

**DIE PERSONENBEZOGENEN MASSENDATENSPEICHER DER DDR IM
BUNDES-ARCHIV**

Die Notwendigkeit, nach der Auflösung von Behörden und Rechenzentren der ehemaligen DDR umfangreiche Datenbestände zu übernehmen, bildete einen wesentlichen Grund dafür, daß am 1. August 1991 ein selbständiges Referat für die Archivierung maschinenlesbarer Daten im Bundesarchiv geschaffen wurde. Die erste Übernahme von Daten nach der Wiedervereinigung war bereits im Januar 1991 erfolgt – es handelt sich um die sogenannte „Datei Grenzzwischenfälle“ des Kommandos der Grenztruppen der DDR. Heute (Stand April 1997) verwahrt das Referat LS 4 im Bundesarchiv einen Bestand von 1.504 Magnetbändern, 1.785 Disketten und 23 CD-ROMs. Nicht eingerechnet sind Duplikate und Sicherungskopien. Etwa 90% der Datenträger stammen aus DDR-Provenienz. Das Referat besteht seit Januar 1997 aus einem Referatsleiter, zwei Sachbearbeitern (einem für die technische Bearbeitung und einem für die archivische Erschließung), einer Bürosachbearbeiterin (zu 30%) und einer weiteren EDV-Fachkraft, die halbtags arbeitet. Die Personalausstattung umfaßt demnach rein mathematisch gesehen 3,8 Personen.

Die technische Ausstattung besteht vor allem aus einer EDV-Anlage für die Sichtung, Analyse und Umkopierung der Datenbestände, die Ende April 1993 – also über zwei Jahre nach Übernahme des ersten DDR-Datenbestands – installiert und im September 1993 abgenommen wurde. Es handelt sich bei diesem DV-System um einen PC 386 DX 25 mit 4 Megabyte Arbeitsspeicher¹. Der Ersatz dieser inzwischen technisch überholten Anlage durch ein neues System wurde bereits für 1996 avisiert, steht aber bislang noch aus.

Einige typische Probleme mit den aus DDR-Provenienz übernommenen Datenträgern sollen hier kurz angesprochen werden.

Zum ersten: Wir mußten feststellen, daß die übernommenen Magnetbänder – es handelt sich um die Fabrikate ORWO und Pyral – von ausgesprochen schlechter Materialqualität waren. Beim Einlesen der Bänder findet ein extrem hoher Materialabrieb statt, der ein sehr häufiges Reinigen der Schreib-Leseköpfe der Magnetbandlaufwerke erfordert, was wiederum zu erhöhtem Verschleiß an den Schreib-Leseköpfen führt. Der schlechte Zustand der Bänder hatte teilweise ganz erhebliche Leseprobleme und in Einzelfällen sogar Datenverluste zur Folge. Glücklicherweise waren die Dateien in den meisten Fällen in doppelter Ausführung, also mit einer zusätzlichen Kopie, übernommen worden. Nach Auskunft von Mitarbeitern von DDR-Rechenzentren wurden wegen der drohenden Gefahr von Datenverlusten sogar bis zu vier Sicherungskopien verwahrt. Im Bundesar-

chiv wurden auf der Umkopierstation seit September 1993 als Erstsicherungen 1:1-Kopien auf Magnetbänder von guter Qualität erstellt. Seit Dezember 1995 werden zusätzlich Zweitsicherungen auf Digital Audio Tape angefertigt. Mit der neu zu beschaffenden DV-Anlage sollen künftig Drittsicherungen auf CD-ROM gebrannt werden.

Zum zweiten: Disketten in den Formaten 8“ und 5 1/4“ von Robotron-Bürocomputern, die unter den Betriebssystemen SCP, SIOS und DCP erstellt wurden, können auf heutigen PCs unter MS-DOS 6.22 oder gar unter WINDOWS 95 nicht gelesen werden. Seit Juli 1996 habe ich mich daher bemüht, eine kleine Sammlung von Robotron-Bürocomputern verschiedener Typen aufzubauen, um die über 1.200 Robotron-Disketten im Bundesarchiv überhaupt sichten und bewerten zu können. Diese Sichtung hat im Januar 1997 begonnen.

Ein drittes, folgenschweres Problem stellt das teilweise oder vollständige Fehlen von Dokumentationsunterlagen zu den übernommenen Datenbeständen dar. Ohne Dokumentation ist ein Datenbestand in der Regel unbrauchbar, da die Daten ohne Erläuterungen zur Datenstruktur, zu Codierungen und Komprimierungsverfahren nicht mehr verstanden und ausgewertet werden können. Besonders beim Zentralen Kaderdatenspeicher (ZKDS) war dies der Fall, zu dem die Mitarbeiter des ehemaligen Rechenzentrums die kompletten Unterlagen vernichtet hatten. Glücklicherweise konnte diese Lücke mit Hilfe der Kollegen in der Abteilung DDR des Bundesarchivs in Berlin mit Unterlagen aus Schriftgutbeständen größtenteils geschlossen werden². In vielen Fällen sind ohnehin zusätzliche Recherchen in den korrespondierenden Schriftgutbeständen erforderlich, was den zeitlichen und organisatorischen Aufwand der archivischen Erschließung erhöht, da die Schriftgutbestände aus DDR-Provenienz in Berlin, die Daten dagegen in Koblenz liegen.

Bei etwa einem Drittel der ins Bundesarchiv übernommenen Datenbestände aus DDR-Provenienz handelt es sich um personenbezogene Daten – genauer um fünf Datenbestände. Die folgende Tabelle faßt einige wichtige Angaben zu diesen fünf Datenbeständen zusammen.

Datum Übernahme	Datenbestand (Kurzbez.)	Anzahl Magnetbänder	Erfassungszeitraum/ Laufzeit	Anzahl der Personen-Datensätze	Maximale Länge Datensätze in Byte
Nov. 1991	GAV	75	1989	ca. 7.250.000	512
Dez. 1991	ZKDS	200	1980, 1985-1989	von 90.000 bis 330.000 je Jahrgang und je nach Datei	2.303
Okt. 1992	NRC	337	1980-1991	Archivdatei bis 1984 ca. 170.000, neuere Dateien je ca. 190.000	680
Dez. 1994	Projekt STUBEW	22	1985-1990	ca. 180.000 je Jahrgang	1.335
Jan. 1995	AKDS Volksbildung	128	1988-1990	ca. 330.000 je Jahrgang	2.303

Der „Zentrale Kaderdatenspeicher“ des Ministerrates der DDR³ wurde im Dezember 1991 ins Bundesarchiv übernommen. Der ZKDS diente der Abteilung Kader des Ministerrates zur Bereitstellung von Informationen für „Auswahl, Entwicklung und Einsatz von Führungskadern und Mitarbeitern in den zentralen und örtlichen Staatsorganen und von Leitungskadern in der zentral und örtlich geleiteten Wirtschaft“⁴. Der ZKDS enthält u. a. Angaben zu folgenden Punkten: Zugehörigkeit zu Parteien und Massenorganisationen, politische und fachliche Aus- und Weiterbildung, gegenwärtige und frühere Arbeitsstellen und Funktionen, berufliche Entwicklung, Wahlfunktionen, Kaderreserve, Reisekader, Auslandserfahrungen. Die Bestands- und Abgangsdateien der Jahre 1985 bis 1989 dürften für sozialwissenschaftliche Fragen, z. B. im Rahmen der Elitenforschung, in Frage kommen.

Das „Projekt NRC“ war eine EDV-Anwendung der Verwaltung Strafvollzug beim Ministerium des Innern über die Häftlinge in der gesamten DDR. Im NRC⁵ sind detaillierte personenbezogene Angaben zu den Häftlingen aus den Jahren 1980 bis 1991 gespeichert. Für sozialwissenschaftliche Auswertungen dürften die Dateien aus dem NRC von besonderem Interesse sein. Sie enthalten insgesamt etwa 170 auswertbare Datenfelder mit auf die Einzelperson bezogenen Angaben. Beispielsweise sind folgende Merkmale enthalten: Disziplinarmaßnahmen, Verlegungen von Häftlingen, Maßnahmen der allgemeinen und politischen

Bildung, Festnahmegrund, Datum der Festnahme, Familien- und Wohnverhältnisse, Zugehörigkeit zu Parteien und Massenorganisationen, Tätigkeit in bewaffneten Organen, Verbindungen ins kapitalistische Ausland, Art des verurteilenden Gerichts, Urteilsdatum, Art des Untersuchungsorgans (z. B. MfS oder Ministerium des Innern), Straftat, Strafmaß, Dauer der U-Haft, Anzahl der Vorstrafen, Entlassungsdatum, Entlassungsgrund, Maßnahmen der Wiedereingliederung usw. Die Datenstruktur des NRC ist sehr kompliziert und die Erschließung daher sehr zeitaufwendig. Die inhaltliche Erschließung steht aber inzwischen vor dem Abschluß. Die Frage der Benutzung des NRC für wissenschaftliche Forschungsprojekte wird jedoch noch große rechtliche und technische Probleme aufwerfen. Es handelt sich teilweise um sensible personenbezogene Angaben, die Dritten selbstverständlich nur in anonymisierter Fassung zur Auswertung verfügbar gemacht werden können. Die schutzwürdigen Belange der Betroffenen müssen berücksichtigt werden, wobei bei diesen Daten eine besondere Qualität der Schutzwürdigkeit gegeben ist. Entsprechende Maßnahmen der Anonymisierung erfordern im Bundesarchiv jedoch einen ganz erheblichen zeitlichen und personellen Aufwand.

Die Daten des sogenannten „Projekts STUBEW“ wurden bei der Zentralstelle für Studienbewerbungen des Ministeriums für Hoch- und Fachschulwesen geführt. Im Rahmen dieses Projektes wurden die Wünsche, Bewerbungen, Zulassungen und Ablehnungen für Studienplätze im Direkt- und Fernstudium zentral gespeichert. Erhalten und ins Bundesarchiv gelangt sind die Dateien aus den Jahren 1985 bis 1990. Erfasst wurden neben Angaben zu Namen und Adressen vor allem Informationen zu Staatsangehörigkeit, Familienstand, Kinderzahl, Mitgliedschaft in Parteien und Massenorganisationen, militärischer Ausbildung, Noten des letzten Zeugnisses, Jahr der Hochschulreife, Zeitpunkt der Studienbewerbung, Zeitpunkt der Zulassung zum Studium, Grund der Nichtzulassung, gewünschte Fachrichtung und Hochschule. Diese Überlieferung kann ein präzises und umfassendes Bild der Sozialstruktur der Studienplatzbewerber und der Studenten in den letzten fünf Jahren des Bestehens der DDR vermitteln. Auch hier sind jedoch datenschutzrechtliche Bestimmungen zu beachten und berechnete Belange Betroffener zu schützen.

Im Ministerium für Volksbildung wurde – wie in jedem anderen DDR-Ministerium – ein sogenannter Arbeitskräfte-datenspeicher – kurz AKDS – geführt, also eine umfangreiche Personaldatenbank. Der „AKDS Volksbildung“ enthält die Daten des gesamten Personals der Schulen, Kindergärten, Kinderheime, Internate und Tagesstätten der DDR, vom Leiter der Einrichtung bis zum Hausmeister oder Reinigungspersonal aus den Jahren 1988 bis 1990. Der Datensatzaufbau entspricht im wesentlichen dem des ZKDS.

Die Problematik des Datenschutzes und die daraus resultierenden Benutzungsbeschränkungen bei den genannten Datenbeständen sollen hier kurz an-

gerissen werden: Im Bundesdatenschutzgesetz § 1 Absatz 1⁶ heißt es: „Zweck dieses Gesetzes ist es, den einzelnen davor zu schützen, daß er durch den Umgang mit seinen personenbezogenen Daten in seinem Persönlichkeitsrecht beeinträchtigt wird.“ Das Persönlichkeitsrecht gehört zu den höchsten vom Grundgesetz geschützten Werten. Das Bundesverfassungsgericht hat in seinem Volkszählungsurteil am 15. Dezember 1983 entschieden: „Das Grundrecht gewährleistet insoweit die Befugnisse des einzelnen, grundsätzlich selbst über die Preisgabe und Verwendung seiner persönlichen Daten zu bestimmen.“⁷ Einschränkungen des Rechts auf „informationelle Selbstbestimmung“ bedürfen jeweils der gesetzlichen Grundlage. Die Bestimmungen des Datenschutzes gelten selbstverständlich auch für das Bundesarchiv. Es hat prinzipiell von der Übernahme an die schutzwürdigen Belange Betroffener zu berücksichtigen, insbesondere bei Unterlagen mit personenbezogenen Daten. So sieht das Bundesarchivgesetz⁸ nach § 5 Absatz 2 vor, daß „Archivgut des Bundes, das sich auf natürliche Personen bezieht, erst 30 Jahre nach dem Tode des Betroffenen durch Dritte benutzt werden darf. Ist das Todesjahr nicht oder nur mit unverhältnismäßigem Aufwand festzustellen, endet diese Schutzfrist 110 Jahre nach Geburt des Betroffenen.“

Diese Schutzfristen, die den Betroffenen dienen und auch von interessierten Wissenschaftlern daher sehr ernst genommen werden sollten, können allerdings unter bestimmten Voraussetzungen verkürzt werden. So ermöglicht § 5 (5) des Bundesarchivgesetzes die Verkürzung von Schutzfristen, „wenn die Benutzung für ein wissenschaftliches Forschungsvorhaben oder zur Wahrnehmung berechtigter Belange unerlässlich ist, die im überwiegenden Interesse einer anderen Person oder Stelle liegen, und eine Beeinträchtigung schutzwürdiger Belange durch angemessene Maßnahmen, insbesondere durch Vorlage anonymisierter Reproduktionen, ausgeschlossen werden kann.“ In der Praxis bedeutet dies, daß bei Benutzungen einer Datei in den Datensätzen bestimmte Datenfelder wie Name, Personenkennziffer, Wohnadresse usw. gelöscht bzw. unkenntlich gemacht werden müssen, so daß eine Einzelperson nur mit einem unverhältnismäßig hohem Aufwand an Zeit, Kosten und Arbeitskraft identifiziert werden kann (faktische Anonymisierung). Im Interesse der Forschung sollen auf der anderen Seite jedoch die Daten so wenig wie möglich verändert werden, um das Analysepotential nicht gravierend einzuschränken, so daß es im Einzelfall erforderlich sein wird, einen Kompromiß zwischen Datenschutz und Wissenschaft zu finden, der die Interessen beider ausreichend befriedigt.

Die erwähnte Personenkennziffer (kurz PKZ) war übrigens nicht eine reine Erfindung der DDR – ihr Aufbau orientierte sich an den Entwürfen des Bundesinnenministeriums in der Amtszeit Hans-Dietrich Genschers und an Vorbildern in Skandinavien und in der damaligen CSSR. Während die PKZ in der Bundesrepublik nicht durchgesetzt werden konnte, wurde sie ab 1971 jedem DDR-Bürger bei seiner Geburt zugewiesen. Sie begleitete ihn durch sein ganzes Leben als persön-

liche Kenn-Nummer, fand sich in sämtlichen persönlichen Dokumenten wieder und diente in der EDV als verbindendes Merkmal zwischen verschiedenen Datenspeichern, was den Datenaustausch von personenbezogenen Daten erheblich vereinfachte. Jeder DDR-Bürger konnte anhand seiner PKZ eindeutig identifiziert werden. Die PKZ war zwölf Stellen lang und enthielt das Geburtsdatum, das Geschlecht, bei vor 1970 Geborenen den Wohnort, bei nach 1970 Geborenen den Geburtsort sowie eine Prüfziffer⁹:

Stellen	Inhalt	Codierung usw.
1 bis 2:	Geburtstag	
3 bis 4:	Geburtsmonat	
5 bis 6:	Geburtsjahr	
7:	Jahrhundert und Geschlecht	2 = vor 1900 und männlich 3 = vor 1900 und weiblich 4 = nach 1900 und männlich 5 = nach 1900 und weiblich
8 bis 11:	Unterscheidungszahl:	• bei vor 1970 Geborenen: Wohnort • bei nach 1970 Geborenen: Geburtsort
12:	Prüfziffer zur Plausibilitätsprüfung	

Bei dem „Datenspeicher Gesellschaftliches Arbeitsvermögen“¹⁰ (GAV) handelt es sich um eine der wenigen erhaltenen und zugleich wohl bedeutendsten Quellen aus dem Bereich Arbeitsmarktstatistik der ehemaligen DDR. Die Daten der letzten Erhebung des GAV vom 31.12.1989 wurden ins Bundesarchiv übernommen. Im GAV sind die Daten von etwa 7,25 Millionen Berufstätigen der DDR enthalten – das sind etwa 75% aller Beschäftigten der DDR. Der GAV diente als Mittel der Planung des Arbeitseinsatzes. Er wurde „zur weiteren Qualifizierung der staatlichen Leitung und Planung des rationellen Einsatzes und der effektiven Nutzung des gesellschaftlichen Arbeitsvermögens“ der DDR aufgebaut¹¹. Das „gesellschaftliche Arbeitsvermögen“ ist ein Begriff aus der sozialistischen Volkswirtschaft: „Das gesellschaftliche Arbeitsvermögen umfaßt die Gesamtheit der im gesellschaftlichen Arbeitsprozeß anwendbaren Kenntnisse, Fähigkeiten und Fertigkeiten der arbeitsfähigen Bevölkerung.“¹² Nach Jürgen Salomon spielte das „gesellschaftliche Arbeitsvermögen“ für die gesellschaftliche und wirtschaftliche Entwicklung der DDR eine wesentliche Rolle¹³. Im Programm der SED heißt es: „Die Sozialistische Einheitspartei Deutschlands geht stets davon aus, daß der Mensch die Hauptproduktivkraft ist, zu dessen Nutzen die Intensivierung der Produktion durchgeführt wird und dessen Arbeit dadurch

erleichtert wird. Sie schenkt der Entwicklung und rationellen Nutzung des gesellschaftlichen Arbeitsvermögens große Bedeutung.“¹⁴

Für den GAV hatten Staatsorgane, Kombinate, wirtschaftsleitende Organe, Betriebe und Genossenschaften die Daten der Berufstätigen zu melden. Nicht erfaßt wurden jedoch z. B. Mitarbeiter der zentralen und örtlichen Staatsorgane, hauptamtliche Mitarbeiter der Parteien und Massenorganisationen und ihrer Betriebe, Mitarbeiter der Bereiche Inneres, Staatssicherheit, Nationale Verteidigung und Zollverwaltung. Umgekehrt konnten Staatsorgane, Kreise und Betriebe aus dem GAV detaillierte Informationen zu qualitativen und quantitativen Aspekten des „Arbeitsvermögens“ der Berufstätigen ihres „Verantwortungsbereiches“ anfordern. Dem GAV kommt aufgrund seiner tiefen Schichtung und seines umfangreichen Merkmalkataloges eine große Bedeutung zu. Differenzierte Bildungs-, Qualifikations- und Tätigkeitsmerkmale standen bei der Datenerhebung im Vordergrund. Von besonderem Interesse dürften beispielsweise die Daten zum Ausbildungs- und Berufsweg oder zur beruflichen Mobilität sein. Eingeschränkt sind die Nutzungsmöglichkeiten jedoch einerseits durch den begrenzten Personenkreis, da nur ca. 75% der Berufstätigen erfaßt sind, andererseits dadurch, daß privatwirtschaftliche Betriebe nicht berücksichtigt wurden. Bei Strukturanalysen dürfte aus diesen Gründen Vorsicht geboten sein, da Verzerrungen zu befürchten sind¹⁵. Da nur der Stichtagsdatenbestand vom 31. Dezember 1989 überliefert ist, sind nur Querschnitt-, nicht aber Längsschnittanalysen durchführbar. Bestimmte Personengruppen wie Lehrlinge, Ausländer, pädagogische Fachkräfte und Handwerker sind zudem nur zum Teil erfaßt.

Für das Institut für Arbeitsmarkt- und Berufsforschung in Nürnberg (IAB) wurden 1990 zwei anonymisierte Versionen des GAV mit reduzierten Merkmalkatalogen erstellt, von denen das Bundesarchiv im März 1994 Kopien übernommen hat. Während die nicht anonymisierte „Original“-Fassung des GAV wegen der starken Komprimierung der Daten und aus datenschutzrechtlichen Gründen nicht ausgewertet werden kann, stehen die anonymisierten Versionen mit reduzierten Merkmalkatalogen bereits jetzt für Forschungszwecke zur Verfügung.

Die im Bundesarchiv archivierten personenbezogenen Massendatenarchive der DDR sind eine wichtige Quelle für die historische und sozialwissenschaftliche Forschung. Ihre Zugänglichmachung wirft jedoch eine Reihe von rechtlichen und technischen Schwierigkeiten auf, die künftig zu bewältigen sein werden.

ANMERKUNGEN

1 Wettengel, Michael: System zur Archivierung maschinenlesbarer Daten im Bundesarchiv. In: Mitteilungen aus dem Bundesarchiv, Jg. 1 (1993), Heft 2, S. 70-72.

- 2 Wettengel, Michael: Zur Rekonstruktion digitaler Datenbestände aus der DDR nach der Wiedervereinigung: Die Erfahrungen im Bundesarchiv. Abdruck vorgesehen in: Der Archivar, Jg. 50, 1997.
- 3 Rathje, Ulf: Der „Zentrale Kaderdatenspeicher“ des Ministerrates der DDR. In: Historical Social Research/Historische Sozialforschung (HSR), Vol. 21 (1996), Heft 3, S. 137-141.
- 4 Gruppe Personalangelegenheiten des Ministerrates in Abwicklung. Vorlage über die Sicherstellung von Unterlagen aus dem Bereich der Personalabteilung des ehemaligen Ministerrates der DDR, 6. November 1990. Bundesarchiv, DC 20/9006.
- 5 Rathje, Ulf: Die Strafgefangenen- und Verhaftetendateien der Verwaltung Strafvollzug des Ministeriums des Innern der DDR. In: Historical Social Research/Historische Sozialforschung (HSR), Vol. 22 (1997), Heft 1, S. 132-140.
- 6 Bundesdatenschutzgesetz (BDSG) vom 20. Dezember 1990 (BGBl. I S. 2954); siehe auch: Der Bundesbeauftragte für den Datenschutz: BfD-Info 1. Bundesdatenschutzgesetz. Text und Erläuterung. Bonn, 3. Auflage 1995.
- 7 Urteil vom 15. Dezember 1983 (1 BvR 209/83). In: Entscheidungen des Bundesverfassungsgerichts. Hrsg. von den Mitgliedern des Bundesverfassungsgerichts. 65. Bd. Tübingen 1984. S. 1.
- 8 Bundesarchivgesetz (BArchG) vom 6. Januar 1988 (BGBl. I S. 62), zuletzt geändert am 13. März 1992 (BGBl. I S. 506).
- 9 Mühlbauer, Holger: Kontinuitäten und Brüche in der Entwicklung des deutschen Einwohnerwesens. Historisch-juristische Untersuchung am Beispiel Berlins. Europäische Hochschulschriften, Reihe III, Band 635. Frankfurt am Main u. a. 1995, S. 156.
- 10 Rathje, Ulf: Der „Datenspeicher Gesellschaftliches Arbeitsvermögen“ der DDR. In: Historical Social Research/Historische Sozialforschung (HSR), Vol. 21 (1996), Heft 2, S. 113-118.
- 11 Staatssekretariat für Arbeit und Löhne: Datenspeicher Gesellschaftliches Arbeitsvermögen. Angebotskatalog. August 1986.
- 12 Kinze, Knop, Seifert: Volkswirtschaftsplanung, Verlag Die Wirtschaft. Berlin 1977, S. 75.
- 13 Salomon, Jürgen: Probleme der Genauigkeit bei der Massenverarbeitung unter besonderer Berücksichtigung der Fehlerbereinigung und der Fortschreibung: dargestellt am Beispiel des Projektes „Gesellschaftliches Arbeitsvermögen“. Berlin, Humboldt-Universität. Diss. 1981, S. 4ff.
- 14 Programm der SED, Dietz-Verlag Berlin, 1976, S. 27.
- 15 Dietz, Friedo und Helmut Rudolph: Berufstätigenerhebung und der Datenspeicher „Gesellschaftliches Arbeitsvermögen“. Statistische Grundlagen zu wichtigen Strukturen der Erwerbstätigen in der vormaligen DDR. In: Mitteilungen aus der Arbeitsmarkt- und Berufsforschung, 23. Jg. (1990), S. 511-518.