computergestützter Inhaltsanalysen<sup>7</sup> zum Ausdruck kommen und mit diesen Anwendungsgebieten bei weitem noch nicht ausgeschöpft sein dürften, bleiben an dieser Stelle unbehandelt.

## 2. Informationssysteme und Datenbanken

Der Anspruch soziologischer Forschung, einen Beitrag zur Gesellschaftsanalyse zu leisten, bedarf eines qualifizierten Vorgehens bei der Erhebung soziologischer Informationen und bei der sekundäranalytischen Nutzung bereits vorhandener Informationen. Die kurzfristige, systematische, komplexe und nutzerfreundliche Bereitstellung von Informationen erhält zur Durchführung von Analysen und die Erarbeitung von Prognosen wachsende Bedeutung. Neben dem Prozeß der unmittelbaren Informationsgewinnung und den damit verbundenen Fragen der Schaffung eines an den inhaltlichen Schwerpunkten soziologischer Forschung orientierten soziologischen Informationssystems zur Sicherung der Erhebung und Bereitstellung benötigter Informationen bietet die auf der Grundlage dieses Systems aufbauende Speicherung von Daten mittels moderner Rechentechnik Voraussetzungen, um den wachsenden Anforderungen an den

Informationsbedarf zu entsprechen. Das betrifft vor allem den Aufbau und die Nutzung von Datenbanken, die sich als Instrument zur Rationalisierung von Informations- und Analysetätigkeiten bewährt haben. Für die Bezeichnung von Datenbanken hat sich international durchgesetzt, daß sie auf der Grundlage von Informationssystemen bestehende Datenbasen darstellen. Sie arbeiten rechnergestützt und zeichnen sich durch hohe Qualität, Aktualität und flexible Nutzungsmöglichkeiten aus. Die in ihnen nach einheitlichen Ordnungsgesichtspunkten rationell und redundanzarm gespeicherten Daten gestatten vielfältige Auswertungsaspekte, einen schnellen Zugriff und den Dialog mit dem Rechner.<sup>8</sup> Der Aufbau und die Nutzung von Datenbanken erfolgt auf verschiedensten Gebieten. Wesentliche Eigenschaften von Datenbanken für die Statistik arbeiten A. Donda, E. Herrde, O. Kuhn, R. Struck<sup>9</sup> heraus, die voll gültig sind für die Anwendung auf die soziologische Informationsproblematik. Datenbanken beinhalten demnach:

"Die Abspeicherung von massenhaft vorhandenen, miteinander in inhaltlicher Beziehung stehender, einheitlich definierter Daten in der Art und Weise, daß sie entsprechend mannigfaltigen Bedürfnissen

- vielfältig zusammengefaßt, das heißt gruppiert werden können;
- sehr schnell verfügbar, das heißt abrufbar sind;
- mit korrespondierenden Daten verglichen und in Beziehung zu setzen sind;
- unter verschiedenen Gesichtspunkten selektiert werden können;
- und schließlich, daß die Datenmassen bzw. Teile von ihnen mit Hilfe von mathematisch-statistischen Verfahren und anderen Methoden analytisch durchdrungen werden können."

Besonders umfangreich berichtet die Literatur über Datenbanken im Bibliothekswesen, medizinischen, technischen und naturwissenschaftlichen Bereichen. Eine Übersicht des wissenschaftlichen Informationszentrums der Akademie der Wissenschaften der DDR weist 1983 eine Zahl von 1600 in der Welt existierenden faktographischen Datenbanken aus, die sich auf folgende ausgewählte Fachgebiete bzw. Fachkomplexe konzentrieren:<sup>10</sup>

<u>Fachgebiet / Problemkomplex</u>	<u>Zahl der Datenbanken</u>
Naturwissen- schaften	703
darunter	232
Chemie	179
Physik	54
Biowissenschaften	226
Geowissenschaften	
Wirtschaft	695
darunter	284
Industrie	43
Verkehrswesen	
Energie	117
Umwelt	81
Rohstoffe	62
Technik	126
Rechtswesen	58
Demographie	81

Allerdings vollzog sich die Entwicklung und der Einsatz von Datenbanken nicht problemlos. Einseitige, aus EDV-Sicht vorangetriebene Lösungen, führten in der Vergangenheit häufig zu technisch perfektionistischen Ansätzen. Überzeugende Fortschritte blieben begrenzt oder standen in einem ungünstigen Aufwand-Nutzen-Verhältnis, so daß eine Phase der Ernüchterung eintrat, die zu Zweifeln am praktischen Nutzen dieser Form der EDV-Anwendung führte.<sup>11</sup> Auch wurde häufig nur unzureichend beachtet, daß Datenbanken keine Informationssysteme darstellen oder ersetzen, sondern nur die rationelle Erfassung, Speicherung, Verwaltung, Verarbeitung und Bereitstellung von Daten eines Informationssystems übernehmen können. In Auswertung vorhandener Erfahrungen wird es als zweckmäßig und notwendig angesehen, die Gewinnung von Informationen im Vorbereich von Datenbanksystemen anzuordnen. Dabei wird betont, daß die eigentliche Informationsgewinnung (im Sinne der Primärerhebung) unabhängig vom Datenbanksystem erfolgen sollte, sie aber gleichzeitig als Basis zu betrachten ist, auf der ein Datenbanksystem operiert.<sup>12</sup>

Entscheidende Voraussetzungen für das Erreichen von hoher Wirksamkeit und praktischem Nutzen von Datenbanken bestehen darin, daß von den inhaltlichen Zielstellungen, den konkreten Anwendungserfordernissen und vorhandenen Bedingungen ausgehend, spezielle Vorstellungen zum Aufbau und zur Funktionsweise der jeweiligen Datenbank entwickelt werden. Dabei stellt die Ausschöpfung der



real vorhandenen EDV-technischen Möglichkeiten, einschließlich der Nutzung umfangreicher rechen- und programmtechnischer Lösungen, zugleich hohe Anforderungen an den Nutzer (Auftraggeber, Anwender) von Datenbanken. Dieser muß bereit sein, seinen Arbeitsprozeß auf die Erfordernisse und Spezifik der Arbeit mit Datenbanken einzustellen. Das erfordert häufig umfangreiche organisatorische und Leitungsentscheidungen zur Sicherung des Informationsflusses und der Informationsbereitstellung, die Regelungen zwischen verschiedenen Abteilungen einer Struktureinheit oder zwischen unterschiedlichen Institutionen und Einrichtungen.

Mit diesen Problemen hat sich die soziologische Forschung im besonderen Maße auseinanderzusetzen. Dafür können folgende Ursachen und Gründe angeführt werden:

- Soziologie und soziologische Forschung stellen an sich kein Leitungs- oder Informationssystem dar und sind nicht Teil eines solchen Systems, welches regelmäßig mit bestimmten Informationen versorgt wird. Die mittels der Materialien der Staatlichen Zentralverwaltung für Statistik (SZS) und anderer staatlicher Institutionen und Einrichtungen erhobenen und herausgegebenen Informationen bilden zwar einen wesentlichen Grundstock, um soziologischen Forschungsansprüchen gerecht zu werden. Sie reichen in der Regel jedoch nicht aus, um entsprechend differenzierte Analysen vornehmen zu können.

- Gegenwärtig verfügt die Soziologie noch nicht über ein ausgereiftes, entsprechend ihren spezifischen Erfordernissen entwickeltes Informationssystem, das den Grundbedarf an soziologischen Informationen nach entsprechendem Umfang und temporärer Regelmäßigkeit abdeckt. Sicher gilt es hierbei, überzogene Ansprüche zu vermeiden, denn hinsichtlich der Themenvielfalt und ihrer sich verändernden Aktualität und Bedeutsamkeit wird es für die Ansprüche einer Wissenschaft stets nur ein relativ ausgearbeitetes Informationssystem geben können. Ausreichende Dynamik und Flexibilität erweisen sich in dieser Hinsicht als notwendige Grundanforderungen, die an ein Informationssystem zur soziologischen Erforschung gesellschaftlicher Gesetzmäßigkeiten zu stellen sind.

Das Informationsangebot, das der Soziologie mittels sekundäranalytischer Nutzung und durch eigene Erhebungen gegenwärtig zur Verfügung steht, kann einerseits als sehr umfangreich bezeichnet

werden, und die Ausschöpfung dieses Informationsfundus reicht gegenwärtig bei weitem nicht aus. In diesem Zusammenhang wird häufig von "Datenfriedhöfen" gesprochen. Andererseits bestehen zu bestimmten Fragestellungen, wie sie beispielsweise soziologische Aspekte sozialer Entwicklung betreffen, gewisse Defizite. Systematisierung und entsprechende Strukturierungen der Informationsanfordernisse und ihre Abdeckung durch spezielle Kennziffern, wie sie von anderen gesellschaftswissenschaftlichen Disziplinen - zum Beispiel der Demographie oder Ökonomie - oft in einer über Jahrhunderte gehende Tradition erreicht wurden, stehen für die Soziologie noch weitgehend aus.

Die Darstellung der gegenwärtigen Situation in der Soziologie zeigt mit Deutlichkeit, daß bei der Nutzung von Datenbanken für die soziologische Forschung keine illusionären Vorstellungen zu entwickeln sind. Mit Datenbanken läßt sich nur das realisieren, was in ein entsprechendes Informationssystem investiert wurde. Gleichzeitig besteht kein Grund zum Pessimismus. Auch wenn es eine universelle Datenbank auf soziologischem Gebiet in nächster Zeit nicht geben wird, in die man alle soziologischen Informationen eingeben kann bzw., was oft wesentlicher erscheint, aus der man zu jedem beliebigen soziologischen Sachverhalt in gewünschter Struktur Informationen abfordert, so bestehen für den Einsatz von Datenbanken auf bestimmten Teilgebieten günstige Voraussetzungen. Diese Auffassung wird vor allem durch vorhandene internationale und erste nationale Erfahrungen zum Einsatz von Datenbanken in der soziologischen Forschung gestützt. An dieser Stelle sei vermerkt, daß sich die Begriffsverwendung "Datenbank" vor allem unter der zunehmenden Anwendung der Mikrorechenstechnik lawinenartig ausgebreitet hat. So werden ganz allgemein gleichartige Datenbestände eines Sachgebietes, die zu Dateien zusammengefaßt sind, wie zum Beispiel das Telefonbuch, eine Materialliste oder Kundenadressen, als Grundlage einer Datenbank angesehen.<sup>13</sup> In diesem Sinne versteht man unter einer Datenbank "eine durch den Computer gespeicherte Menge von Informationen (eines betrachteten Gegenstandsbereichs) und deren Beziehungen untereinander".<sup>14</sup> Dieser Hinweis auf eine zunehmend breitere Fassung des Datenbankbegriffs erscheint vor allem deshalb angebracht, weil mit den Vorstellungen über Datenbanken aus den 70er Jahren relativ komplexe, umfassende

Informationssysteme assoziiert werden.

Aufbau, Gestaltung und Nutzung der Datenbank folgen in der Regel gleichen Kriterien und Prinzipien, unabhängig vom Umfang und von der Komplexität.

Einige dieser Kriterien und Prinzipien, wie sie beim Aufbau von Datenbanken in anderen Bereichen erstellt wurden bzw. die sich allgemein als notwendig zu berücksichtigend erwiesen haben, sind zur Nutzung von Datenbanken für die soziologische Informationsgewinnung von grundlegender Bedeutung. Dabei sollten rechen- und programmtechnische Lösungen bei der Behandlung der Datenbankproblematik im Rahmen dieses Beitrags ausgespart bleiben. Schwerpunkt bilden hingegen die Aspekte, die sich nutzerseitig an die Rechentechnik stellen. Datenbanken, die einen entsprechenden Beitrag für die soziologische Informationsgewinnung leisten sollen, müssen vor allem folgende Erfordernisse erfüllen (entsprechende Kriterien für den Einsatz von Datenbanken zur Leitung und Planung ökonomischer Prozesse, die in diesem Zusammenhang von Interesse sind, finden wir unter anderem bei W. Lorenz, U. Scherzer und J. Bergmann).<sup>15</sup>

1. Datenbanken müssen zielgerichtete, zuverlässige und aktuelle Informationen zu den Forschungsaufgaben bereitstellen.
2. Sie haben eine möglichst komplexe Darstellung der Zusammenhänge nach inhaltlichen Gesichtspunkten zu gewährleisten.
3. Der Aufbau der Datenbank hat genügend Spielraum für Flexibilität und Reaktionsfähigkeit zu lassen. Gleichzeitig muß die langfristige Orientiertheit der Datenbasis (zum Beispiel zur Sicherung von Zeitreihen) gewährleistet werden.
4. Durch gleiche Strukturiertheit der Datenbasen ist ein einheitlicher methodischer und organisatorischer Aufbau der Datenbank anzustreben. Ein klar strukturierter, leicht überschaubarer Aufbau des Datenbestandes hat den Überblick über die verfügbaren Informationen und das Verständnis für die Nutzungsmöglichkeiten zu sichern sowie zur leichten Handhabbarkeit beizutragen.
5. Durch die richtige (optimale) Bemessung des Datenbestandes sind die ablaufenden Prozesse mit ausreichender Genauigkeit zu charakterisieren. Gleichzeitig sind Übersichtlichkeit, leichte Handhabung und schnelle rechentechnische Bearbeitung zu berücksichtigen.

sichtigen.

6. Eine weitgehend redundanzfreie Erfassung und Speicherung und damit Minimierung von Mehrfachspeicherung ist zu gewährleisten.
7. Eine entsprechende Datenspeicherungsstruktur hat eine vielseitige Auswertung und problembezogene Bereitstellung der Daten zu ermöglichen. Durch den Direktzugriff mittels Bildschirm-Dialogsystem und entsprechende leistungsfähige Betriebssysteme sind effektive und nutzerfreundliche Formen der Arbeit mit der Datenbank zu gewährleisten.

### 3. Nutzung vorhandener Datenbanken zur soziologischen Informationsgewinnung

Die sekundäranalytische Verwertung von Informationen aus anderen Quellen als empirisch-soziologischen Erhebungen ist von immanenter Bedeutung für die soziologische Informationsgewinnung. Unter dem wachsenden Einfluß der EDV-Anwendung für Informationsprozesse auf allen Ebenen der gesellschaftlichen Praxis treten uns dabei zunehmend EDV-gestützte Register und Datenbanken entgegen. Ihre Nutzung für die soziologische Informationsgewinnung hat zwei Aspekte zu berücksichtigen:

1. Hinsichtlich der Einordnung in den soziologischen Erkenntnisprozeß erfüllen die in Datenbanken enthaltenen Informationen eine ebensolche erkenntnismäßige Funktion wie Informationen, die nicht in solchen hochentwickelten technischen Systemen vorliegen. Es ist folglich auch hier stets einzuschätzen, welche Relevanz die Daten für die jeweilige Problemstellung besitzen, wie ihre Aussagekraft zu beurteilen ist, welche methodischen Fragen, die mit der Erfassung der Informationen im jeweiligen Informationssystem verbunden sind, bei der Verwertung unter soziologischer Problemstellung Berücksichtigung finden müssen. Die Prüfung, ob und wie die entsprechenden Informationen in den soziologischen Forschungsprozeß einzubinden sind, wird nicht durch das Vorhandensein bzw. die Möglichkeit der Verfügbarkeit in einer Datenbank aufgehoben oder eingeschränkt.

2. Durch die Einbeziehung von Informationen aus vorhandenen Datenbanken bieten sich in der Regel für die soziologische Forschungstätigkeit eine Reihe günstiger Bedingungen bzw. Vorteile. Das betrifft Möglichkeiten des schnellen Zugriffs, der einheit-

lichen Strukturiertheit, des hohen Grades der Systematisiertheit. Gleichzeitig wird häufig der Einsatz statistischer Verfahren einfacher und rationeller gestaltet, so daß durch die Bildung von Aggregationen oder Desaggregationen, durch Zusammenhänge, Gruppenbildungen usw. der Informationsgehalt einfacher, schneller und in größerem Umfang erweitert werden kann.

Die Vorteile der in Datenbanken anderer Informationssysteme oder Bereiche enthaltenen Informationen sind folglich für die soziologische Informationsgewinnung wirksam auszunutzen. Neben den erkenntnismäßigen Voraussetzungen, die über die Geeignetheit und Verwendbarkeit der Informationen für die Lösung soziologischer Fragestellungen entscheiden, sind für die Einbeziehung von Datenbanken zur soziologischen Informationsgewinnung gegenwärtig vor allem zwei Fragen zu lösen:

1. Kenntnis bzw. Übersicht zu den vorhandenen Datenbanken, die für eine Nutzung im Rahmen soziologischer Forschungsaufgaben bedeutsam sein können. Im Rahmen des vorliegenden Beitrages kann dieses Ziel durch den damit verbundenen hohen analytischen Aufwand nicht erreicht werden. Statt dessen ist nur demonstrativ auf einige Datenbankprojekte, die von besonderer Bedeutung sind, einzugehen.

2. Die Einbeziehung und Nutzung der Informationsinhalte von Datenbanken für soziologische Analysen setzt leitungsmäßige Entscheidungen bei den Datenbank aufbauenden und führenden Institutionen bzw. Einrichtungen voraus. Diese relativ simpel erscheinende Voraussetzung gestaltet sich gegenwärtig unter den Anstrengungen zur Sicherung eines verantwortungsvollen Umgangs mit Datenmaterial häufig als besonders problematisch. Ein Ausweg zur Lösung dieses Problems kann gegenwärtig nur darin gesehen werden, daß es einerseits den soziologisch Forschenden besser gelingen muß, ihre Informationsbedürfnisse konkreter zu fassen und daß in entsprechenden Anforderungen klarer zum Ausdruck kommt, welche Daten zu welchem Zweck benötigt werden. Andererseits ist die Qualität soziologischer Forschungsergebnisse so zu verbessern, daß sie als überzeugendes Argument für die Zuverfügungstellung von Datenbanken wirken. Dabei spielt das Eingehen auf besondere Erfordernisse und Interessen des jeweiligen Praxispartners zur Auswertung bestimmter Fragestellungen eine besondere Rolle.

Praktische Anwendungen zur Nutzung von Informationen, die mittels Datenbanken gespeichert sind, werden in letzter Zeit in Arbeiten von H. Wendt und D. Lindig<sup>16</sup> vorgestellt. Wendt untersucht in seiner Arbeit die Zusammenhänge zwischen Reproduktionsverhalten in unterschiedlich strukturierten Territorien (Kreistypen bzw. Gebietstypen) und das territoriale Bedingungsgefüge für das Reproduktionsverhalten auf der Grundlage von Daten der Territorialstatistik. Die in der Datenbank verfügbaren Informationen der SZS konnten dabei allerdings noch nicht mittels optimaler technischer Möglichkeiten bearbeitet werden, sondern waren in einer gesonderten Vorgehensweise nach Entnahme aus den Datenfonds der SZS erneut einzugeben und statistisch auszuwerten.

Lindig nutzt bei seiner Sekundäranalyse statistischer Dokumente im Rahmen von soziologischen Untersuchungen zum Leistungsstand wissenschaftlich-technischer Kader der industriellen Forschung und Entwicklung statistische Dokumente, die Träger von Einzeldaten (bis zu personenbezogenen Daten) und Träger von bereits zusammengefaßten Daten sind. Dabei handelt es sich bei den Materialien auf zentraler Ebene unter anderem um solche wie das Patentregister und die Patentakten, Erfinderkarteien und Jahresberichte des Europäischen Patentamtes. Diese liegen im Patentamt und im Zentralen Datenspeicher der Staatlichen Zentralverwaltung für Statistik vor. Zugleich werden Datenspeicher zum gesellschaftlichen Arbeitsvermögen aus der Fachberichterstattung, Kaderprojekten der Ministerien und ausgewählter Kombinate sowie Datenspeicher der Wohnungspolitik von kommunalen Einrichtungen genutzt. Von besonderem Interesse für die soziologische Informationsgewinnung sind EDV-gestützte Register und Datenbanken der SZS.

Zur ersten Gruppe gehört das statistische Betriebsregister. Es enthält für die 50 000 Betriebe und Einrichtungen der DDR die wichtigsten konstanten Merkmale wie zum Beispiel:

- Betriebsnummer,
- territoriale Zuordnung,
- Eigentumsform,
- Kurzbezeichnung des Betriebes,
- Schwerpunkt der ökonomischen Tätigkeit,
- Berichtspflicht,

- Name und postalische Anschrift.<sup>17</sup>

Für soziologische Analysen besitzt das statistische Betriebsregister neben Aussagen zu den Eigentumsformen der Betriebe, ihrer territorialen Zuordnung, den Schwerpunkten der ökonomischen Tätigkeit und bestimmten Charakteristika der beschäftigten Werktätigen vor allem als Zugang für durchzuführende empirische Erhebungen besondere Relevanz. Mittels des Betriebsregisters kann eine Zufallsauswahl von Betrieben erfolgen, die in empirisch-soziologischen Erhebungen einzubeziehen sind. Es ist ebenfalls auf der Grundlage der im statistischen Betriebsregister enthaltenen Merkmale eine geschichtete Zufallsauswahl oder eine selektive Auswahl von Betrieben zur Durchführung von Fallstudien möglich.

Von weiterer Bedeutung sind die im Bereich der SZS betriebenen bzw. gegenwärtig im Aufbau befindlichen Datenbanken:

- Datenbank Statistik.

Sie enthält wichtige Plan- und Ist-Daten der Betriebe der Industrie, des Bauwesens, der Landwirtschaft und des Handels, die für die Leitung und Planung der Volkswirtschaft von Interesse sind. Ausgehend von dieser zentralen Datenbank werden Bezirksdatenbanken aufgebaut und genutzt. Sie orientieren vor allem auf den territorialen Aspekt des jeweiligen Bezirks.

- Wohnungs- und Wohngebäudedatenspeicher (Datenbank "Wohnungspolitik")

Für ca. siebzig Groß- und Mittelstädte bzw. Ballungszentren werden Wohnungsbestand, Wohnungsgröße, Ausstattungskomfort, Bauzustand, Belegung und Angaben zu Wohngebieten erfaßt.

- Einwohnerdatenspeicher

Sie werden schrittweise für alle Kreise und Bezirke nach einheitlicher Konzeption, gleicher inhaltlicher Struktur und identischer Technologie aufgebaut. Die Einwohnerdatenspeicher stützen sich auf Daten, die im Personalausweis erfaßt sind.

A. Donda<sup>18</sup> stellt durch die Nutzung von Datenbanken folgende Effekte anhand des Einwohnerdatenspeichers für die SZS fest:

"Die traditionellen statistischen Berichterstattungen über die natürliche und räumliche Bevölkerungsbewegung (werden) wesentlich durch systematische Abfrage der Datenbanken ersetzt." Es wird "eine inhaltliche Qualifizierung der Bevölkerungsstatistik erreicht".

Entsprechende Auswirkungen sind bei Nutzung der Datenbanken der SZS ebenfalls für die soziologische Informationsgewinnung bzw. -verwendung zu erwarten. Neben effektiverem Vorgehen bietet sich die Einbeziehung weiterer Merkmale, ihre Verfügbarkeit als Zeitreihen sowie ihre Kombinerfähigkeit mit anderen Merkmalen an.

Einen interessanten Ansatz zur Nutzung betrieblicher EDV-Projekte für die soziologische Sozialstrukturanalyse stellen P. Glotz und H. Brandt vor.<sup>19</sup>

#### 4. Zum Aufbau und zur Nutzung von soziologischen Forschungsdatenbanken

Solchermaßen faktographische Datenbanken sind stets als ein Komplex soziologischer, informationstechnologischer und mathematisch-statistischer Methoden aufzubauen. Mit Hilfe der Rechentech-  
nik und datenbankspezifischer Technologien für die automatisierte Datenverarbeitung können große Komplexe von empirischen Daten erfaßt, geordnet, gespeichert, abgerufen und analysiert werden. Durch die Einbeziehung verschiedener Untersuchungen und unterschiedlicher Informationsquellen beschränkt sich die Forschungsarbeit nicht auf einzelne Untersuchungen, sondern bezieht eine umfassende Informationsbasis ein. Allerdings ermöglicht diese es nicht, die Informationen zu erläutern. Die systematische Sammlung einer großen Anzahl von Daten führt nicht automatisch zu verallgemeinernden Erkenntnissen. Ihre Nutzung erfordert stets die Eingliederung als funktioneller Bestandteil (Mittel) in den soziologischen Erkenntnisprozeß.

In der Regel verfügen soziologische Forschungsrichtungen, Forschungskollektive oder Soziologen allgemein über einen gewissen Fundus von Informationen. Sie bilden ein mehr oder weniger entwickeltes Informationssystem und können Kennziffern verschiedener Statistiken, Angaben aus Publikationen, Forschungsberichten und anderen wissenschaftlichen Materialien oder die ausgewerteten Ergebnisse empirischer Untersuchungen beinhalten. Struktur, Inhalt und Umfang dieses Informationssystems richten sich in starkem Maße nach der Tiefe des Eindringens in eine Thematik, den gegenwärtig und zukünftig zu bearbeitenden Problemstellungen sowie der empirischen Orientierung und Ausrichtung. Entsprechend sind die empirischen Informationen auch systematisch aufbereitet.



Informationssysteme, die sich bisher in dieser Hinsicht als gut bewährt haben, sind in der Regel Sammlungen von Kennziffern und weiteren empirischen Daten auf Karteikarten nach einem bestimmten System. Dieses System kann an bestimmte Instrumentarien der Leitungs- und Planungspraxis oder der statistischen Berichterstattung angelehnt sein. Beispielsweise bieten der Planteil Arbeits- und Lebensbedingungen der Rahmenrichtlinie für die Planung in den Kombinat und Betrieben<sup>20</sup> und die im Statistischen Jahrbuch der DDR oder in Spezialstatistiken enthaltenen Kennziffern zur Bevölkerungsentwicklung gute Grundlagen, um entsprechende systematische Datensammlungen zu entsprechenden Themenstellungen anlegen zu können. Für spezielle soziologische Fragestellungen reichen die damit gegebenen Kennziffern jedoch oft nicht aus bzw. bringen den spezifischen Informationsbedarf nicht ausreichend zum Ausdruck. Deshalb werden Systeme für die systematische Informationssammlung erstellt, die den speziellen Ansprüchen folgen und auf entsprechenden soziologisch-theoretischen Überlegungen basieren. Solche systematischen Gliederungen für die Informationssammlung wurden beispielsweise mit dem Indikatoren-system zur Erforschung der Arbeitsbedingungen<sup>21</sup> oder dem System sozialer Indikatoren der sozialistischen Lebensweise<sup>22</sup> entwickelt. Häufig findet man eine noch mehr oder weniger systematische Sammlung von Daten, die auf folgenden miteinander verbundenen pragmatischen Kriterien aufgebaut sind: a) Daten, die einer Thematik zugehörend bzw. sie tangierend eingeschätzt werden und b) Verfügbarkeit der Daten aus einer der in Frage kommenden Informationsquellen. Dabei finden nur solche Kennziffern in die Sammlung Eingang, über die momentan verfügt werden kann bzw. zu denen man im Verlauf der Forschung gesicherten Zugang erhält.

Die verschiedenen Formen der Datensammlung haben das Gemeinsame, daß ihnen Bedeutsamkeit und Aussagekraft für die Lösung aktueller oder zukünftiger Forschungsaufgaben beigemessen wird. In diesem Sinne bilden sie einen, wenn nicht den wesentlichen, Teil der Informationsbasis der Forschung.

Mittels der Nutzung von Datenbanktechnologien können diese Informationen rationell gesammelt, gespeichert und abgefragt werden. Insbesondere die sich gegenwärtig erweiternden Möglichkeiten der Mikrorechentechnik bieten für den Aufbau solcher Forschungs-

datenbanken günstige Voraussetzungen. In dieser Hinsicht sind vor allem folgende Faktoren bedeutsam:

- Die dezentrale Stationierung von Mikrorechentchnik ermöglicht einen direkten, arbeitsplatzbezogenen Zugriff zum Datenbestand. Aufbau, Vervollständigung und Verwaltung der Datenbank sind integrativ in den Forschungsprozeß einzubeziehen, ohne daß zusätzliches Kaderpotential erforderlich wird.

- Der bisher betriebene Aufwand zur Erstellung von Datensammlungen läßt sich langfristig effektiver gestalten. Dabei ist allerdings nicht zu unterschätzen, daß in der Aufbauphase ein hoher arbeitszeitlicher Aufwand erforderlich ist. Dieser gestaltet sich jedoch durch die spätere günstigere Verwaltung des Datenbestandes insgesamt effektiver.

- Für die gegenwärtigen Spezifika der Datensammlungen (Umfang, Zielstellungen) bietet die Mikrorechentchnik vor allem in folgender Hinsicht günstige technische Parameter:<sup>23</sup>

"einfache Algorithmen der Datenbankorganisation,  
 . rationelle Speicherungsverfahren,  
 . nutzerfreundliche, problemflexible Datenbanksprache,  
 . Minimum von Dienstfunktionen für das Datenbankbetriebssystem,  
 . Orientierung auf eigenständige Anwendungsgebiete im Sinne einer Arbeitsplatzunterstützung".

- Mittels des Datenverwaltungssystems REDABAS (bzw. dBASE II für PC, dBASE III für AC) liegt ein anwendungsbereites Datenverwaltungssystem vor. Des weiteren sind Tabellenkalkulationsprogramme vorhanden, die entsprechend den konkreten Dateistrukturen und Zielstellungen zu nutzen sind.

- Entsprechend der bisherigen Orientierung auf themengebundene Datensammlungen bietet sich der Aufbau entsprechender Subdatenbanken an, die durch einen entsprechenden einheitlichen Struktur- aufbau der Dateien miteinander verknüpfbar gestaltet werden können.

Der Aufbau von soziologischen Forschungsdatenbanken kann zu folgenden Vorteilen führen:

1. Die einmalige Abspeicherung der Daten sichert eine, modernen technischen Möglichkeiten entsprechende, Nutzung. Dazu zählen vielfältige Auswertungen und Verknüpfungen, die eine verstärkte disziplinäre und interdisziplinäre Zusammenarbeit unter-

stützen.

2. Der Aufwand für Verwaltung und Aktualisierung des Datenbestandes verringert sich.
3. Daten können weitgehend redundanzfrei, das heißt sie müssen nicht mehrfach gespeichert werden.
4. Trotz dezentralem Aufbau ist bei Berücksichtigung einheitlicher struktureller Gestaltungsmerkmale für den Dateiaufbau innerhalb einer Forschungsrichtung oder Institution ein zentraler Zugriff zu den Daten der Datenbank bzw. der Subdatenbanken möglich.

#### 5. Anwendung der Datenbanktechnologie auf methodische Speicher

In der Regel wird beim Aufbau von Datenbanken gleichzeitig auf die Nutzung von Methodenbanken orientiert.<sup>24</sup> Sie werden teilweise als integrativer Bestandteil einer Datenbank gefaßt. Ausgangspunkt sind Überlegungen der Art, daß mit Methodenbanken eine Übersicht zu den im Rahmen der eigentlichen Datenbank anwendbaren Methoden zu geben ist. In der Regel handelt es sich hierbei um mathematisch-statistische Methoden, die im Rahmen der jeweiligen Datenbankkonstruktion möglich und sinnvoll sind. Besitzt damit die Methodenbank häufig einen engen Bezug zur Datenbank, erfordern ihre Ansprüche und Kriterien gleichzeitig eine relativ selbständige Bearbeitung. Zielstellung, Vorgehensweisen und Aufbau von Methodenbanken lassen sich in prinzipieller Hinsicht auf in den letzten Jahren zunehmend an Aktualität gewinnende methodische Speicher der Soziologie anwenden. Anfang der 80er Jahre wurde im Institut für Soziologie und Sozialpolitik der Akademie der Wissenschaften der DDR der Aufbau verschiedener Methodenspeicher erprobt. Auf erhebungsmethodischem Gebiet betrifft das den Aufbau eines Indikatorenspeichers,<sup>25</sup> eines Methodenspeichers<sup>26</sup> und in den letzten Jahren den Aufbau einer Verfahrenssammlung<sup>27</sup>. Eine Übersicht von mathematisch-statistischen Verfahren und Methoden zur Auswertung empirischer Untersuchungen, die als Entscheidungs- und Orientierungshilfe konstruiert wurde und den Charakter einer Methodenbank trägt, wurde von E. Lautsch<sup>28</sup> entwickelt. Von den Speichern auf erhebungsmethodischem Gebiet bewährt sich die Anwendung der Datenbanktechnologie besonders für den Methodenspei-

cher und die Verfahrenssammlung. Der Methodenspeicher stellt eine Sammlung von Erhebungsunterlagen durchgeführter empirisch-soziologischer Untersuchungen der DDR und des Auslandes dar. Neben den eigentlichen Erhebungsinstrumenten wie zum Beispiel hauptsächlich Fragebögen sind eine Reihe wesentlicher Charakteristika der jeweiligen Untersuchung erfasst. Durch die Nutzung der Datenbanktechnologie kann ein schneller Zugriff zu diesen Informationen erfolgen, die bei der Konstruktion von empirisch-soziologischen Erhebungsunterlagen von besonderem Interesse sind.

Sie gewährleisten, daß bisherige Erfahrungen und Ergebnisse sowie erhebungsmethodische Anregungen für Vergleichs- und Ergänzungsuntersuchungen vermittelt werden. Ausgangspunkt für den Zugang zum Methodenspeicher ist eine Schlagwortliste, die die wesentlichen Sachverhalte, mit denen sich empirisch-soziologische Untersuchungen bisher befaßten, beinhaltet. Mittels Eingabe eines bestimmten Schlagwortes in den Computer erhalten wir alle Untersuchungen, die im Methodenspeicher enthalten sind und diesen Sachverhalt untersucht haben, angezeigt. Gleichzeitig sind charakteristische Merkmale der Untersuchungen, der Standorte der Erhebungsunterlagen sowie weiterer Informationen ersichtlich.

Nach einem ähnlichen Prinzip arbeitet die Anwendung der Datenbanktechnologie auf die Verfahrenssammlung. Statt der empirischen Untersuchungen in ihrer gesamten Komplexität handelt es sich hierbei um Instrumentarien, die der Erfassung relativ eng abgegrenzter Sachverhalte dienen und für die sowohl eine entsprechende theoretische Begründung als auch empirisch geprüfte Güterwerte vorliegen.<sup>29</sup> Mittels Nutzung des Computers über einen entsprechenden Schlagwortkatalog bieten sich hier ebenfalls günstige Möglichkeiten für den schnellen Zugriff und ein hohes Maß an Übersichtlichkeit an. Die computergestützte Umsetzung des Methodenspeichers und der Verfahrenssammlung wurden beide im Institut für Soziologie und Sozialpolitik 1986 begonnen und bestätigten prinzipiell die Zweckmäßigkeit und Rationalität der Anwendung der Datenbanktechnologie auf methodische Speicher.

6. Datenbanken zur Speicherung und sekundäranalytischen Nutzung empirisch-soziologischer Untersuchungen  
(Archive empirisch-soziologischer Untersuchungen)

Neben dem forschungsthemenorientierten Aufbau faktographischer Datenbanken haben sich Datenbanken zur Anlage und rationellen Nutzung von Archiven empirisch-soziologischer Untersuchungen bewährt. Die Erhebungen einer oder, bei größeren Systemen, mehrerer wissenschaftlicher Institutionen werden dabei nach einem entsprechenden inhaltlichen Ordnungs- und Beschreibungssystem sowie entsprechenden datenbanktechnologischen Kriterien gespeichert. Die entsprechend dokumentierten Daten liegen in einer für Sekundär- und Vergleichsanalysen schnell abrufbaren und nutzerfreundlichen Form vor, die sowohl den einfachen Informationsabruf als auch tiefergehende Analysen optimal unterstützen.

In den 50er Jahren entstanden die ersten Archive für Sozialforschung vor allem in hochentwickelten kapitalistischen Ländern. Das älteste und bedeutendste Archiv, das The Roper Center in den USA, wurde 1957 geschaffen. Zu den bekanntesten Einrichtungen dieser Art gehört das Zentralarchiv für empirische Sozialforschung an der Universität Köln. Die meisten Archive sind heute in der internationalen Organisation "International Federation of Data Organization for the Social Science" (IFDO) vertreten. Unter einem Archiv der Sozialforschung wird "eine Institution verstanden, die die durch Sozialforschung gewonnenen empirischen Daten erfaßt, gewinnt und in der Regel mit Hilfe der Rechentechnik speichert, bearbeitet, zusammenfaßt und sie zu nutzen hilft".<sup>30</sup>

Möglichkeiten zum Aufbau und zur Nutzung von Datenbanken bei empirischen Untersuchungen finden seit längerem in der DDR Aufmerksamkeit. R. Ludwig<sup>31</sup> weist unter dem Gesichtspunkt der EDV-Anwendung für die empirische Forschung darauf hin, daß der Nutzungsgrad empirischer Untersuchungen, insbesondere der von Intervallstudien, sich durch die Verwendung einer Datenbank wesentlich erhöhen kann.

Neben einem schnellen Zugang zum aufbereiteten Material aller durchgeführten Untersuchungen wird besonders die Möglichkeit zum Aufzeigen von Tendenzen und die Zusammenstellung der Daten auf neue Art (Daten der Versuchspersonen zum gleichen Themenkomplex, aber auch verschiedenen Untersuchungen) hervorgehoben.

In mehr oder weniger systematischer Weise verfügen die soziologischen Forschungseinrichtungen der DDR, die in größerem Umfang empirische Untersuchungen durchführen, über entsprechende Regelungen zur Archivierung ihrer Erhebungen. Diese Archive bleiben bisher stets auf den engen Rahmen der eigenen Einrichtungen begrenzt. Fragen des Datentransfers an andere Institutionen sind unvollkommen geregelt, und auf dieser Grundlage ist aus unterschiedlichen Gesichtspunkten (Urheberrechte, finanzielle Aufwendungen, Datenschutz usw.) nur schwer eine Zusammenführung der empirischen Untersuchungen in ein nationales Archiv realisierbar. Für eine umfassende Verwertung empirischer Erhebungen ergeben sich dadurch wesentliche Einschränkungen. Gleichzeitig läßt sich konstatieren, daß, wie bereits dargestellt, die Sekundäranalyse von empirisch-soziologischen Erhebungsdaten als Informationsgewinnungsstrategie noch unzureichend praktiziert wird. Die Sammlungen bzw. Archive empirischer Untersuchungen stellen in dieser Hinsicht ein noch nicht genügend genutztes Potential dar.

Im folgenden soll das Archiv soziologischer Untersuchungen des ISS kurz vorgestellt werden, da es einerseits der gegenwärtig in der DDR-Soziologie typischen Archivierungen entspricht. Andererseits ist mit dem Systematisierungsgrad und den potentiell vorhandenen technischen Nutzungsmöglichkeiten ein Standard erreicht worden, auf den entsprechende Archive in anderen soziologischen Forschungseinrichtungen noch nicht verweisen können.

Bisher sind in das Archiv achtzehn empirische Untersuchungen eingegangen, die durch unterschiedliche thematische Schwerpunktssetzungen, Zielstellungen, Zeiträume der Erhebung, Probandenumfänge und anderen Spezifika gekennzeichnet sind. Das Archiv geht von zwei grundlegenden Komponenten aus.<sup>32</sup> Einerseits ist der einheitliche Aufbau der Datenbasen zu sichern und andererseits muß eine einheitliche Beschreibung der Erhebungen nach inhaltlichen und erhebungsmethodischen Gesichtspunkten erfolgen. Der erste Aspekt beinhaltet vor allem EDV-technische Problemlösungen. Diese gewährleisten, daß die auf Magnetband gespeicherten Daten der Erhebung auf einem gleichen Datenmodell basieren und die gleiche Struktur aufweisen. Die zugrunde gelegte Software sichert die einheitliche Behandlung der Datenbasen hinsichtlich der Datenstrukturen, der Datenkompatibilität, der Datenauswertung und

-bereitstellung. Die Datenbasen der einzelnen Erhebungen sind dabei einheitlich als SPSS-Dateien aufgebaut. Durch die SPSS-Steuersprache sind Datenselektionen, Vergleiche und statistische Analysen durchführbar. Die Kompatibilität zu anderer für die Auswertung bedeutsamer Software (SAS, CLUSTAN und andere) wird gewährleistet. Der einheitliche rechentechnische Aufbau sichert, daß Daten verschiedener Untersuchungen miteinander verknüpfbar sind. Zur Archivierung der Magnetbänder wird eine Datei mit folgenden Angaben geführt:

- Bezeichnung der Untersuchung,
- Zeitraum der Erhebung,
- Anzahl der Probanden,
- Anzahl der Merkmale,
- Kurzbezeichnung der Merkmale und synthetisch gebildeter Merkmale sowie Zuordnung der Literals.

Eine differenziertere inhaltliche und erhebungsmethodische Beschreibung erfolgt durch die Angaben im Methodenspeicher. Sie beinhalten:

- die methodensystematische Einordnung der Untersuchung (angewandte Erhebungsmethoden);
- die inhaltlich-thematische Charakteristik der Untersuchung;
- eine Beschreibung grundlegender inhaltlicher Merkmale der Untersuchung durch die Angabe der angezielten Indikata;
- den Nachweis topologischer Daten hinsichtlich:
  - a) der Population (strukturelle Charakteristik der Probanden, Probandenumfang, Auswahl der Stichprobe, Repräsentativität der Stichprobe);
  - b) der Durchführung der Erhebung (Untersuchungszeitraum, Spezifika der Durchführung, veranlassende bzw. durchführende Abteilung des Instituts, Auftraggeber und Praxispartner);
  - c) der Charakteristika der erhobenen Daten (Adäquanz und Niveau der statistischen Analyse, Möglichkeiten des Datenzugriffs, bisherige Verwendung der Daten);
  - d) Angaben zur instrumentellen Güte (problembezogene Angemessenheit, Objektivität, Reliabilität, Validität, instrumentelle Handhabbarkeit).

Mit diesen Angaben wird nicht nur eine umfassende Beschreibung der Untersuchung und des bereitgestellten Datenfonds ermöglicht,

sondern es sind zugleich wesentliche Kriterien für deren Einordnung und Bewertung gegeben. Dabei ist zu berücksichtigen, daß die Angaben entsprechend diesem Beschreibungsschema für die einzelnen Untersuchungen in Abhängigkeit von den methodischen Ansprüchen und Zielkriterien unterschiedlich vorhanden sind. Aber auch fehlende Angaben ermöglichen in gewisser Weise bereits entsprechende Einordnungen.

Durch die Implementierung des Methodenspeichers auf die Mikrorechentechnik werden die Benutzerfreundlichkeit und Zugriffszeiten zu den Informationen weiter verbessert.

#### Anmerkungen

- 1 H. Berger, H. Jetzschmann, Der soziologische Forschungsprozeß, Berlin 1973, S. 177, 187; R. Ludwig, Statistische Auswertungsverfahren, in: W. Friedrich, W. Hennig (Hrsg.), Der sozialwissenschaftliche Forschungsprozeß, Berlin 1975, S. 707; R. Stollberg, Warum und wozu Soziologie?, Berlin 1983, S. 108.
- 2 H. Berger, H. Jetzschmann, Der soziologische Forschungsprozeß, a.a.O., S. 173.
- 3 Autorenkollektiv unter Leitung von W. Meske, Technik für die Wissenschaft, Studien und Forschungsberichte, Heft 20, Akademie der Wissenschaften der DDR, Institut für Theorie, Geschichte und Organisation der Wissenschaft, Berlin 1986, S. 158.
- 4 E. Banse, H. Wendt (Hrsg.), Erkenntnismethoden in den Technikwissenschaften, Berlin 1986, S. 150.
- 5 Vgl. C. W. Türschmann, Computergestützte Befragung in der empirischen Sozialforschung, Berlin(West) 1981; vgl. E. Mochmann, Computerstrategie für die Kommunikationsanalyse, Frankfurt am Main 1980.
- 6 Vgl. E. Brückner, Telefoninterviews - ein alternatives Verfahren? Ergebnisse einer Pilotstudie, in: Zusanachrichten 11/1982 (Mannheim).
- 7 H. D. Klingemann, Computergestützte Inhaltsanalyse in der empirischen Sozialforschung, Königstein 1982.
- 8 J. Bergmann, Konzeption zur Entwicklung eines Systems der ökonomischen Information auf der Grundlage einer Datenbank, Dissertation B (unveröffentlicht), Berlin 1985, S. 103.
- 9 A. Donda, E. Herrde, O. Kuhn, R. Struck, Statistik, Berlin 1986, S. 55.



- 10 W. Richter, Zu qualitativ neuen Anforderungen und Aufgaben auf dem Gebiet der Wissenschaftsinformation, in: Zur Bedeutung der Information für Individuum und Gesellschaft, Berichtsband der Wissenschaftlichen Konferenz zum Leibniz-Tag der Akademie der Wissenschaften der DDR, Berlin 1983, S. 332.
- 11 J. Bergmann, Konzeption zur Entwicklung eines Systems der ökonomischen Information auf der Grundlage einer Datenbank, Dissertation B (unveröffentlicht), a.a.O., S. 72.
- 12 K. H. Dreckmann, Aufbau und Arbeitsweise eines Datendokumentationssystems, in: R. Kuhlen (Hrsg.), Datenbasen, Datenbanken, Netzwerke, Bd. 2: Konzepte von Datenbanken, München, New York, London, Paris 1979, S. 102.
- 13 U. Hempel, H. Loley, Datenbanken mit Personalcomputern, Berlin 1987, S. 12.
- 14 Ebenda, S. 13.
- 15 W. Lorenz, U. Scherzer, Inhalt und Methodik der Projektierung von Informationssystemen auf der Basis von Systemunterlagen (Programmbausteine) und Datenbanken, Leipzig 1981, S. 102; Vgl. J. Bergmann, Konzeption zur Entwicklung eines Systems der ökonomischen Information auf der Grundlage einer Datenbank, Dissertation B (unveröffentlicht), a.a.O.
- 16 Vgl. H. Wendt, Reproduktionsverhalten in territorialer Differenzierung. Ein Beitrag zu den Grundlagen sozialen Verhaltens in Territorialeinheiten unterschiedlicher sozialökonomischer Struktur in der DDR, Dissertation B (unveröffentlicht), Berlin 1986; vgl. D. Lindig, Zur Anwendung der Sekundäranalyse statistischer Dokumente in der Untersuchung zum Leistungsstand wissenschaftlich-technischer Kader (Erfinder) der industriellen Forschung und Entwicklung, Dissertation B (unveröffentlicht), Berlin 1986.
- 17 A. Donda, Parastatistik?, in: Zum 80. Geburtstag von J. Kuczynski, Sitzungsberichte der AdW der DDR, Reihe Gesellschaftswissenschaften 4 G/1985, S. 51.
- 18 Ebenda, S. 52.
- 19 Vgl. P. Glotz, H. Brandt, Die Sozialstruktur des Kombinars, Leipzig 1983.
- 20 Gesetzblatt der DDR 1980, Sonderdruck Nr. 1021, S. 188.
- 21 Vgl. E. Priller, Die Gewinnung von Indikatoren zur soziologischen Erforschung der Arbeitsbedingungen, in: Jahrbuch für Soziologie und Sozialpolitik 1980, Berlin 1980.
- 22 Vgl. H. Berger, Th. Hanf, W. Hinrichs, E. Priller, D. Rentzsch, System sozialer Indikatoren der sozialistischen Lebensweise, Schriftenreihe des ISS. Beiträge aus der Forschung 1/1984.
- 23 Vgl. E. Priller, B. Krüger, I. Piens, Zu den Grundlagen und Voraussetzungen des Aufbaus und der Arbeit mit Datenbanken in der soziologischen Forschung (unveröffentlicht), Berlin 1987, S. 17.
- 24 Vgl. J. Bergmann, Konzeption zur Entwicklung eines Systems der ökonomischen Information auf der Grundlage einer Datenbank, Dissertation B (unveröffentlicht), a.a.O., S. 118;

- vgl. W. Lorenz, U. Scherzer, Inhalt und Methodik der Projektierung von Informationssystemen auf der Basis von Systemunterlagen (Programmbausteine) und Datenbanken, a.a.O., S. 40.
- 25 Vgl. E. Priller, Erfahrungen zum Aufbau und zur Arbeit eines Indikatorenspeichers und zur Entwicklung von Standardindikatoren, in: Schriftenreihe des ISS. Beiträge aus der Forschung 4/1982.
- 26 Vgl. M. Häder, E. Priller, R. Schubert, E. Brislinger, Die Nutzung des Methoden- und Indikatorenspeichers, Arbeitsblatt Nr. B 06-82, in: Autorenkollektiv, Methodischer Leitfaden für empirische Untersuchungen - Arbeitsblätter, Schriftenreihe des ISS. Beiträge aus der Forschung, 1982.
- 27 Vgl. B. Hausstein, Zielstellung, Aufbau und beispielhafte Darstellung einer Verfahrenssammlung erhebungsmethodischer Instrumente, in diesem Heft, S. 5.
- 28 Vgl. E. Lautsch, Methodologische und methodische Betrachtung zu einer möglichen Systematisierung mathematisch-statistischer Auswertungsmethoden in der empirisch-soziologischen Forschung, Dissertation B (unveröffentlicht), Berlin 1985; ders., Auswertungsprogrammplanung - Voraussetzungen und Arbeitsschrittfolge, in: Informationen zur soziologischen Forschung in der DDR, 22(1986) Nr. 5, Berlin 1986.
- 29 Vgl. B. Hausstein, Zielstellung, Aufbau und beispielhafte Darstellung einer Verfahrenssammlung erhebungsmethodischer Instrumente, a.a.O., S. 5.
- 30 M. Illner, M. Foret, Socialni ukazatele, Praha 1980, S. 303.
- 31 R. Ludwig, Statistische Auswertungsverfahren, in: W. Friedrich, W. Hennig (Hrsg.), Der sozialwissenschaftliche Forschungsprozeß, a.a.O., S. 707/708.
- 32 Vgl. E. Priller, B. Krüger, I. Piens, Zu den Grundlagen und Voraussetzungen des Aufbaus und der Arbeit mit Datenbanken in der soziologischen Forschung (unveröffentlicht), a.a.O., S. 31 - 33.