

Inhaltsübersicht

Abkürzungsverzeichnis	11
A Kurzfassung	15
A.1 Wichtige Ergebnisse der Bestandsaufnahme	15
A.2 Empfehlungen der Kommission	25
B Einleitung	37
B.1 Aufgabe der Kommission	37
B.2 Aufbau des Gutachtens	39
C Probleme der gegenwärtigen Situation	45
C.1 Bedarf der Gesellschaft an zugänglichen statistisch-wissenschaftlichen Informationen	45
C.2 Übergreifende Darstellung der informationellen Infrastruktur in Deutschland	83
C.3 Kooperation und Kommunikation	137
C.4 Zugang der Wissenschaft zu Mikrodaten	149
D Lösungsmöglichkeiten für eine bessere informationelle Infrastruktur zwischen Wissenschaft und Statistik – Beispielhafte Lösungen in anderen Ländern	195
D.1 Internationaler Vergleich der Kooperation und Kommunikation	195
D.2 Situation der Fächer Statistik, Ökonometrie und sozialwissenschaftliche Methodenlehre an den Universitäten in anderen Ländern	207
D.3 Situation des Zugangs zu Mikrodaten in ausgewählten Ländern	209
E Empfehlungen der Kommission	247
E.1 Notwendigkeit einer besseren Kooperation von Wissenschaft und Statistik für Datennutzer und Datenproduzenten	247
E.2 Empfehlungen zur Mitwirkung der Wissenschaft bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen	251
E.3 Empfehlungen zur Fortführung und zum Ausbau wichtiger Statistiken	255
E.4 Empfehlungen zur Förderung der Forschung zur Datenerhebung, -aufbereitung und -archivierung	257
E.5 Empfehlungen zur Aus- und Weiterbildung	259
E.6 Empfehlungen zum Datenzugang	263
E.7 Empfehlungen zur Datenverknüpfung	283
E.8 Empfehlungen zum gesetzlichen Rahmen	285
E.9 Institutionelle Schlussfolgerungen	289

Inhaltsverzeichnis

Abkürzungsverzeichnis	11
A Kurzfassung	15
A.1 Wichtige Ergebnisse der Bestandsaufnahme	15
A.2 Empfehlungen der Kommission	25
B Einleitung	37
B.1 Aufgabe der Kommission	37
B.2 Aufbau des Gutachtens	39
C Probleme der gegenwärtigen Situation	45
C.1 Bedarf der Gesellschaft an zugänglichen statistisch-wissenschaftlichen Informationen	45
C.1.1 Bestimmungsfaktoren des Bedarfs an Informationen	45
C.1.2 Die Deckung des Datenbedarfs in den verschiedenen Informationsfeldern	52
C.1.2.1 Das Informationsfeld Gesundheits- und Sozialwesen als Beispiel	53
C.1.2.2 Andere Informationsfelder	64
C.2 Übergreifende Darstellung der informationellen Infrastruktur in Deutschland	83
C.2.1 Amtliche Statistik	83
C.2.1.1 Bundesstatistik	83
C.2.1.2 Deutsche Bundesbank	90
C.2.1.3 Landesstatistik	90
C.2.1.4 Kommunalstatistik	92
C.2.2 Prozessproduzierte Statistiken aus Verwaltungen	93
C.2.2.1 Arbeitsmarkt	93
C.2.2.2 Alterssicherung	95
C.2.2.3 Bundesamt für Güterverkehr	98
C.2.2.4 Kraftfahrt-Bundesamt	99
C.2.2.5 Gesundheitswesen	99
C.2.2.6 Weitere öffentliche Datenproduzenten	100
C.2.3 Ressortstatistik	100
C.2.4 Wirtschaftswissenschaftliche Einrichtungen der Datenproduktion	102
C.2.4.1 Institut der deutschen Wirtschaft (IW), Köln	102
C.2.4.2 Wirtschafts- und Sozialwissenschaftliches Institut (WSI) in der Hans-Böckler-Stiftung	103
C.2.4.3 DIW Berlin	103
C.2.4.4 Zentrum für Europäische Wirtschaftsforschung (ZEW), Mannheim	107
C.2.4.5 Institut für Arbeitsmarkt- und Berufsforschung (IAB)	107
C.2.4.6 Institut für Wirtschaftsforschung Halle (IWH)	108
C.2.4.7 Institut für Wirtschaftsforschung, Hamburg (HWWA)	109

C.2.4.8	Institut für Mittelstandsforschung (IfM), Bonn	109
C.2.4.9	Institut für Handelsforschung an der Universität Köln (IFH)	109
C.2.4.10	ifo Institut für Wirtschaftsforschung, München	109
C.2.5	Sozialwissenschaftliche Einrichtungen der Datenproduktion	110
C.2.6	Interdisziplinäre Datenproduktion	128
C.2.7	Privatwirtschaftlich verfasste Markt- und Sozialforschung	128
C.2.8	Kommerzielle Informationsanbieter	130
C.3	Kooperation und Kommunikation	137
C.3.1	Verflechtungen der amtlichen und der nichtamtlichen Statistik sowie der Ressortstatistik mit der Wissenschaft	137
C.3.2	Aufstellung von Erhebungs- und Aufbereitungsprogrammen	139
C.3.3	Methodenwahl	143
C.3.4	Datenerhebung, -aufbereitung und -archivierung	144
C.3.5	Ausbildungsdefizite in und außerhalb von Universitäten	146
C.4	Zugang der Wissenschaft zu Mikrodaten	149
C.4.1	Rechtliche Rahmenbedingungen in Deutschland	149
C.4.2	Organisatorisch-technische Ausgestaltung und faktische Situation des Zugangs zu aggregierten Daten in Deutschland	159
C.4.3	Organisatorisch-technische Ausgestaltung und faktische Situation des Zugangs zu Mikrodaten in Deutschland	161
C.4.3.1	Zugang zu Mikrodaten der amtlichen Statistik	161
C.4.3.2	Zugang zu Mikrodaten in der nicht-kommerziellen Forschung	183
C.4.3.3	Nutzung kommerzieller Daten und ihre Kosten	189
C.4.4	Faktische Situation der Datenverknüpfung in Deutschland	193
D	Lösungsmöglichkeiten für eine bessere informationelle Infrastruktur zwischen Wissenschaft und Statistik – Beispielhafte Lösungen in anderen Ländern	195
D.1	Internationaler Vergleich der Kooperation und Kommunikation	195
D.2	Situation der Fächer Statistik, Ökonometrie und sozialwissenschaft- liche Methodenlehre an den Universitäten in anderen Ländern	207
D.3	Situation des Zugangs zu Mikrodaten in ausgewählten Ländern	209
D.3.1	Problemstellung	209
D.3.2	Situation des Datenzugangs in den einzelnen Ländern	211
D.3.2.1	Frankreich	211
D.3.2.2	Vereinigtes Königreich	214
D.3.2.3	Kanada	218
D.3.2.4	Niederlande	221
D.3.2.5	Norwegen	224
D.3.2.6	Österreich	227
D.3.2.7	USA	229
D.3.2.8	Dänemark	233
D.3.2.9	Schweden	235

D.3.3	Zusammenfassung der verschiedenen institutionellen Lösungen und Modelle	238
D.3.4	Family and Fertility Survey (FFS)	242
D.3.5	Luxembourg Income Study-Projekt (LIS)	243
D.3.6	Datenverknüpfung	245
E	Empfehlungen der Kommission	247
E.1	Notwendigkeit einer besseren Kooperation von Wissenschaft und Statistik für Datennutzer und Datenproduzenten	247
E.2	Empfehlungen zur Mitwirkung der Wissenschaft bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen	251
E.3	Empfehlungen zur Fortführung und zum Ausbau wichtiger Statistiken	255
E.4	Empfehlungen zur Förderung der Forschung zur Datenerhebung, -aufbereitung und -archivierung	257
E.5	Empfehlungen zur Aus- und Weiterbildung	259
E.6	Empfehlungen zum Datenzugang	263
E.6.1	Wirtschaftliche Gesichtspunkte des Datenzugangs	263
E.6.2	Zugang zu aggregierten Daten	264
E.6.3	Zugang zu Mikrodaten	264
E.6.3.1	Verpflichtung zur Datenweitergabe	265
E.6.3.2	Datenschutz	266
E.6.3.3	Nutzung internationaler Mikrodaten	275
E.6.3.4	Bedarf an Serviceleistungen und Vermittlungsstelle für Mikrodaten	278
E.7	Empfehlungen zur Datenverknüpfung	283
E.8	Empfehlungen zum gesetzlichen Rahmen	285
E.8.1	Forschungsdatengeheimnis	285
E.8.2	Forschungsdatenzentren	286
E.8.3	Zertifizierung und Lizenzierung	286
E.8.4	Datenzugang und -verarbeitung	286
E.8.5	Die gesetzliche Verankerung des Forschungsauftrags für die amtliche Statistik	287
E.8.6	Zugangsregelungen für Daten in Forschungsdatenzentren	287
E.8.7	Verfahren	287
E.9	Institutionelle Schlussfolgerungen	289
	Anhang: Entwicklungsaufgaben zur Verbesserung des Datenzugangs bei ausgewählten Datenbasen	295

Abkürzungsverzeichnis

aba	Arbeitsgemeinschaft für betriebliche Altersversorgung e.V.
Abs.	Absatz
ABV	Arbeitsgemeinschaft berufsständischer Versorgungseinrichtungen e.V.
ADM	Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.
AentG	Arbeitnehmer-Entsendegesetz
AKA	Arbeitsgemeinschaft kommunale und kirchliche Altersversorgung e.V.
ALLBUS	Allgemeine Bevölkerungsumfrage der Sozialwissenschaften
APDU	Association of Public Data Users
ARD	Arbeitsgemeinschaft der öffentlich-rechtlichen Rundfunkanstalten der Bundesrepublik Deutschland
ASA	American Statistical Association
ASI	Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V.
ASID	Alterssicherung in Deutschland
ASR	American Sociological Review
AÜG	Arbeitnehmerüberlassungsgesetz
AVID	Altersvorsorge in Deutschland
AZ	Arbeitsstättenzählung
AZR	Ausländerzentralregister
BA	Bundesanstalt für Arbeit
BAFl	Bundesamt für die Anerkennung ausländischer Flüchtlinge
BAG	Bundesamt für Güterverkehr
BBR	Bundesamt für Bauwesen und Raumordnung
Bd.	Band
BDSG	Bundesdatenschutzgesetz
BGBI.	Bundesgesetzblatt
BHPS	British Household Panel Study
BIJU	Bildungsverläufe und psycho-soziale Entwicklung im Jugendalter
BKGG	Bundeskindergeldgesetz
BLS	Bureau of Labor Statistics
BMA	Bundesministerium für Arbeit und Sozialordnung
BMBF	Bundesministerium für Bildung und Forschung
BMVBW	Bundesministerium für Verkehr, Bau- und Wohnungswesen
BMWi	Bundesministerium für Wirtschaft und Technologie
BOC	Bureau of the Census
BStatG	Bundesstatistikgesetz
BVerfG	Bundesverfassungsgericht
BVerfGE	Bundesverfassungsgerichtsentscheidung
bzw.	beziehungsweise
CASS	Centre for Applied Social Surveys
CBS	Centraal Bureau voor de Statistiek (Statistics Netherlands)
CCS	Central Commission for Statistics
CEPS/INSTEAD	Centre d'Études de Populations, de Pauvreté et de Politiques Socio-Économiques/Center for Population, Poverty, and Policy Studies

CIDSP/BDSP	Centre d'Information des Données Socio-Politiques/Banque de Données Socio-Politiques
CLS	Centre for Labour Market and Social Research
CNIL	Commission National de l'Informatique et des Libertés
CNIS	Conseil National d'Information de Statistique
CNRS	Centre National de la Recherche
CPS	Current Population Survey
CREST	Centre de Recherche en Économie et Statistique
DDI	Data Documentation Initiative
DELTA	Département et laboratoire d'économie théorique et appliquée
DFG	Deutsche Forschungsgemeinschaft
d.h.	das heißt
DIPF	Deutsches Institut für Internationale Pädagogische Forschung, Frankfurt am Main
DIW	Deutsches Institut für Wirtschaftsforschung Berlin
DLI	Data Liberation Initiative
DS	Danmarks Statistik
ECHP	European Community Household Panel
ENSAE	École Nationale de la Statistique et de l'Administration Économique
ENSAI	École Nationale de la Statistique et de l'Analyse de l'Information
ESI	Economic Survey International
ESRC	Economic and Social Research Council
ESVG	Europäisches System Volkswirtschaftlicher Gesamtrechnungen
EU	Europäische Union
Eurostat	Statistisches Amt der Europäischen Gemeinschaften
EVS	Einkommens- und Verbrauchsstichprobe
FFS	Family and Fertility Survey
FPStatG	Finanz-Personal-Statistikgesetz
FTC	Federal Trade Commission
GBE	Gesundheitsberichterstattung des Bundes
GESIS	Gesellschaft sozialwissenschaftlicher Infrastruktureinrichtungen e.V.
GfK	Gesellschaft für Konsumforschung, Nürnberg
GG	Grundgesetz
ggf.	gegebenenfalls
GIS	geographische Informationssysteme
GSS	General Social Survey
GWSt	Gebäude- und Wohnungsstichprobe
GWZ	Gebäude- und Wohnungszählung
HIS	Hochschul-Informationssystem
HWWA	Hamburgisches Welt-Wirtschafts-Archiv
IAB	Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit
ICPSR	Inter-university Consortium for Political and Social Research
IDG	Informations- und Dokumentationszentrum „Gesundheitsdaten“
IEA	International Energy Agency

IFH	Institut für Handelsforschung an der Universität Köln
IfM	Institut für Mittelstandsforschung, Bonn
IMF	International Monetary Fund
INSEE	Institut National de la Statistique et des Études Économiques
IRESO	Institut des Recherches sur les Sociétés Contemporaines
IS-GBE	Informationssystem der Gesundheitsberichterstattung des Bundes
ISSP	International Social Survey Program
IW	Institut der deutschen Wirtschaft, Köln
IWH	Institut für Wirtschaftsforschung Halle
IZ	Informationszentrum Sozialwissenschaften
KBA	Kraftfahrt-Bundesamt
KOSTAT-DST	Kommunalstatistik Deutscher Städtetag
LASMAS-IdL	Laboratoire d'Analyse Secondaire et de Méthodes appliquées à la Sociologie – Institut du Longitudinal
LECEPE	Centre de Formation
LEED	Longitudinal Employee-Employer Data
LeLa	Lebenslagenumfrage
LES	Luxembourg Employment Study
LIAB	Linked-employer-employee-Datensatz
LIS	Luxembourg Income Study
MIMAS	Manchester Information & Associated Services
MIP	Mannheimer Innovationspanel
MPG	Max-Planck-Gesellschaft
MPI	Max-Planck-Institut
MPLV	Max-Planck-Lebensverlaufsstudie
MUP	Mannheimer Unternehmenspanel
MZ	Mikrozensus
MZES	Mannheimer Zentrum für Europäische Sozialforschung
NCES	National Center for Education Statistics
NCHS	National Center for Health Statistics
N.F.	Neue Folge
NIEP	Niedrigeinkommenspanel
NSD	Norsk samfunnsvitenskapelig datatjeneste (Norwegian Social Science Data Services)
NSF	National Science Foundation
NWO	Nederlandse Organisatie voor Wetenschappelijk Onderzoek
OCG	Occupational Change in one Generation
OECD	Organisation for Economic Co-Operation and Development
ONS	Office for National Statistics
PSID	Panel Study of Income Dynamics
RückHG	Rückkehrhilfegesetz
S.	Satz/Seite
SCB	Statistiska centralbyrån (Statistics Sweden)
SchwBG	Schwerbehindertengesetz
Sfb 3	Sonderforschungsbereich 3
SGB	Sozialgesetzbuch
SOEP	Sozio-oekonomisches Panel

SPES	Sozialpolitisches Indikatoren- und Entscheidungssystem
SSA	Social Security Administration
SSB	Statistisk Sentralbyrå (Statistics Norway)
SSHRC	Social Sciences and Humanities Research Council
STADA	Statistische Datenbank der Bundesanstalt für Arbeit
STATIS-BUND	Statistisches Informationssystem des Bundes
StGB	Strafgesetzbuch
u.a.	unter anderem
UNIDO	United Nations Industrial Development Organisation
UNSD	United Nations Statistics Division
usw.	und so weiter
u.U.	unter Umständen
VASMA	Vergleichende Analysen der Sozialstruktur mit Massendaten
VBL	Versorgungsanstalt des Bundes und der Länder
VDR	Verband Deutscher Rentenversicherungsträger
VGR	Volkswirtschaftliche Gesamtrechnung
WGL	Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz
WISDOM	Wiener Institut für Sozialwissenschaftliche Dokumentation und Methodik
WSA	Wetenschappelijk Statistisch Agentschap
WSI	Wirtschafts- und Sozialwissenschaftliches Institut in der Hans-Böckler-Stiftung
WZB	Wissenschaftszentrum Berlin für Sozialforschung
ZA	Zentralarchiv für empirische Sozialforschung, Köln
z.B.	zum Beispiel
ZDF	Zweites Deutsches Fernsehen
ZEW	Zentrum für Europäische Wirtschaftsforschung, Mannheim
ZUMA	Zentrum für Umfragen, Methoden und Analysen, Mannheim

A **Kurzfassung**

A.1 Wichtige Ergebnisse der Bestandsaufnahme

Aufgabe der Kommission

Nur auf der Grundlage gültiger und verlässlicher Daten können die Sozial- und Wirtschaftswissenschaften realitätsgerechte Untersuchungen durchführen und überzeugende Handlungsempfehlungen entwickeln. Die Verfügbarkeit und Qualität von Daten sind dafür eine zentrale Voraussetzung. Die Leistungsfähigkeit der Dateninfrastruktur ist damit eine entscheidende Grundlage für die Leistungsfähigkeit einer Gesellschaft, aber auch für eine im internationalen Maßstab innovationsfähige sozial- und wirtschaftswissenschaftliche Forschung.

Mit statistischen Erhebungen sind Kosten und Belastungen der Bürgerinnen und Bürger sowie der Firmen verbunden. Deswegen müssen vorhandene Daten, auch solche, die für andere Zwecke im Rahmen staatlichen Handelns entstehen, so effizient wie möglich genutzt werden.

Die Art der Nutzung statistischer Daten hat sich in der Informationsgesellschaft verändert. Wirtschaftlicher und sozialer Wandel kann nur untersucht werden, wenn Informationen auf der Mikroebene von Beobachtungseinheiten (Personen, Haushalte, Unternehmen) zu verschiedenen Zeitpunkten für statistische Zwecke verfügbar und zugänglich sind. Moderne statistische Methoden erfordern den Zugang zu Mikrodaten, um Verhaltensregelmäßigkeiten und -zusammenhänge aufzudecken.

Die Bundesministerin für Bildung und Forschung hat in ihrer Verantwortung für die Wissenschaft und die wissenschaftliche Infrastruktur eine Kommission gebeten zu klären, ob die informationelle Infrastruktur der Bundesrepublik Deutschland dieser veränderten Aufgabenstellung noch gerecht wird, und sie beauftragt, Vorschläge zur Verbesserung der informationellen Infrastruktur zwischen Wissenschaft und Statistik zu erarbeiten. Der Auftrag der Kommission zielt auf die generelle Verbesserung der Kooperation zwischen Statistik und Wissenschaft. Er schließt alle Datenquellen ein und soll vor allem internationale Erfahrungen nutzen und entsprechende Beispiele guter Praxis auf ihre Übertragbarkeit prüfen.

Bedarf der Gesellschaft an zugänglichen statistisch-wissenschaftlichen Informationen

Betrachtet man im internationalen Vergleich die besten Statistiken und die besten statistischen Analyseverfahren, so fallen für die Bundesrepublik Deutschland eine Reihe von Verbesserungsmöglichkeiten der informationellen Infrastruktur auf:

- In Bereichen von Wirtschaft und Gesellschaft, die einem raschen Wandel unterworfen sind, gibt es gravierende Informationslücken.
- Die wiederholte Befragung derselben Erhebungseinheiten durch die amtliche Statistik wird nur in wenigen Bereichen genutzt. Dies führt zu einem Defizit an Längsschnittdaten und damit möglichen dynamischen Analysen.
- Wissenschaftsgetragene Längsschnittstudien stehen auf einer fragilen finanziellen und institutionellen Grundlage, während diese in anderen Ländern auf Dauer öffentlich finanziert werden.
- Registerdaten werden überwiegend nur unzureichend differenziert statistisch-wis-

senschaftlich analysiert, sie sind in der Regel für die Wissenschaft nicht zugänglich. Die statistische Verknüpfung der Daten verschiedener Register ist nur schwer oder gar nicht möglich.

- Viele relevante Fragestellungen erfordern den Zugang zu Mikrodaten, der aber für zahlreiche Datenbestände nicht befriedigend gegeben ist.

Zur Deckung des Datenbedarfs in den verschiedenen Informationsfeldern

Die Kommission hat sich im Rahmen ihrer Bestandsaufnahme mit der Deckung des Datenbedarfs in den verschiedenen Informationsfeldern beschäftigt. Sie hat zu diesem Zweck Expertisen zu den verschiedenen Informationsfeldern vergeben und die dort enthaltenen Wahrnehmungen und Bewertungen von Datendefiziten und Datenbedarf ausgewertet (Abschnitt C.1.2). Untersucht wurden die folgenden *Informationsfelder*: Bevölkerung; Staat und Politik; Erwerbstätigkeit; Einkommen und Vermögen; Gesundheits- und Sozialwesen; Konsum und Wohnen; Bildung und Kultur; Wirtschaft und Verkehr; Umwelt und Energie; Wissenschaft und Forschung; subjektive Einstellungen.

Diese Expertisen zeigen ein außerordentlich umfangreiches und vielfältiges Informationsangebot. Es ist aber einerseits so unübersichtlich, dass sogar ausgewiesene Expertinnen und Experten Schwierigkeiten haben, in ihren eigenen Bereichen den Überblick zu behalten, andererseits bestehen erhebliche Informationsdefizite.

Übergreifende Darstellung der informationellen Infrastruktur in Deutschland

Das Gutachten gibt einen Überblick über die wichtigsten Anbieter statistischer Informationen in Deutschland. Nach einer kurzen Darstellung der Informationsanbieter wird jeweils der vorhandene Datenbestand und die Weitergabe von anonymisierten Einzeldaten beschrieben.

Die öffentlich finanzierte statistische Infrastruktur reicht weit über die amtliche Statistik des Statistischen Bundesamtes, der Statistischen Landesämter und der kommunalstatistischen Ämter hinaus. Statistische Daten werden auch von Bundes- und Landesministerien im Rahmen der Ressortforschung und von anderen staatlichen Stellen sowie von parastaatlichen Einrichtungen (z.B. Sozialversicherungen) erhoben und ausgewertet (amtliche originärstatistische Daten). Außerdem entstehen statistische Datenbestände im Zuge des Verwaltungsablaufs (amtliche prozessproduzierte Daten). Auch die Einrichtungen der staatlich finanzierten Sozial- und Wirtschaftswissenschaften (Universitäten, Fachhochschulen, Max-Planck-Institute [MPI], Institute der Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz [WGL], andere staatlich finanzierte Forschungsinstitute) führen vielfältige Umfragen durch, die den Gesamtbestand an verfügbaren empirischen Daten vergrößern (wissenschaftsgetragene originärstatistische Daten). Daneben werden privatwirtschaftlich verfasste Markt- und Sozialforschung sowie kommerzielle Informationsanbieter einbezogen.

Kooperation und Kommunikation zwischen der amtlichen und der nichtamtlichen Statistik sowie der Ressortstatistik mit der Wissenschaft

Das Konglomerat verschiedener Datenproduzenten und Datenbestände ist historisch gewachsen. Eine koordinierte Festlegung der Erhebungs- und Aufbereitungsprogramme, die allgemeine Zugänglichkeit der Datenbestände und eine systematische Archivierung ist bisher nicht gegeben.

Kooperation und Kommunikation zwischen Wissenschaft und amtlicher Statistik sind zwar auf Seiten der amtlichen Statistik teilweise institutionalisiert – insbesondere durch den Statistischen Beirat beim Statistischen Bundesamt –, nicht jedoch auf Seiten der Wissenschaft, die kein ausreichendes Legitimationsverfahren für die institutionelle Mitsprache bei der Gestaltung der amtlichen Statistik entwickelt hat. Regelmäßige Möglichkeiten der Einflussnahme durch die Wissenschaft auf die Ressorterhebungen und die Gestaltung der prozessproduzierten Daten fehlen fast völlig. Das statistische Selbstverständnis öffentlich-rechtlicher Körperschaften, die faktisch zur statistischen Infrastruktur gehören, aber andere Schwerpunktaufgaben haben, ist oft ungeklärt (z.B. Bundesbank, Umweltbundesamt, Bundesanstalt für Arbeit [BA] oder die öffentlich-rechtlichen Rundfunkanstalten). Die Datenarchivierung ist zwar breit gestreut, aber lückenhaft und ohne geschlossenes Konzept.

Die Vorteile einer einheitlichen und systematisch abgestimmten Konzeption der öffentlich finanzierten Datenerhebung, -bereitstellung und -archivierung sind um so größer und die unkoordinierte Durchführung von Doppelerhebungen ist um so leichter vermeidbar, je breiter, billiger und schneller vorhandene Datenbestände – unter Beachtung des Datenschutzes für Einzeldaten – für Wissenschaftler¹ außerhalb der datenproduzierenden Einrichtung zugänglich sind.

Die Kommission stellt würdigend fest, dass sich die Zusammenarbeit zwischen der amtlichen Statistik und der Wissenschaft an Hochschulen und Forschungseinrichtungen in den letzten Jahren bereits deutlich verbessert hat. Dies kommt in vielfältigen Aktivitäten zum Ausdruck: Zu nennen sind hier insbesondere die seit 1983 von Statistischem Bundesamt, dem Wissenschaftszentrum Berlin für Sozialforschung (WZB) und dem Zentrum für Umfragen, Methoden und Analysen, Mannheim (ZUMA) gemeinsam erarbeitete Publikation eines umfangreichen Datenreports, die Arbeiten der Abteilung Mikrodaten bei ZUMA, die 1999 erfolgte Neugründung eines Gesprächskreises mit Vertretern der Mitgliedsinstitute deutscher wirtschaftswissenschaftlicher Forschungsinstitute, die Veranstaltung von Kolloquien und Workshops gemeinsam mit der Deutschen Statistischen Gesellschaft, mit der Arbeitsgemeinschaft Sozialwissenschaftlicher Institute e.V. (ASI) und dem Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V. (ADM) und mit einzelnen Wissenschaftlern.

Verfahren der Datenweitergabe an die Wissenschaft außerhalb der amtlichen Statistik sind durch die Bereitstellung von faktisch anonymisierten Mikrodatenfiles aus dem Mikrozensus (MZ), der Einkommens- und Verbrauchsstichprobe (EVS), der Beschäftigtenstichprobe, der Zeitbudgeterhebung und des Europäischen Haushaltspanels erprobt worden. Praktika und Forschungsaufenthalte von außenstehenden Wissenschaftlern im Statistischen Bundesamt und beim Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) (zum Teil mit Hilfe von Pilotprojekten des Bundesministeriums für Bildung und Forschung [BMBF] und einer Förderung durch die Deutsche Forschungsgemeinschaft [DFG]) waren weitere begrüßenswerte Schritte zu einer besseren Vernetzung von amtlicher Statistik und Wissenschaft.

¹ Wenn im Folgenden von Wissenschaftler, Forscher, Professor, Mitarbeiter usw. gesprochen wird, ist stets auch die weibliche Form des Wortes gemeint.

Aufstellung von Erhebungs- und Aufbereitungsprogrammen

Die *Entscheidung*, welche amtlichen Statistiken erhoben werden, liegt – im Gegensatz zu den meisten anderen Industrieländern – nicht bei der Regierung oder bei den Statistikämtern, sondern beim Gesetzgeber. Dies erschwert es nicht nur der amtlichen Statistik, die Statistiken an veränderte Bedingungen anzupassen, sondern hat in der Vergangenheit auch dazu geführt, dass wissenschaftlicher Sachverstand nicht in einem optimalen Ausmaß in die Gestaltung von Erhebungsprogrammen einfließt.

Archivierung

Ein für die Wissenschaft besonders bedeutsames Problem ist das der dauerhaften Archivierung von Daten. Historische Statistikdaten sind oftmals nicht mehr lesbar (aus physikalischen Gründen oder aus Softwaregründen). Die Archivierung der ersten EVS in einem universitären Forschungsprojekt (Universität Frankfurt) zeigt, dass die Langfrist-Archivierung bislang von (glücklichen) Zufälligkeiten gekennzeichnet ist.

Ausbildungsdefizite in und außerhalb von Universitäten

Sorge bereitet eine Reihe von Defiziten in der Ausbildung, die zu geringer bzw. nicht sachgerechter Nutzung der bereits vorhandenen Infrastruktur führen. Tragfähige empirische Analysen erfordern die Verbindung von substanzwissenschaftlicher Fachkenntnis, genauem Verständnis der Erhebung und Aufbereitung von Daten und problemgerechter logischer und statistischer bzw. ökonometrischer Analyse. Die Grundkenntnisse der Studierenden in den Bereichen deskriptive Statistik, einfache statistische Modellierung, einfache Inferenzstatistik und Ökonometrie sowie in Grundzügen der angewandten Informatik, insbesondere bei Datenbanksystemen sind an vielen Hochschulen mangelhaft. Darüber hinaus fehlt vielfach die Kenntnis von Standardauswertungssystemen der Statistik.

Viele Lehrende in den Bereichen Volks- und Betriebswirtschaft, Sozialwissenschaft, aber auch in mathematischer Statistik und Ökonometrie haben selbst keinen Bezug zu empirischen Fragestellungen und können daher entsprechende Arbeiten nicht anleiten. In den volks- und betriebswirtschaftlichen sowie in den sozialwissenschaftlichen Fakultäten werden Stellen für Statistik, Ökonometrie sowie empirische Sozial- bzw. Wirtschaftsforschung abgebaut, in geringer dotierte C3-Positionen eingestuft bzw. mit einer sehr geringen Zahl von Mitarbeiterinnen und Mitarbeitern bzw. Hilfskräften ausgestattet. Dies verschärft die oben beschriebenen Probleme.

Die Erhebung, Aufbereitung, Kodierung und Transformation von Daten wird insbesondere in den Wirtschaftswissenschaften nicht systematisch unterrichtet. Bereits vorhandene Datenquellen wie die Bezugsmöglichkeit von Daten des Statistischen Bundesamtes oder der Daten des Zentralarchivs für empirische Sozialforschung (ZA), Köln werden nicht ausreichend genutzt. Darüber hinaus besteht eine mangelnde Verfügbarkeit von Public Use Micro Data Files, mit denen Studierende an die Nutzung solcher Daten herangeführt werden können.

Zugang der Wissenschaft zu (Mikro-)Daten

Rechtliche Rahmenbedingungen in Deutschland

Zum Regelungsinstrumentarium des Datenschutzes auf *internationaler* Ebene gehören z.B. die Konvention des Europarates und auf *Gemeinschaftsebene* die Datenschutzrichtlinie der Europäischen Union (EU), auf *nationaler* Ebene das Bundesdatenschutzgesetz (BDSG) und die Landesdatenschutzgesetze.

Zu den wichtigsten Grundprinzipien, die auf die besonderen Rahmenbedingungen von Wissenschaft und Forschung eingehen, gehören das *Einwilligungsprinzip* und das *Zweckbindungsprinzip*.

Das *Einwilligungsprinzip* bestimmt, dass personenbezogene Daten nicht ohne Einwilligung der Betroffenen be- oder verarbeitet werden dürfen, es sei denn, die Verarbeitung erfolgt aufgrund und im Rahmen einer Rechtsnorm.

Das *Zweckbindungsprinzip* verlangt, dass personenbezogene Daten nur im Rahmen vorher bestimmter (und rechtlich zulässiger) Zwecke verarbeitet werden dürfen.

Diese Regelungsprinzipien können mit einzelnen Forschungsdesigns, aber auch mit allgemeinen Forschungsmethoden kollidieren. Beide Prinzipien kennen Ausnahmen. Den spezifischen Interessen von Wissenschaft dienen *drei Ansätze*:

- die Modifikation des Personenbezugs,
- die Modifikation des Einwilligungsprinzips und
- die Modifikation des Zweckbindungsprinzips.

Der *Personenbezug* kann durch Anonymisierung beseitigt werden. Aufgrund des Zusatzwissens kann allerdings eine vollständige Anonymisierung kaum erreicht werden. Ergibt jedoch eine Risikobeurteilung ein geringes Risiko der Re-Identifikation, so entfällt das normative Merkmal „Personenbezug“, wenn zugleich weitere dieses Risiko absichernde Maßnahmen ergriffen werden. Der Personenbezug kann auch durch „Treuhandlöserungen“ entflechtet werden.

Eine *Einwilligung* ist nicht erforderlich, wenn (u.a.) eine gesetzliche Regelung die Datenverarbeitung erlaubt.

Einige Datenschutzgesetze benennen wissenschaftliche Forschung als eigenständig legitimiert. Solche Lösungen verlangen als Kompensation eine Abwägung mit schutzwürdigen Interessen der Betroffenen.

Die bereichsspezifischen Datenschutzregelungen der statistischen Geheimhaltung in der Bundesrepublik sind sowohl europarechtlich abgesichert wie auch durch das Volkszählungsurteil des Bundesverfassungsgerichts (BVerfG) (BVerfGE 65, 1 ff) weitgehend determiniert. Danach gilt:

- a) Die Möglichkeiten der Übermittlung müssen gesetzlich angeordnet sein.
- b) Die Einzelangaben müssen faktisch anonymisiert sein.
- c) Die Empfänger der Daten müssen Amtsträger, für den öffentlichen Dienst besonders Verpflichtete oder Verpflichtete nach dem Verpflichtungsgesetz sein.
- d) Die Übermittlungsadressaten müssen in Hochschulen oder anderen Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung inkorporiert sein.
- e) Die Einzelangaben sind zu löschen, sobald das wissenschaftliche Vorhaben durchgeführt ist (§ 16 Abs. 8 S. 2 Bundesstatistikgesetz [BStatG]).

Diesen Grundsätzen entspricht die Regelung des § 16 Abs. 6 und 7 des BStatG vom 22. Januar 1987.

In der Praxis der Kooperation zwischen amtlicher Statistik und Wissenschaft haben sich hierbei im Hinblick auf die Bereitstellung von Mikrodaten folgende Strategien entwickelt:

1. die Entwicklung faktisch anonymisierter Scientific Use Files,
2. Methodenintegration statt Datenübermittlung,
 - a) durch Inkorporation der Forscher in das Statistische System,
 - b) durch „Fernrechnen“ mit Einzelangaben nach einem getesteten Methodendesign des Forschers.

Die grundsätzlich mögliche Lösung einer Datenübermittlung aufgrund von Einwilligungen hat in der Praxis bisher keine wesentliche Rolle gespielt.

Darüber hinaus umfasst die bereichsspezifische Datenschutzregelung des Statistikgeheimnisses alle Einzelangaben, nicht nur personenbezogene Daten. Dies betrifft vor allem unternehmensbezogene Einzeldaten.

Das Auslegungsprinzip der Einheit der Verfassung verlangt gesetzliche Regelungen, die eine normative Konkretisierung der Abwägung zwischen Datenschutz und Datenzugang der Forschung enthalten. Die Existenz der Kommission selbst belegt, dass trotz erheblicher Fortschritte die Regelungen nicht ausreichen. Dies liegt im Wesentlichen am Regelungsmodell der allgemeinen Datenschutzgesetze in der Bundesrepublik, das als *Verbot mit Erlaubnisvorbehalt* konzipiert ist.

Ein erweiterter Datenzugang für die Forschung wird unter dem Aspekt der Kreierung eines „Forschungsdatengeheimnisses“ rechtspolitisch diskutiert. Dieses müsste über § 16 Abs. 6, 7, 10 BStatG hinaus

- eine institutionelle Abschottung,
- eine beschlagnahmeresistente Datenverarbeitung,
- ein Herausgabe- und Übermittlungsverbot außerhalb dieses Sektors,
- eine eindeutige Definition des Forschungsbereichs sowie
- weitere grundrechtssichernde Maßnahmen, wie Auskunftsrechte der Betroffenen, Codes of Conduct, Codes of Ethics sowie eine unabhängige Kontrolle der Datenverarbeitung garantieren.

Der internationale Datenaustausch – das Safe-Harbor-Prinzip

Es ist ein Grundsatz des internationalen Datenschutzes, dass bei Wirkungsähnlichkeit der Regelungen in den betroffenen Ländern der Datenaustausch zwischen diesen Ländern nicht allein wegen der Grenzüberschreitung zusätzlichen Datenschutzerfordernungen unterworfen werden soll. So gilt für die EU, dass personenbezogene Daten nur in solche Nicht-EU-Länder geliefert werden können, die einen angemessenen Schutz der Privatsphäre gewährleisten.

Die Europäische Kommission hat nach über zweijähriger Verhandlung mit dem US-Handelsministerium unter Beteiligung des Europäischen Parlaments am 26. Juli 2000 eine Entscheidung getroffen, dass eine vom US-Handelsministerium getroffene Regelung, der sog. Safe-Harbor-Mechanismus einen angemessenen Schutz für personenbezogene Daten darstellt, die aus der EU nach den USA übermittelt werden.

Diese Safe-Harbor-Konzeption, die bereits seit dem 1. November implementiert ist, könnte analog ein Modell für den Datenaustausch von Forschungseinrichtungen aus Mitgliedstaaten mit Drittländern wie z.B. USA, Kanada, Australien, Japan sein.

Organisatorisch-technische Ausgestaltung und faktische Situation des Zugangs zu aggregierten Daten in Deutschland

Die statistischen Ämter des Bundes und der Länder bieten im Rahmen ihres gesetzlichen Auftrags zur Veröffentlichung ihrer statistischen Daten und aufgrund ihres Selbstverständnisses, statistische Informationen nutzergerecht zu verbreiten, Daten aller Bundesstatistiken in aggregierter Form auf vielfältige Weise an. Zu unterscheiden sind drei Angebotsformen:

1. Informationelle Grundversorgung als kostenloses Grundangebot;
2. Statistische Produkte und Dienstleistungen;
3. Maßgeschneiderte Lösungen für spezielle Nutzer.

Das letzte Segment umfasst problemorientierte maßgeschneiderte Produkte und Serviceleistungen, die auf spezielle Nachfrage individuell für die Kunden gegen Vollkostenerstattung entwickelt werden.

In der Vergangenheit waren die typischen Nutzer statistischer Daten die Endverbraucher, die Daten zu eigenen Zwecken benötigten. Zunehmend wird das Datenangebot der statistischen Ämter auch für gewerbliche Informationsanbieter attraktiv. So werden die statistischen Daten gebührenpflichtig über Lizenzverträge zur Vermarktung durch Dritte freigegeben.

Organisatorisch-technische Ausgestaltung und faktische Situation des Zugangs zu Mikrodaten in Deutschland

Die Weitergabe von Datensätzen auf der Mikroebene der Beobachtungseinheiten (Personen, Haushalte, Unternehmen) gehört in Deutschland nicht zu den traditionellen Aufgaben der amtlichen Statistik. Der Gesetzgeber hat der Wissenschaft allerdings eine Sonderstellung unter den Datennutzern der amtlichen Statistik eingeräumt, das sogenannte *Wissenschaftsprivileg*.

Dementsprechend werden der Wissenschaft Mikrodatensätze aus amtlichen Statistiken zur Nutzung angeboten. Die Bandbreite der Datennutzungsmöglichkeiten wird anhand von „Grundtypen“ dargestellt.

(1) Wissenschaftler erhalten „faktisch“ anonymisierte Mikrodaten

Seit der Novellierung des BStatG im Jahre 1987 hat die Wissenschaft – anders als die übrigen Nutzer der amtlichen Statistik – neben der Nutzung „absolut“ anonymisierter Einzelangaben auch die Möglichkeit, Daten in „faktisch“ anonymisierter Form zu beziehen (Wissenschaftsprivileg). Der Zugang zu faktisch anonymisierten Einzeldaten ist bereits vielfach ermöglicht worden. Die prozessproduzierten Daten der Verwaltungen unterliegen in vielen Fällen den besonderen Regelungen und den Beschränkungen des Sozialdatenschutzes.

Bei der Bereitstellung von faktisch anonymisierten Mikrodaten erstellen die statistischen Ämter das Material fallweise aufgrund einer spezifischen Datenanfrage als Sonderaufbereitung. Bei großer Nachfrage stellen die statistischen Ämter faktisch anonymisierte Mikrodatenfiles (Scientific Use Files) als Standardprodukte zur Verfügung. In beiden Fällen erhalten die Wissenschaftler die anonymisierten Einzeldaten und können diese für ihr Forschungsvorhaben nach Belieben auswerten.

Faktisch anonymisierte Grundfiles, die Daten aller Bundesländer umfassen, hat das Statistische Bundesamt bisher in Kooperation mit den statistischen Ämtern der Länder für folgende haushalts- und personenbezogenen Statistiken erstellt:

- MZ 1989, 1991, 1993, 1995, 1996, 1997 (1998 für April 2001 geplant);
- EVS 1993, 1998;
- deutsche Stichprobe des Europäischen Haushaltspanels 1994 bis 1996;
- Zeitbudgeterhebung 1991/92.

Da die Anonymisierung von Einzelmaterial im Allgemeinen erheblichen Aufwand verursacht, werden solche Daten nur gegen Kostenerstattung gemäß den jeweils gültigen Kostenregelungen des Statistischen Bundesamtes bereitgestellt. Wissenschaftler in Deutschland können faktisch anonymisierte haushalts- und personenbezogene Statistiken vom Statistischen Bundesamt gegen geringe Gebühren erhalten. Grundlage dafür war ein Pilotprojekt, das in Kooperation mit dem BMBF und der Gesellschaft sozialwissenschaftlicher Infrastruktureinrichtungen (GESIS) durchgeführt wurde. Auch die Bundesanstalt für Arbeit ermöglicht den Zugang zu faktisch anonymisierten Einzeldaten unter den gesetzlichen Voraussetzungen des § 75 Sozialgesetzbuch X (SGB X). Hervorzuheben sind hier insbesondere die 1 %-Stichproben der Beschäftigtenstatistik, die der Forschung in Zusammenarbeit mit dem ZA als anonymisierte Daten zur Verfügung gestellt werden. Darüber hinaus ist die Bereitstellung anonymisierter Daten aus den Primärstatistiken der Bundesanstalt für Arbeit derzeit nur nach Prüfung im Einzelfall im Rahmen ressourcentechnischer Möglichkeiten und gegen Kostenerstattung möglich. Die Deutsche Bundesbank, die neben den annähernd gleichwertigen Vorschriften des Bankgeheimnisses ebenfalls die Regelungen des BStatG zu Einzeldaten anwendet, hat bislang keine solchen Daten zur Durchführung wissenschaftlicher Vorhaben weitergegeben.

(2) Wissenschaftler stellen eigene Analyseprogramme zur Verfügung

Bei dieser Form der Zusammenarbeit erstellen Wissenschaftler zu ihrem Forschungsvorhaben ein Auswertungsprogramm und übermitteln dies (u.U. im Wege der Ferndatenverarbeitung) dem zuständigen Amt, das über die benötigten Einzeldaten verfügt. Das Amt stellt die benötigten internen Einzeldaten in der notwendigen Form bereit und wendet das Programm hierauf an. Die Wissenschaftler haben somit keinen Kontakt mit den geheimhaltungsbedürftigen Einzeldaten, eine Anonymisierung der Einzeldaten wird nicht vorgenommen.

(3) Wissenschaftler analysieren nicht-anonymisierte Einzeldaten

Eine weitere Form der Nutzung der Einzeldatenbestände der amtlichen Statistik ist die Forschung mit *Originaldaten*. Allerdings entsteht dadurch für den Bearbeiter u.U. die Möglichkeit, die Identität der Datengeber relativ leicht identifizieren zu können. Unter den gegebenen rechtlichen Voraussetzungen ist deshalb nach Auffassung der *statistischen Ämter* eine solche Datennutzung nach diesem Verfahren nur möglich, wenn das statistische Amt das Forschungsvorhaben selbst initiiert oder im Einzelfall ein so starkes eigenes Interesse an der Fragestellung des externen Wissenschaftlers hat, dass es sich das Vorhaben zu eigen macht.

Die hier beschriebene Form der Zusammenarbeit geht von den Regelungen des BStatG aus. Dessen Auslegung in den *statistischen Ämtern* ist in diesem Punkt allerdings nicht einheitlich. Nach der mehrheitlichen Auffassung wäre es nur unter anderen

rechtlichen Voraussetzungen möglich, dass Wissenschaftler, die auf die Wahrung des Statistikgeheimnisses verpflichtet wurden, in den statistischen Ämtern mit nicht-anonymisierten Einzeldaten arbeiten, *ohne* dass sie im Auftrag des statistischen Amtes forschen. Ohne eine solche Änderung im BStatG wird eine einheitliche und für die Wissenschaft vorteilhaftere Praxis nicht zu erreichen sein.

(4) Durchführung von Sonderauswertungen im Auftrag der Wissenschaft

Die Sonderauswertung ist sicher nicht die Form der Datennutzung, die externe Wissenschaftler vorrangig anstreben. Sie ist auch keine auf wissenschaftliche Nutzer beschränkte Form der Datennutzung. Dennoch soll auch sie hier erwähnt werden, denn sie kann ein Weg sein, das Informationspotential von Mikrodaten in Fällen zur Beantwortung wissenschaftlicher Fragestellungen zu nutzen, in denen andere Verfahren nicht zum Einsatz kommen können.

Datenzugang zu Mikrodaten in der nicht-kommerziellen Forschung

Die replikativen Datenerhebungsprogramme der sozial- und wirtschaftswissenschaftlichen Forschung folgen dem Prinzip, dass die Daten entweder sofort oder nach ersten Analysen durch die datenerhebenden Einrichtungen und Forscher der wissenschaftlichen Gemeinschaft zur weiteren Nutzung unentgeltlich oder gegen eine geringe Schutzgebühr zur Verfügung gestellt werden. Mit Ausnahme des Sozio-oekonomischen Panels (SOEP) werden die Daten von den Datenproduzenten in aller Regel dem ZA übergeben. Das ZA archiviert die Daten und gibt sie an wissenschaftliche und andere Nutzer weiter. Die Daten des SOEP werden direkt vom Produzenten (beim Deutschen Institut für Wirtschaftsforschung [DIW] Berlin) an Nutzer weitergegeben. Insgesamt hat sich damit in der sozial- und wirtschaftswissenschaftlichen empirischen Forschung das Prinzip weitgehend durchgesetzt, erhobene Daten – teilweise nach einer Schutzfrist für eigene Analysen – anderen Wissenschaftlern in geeigneter Form zugänglich zu machen. Bei vielen Datensätzen führt dies zu einer sehr intensiven Datenverwertung. Es schafft die zentrale Voraussetzung, durch Re-Analysen Ergebnisse einzelner Forscher der datenfundierten Kritik durch andere Forscher auszusetzen, alternative Modelle der Wirklichkeitserklärung zu entwickeln und zu überprüfen sowie vor allem auch sozialen und wirtschaftlichen Wandel im Zeitverlauf abzubilden. Servicefunktionen, wie sie das ZA ausübt, sind dazu unverzichtbar.

Freilich gibt es immer noch Fälle, in denen Wissenschaftler nicht bereit sind, erhobene Daten in diesem Sinne der wissenschaftlichen Gemeinschaft zur Verfügung zu stellen. Deswegen sollte für Wissenschaftler, deren Datenerhebungen durch öffentliche Mittel, insbesondere durch die DFG, gefördert werden, die grundsätzliche Verpflichtung bestehen, die gewonnenen Mikrodaten für Re-Analysen anderen Wissenschaftlern in geeigneter Weise zugänglich zu machen.

Nutzung kommerzieller Daten und ihre Kosten

In den letzten zwei Jahrzehnten hat sich ein bedeutendes privatwirtschaftliches Angebot an statistischen Daten etabliert, das aufgrund seiner Nachfragestruktur überwiegend dem Bereich Wirtschaftsinformationen zuzuordnen ist. Unternehmensberatungen, Marktforschungsinstitute und Unternehmen der Informationsindustrie vermarkten kommerziell selbst produzierte Datenbasen. Die Quellen für die Datenbasen

bilden einerseits die amtlichen Statistiken, andererseits eigene Erhebungen. Die Daten des Statistischen Bundesamtes, der Bundesbank und sonstiger öffentlicher Einrichtungen, die selbst auch als Anbieter ihrer Datenbasen fungieren, werden ggf. durch Homogenisierung (Angleichung unterschiedlich definierter Daten) und Aggregation (Verdichtung) aufbereitet. Die kommerziellen Produzenten und Anbieter orientieren sich am Markt und dem Bedarf ihrer Kunden.

Im Wissenschaftsbereich haben diese Datenbasen bisher eine eher geringe, aber vermutlich wachsende Bedeutung.

Faktische Situation der Datenverknüpfung in Deutschland

Die Verknüpfung von statistischen Daten ist in Deutschland wenig befriedigend gelöst. Zwar ist die Verknüpfung von Daten kein datenschutzrechtliches Problem, wenn jeder einzelne Befragte bzw. „Erhobene“ dieser Verknüpfung zustimmt. Dies ist aber ein weitgehend inpraktikables bzw. ineffizientes Verfahren, da bei expliziter Zustimmungspflicht viele Betroffene nicht zustimmen, obwohl die verknüpften Daten nicht individuell zuzuordnen sind.

Internationaler Vergleich zum Zugang der Wissenschaft zu Mikrodaten der amtlichen Statistik

Die Kommission hat einen breit angelegten internationalen Vergleich durchgeführt, der die gegenwärtige Situation in Frankreich, im Vereinigten Königreich, in Kanada, den Niederlanden, Norwegen, Österreich, den USA, Dänemark und Schweden beschreibt. Für die untersuchten Länder werden jeweils die folgenden Themen diskutiert:

- Für die Forschung verfügbare Mikrodaten der amtlichen Statistik;
- Regelung des Zugangs der Wissenschaft;
- Weitergabeverfahren, Nutzungskonditionen und Gebührengestaltung;
- Benutzerunterstützung;
- Kooperation zwischen amtlicher Statistik und Forschung.

Insgesamt ergibt sich aus der Darstellung der Situation in den ausgewählten Ländern ein sehr vielfältiges Bild der Kooperation zwischen Wissenschaft und statistischen Behörden. Die Voraussetzungen für die informelle und institutionelle Kooperation variieren zwischen den Ländern in hohem Grad. Die Nähe zwischen Statistik und Wissenschaft und entsprechend die Kooperation zwischen beiden Bereichen ist in einer Reihe von Ländern deutlich stärker ausgeprägt als in Deutschland. Die Beispiele in den untersuchten Ländern verdeutlichen auch, wo und wie diese Kooperation – selbst bei unterschiedlichen institutionellen Voraussetzungen – für beide Seiten mit Gewinn und zum Nutzen für die Gesellschaft ausgebaut werden kann.

A.2 Empfehlungen der Kommission²

Notwendigkeit einer besseren Kooperation von Wissenschaft und Statistik für Datennutzer und Datenproduzenten

1. Die Kommission empfiehlt, an der grundsätzlich bewährten Arbeitsteilung von amtlicher Statistik und wissenschaftsgetragener Datenproduktion festzuhalten, aber die Koordination zwischen Wissenschaft und Statistik durch institutionelle Regelungen zu verbessern.

Die Sozial- und Wirtschaftswissenschaften in der Bundesrepublik Deutschland haben in Kooperation mit der kommerziellen Markt- und Meinungsforschung im Verlauf der letzten zwanzig Jahre eine hochdifferenzierte, auf repräsentativen Stichproben basierte Mikrodateninfrastruktur aufgebaut. Diese Entwicklung war zum Teil motiviert durch den schwierigen Zugang zu Mikrodaten der amtlichen Statistik, überwiegend war sie aber auch motiviert durch die besonderen Chancen wissenschaftsbasierter Umfragedaten.

Für die wissenschaftliche Forschung muss es daher ein vordringliches Ziel sein, die Vorteile der Daten aus der bisherigen Umfrageforschung (Theoriebezug, Messsensibilität, Aktualität der Fragestellungen und des Datenzugangs, multivariate Mikrodatenanalyse, Längsschnitte) mit den Vorteilen der Daten der amtlichen Statistik (Vollerhebungen bzw. große Stichproben, Kohortenserien, regionale Gliederung, kleine Gruppen, zuverlässige historische Vergleiche, Kontextbezug zu Haushalten, Wohnbezirken, Wohngemeinden und Arbeitsstätten, Nutzbarkeit als Hochrechnungsrahmen) verknüpfen zu können.

Die Notwendigkeit, öffentliche Mittel so effizient wie möglich einsetzen zu müssen, erfordert eine wissenschaftliche und öffentliche Koordination sowohl der amtlichen als auch der wissenschaftsgetragenen Statistik. Ein Teil der Koordination könnte sich erübrigen, wenn alle öffentlich finanzierten statistischen Mikrodaten für Re-Analysen uneingeschränkt (bei Beachtung des Datenschutzes) der gesamten Scientific Community zur Verfügung gestellt werden.

2. Die Kommission empfiehlt die Einrichtung eines Rates für Sozial- und Wirtschaftsdaten zur Weiterentwicklung und Bewertung der informationellen Infrastruktur.

Die Kommission empfiehlt als bundesweites Steuerungs- und Initiativgremium die Einrichtung eines Rates für Sozial- und Wirtschaftsdaten, in welchem die wichtigsten Datenproduzenten und Datennutzer sowie Bund und Länder als Mittelgeber und Sitzländer vertreten sind.

- Ein solcher Rat für Sozial- und Wirtschaftsdaten hätte die Aufgabe,
- die Dateninfrastruktur in Deutschland zu bewerten, weiterzuentwickeln sowie für Bund und Länder Empfehlungen zum Programm der wissenschaftsgetragenen Statistik und dessen Finanzierung zu erarbeiten,
 - die Sozial- und Wirtschaftsberichterstattung sowie die gesellschaftliche Dauerbeob-

² Diese Kurzfassung enthält nicht alle Detailempfehlungen. Die Kommission verweist ausdrücklich auf das ausführliche Empfehlungskapitel E.

- achtung zu fördern,
- die Einrichtung und Arbeit von Forschungsdaten- und Servicezentren (Empfehlungen 29 und 33) zu beraten, zu empfehlen und zu evaluieren,
 - Projektmittel zu vergeben oder zur Vergabe vorzuschlagen.
- Die Vollversammlung des Rates für Sozial- und Wirtschaftsdaten sollte einen Arbeitsausschuss bestellen, der von einem hauptamtlichen Stab unterstützt wird.

Mitwirkung der Wissenschaft bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen

3. Die Kommission empfiehlt dem Rat für Sozial- und Wirtschaftsdaten in Zusammenarbeit mit den einschlägigen Fachverbänden (insbesondere der Deutschen Gesellschaft für Soziologie, dem Verein für Socialpolitik, der Deutschen Gesellschaft für Bevölkerungswissenschaft, der Deutschen Statistischen Gesellschaft, der Deutschen Vereinigung für Politische Wissenschaft, der Schmalenbach Gesellschaft) und den übergreifenden Organisationen (insbesondere der DFG, Max-Planck-Gesellschaft [MPG], WGL) Gespräche mit dem Ziel, eine systematische Mitwirkung bei der Aufstellung von Erhebungs- und Aufbereitungsprogrammen der amtlichen Statistik und den einschlägigen Anhörungen der Parlamente (einschließlich der Datenschutzgesetzgebung) zu organisieren.

Im Hinblick auf die Festlegung der Erhebungs- und Aufbereitungsprogramme der amtlichen Statistik wird zu Recht beklagt, dass der Einfluss wissenschaftlicher Überlegungen zu gering ist. Freilich liegt das keineswegs nur daran, dass die amtliche Statistik nicht offen genug für Anregungen wäre. Zum einen sind die Erhebungs- und Aufbereitungsprogramme der amtlichen Statistik teilweise nur deswegen unflexibel, weil sie zu stark durch gesetzliche Details bestimmt werden. Darüber hinaus ist auch die innerwissenschaftliche Diskussion zur Aufstellung von Erhebungs- und Aufbereitungsprogrammen der amtlichen Statistik und über die optimale Bereitstellung wissenschaftstragender statistischer Daten verbesserungsbedürftig.

4. Die Kommission empfiehlt dem Gesetzgeber, die gesetzlichen Vorgaben in den einzelstatistischen Gesetzen auf den verfassungsrechtlich gebotenen Mindestumfang zu beschränken und Detailfestlegungen den statistischen Ämtern und ihren Beiräten zu überlassen.

Die statistischen Ämter können neue Problemstellungen und Anregungen aus der Wissenschaft nur in dem Maße berücksichtigen, wie ihnen bei der Ausgestaltung der Erhebungsprogramme auch entsprechender Spielraum verbleibt.

5. Die Kommission empfiehlt dem Gesetzgeber, die Aufgabendefinition für den Statistischen Beirat im BStatG in Richtung einer mittelfristig ausgerichteten Programmplanung zu erweitern und eine formelle Anhörungspflicht vor der Einführung, Abschaffung oder vor wesentlichen Änderungen amtlicher Statistiken vorzusehen. Die Zahl der Vertreter der Wissenschaft sollte erhöht und es sollte sichergestellt werden, dass Vertreter der empirischen Sozial- und Wirtschaftsforschung entsandt werden.

Mit einer solchen Regelung wäre zu gewährleisten, dass die Wissenschaft im Statistischen Beirat angemessen und sachkundig vertreten ist, rechtzeitig vor der Verabschiedung von Gesetzen über geplante Programmänderungen informiert ist und ihre Interessen, auch hinsichtlich des Datenschutzes, im Verlaufe des Gesetzgebungsprozesses artikulieren und vertreten kann.

6. Die Kommission empfiehlt dem Gesetzgeber, für Erhebungen der Ressorts und von nicht-statistischen Behörden bzw. Körperschaften und Institutionen wie der Bundesbank, der Bundesanstalt für Arbeit oder aus der Sozialversicherung regelhafte wissenschaftliche Beratungsmöglichkeit zu schaffen.

Prioritäten bei Fortführung und Ausbau wichtiger Statistiken

Die Kommission hat es als ihre Aufgabe angesehen, neben Verfahrensvorschlägen auch inhaltliche Aussagen zu Fortführung und Ausbau der informationellen Infrastruktur zu machen. Sie hält die folgenden Vorhaben für besonders dringlich.

Im Bereich der *amtlichen Statistik*:

7. Die Kommission empfiehlt die Durchführung einer Volkszählung, um neben den unmittelbar interessierenden Grunddaten eine Basis für verlässliche Differenzierung und Hochrechnungen aller Stichprobenerhebungen zu gewinnen.

8. Die Kommission empfiehlt die Weiterentwicklung des MZ. Hier geht es einmal um den Übergang auf ein Konzept der unterjährigen Erhebung, die eine Aufbereitung als rotierendes Panel erlauben sollte. Die Nutzbarkeit würde erheblich erhöht, wenn exakte Angaben über Bruttoerwerbseinkommen erfragt würden. Auf dieser Basis könnte ein Access-Panel für die Ziehung freiwilliger Haushaltsstichproben entwickelt werden. Ihren vollen Nutzen entfaltet diese Aktivität allerdings erst, wenn aufbauend auf dem Erfolg der Pilotprojekte die Weiterführung der Mikrozensus-Scientific Use Files realisiert wird.

9. Die Kommission empfiehlt eine Verbesserung der EVS – methodisch durch Verringerung der Zeitabstände und Einführung eines rotierenden Stichprobenkonzepts, inhaltlich durch den Nachweis differenzierter Vermögensangaben, um Aussagen über die private Absicherung sozialer Risiken zu gewinnen.

10. Die Kommission empfiehlt die Schließung gravierender Lücken in der Statistik des Unternehmenssektors.

Von besonderer Dringlichkeit ist hier der weitere Ausbau der Dienstleistungsstatistik. Aber auch Veränderungsprozesse bei den Unternehmen wie Unternehmensgründungen und Konzentrationsprozesse bedürfen einer besseren statistischen Erfassung.

Im Bereich der *wissenschaftsgetragenen Statistik*:

11. Die Kommission empfiehlt die dauerhafte Institutionalisierung des SOEP als breit angelegte wissenschaftsgetragene und ausreichend große Längsschnitt-Stichprobe.

Das SOEP ist wegen seines für die wissenschaftliche Forschung wie für die Politikberatung gleichermaßen reichen Informationsgehalts und wegen seiner breiten Nutzung für die informationelle Infrastruktur in Deutschland unverzichtbar. Dringlich ist die Umsetzung der Empfehlung des Wissenschaftsrates hinsichtlich einer dauerhaften Institutionalisierung und Finanzierung. Notwendig ist auch eine Vergrößerung der Stichprobe, da sie die Nutzbarkeit für Forschungen über kleinere Teilgruppen erheblich verbessern würde. Dabei wird auf die Einbindung und Nutzung des SOEP in die von der EU angestrebte Sozialberichterstattung auf europäischer Ebene zu achten sein.

12. Die Kommission empfiehlt die Fortführung von ALLBUS (Allgemeine Bevölkerungsumfrage der Sozialwissenschaften), ISSP (International Social Survey Program) und Wohlfahrtssurveys, da sie in den Sozialwissenschaften wichtige Informationsquellen darstellen.

ALLBUS und ISSP sind die am häufigsten benutzten Umfragedaten in den Sozialwissenschaften und haben vor allem auch eine hohe Bedeutung für die Ausbildung an den Hochschulen. Mit den Wohlfahrtssurveys ist ein wichtiges Instrument der gesellschaftlichen Dauerbeobachtung entstanden, das in geeigneter Weise fortgeführt werden sollte.

13. Die Kommission empfiehlt die verstärkte Förderung von Kohortenstudien, wie z.B. eines „Entwicklungsbus“.

Prospektive und retrospektive Kohortenstudien sind als jahrgangs- und personenbezogene Längsschnittdaten in besonderer Weise geeignet, Prozesse der menschlichen Entwicklung und des Lebensverlaufs im Kontext des sozialen und wirtschaftlichen Wandels zu erfassen und zu analysieren. Aufgrund der damit verfügbaren, hoch differenzierten Individualdaten sind besondere Datenorganisationsarbeiten notwendig, um solche Studien für die breitere wissenschaftliche Forschung zugänglich zu machen. Förderungsbedürftig ist sowohl die Fortführung bestehender Kohortenstudien als auch die Generierung neuer Kohortenstudien, die frühe Kindheit, Adoleszenz und das frühe Erwachsenenalter umfassen.

Die Kommission möchte sich in ihren Empfehlungen auf diese wenigen, besonders dringlichen Vorhaben beschränken. Wie in Abschnitt C.1.2 dargestellt, gibt es darüber hinaus eine große Anzahl von Bereichen, bei denen eine Verbesserung der informationellen Infrastruktur sinnvoll ist. Hierüber sollte im Rahmen der von der Kommission vorgeschlagenen Verfahren entschieden werden.

Förderung der Forschung zur Datenerhebung, -aufbereitung und -archivierung

14. Die Kommission empfiehlt dem Bund, eine Kommission des Wissenschaftsrates zum Stand der Ausbildung und Forschung in Methoden der empirischen Sozial- und Wirtschaftsforschung zu beantragen bzw. diesen Bereich explizit zum Aufgabebereich der Kommission zu machen, die gegenwärtig die Wirtschaftswissenschaft untersucht (Kommission „Empirische Wirtschaftsforschung“).

Wie der Vergleich mit dem Ausland zeigt, gibt es in Deutschland ein Defizit im Stand der Ausbildung und Forschung in Methoden der empirischen Sozial- und Wirtschaftsforschung. Die Kommission empfiehlt die Einrichtung von Lehrstühlen oder Zentren an Hochschulen, die in Forschung und Lehre in besonderer Weise auf methodische Probleme der Umfrageforschung und der amtlichen Statistik ausgerichtet sind.

15. Die Kommission empfiehlt, zur Förderung der Methodenforschung in der amtlichen Statistik die Wissenschaft stärker in die Weiterentwicklung des methodischen Instrumentariums einzubeziehen und die budgetären Voraussetzungen für eine nachhaltige Methodenforschung in der amtlichen Statistik zu schaffen.

Die statistischen Ämter sind nicht in erster Linie Verwaltungsbehörden, sondern zuallererst Wissenschaftsbetriebe. Diese Komponente ihrer Arbeit wird in ihrer Aufgabenstellung und in ihren Finanzhaushalten nicht sichtbar. Um den wissenschaftlichen Sachverstand für die amtliche Statistik zu nutzen und neue Anstöße für methodische Verbesserung zu geben, sollten in weit stärkerem Umfang als bisher gemeinsame Forschungsprojekte von Wissenschaft und amtlicher Statistik durchgeführt werden.

16. Die Kommission empfiehlt, den neu zu schaffenden Rat für Sozial- und Wirtschaftsdaten mit der Behandlung des Archivierungsproblems bei statistischen Daten zu beauftragen.

17. Die Kommission empfiehlt der DFG die Etablierung des Faches „Empirische Wirtschaftsforschung“ (analog zur „Empirischen Sozialforschung“ als Teildisziplin [bzw. als Erweiterung der Teildisziplin „Statistik“]).

Dies bedeutet die explizite Wahl von Fachgutachtern für „Empirische Wirtschaftsforschung“. Dies würde im DFG-Normalverfahren und darüber hinaus dazu beitragen, eine angemessene Repräsentanz der Empirischen Wirtschaftsforschung in den Selbstverwaltungsorganen der (deutschen) Wissenschaft zu erreichen und mittelfristig den Ausbau des Faches in den Universitäten zu fördern und zu sichern.

Aus- und Weiterbildung

18. Die Kommission empfiehlt den Universitäten und Fachbereichen, eine Verbesserung der Ausbildung der Studierenden im Grundstudium in Richtung Statistik, Ökonometrie und angewandte Informatik unter Verwendung realistischer Datensätze anzustreben. Im Hauptstudium ist eine Vertiefung der Ausbildungsmöglichkeiten in Statistik und Ökonometrie und in den Bereichen Datenerhebung, Aufbereitung der Daten, Datenschutz und Auswertung notwendig. Die Etablierung einer regulären Graduiertenausbildung nach amerikanischem Muster ist für wissenschaftliche Spitzenleistungen von zentraler Bedeutung, wobei für Wirtschafts- und Sozialwissenschaftler das Erlernen neuer empirischer Methoden und die Vertiefung in Statistik und Ökonometrie obligatorisch sein sollte.

Die Präsenz der amtlichen Statistik sollte im Lehrangebot der Hochschulen durch vermehrte Vergabe von Lehraufträgen zur Verbesserung der Kenntnis in spezifischen Methoden und Datenbanken der amtlichen Statistik verstärkt werden.

19. Die Kommission empfiehlt Universitäten und Wissenschaftsministerien die Rücknahme der Maßnahmen zur Verringerung der Zahl der Professuren in den Bereichen empirische Sozial- bzw. Wirtschaftsforschung, Statistik und Ökonometrie in den einschlägigen Fakultäten und die Hebung vieler der vorhandenen C3-Professuren auf C4-Professuren, um die Attraktivität von Arbeiten in diesem Bereich zu erhöhen.

20. Die Kommission empfiehlt die Einrichtung von Kontaktseminaren und Weiterbildungskursen in Zusammenarbeit von Universitäten und statistischen Behörden bis hin zu einem kontinuierlichen Aus- und Weiterbildungsprogramm oder Ergänzungsstudium für junge Wissenschaftlerinnen und Wissenschaftler, aber auch für Nachwuchskräfte der amtlichen Statistik. Dies könnte zu den spezifischen Aufgaben der außeruniversitären Institute gehören.

Zum Beispiel sollte GESIS den expliziten Auftrag erhalten, ihr Fortbildungsangebot z.B. in der Form interdisziplinärer Summer Schools in Zusammenarbeit mit geeigneten Hochschullehrerinnen und Hochschullehrern auszuweiten mit Lehrangeboten für Stichprobenverfahren, Erhebungsmethoden sowie Datenanalyse für die Bereiche amtlicher und nichtamtlicher Mikrodaten.

Wirtschaftliche Gesichtspunkte des Datenzugangs

21. Die Kommission empfiehlt, aggregierte Daten der amtlichen Statistik via Internet zumindest für die Wissenschaft weitgehend kostenfrei oder gegen eine geringe Schutzgebühr zur Verfügung zu stellen.

Durch zusätzliche Abrufe von Internet-Informationen entstehen praktisch keine zusätzlichen Kosten. Zum Forschen gehört, dass man zunächst nicht weiß, welche Daten für eine bestimmte Aufgabe geeignet sind. Erst im Forschungsprozess stellt sich heraus, welche Daten wirklich benötigt werden. Wird aber jede einzelne Reihe mit einem Preis belegt, ergeben sich schnell erhebliche Summen.

22. Die Kommission empfiehlt, in Anlehnung und Weiterführung an die vom BMBF geförderten Pilotprojekte den Datenproduzenten die Kosten für die Durchführung der Anonymisierung (einschließlich Dokumentation, also für die Erstellung von Public Use Files für die Wissenschaft und Scientific Use Files) pauschal über Mittel der Forschungsförderung zu erstatten. Von den einzelnen Forschern wäre dann noch ein Entgelt in Höhe der marginalen Kosten der Abwicklung der Datenlieferung zu zahlen.

Da für die weit überwiegende Mehrzahl der Wissenschaftler die auf den einzelnen Nachfrager umgelegten Bereitstellungskosten für Scientific und Public Use Files prohibitiv wirken würden, werden die fixen Kosten für alle Nutzer gemeinsam durch die Forschungsförderung übernommen. Ein solches Preismodell liegt auch im Interesse der Datenproduzenten, da es sicherstellt, dass ihnen die Fixkosten der Erstellung solcher Files erstattet werden.

Zugang zu aggregierten Daten

23. Die Kommission empfiehlt ein gemeinsames Datenbanksystem der amtlichen Statistik, auf das via Internet zugegriffen werden kann und das die Daten aller Bundesstatistiken in regionalisierter Form enthält.

Hinsichtlich des Zugangs zu den Datenbanken der statistischen Ämter besteht erheblicher Modernisierungsbedarf. Es ist wenig nutzerfreundlich, wenn die einzelnen statistischen Ämter über jeweils eigene Datenbanken mit unterschiedlichen Benutzeroberflächen verfügen und regional tief gegliedertes Material nicht an einer Stelle zugänglich ist.

Zugang zu Mikrodaten und Datenschutz

24. Die Kommission empfiehlt, je nach der Art der Daten unterschiedliche Zugangswege zu Mikrodaten zu verwenden.

Mikrodaten sind Daten einzelner Personen, Unternehmen und Organisationen. Die Daten werden im Vertrauen auf Anonymität und Geheimhaltung zur Verfügung gestellt und bedürfen deshalb effektiver Maßnahmen zu ihrem Schutz.

Konstitutiver Bestandteil der wissenschaftlichen Methodik ist das Prinzip der Überprüfbarkeit vorgelegter Ergebnisse. Konsequenterweise kann dieses nur verfolgt werden, wenn die ursprünglichen Mikrodaten, aus denen Ergebnisse gewonnen werden, zum Zwecke der Re-Analyse verfügbar sind.

25. Die Kommission empfiehlt, Verfahren zur Klärung der Nutzungsberechtigung und der Voraussetzungen auf der Nutzerseite zu entwickeln. Wünschenswert ist die Entwicklung eines sanktionsbewehrten Verhaltenscodex (Code of Conduct).

Dem Datenschutz dienen technische und weitere Vorkehrungen bei den Datennutzern, durch die erreicht wird, dass die Daten nicht in unbefugte Hände geraten, dass keine Versuche unternommen und erfolgreich durchgeführt werden können, die Identität der Datengeber zu rekonstruieren. Erforderlich ist z.B. die periodische Revision des im Rahmen des deutschen Anonymisierungsprojektes entwickelten Kataloges von technischen Maßnahmen. Ein Ethikcodex, der inzwischen in vielen Disziplinen besteht, sollte im Einzelnen die Verpflichtungen der Wissenschaftler und Institutionen im Hinblick auf ein datenschutzgerechtes Verhalten bei der Nutzung von Mikrodaten beschreiben. Er sollte in Zusammenarbeit der betroffenen Disziplinen erarbeitet werden.

Zu klären ist auch, welche Nutzer in den Genuss des Wissenschaftsprivilegs kommen. Für Einrichtungen, bei denen die Zugehörigkeit nicht eindeutig institutionell abgegrenzt werden kann, ist das Instrument der *Zertifizierung* zu schaffen.

26. Die Kommission empfiehlt, die Entwicklung von Scientific Use Microdata Files als wichtigstes Instrument des Mikrodatenzugangs weiter voranzutreiben. Zur Ermöglichung von Analysen des sozialen Wandels wird empfohlen, für die Erhebungen, bei denen jetzt aktuelle Scientific Use Files geschaffen wurden, auch geeignete ältere Daten in dieser Weise zugänglich zu machen und weitere derartige Files wie z.B. Mikrozensus-Regionalfiles oder Mikrozensus-Panelfiles bereitzustellen.

Faktisch anonymisierte Mikrodaten, die in einer standardisierten Weise den wissenschaftlichen Einrichtungen für Forschungszwecke zur Verfügung gestellt werden, stellen ohne Zweifel die kosteneffizienteste Variante der Datenweitergabe dar. Zugleich ist sie in aller Regel auch die für den Forscher günstigste und flexibelste Lösung. Bisher sind Scientific Use Files vor allem für Personen- und Haushaltsdaten erstellt worden. Es gibt allerdings noch einen erheblichen Forschungsbedarf für Anonymisierungsuntersuchungen bei Organisations- und Unternehmensdaten.

27. Die Kommission empfiehlt die Durchführung eines gemeinsamen Forschungsprojekts von Wissenschaft und amtlicher Statistik zu Möglichkeiten der faktischen Anonymisierung von Unternehmens- und Betriebsdaten.

Das Projekt soll im Lichte der neueren methodischen und technischen Möglichkeiten prüfen, ob die häufig sehr pauschal geäußerte Auffassung, wirtschaftsstatistische Einzeldaten seien generell nicht für eine Anonymisierung geeignet, gerechtfertigt ist oder ob trotz der bekannten Schwierigkeiten Anonymisierungsstrategien entwickelt werden können, um auch Unternehmens- und Betriebsdaten an die Wissenschaft übermitteln zu können.

28. Die Kommission empfiehlt die Entwicklung von Public Use Microdata Files.

Public Use Files unterscheiden sich von Scientific Use Files im Wesentlichen nur durch eine weitergehende, sogenannte absolute Anonymisierung. Dafür ist ihre Weitergabe nicht an das Wissenschaftsprivileg gebunden. Sie könnten vorteilhaft in der Ausbildung von Studierenden eingesetzt werden; sie könnten einen wesentlichen Teil kommerzieller Nutzerinteressen abdecken und könnten zumindest teilweise einige Probleme auffangen, die mit dem bisherigen Ausschluss ausländischer Wissenschaftler von der Nutzung von Mikrodaten verbunden sind.

29. Die Kommission empfiehlt die Einrichtung von Forschungsdatenzentren mit der Möglichkeit der kontrollierten Ferndatenverarbeitung.

Die entscheidende Grenze für die Nützlichkeit von Scientific Use Files sind Datenfiles, bei denen schon die faktische Anonymisierung die Aussagekraft der Daten in hohem Maße einschränkt. Ähnliches gilt für Untersuchungen, die das Zusammenführen verschiedener Datensätze erfordern. Für solche Daten und Arbeiten wird der Aufbau von Forschungsdatenzentren mit besonderen Sicherungsmaßnahmen empfohlen, mit denen in den USA bereits gute Erfahrungen gemacht werden. Wissenschaftler können dort ihre Untersuchungen durchführen. Der Aufbau solcher Forschungsdatenzentren erscheint als einzig gangbarer Weg, um auch besonders sensible und kaum anonymisierbare Daten vor allem auf Betriebs- oder Unternehmensebene für Forschungszwecke zu öffnen.

Forschungsdatenzentren bieten im Grundsatz Daten unterschiedlicher Datenbesitzer und Datenproduzenten an. Insbesondere sind sie auch nicht ausschließlich auf Daten der amtlichen Statistik begrenzt. Sie sollten auch besonders zu schützende und nicht hinreichend anonymisierbare Daten aus der wissenschaftlichen Forschung zugänglich machen. Forschungsdatenzentren sollten bei den entsprechenden Datenbesitzern und unter geeigneter Kontrolle durch die Datenbesitzer auch in einzelnen wissenschaftlichen Ein-

richtungen eingerichtet werden. Dies darf aber nicht zu einer Aufsplitterung von Forschungsdatenzentren nach Datenbesitzern führen.

In Forschungsdatenzentren sollte die Möglichkeit kontrollierter Datenfernverarbeitung bestehen. Wissenschaftler können ihre Datenanalyseaufträge auf elektronischem Weg an ein Forschungsdatenzentrum übermitteln. Das Personal des Forschungsdatenzentrums bringt nach datenschutzbezogener Kontrolle den Analyseauftrag zur Ausführung und übermittelt das ebenfalls datenschutzbezogen kontrollierte Analyseergebnis an den Auftraggeber.

30. Die Kommission empfiehlt den Datenproduzenten, ein Gastwissenschaftler-Modell zu praktizieren.

Es ist absehbar, dass die Entwicklung bis zur vollen Funktionsfähigkeit von Forschungsdatenzentren erhebliche Zeit in Anspruch nehmen wird. Es ist sinnvoll, schon jetzt Möglichkeiten zu schaffen, in denen Forscher nach entsprechenden Datenschutzverpflichtungen Mitarbeiterinnen und Mitarbeitern von statistischen Ämtern oder anderen datenhaltenden Einrichtungen gleichgestellt werden und im Rahmen eines Aufenthaltes als Gastwissenschaftler Auswertungen von Mikrodaten vornehmen können. Bezüglich der Auswahl von Gästen empfiehlt die Kommission die Erarbeitung eines transparenten Verfahrens.

Nutzung internationaler Mikrodaten

31. Die Kommission empfiehlt, die Situation für die international ausgerichtete wirtschafts- und sozialwissenschaftliche Forschung zu verbessern. Hierzu sind zahlreiche Maßnahmen dringlich.

Insbesondere ist die Entwicklung und Weitergabe an die Wissenschaft der bei Eurostat vorhandenen und für die Mitgliedstaaten der EU harmonisierten Datenbasen einschließlich der nationalen Scientific Use Files in der Form von anonymisierten Scientific Use Files notwendig. Die Bundesregierung und die statistischen Ämter Deutschlands sollten entsprechende Schritte nachdrücklich fordern, unterstützen und durchsetzen.

Die Harmonisierung von Daten aus verschiedenen Ländern ist selbst ein dringliches Forschungsproblem.

32. Die Kommission empfiehlt der Bundesregierung, zur Förderung des Datenaustausches mit Forschungseinrichtungen in Nicht-EU-Staaten, insbesondere USA, Kanada, Australien, Japan, Verhandlungen mit dem Ziel der Implementation eines Safe-Harbor-Mechanismus aufzunehmen.

Die Kommission sieht diese Vereinbarung zum angemessenen Datenschutz in Forschungseinrichtungen von Drittstaaten in Form der Selbstregulierung und Selbstzertifizierung der Scientific Community für dringlich und geeignet, den Datenaustausch und die Zusammenarbeit der empirischen Wirtschafts- und Sozialwissenschaften nachhaltig zu verbessern.

33. Die Kommission empfiehlt, in Deutschland auch in Zukunft Serviceeinrichtungen der Forschung als Teil der informationellen Infrastruktur vorzuhalten.

Die Effizienz der Nutzung von Mikrodaten in der Forschung kann erheblich gesteigert werden, wenn verschiedene Grundlagenarbeiten nicht von den einzelnen Forschern oder Forschergruppen jeweils ad hoc erledigt werden müssen, sondern diese kontinuierlich und systematisch für die gesamte Wissenschaftsgemeinschaft in zentralen Serviceeinrichtungen bearbeitet werden.

Es ist wichtig, dass die Organisation und Erledigung der Servicefunktionen wissenschaftsnah erfolgt, damit die Entwicklungen in der Wissenschaft rechtzeitig erkannt und die daraus sich ergebenden Daten- und Servicebedarfe daran ausgerichtet werden. Gleichzeitig ist eine Unabhängigkeit von den Datenproduzenten notwendig, um die Entwicklungen dort mit einer gewissen kritischen Distanz zu verfolgen, die Interessen der Wissenschaft zu bündeln und gegenüber den Datenproduzenten zu vertreten.

Die Kommission unterstützt den Ausbau dieser Dienste u.a. für bessere Datendokumentation, Entwicklung datenschutz- und nutzergerechter Anonymisierungskonzepte für neue Datenfiles, verstärkte Nutzerbetreuung und Kooperation mit den Datenbesitzern, Erstellung von Metadatenbanken, Mitwirkung bei der Fortentwicklung des Datenschutzes für Daten im Wissenschaftsbereich sowie in begrenztem Umfang auch Durchführung eigener Forschung.

Datenverknüpfung

34. Die Kommission empfiehlt, die Möglichkeit exakter Datenverknüpfungen – ohne ausdrückliche Zustimmung aller einzelnen Befragten – für rein statistische Zwecke gesetzlich zu regeln. Eine exakte Verknüpfung von Erhebungsdaten mit Registerdaten sollte grundsätzlich ebenso beurteilt und gesetzlich geregelt werden.

Um die Kosten der Datenbeschaffung sowie die Belastung der Befragten zu senken, sollten die Anbieter der statistischen Infrastruktur ihr Augenmerk in Zukunft noch stärker als bisher auf eine volle Ausschöpfung des Informationsgehalts von vorhandenem Zahlenmaterial durch statistische Mikrodatenverknüpfung richten.

Voraussetzung für eine datenschutzrechtlich unbedenkliche exakte Verknüpfung von Daten ist die Sicherstellung der Anonymität der statistischen Einheiten im Hinblick auf staatliches Handeln (im Verwaltungs- und Justizbereich) und kommerzielle Interessen. Eine Zusammenführung von Dateien sollte deswegen nur in einem explizit zu definierenden, abgeschotteten Forschungs- und Statistikbereich möglich sein.

Forschungsdatengeheimnis

35. Die Kommission empfiehlt dem Gesetzgeber, ein Forschungsdatengeheimnis einzuführen.

Problematische Abwägungen im Statistikrecht wie im allgemeinen Datenschutzrecht über die „Erforderlichkeit“ der Verarbeitung bestimmter Daten zu Forschungszwecken durch staatliche Stellen werden durch ein Forschungsdatengeheimnis vermieden und prekäre Eingriffe in die Wissenschaftsfreiheit können entfallen. Ein Forschungsdaten-

geheimnis muss so ausgestaltet werden, dass sowohl das Zeugnisverweigerungsrecht des Wissenschaftlers im Hinblick auf Forschungsdaten als auch ein entsprechendes Beschlagnahmeverbot in der Regelung verankert werden. Ebenso ist das Forschungsdatengeheimnis und seine Verletzung in den strafrechtlichen Rechtsgüterschutz des § 203 Strafgesetzbuch (StGB) aufzunehmen.

Nach Auffassung der Kommission sollte das Forschungsdatengeheimnis in der zweiten Stufe der Anpassung der nationalen Datenschutzgesetzgebung sowie anderer Regelungen des bereichsspezifischen Datenschutzes an die Europäische Datenschutzrichtlinie realisiert werden.

Umsetzung und Finanzierung

36. Die Kommission empfiehlt den zuständigen Institutionen der Forschungs- und Wissenschaftspolitik, die nicht unbeträchtlichen finanziellen Mittel, die zur Verwirklichung dieser Empfehlungen benötigt werden, bereitzustellen.

Um möglichst schnell zu spürbaren Verbesserungen in der Kooperation zwischen Statistik und Wissenschaft zu kommen, empfiehlt die Kommission dringlich neben der Fortführung und dem Ausbau wichtiger Statistiken die Einsetzung des Rates für Wirtschafts- und Sozialdaten (Empfehlung 2), die zügige Finanzierung von Scientific Use Files (Empfehlung 22), die Einrichtung von Forschungsdatenzentren (Empfehlung 29) bei den Datenproduzenten aus amtlicher Statistik und Wissenschaft und von Servicezentren (Empfehlung 33). Die Kommission empfiehlt, hierzu vorhandene Ansätze aufzugreifen. Die Kommission hält es für dringlich, die von ihr vorgeschlagenen legislativen Schritte bei der von der Bundesregierung bereits angekündigten weiteren Novellierung des BDSG und anderer Regelungen des bereichsspezifischen Datenschutzes noch in dieser Legislaturperiode einzuleiten.

