

Ulrich Schlungbaum

## Online-Datenerhebung – Momentaufnahme eines laufenden Verfahrens

### Vorbemerkungen

Im April 2003 startete das Statistische Landesamt des Freistaates Sachsen im Rahmen der Initiative „Bund Online 2005“ den Produktionsbetrieb einer Datenerfassungslösung per Internet. Dieses Verfahren „Statistik-online“ wurde in dieser Zeitschrift in seinem technischen Konzept ausführlich vorgestellt. [1] Hier soll nun nach gut eineinhalb Jahren Produktionsbetrieb eine erste Bestandsaufnahme erfolgen.

Der nachfolgende Bericht bezieht sich dabei nur auf dieses Verfahren<sup>1)</sup>, auch wenn es zuvor einige einzelne Vorläufer gegeben hat, so zum Beispiel eine erste Online-Registerumfrage im Sommer 2000 oder das erste Dateiapload-Verfahren für die Pflegestatistik Anfang 2002. Der quantitative Erfolg dieser Vorgängerverfahren war zwar aus heutiger Sicht gering, aber es wurde eine ganze Reihe wertvoller technischer und inhaltlicher Erfahrungen gesammelt, die letztlich auch heute noch nachwirken.

### Bilanz nach achtzehn Monaten

Der April 2003 gilt deshalb als Startzeitpunkt für das Verfahren, weil damals das erste Formular, das für die „Vierteljährliche Konjunkturstatistik in bestimmten Dienstleistungsbereichen im ersten Quartal 2003“, frei geschaltet werden konnte. Mehr als 2 200 Berichtspflichtige waren mit ihren personalisierten Zugangsdaten versorgt worden, das Formular war mit großer Sorgfalt programmiert, es gab eine Webanwendung mit „am grünen Tisch“ erdachten Funktionalitäten und ein für alle möglichen Erhebungen einheitliches definiertes Eingangsformat. Niemand wusste aber wirklich, was nun geschehen würde.

Sicher war, die Webanwendung würde stabil laufen, aber hatte man wirklich alle Probleme vorab bedacht, und die eigentlich spannendste Frage war, wie wird die Beteiligung am Verfahren ausfallen? – Schon nach wenigen Tagen bestätigte es sich, alle wesentlichen Eckpunkte des Verfahrens funktionieren wie geplant. Die ersten 100 Internet-Datenmeldungen waren bereits nach zehn Tagen eingegangen. Sie konnten alle ordnungsgemäß registriert und ihre Inhalte in das Verarbeitungsprogramm weitergeleitet werden. Größere Probleme bei den Erhebungspflichtigen

wurden nicht bekannt, der angebotene Telefonsupport wurde so gut wie nicht genutzt.

Nachdem dann auch das zweite Quartal für die Dienstleistungsstatistik erfolgreich durchlaufen werden konnte, wurde es im Oktober 2003 möglich, die Pilotphase zu beenden und mit der Anwendung in die Breite zu gehen. Schritt für Schritt wurden jetzt für weitere Statistiken die Online-Erhebungen vorbereitet und im Internet frei geschaltet.

Ein Jahr später, zum 1. November 2004, sind für 23 Statistiken insgesamt 105 Erhebungen im Internet durchgeführt worden oder sind gerade aktiv. Diese 105 Erhebungen umfassen insgesamt rund 109 700 Berichtsfälle. Für diese sind bis zum Stichtag 1. November 2004 insgesamt 95 700 Meldungen eingegangen. Die fehlenden Meldungen resultieren zum allergrößten Teil aus den am 1. November 2004 noch nicht abgeschlossenen Erhebungen. Unter den eingegangenen Meldungen sind 13 353 Internetdateneingänge, also knapp 14 Prozent.

### Vergleich der Nutzergruppen

Da im Online-Verfahren des Statistischen Landesamtes Sachsen unterschiedslos alle Berichtspflichtigen einen Zugang zu ihrem Formular erhalten, sind diese Nutzergruppen untereinander vergleichbar. Man erkennt schnell gravierende Unterschiede zwischen verschiedenen Statistiken. Es gibt drei Gruppen. Den höchsten Online-Anteil an den eingegangenen Meldungen mit deutlich über 25 Prozent haben die Berufsschulstatistik (60 Prozent), die „Standesamts“-Statistiken – Geborene, Gestorbene, Eheschließungen (50 Prozent), die Personenbeförderungsstatistik (35 Prozent) und die laufende Wirtschaftsrechnung (26 Prozent).

1) Neben dem Verfahren „Statistik-online“ existiert noch das Verfahren PEGASUS, mit dessen Hilfe der Dateneinzug für die Statistiken der allgemein bildenden und Förderschulen erfolgt. Statistik-online unterscheidet sich von PEGASUS u. a. dadurch, dass die Teilnahme für die Auskunftspflichtigen freiwillig ist. Außerdem beinhaltet PEGASUS weitere Funktionalitäten. Damit werden bei Statistik-online keine so hohen Teilnehmerzahlen erreicht wie bei den Schulstatistiken. Zu PEGASUS vgl. ausführlich [2].

Auf der anderen Seite haben sich an der Pflegestatistik (3 Prozent), Statistik der Bodennutzung (5 Prozent) bis hin zur Beherbergungsstatistik (7 Prozent) jeweils nur relativ wenige Berichtspflichtige, d. h. durchschnittlich 5 Prozent beteiligt. Im Mittelfeld liegen dann die restlichen Statistiken mit einem Anteil zwischen 10 und 20 Prozent.

In der Spitzengruppe befinden sich nur Statistiken, die einen relativ kleinen und in seiner Zusammensetzung sehr homogenen Berichtskreis haben. Am deutlichsten wird das bei den Berufsschulen und Standesämtern. Hier handelt es sich jeweils um Berichtspflichtige, die sich einander durch die Art Ihrer Tätigkeit und auch durch die damit verbundene technische Ausstattung sehr stark ähneln.

Das Mittelfeld wird bestimmt von Statistiken mit relativ großen Berichtskreisen, in denen sich sehr unterschiedliche Berichtspflichtige befinden. Die Unterschiede bestehen in erster Linie in der Größenklasse dieser Betriebe und Einrichtungen, oft aber auch noch zusätzlich im Wirtschaftszweig.

Für die „5-Prozent-Statistiken“ lässt sich schwer eine für alle geltende Eigenschaft erkennen. Allerdings gibt es auch hier mit den Pflegeeinrichtungen, insbesondere den ambulanten, eine Gruppe, die sehr homogen zusammengesetzt ist.

Versucht man zunächst einmal die unterschiedlichen angesprochenen Berichtskreise nach ihren Mitgliedern zu klassifizieren, so erkennt man im Wesentlichen fünf verschiedene. Das sind die öffentlich-rechtlichen Einrichtungen und Unternehmen, der Dienstleistungsbereich, die Landwirtschaft, das Produzierende Gewerbe und die privaten Haushalte. Augenscheinlich ist die Tendenz, dass Erhebungen im öffentlich-rechtlichen Umfeld einen besonders

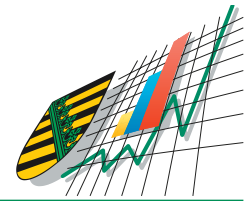
hohen Beteiligungsgrad am Online-Verfahren aufweisen. Von den sechs Statistiken dieser Gruppe befinden sich vier in der Spitzengruppe und zwei im Mittelfeld der zweiten Gruppe.

Nun lag es nahe zu untersuchen, ob in den inhomogener zusammengesetzten Nutzergruppen möglicherweise Unterschiede bestehen. Dazu wurden zunächst die sechs bereits im Wesentlichen abgeschlossenen „Erhebungen in der Konjunkturstatistik in bestimmten Dienstleistungsbereichen“ näher untersucht. Der Berichtskreis teilt sich in Nutzer aus den zwei Wirtschafts-

**Tab. 1 Übersicht über die verfügbaren Online-Verfahren**

EVAS-Nr.	Bezeichnung	Art des Verfahrens	Periodizität <sup>1)</sup>	Anzahl der durchgeführten Erhebungen	Durchschnittliche Anzahl der Berichtspflichtigen	Anteil der Online-Meldungen an den Meldungen insgesamt in %
21121	Statistik der beruflichen Schulen	Upload	j	2	280	61,4
12611	Statistik der Eheschließungen	Upload	m	16	269	47,2
12613	Statistik der Sterbefälle	Upload	m	16	269	47,2
12612	Statistik der Geburten	Upload	m	16	269	45,0
46211	Personenbeförderung	Formular	vj	3	40	34,2
63111	Laufende Wirtschaftsrechnung	Formular	j	1	580	26,1
47414	Konjunkturstatistik in bestimmten Dienstleistungsbereichen	Formular	vj	7	2 200	19,3
42111	Monatsbericht für Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden	Formular	m	8	2 950	18,9
62321	Verdienstenerhebung im Produzierenden Gewerbe, Handel, Kredit- und Versicherungsgewerbe	Formular	vj	4	2 000	16,5
74121	Personalstandsstatistik der Länder, Gemeinden und Gemeindeverbände	Formular	j	1	7 450	16,4
44111	Monatsbericht im Bauhauptgewerbe	Formular	m	5	700	15,9
72111	Jahresabschlussstatistik der staatlichen und kommunalen Fonds, Einrichtungen, Betriebe und Unternehmen	Formular	j	1	1 150	14,7
62351	Bruttojahresverdienstenerhebung im Produzierenden Gewerbe, Handel, Kredit- und Versicherungsgewerbe	Formular	j	1	2 100	12,0
44141	Statistik über den Auftragsbestand im Bauhauptgewerbe	Formular	vj	1	700	11,6
45511	Monatserhebung im Tourismus	Formular	m	10	2 200	6,7
52111	Registerumfrage	Formular	u	2	7 000	6,6
53211	Handwerksberichterstattung	Formular	vj	3	2 700	6,5
41311	Repräsentative Erhebung über die Viehbestände	Formular	j	2	1 900	5,5
62331	Verdienstenerhebung im Handwerk	Formular	j	1	1 250	5,0
41131	Integrierte Erhebung über die Bodennutzung und Viehbestände	Formular	j	1	3 500	4,6
22412	Statistik über stationäre Pflegeeinrichtungen	Upload	2j	2	600	4,1
22411	Statistik über ambulante Pflegeeinrichtungen	Upload	2j	2	900	3,2

1) m - monatlich; vj - vierteljährlich; j - jährlich; 2j - zweijährlich; u - unregelmäßig



abschnitten I und K gemäß Klassifikation der Wirtschaftszweige Ausgabe 2003. Das bedeutet:

- I : Verkehr und Nachrichtenübermittlung,
- K: Grundstücks- und Wohnungswesen, Vermietung beweglicher Sachen, Erbringung von wirtschaftlichen Dienstleistungen.

Im Wirtschaftsabschnitt K sind unter dem Sammelbegriff „Erbringung von wirtschaftlichen Dienstleistungen“ eine ganze Reihe von Dienstleistern zusammengefasst, die alle schon aus der Art ihrer Tätigkeit heraus als besonders IT-nah einzustufen sind. Das sind z. B. alle Software- und Systemhäuser, aber auch Forschungsbüros und Institute unterschiedlichster Fachrichtungen. Hypothetisch sollte also diese Gruppe einen besonders hohen Anteil bei der Nutzung des Online-Verfahrens aufweisen.

In der Tat ist ein leicht erhöhter Anteil der Online-Nutzer im Wirtschaftszweig K erkennbar (vgl. Abb. 1). Während der Anteil des Wirtschaftszweiges K am Gesamtberichtsreis nur bei rund 70 Prozent liegt, sind diese Nutzer mit rund 80 Prozent unter den Online-Meldern vertreten. In der Tendenz über alle sechs zur Verfügung stehenden Erhebungen schrumpft diese Differenz aber bereits deutlich. Offenbar ist die Fähigkeit und Bereitschaft zur Nutzung von Internetdiensten, die in dieser Gruppe von Statistiken mit rund 15 Prozent letztlich immer noch niedrig ist, trotzdem relativ gleichmäßig über alle Beteiligten verteilt. Die technische

Ausstattung und das Know-how im IT-Bereich beim Nutzerkreis sind also nicht das entscheidende Argument, wenn man über Prioritäten bei der Beteiligung an der Online-Datenübertragung bei Statistiken nachdenkt.

Weiterhin gibt es Nutzergruppen, in denen deutliche Unterschiede in der Größenklasse der Unternehmen vorhanden sind, wobei hier eine Klassierung nach der Anzahl der Beschäftigten gemeint ist. Dazu wurden die Beteiligungszahlen am Monatsbericht für Betriebe des Verarbeitenden Gewerbes und des Bergbaus sowie der Gewinnung von Steinen und Erden näher untersucht. Das Ergebnis ist schon erstaunlich: Es gibt keinen wirklichen Unterschied zwischen den einzelnen Größenklassen. Lediglich bei den kleinsten Berichtseinheiten weicht die Beteiligung etwas nach oben ab (vgl. Abb. 2).

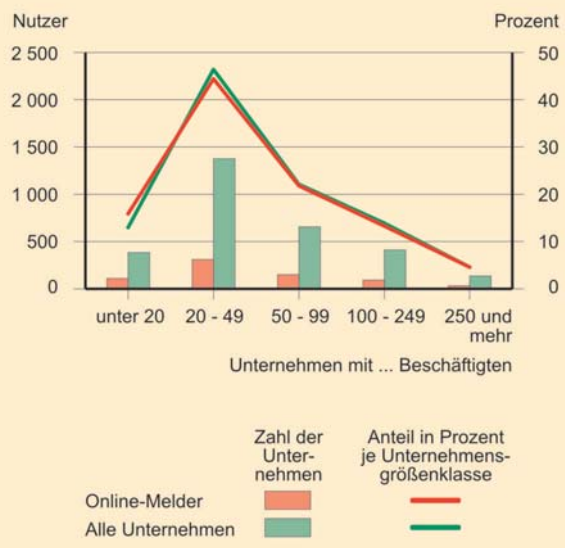
### Tendenzen im Nutzerverhalten

Für die monatlichen und vierteljährlichen Erhebungen kann zum jetzigen Zeitpunkt auch schon eine erste Aussage zur Entwicklung des Nutzerverhaltens getroffen werden. Bis auf den Monatsbericht für Betriebe im Verarbeitenden Gewerbe zeigen hier alle Statistiken ein einheitliches Verhalten. Die Beteiligung am Online-Verfahren steigt stetig und dabei relativ gleichmäßig an (vgl. Tab. 2). Auf Grund dieser Tendenz kann man davon ausgehen, dass man ohne zusätzliche Werbemaßnahmen im Laufe des Jahres 2005 im Durchschnitt auf eine 20-prozentige Teilnah-

**Abb. 1 Einfluss des Wirtschaftszweiges auf die Teilnahme am Online-Verfahren**



**Abb. 2 Einfluss der Größenklassen auf die Teilnahme am Online-Verfahren in der Dienstleistungsstatistik**



mequote am Online-Verfahren kommen kann. Unternimmt man zusätzliche geeignete Maßnahmen, um bei bestimmten Nutzerkreisen auf das Verfahren aufmerksam zu machen, so kann man damit die Teilnehmerzahlen nachweislich deutlich erhöhen.

Im Falle der drei Statistiken zur natürlichen Bevölkerungsbewegung ist dieses bisher am deutlichsten gelungen. Bei diesen Statistiken wird in der Hauptsache jeweils eine Datei an das Statistische Landesamt übermittelt. Diese Datei wird dabei mit einem der beiden Programme Autista und Elvis in den entsprechenden Standesämtern erzeugt und anschließend gesendet. Zum Jahreswechsel 2003/2004 wurden die Mitarbeiter der Standesämter im Zuge eines Versionswechsels der Autista-Software neu geschult, und bei diesen Schulungen wurde eine zusätzliche Unterlage zum bestehenden Online-Verfahren verteilt. Danach stiegen die Teilnehmerzahlen am Online-Verfahren von knapp 20 Prozent im Herbst 2003 auf deutlich über 40 Prozent im Frühjahr 2004 (vgl. Abb. 4).

Auch für die Erhebungen zu den anderen Statistiken ist bereits begonnen worden, entsprechende Informationsmaßnahmen durchzuführen. Eine gute Möglichkeit liegt hier in den bereits lange bestehenden Kontakten des Statistischen Landesamtes zu den entsprechenden Fachverbänden. Auf dieser Basis wurde z. B. im Frühjahr 2004 durch die Industrie- und Handelskammer Dresden eine Informationsveranstaltung nur zum Online-Dateneinzugsverfahren des Statistischen Landesamtes organisiert, in deren Folge dann auch entsprechende Informationen in den Printmedien der Fachverbände der Industrie und des Handwerkes erschienen. Der Erfolg dieser Maßnahmen ist nicht so deutlich messbar wie bei den Bevölkerungsstatistiken, aber richtig ist in jedem Fall, dass eine fundierte Information über ein neues Verfahren zur Akzeptanzsteigerung beiträgt.

## Der Monatsbericht im Verarbeitenden Gewerbe – eine Ausnahme?

Das Online-Verfahren für den Monatsbericht für Betriebe des Verarbeitenden Gewerbes ist das einzige mit einem etwas abweichenden Nutzerverhalten. Der Start lag sofort auf dem relativ hohen Niveau von gut 500 Online-Meldungen, das sind etwa 18 Prozent. Danach stagnierte die Beteiligung. Ein Grund dafür

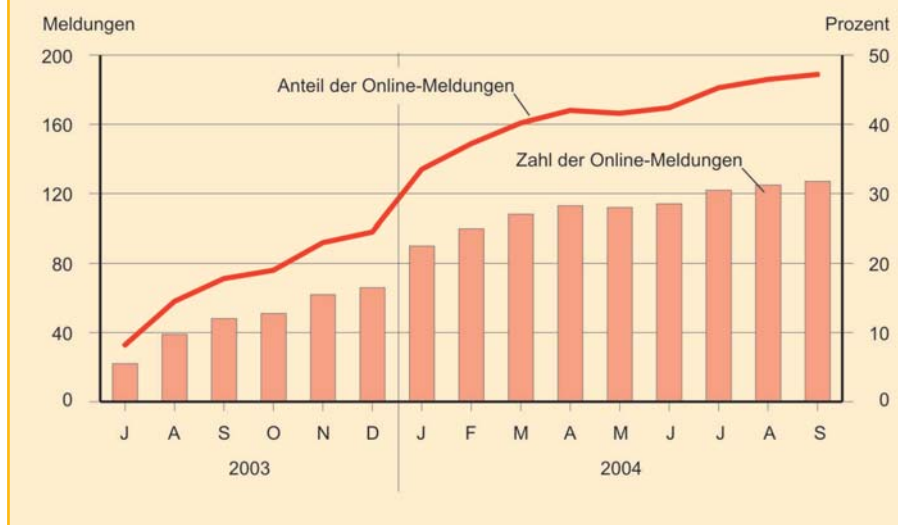
**Tab. 2 Entwicklung der Teilnahme am Online-Verfahren<sup>1)</sup>**

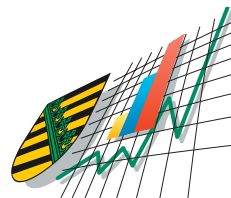
Quartal	Nutzer	Online-Melder	
		absolut	%
1/2003	2 328	265	11,4
2/2003	2 213	272	12,3
3/2003	2 093	268	12,8
4/2003	2 041	275	13,5
1/2004	2 371	344	14,5
2/2004	2 271	380	16,7

1) Konjunkturstatistik in bestimmten Dienstleistungsbereichen

liegt mit Sicherheit darin, dass der Starttermin in Sachsen mit dem Mai 2004 für den Berichtsmonat April 2004 im bundesweiten Vergleich sehr spät lag. Der Monatsbericht war durch die Arbeitsgruppe „Nutzung der Internettechnologie für die Datenerhebung im Statistischen Verbund“ als eine von drei Piloterhebungen definiert worden und zu diesem Zeitpunkt schon in zehn anderen Bundesländern über das Verfahren „Statistik Online“ [3] aus Nordrhein-Westfalen verfügbar. Damit war bei den in Sachsen Berichtspflichtigen eine hohe Erwartungshaltung entstanden, die sich auch schon in entsprechenden Anfragen (telefonisch und per E-Mail) nach dem Starttermin niederschlug. Dieser Umstand erklärt die sofortige hohe Startbeteiligung. Ein genauerer Blick auf das Nutzerverhalten zeigt, dass nun in den folgenden Monaten wie bei allen anderen Statistiken Nutzer hinzukamen, aber eben auch in etwa gleichem Maße Nutzer das Online-Verfahren nicht mehr bedienten (vgl. Tab. 3).

**Abb. 3 Online-Meldungen in der Statistik der Eheschließungen Juli 2003 bis September 2004**





**Tab. 3 Stetigkeit der Nutzung der Internet-Formulare  
April bis August 2004**

Monat	Mel- dungen	Ab- gänge	Zugänge gesamt	Davon mit einer Dauer über ... Monate				
				1	2	3	4	5
April	542	61	542	61	20	18	22	421
Mai	566	40	97	20	5	4	68	x
Juni	540	27	30	4	1	25	x	x
Juli	549	32	29	5	24	x	x	x
August	507	.	15	15	x	x	x	x

Abgesehen davon, dass das Online-Verfahren freiwillig ist, und somit jeder Nutzer selbst über seine Teilnahme entscheidet, besteht natürlich Interesse an einer möglichst hohen Teilnehmerzahl. Durch entsprechende E-Mails bzw. telefonische Rückfragen wurde hier eine mögliche Ursache für die Rückkehr zum Papierformular bekannt. Jeder Nutzer erhält nach erfolgreicher Datenübermittlung als Bestätigung eine entsprechende Quittung mit den übermittelten Daten. Diese Quittung wurde konzeptionell nach dem wysiwyg-Prinzip entwickelt. Wenn man eine solche Quittung ausdruckt, entsteht dadurch zwangsläufig pro Formularreiter – die Internetformulare sind nach dem Karteikarten-Prinzip aufgebaut – eine neue Druckseite. Dieser Umstand war billigend in Kauf genommen worden, da es auch die Möglichkeit gibt, diese Quittung zu downloaden und als Datei zu speichern, nämlich so wie man das bei einem elektronischen Verfahren auch erwartet. Offensichtlich gibt es ein Interesse an einer Papierform der Quittung, das in betriebsinternen Abläufen begründet liegt und das durch die Statistik somit nicht weiter zu beurteilen ist. Deswegen wird in der nächsten Version der Formularserversoftware die Quittung konfigurierbar sein und damit in einer verkürzten Form angeboten.

## Ein Blick voraus

Zum jetzigen Zeitpunkt (November 2004) sind die Online-Datenerhebungen für 23 Statistiken fertig gestellt und eingeführt worden. Für weitere sieben Statistiken sind die Vorbereitungsmaßnahmen zum aktuellen Zeitpunkt weitgehend beendet. Setzt man einmal für 2005 eine konstante Weiterentwicklung voraus, so ist die Hypothese realistisch, zum Jahresende für etwa 50 Statistiken die Verfahren auf der Basis der bisher verwendeten Komponenten erarbeitet zu haben.

Bei der bisher beobachteten Entwicklung in der Nutzerakzeptanz und den offensichtlichen Möglichkeiten ihrer Steigerung gibt das berechnete Anlass zu der Annahme, zum Ende des Jahres 2005 etwa 5 000 Online-Datenmeldungen pro Monat erreichen zu können. Unter der Voraussetzung, dass die neuen Online-Erhebungen alle über das Erhebungsmanagementsystem EMSy

[4] organisiert werden, kann man davon ausgehen, dass kein wesentlicher zusätzlicher Betreuungsaufwand entsteht. Diese Steigerung ist also mit den zur Verfügung stehenden Ressourcen zu bewältigen.

Intensiviert man gleichzeitig die Bemühungen um eine medienbruchfreie Übergabe dieser Dateneingänge in die entsprechenden Produktionssysteme (Verbundprogramme), ein Themenkomplex der im Rahmen dieses Beitrages nicht weiter beleuchtet werden soll, so steckt hier ein weiteres bedeutendes Potenzial für eine nachhaltige Effektivierung des statistischen Produktionsprozesses.

Ulrich Schlungbaum, Dipl.-Math.,  
Sachgebietsleiter WebRedaktion

### Literatur- und Quellenverzeichnis:

- [1] Schlungbaum, U.: Online-Datenerfassung – ein Beitrag zur Initiative E-Government 2005, In: Statistik in Sachsen 3/2003, S. 49 - 55.
- [2] Lehmann, C., S. Prittmann: PEGAUS – Projekt der Online-Erfassung der Daten der amtlichen Schulstatistik der allgemein bildenden Schulen und Förderschulen. In: Statistik in Sachsen, 3/2004, S. 66 - 75.
- [3] Klug, S.: Einführung des Verbundverfahrens. In: Statistische Monatshefte Rheinland-Pfalz, 6/2003, S. 99 - 100.
- [4] Lenk, K., J. Lesch: Einführung eines allgemeinen Erhebungs-Management-Systems. In: Statistik in Sachsen, 3-4/2001, S. 30 - 33.