

Ulrich Schlungbaum

Online-Datenerfassung – ein Beitrag zur Initiative „E-Government 2005“

Vorbemerkungen

Im Jahr 2000 wurde durch die Bundesregierung die Initiative „Bund Online 2005“ gestartet. Ziel ist es, onlinefähige Dienstleistungen der Verwaltung für alle Bürger und Unternehmen bis 2005 im Internet zur Verfügung zu stellen. In diesem Kontext sind natürlich auch das Statistische Landesamt des Freistaates Sachsen (StLA) ebenso wie seine 15 Partnerämter in den anderen Bundesländern und das Statistische Bundesamt (StBA) aufgefordert, einen entsprechenden Beitrag zu leisten. Deshalb wurde im Spätherbst 2001 eine Arbeitsgruppe im Verbund der Statistischen Ämter ins Leben gerufen, die unter dem Titel „Nutzung der Internettechnologie für die Datenerhebung im Statistischen Verbund“ (AG Internet) entsprechende Aktivitäten bündelt und steuert. Die Betonung der Datenerhebung in der Aufgabe dieser E-Government-Arbeitsgruppe lässt bereits erkennen, dass hier für die amtliche Statistik ein wesentlicher Schwerpunkt gesetzt ist. Betrachtet man die Tatsache, dass allein in Sachsen über 100 000 Unternehmen, Behörden und auch Personen zu mehr als 250 Statistiken in unterschiedlichen Abständen befragt werden, so wird die eigentliche Dimension dieser Aufgabe aus dem doch wesentlich breiteren Spektrum von E-Government schnell klar.

Das Konzept

Für jede Erhebung mit Online-Dateneintrag müssen im Grunde drei wesentliche Bestandteile analysiert und beschrieben werden. Das sind die Nutzer eines solchen Angebotes, die eigentliche Dateneingabe und Übertragung und drittens die Weiterverarbeitung der Daten.

Die Nutzer

Um den erstrebten Rationalisierungseffekt wirklich erzielen zu können, braucht man eine hohe Akzeptanz der Lösung bei den potentiellen Anwendern. Klar ist, es werden nicht sofort alle Berichtspflichtigen an einer neuen Lösung teilnehmen, aber der Kreis der dem Neuen gegenüber aufgeschlossenen Nutzer sollte möglichst vollständig erreicht werden. Das verlangt eine einfache und übersichtliche Anwendung, die ohne aufwändige Einweisung

gen sofort und intuitiv bedient werden kann. Trotzdem darf das Vertrauen in die Funktionalität nicht beeinträchtigt werden. Und ganz wichtig ist, der Nutzer möchte einen Mehrwert erzielen und nicht etwa noch mit Zusatzkosten belastet werden.

Andererseits spielt der Nutzer auch auf der Anbieterseite eine wichtige Rolle. Jederzeit muss Klarheit über die Authentizität jedes Teilnehmers herrschen, da sonst eine ordnungsgemäße Abarbeitung des gesamten Verfahrens in Frage gestellt ist. Werden darüber hinaus nur berechtigte Nutzer zugelassen, spart dieses eine Menge Mehraufwand bei der Weiterverarbeitung eingegangener Daten.

Eingabe und Übertragung der Daten

Für die Dateneingabe kristallisieren sich bei genauerer Betrachtung der Anwenderseite drei unterschiedliche Möglichkeiten heraus. Die erste und heute wohl auch am weitesten verbreitete ist das Angebot eines Webformulars, in dessen Eingabefelder durch den Nutzer alle Angaben per Tastatur eingegeben werden. Dabei kann die Eingabe durch eine entsprechende grafische Gestaltung und durch zusätzliche Nutzerführung (Scripte) erleichtert und gesteuert werden.

Eine weitere Möglichkeit ergibt sich immer dann, wenn der Anwender entweder ein Programm hat oder man es ihm per Download zur Verfügung stellen kann, mit dem er selbst seine Daten erfassen und gegebenenfalls berechnen kann, und das am Ende der Bearbeitung eine entsprechende Datei erzeugt. Diese Datei wird nun mit Hilfe eines Upload-Vorganges zum Statistischen Landesamt übertragen.

Das dritte Angebot ist eine Kombination der beiden ersten. Es wird ein Webformular zur Verfügung gestellt, das über eine geeignete Schnittstelle aus einer Datei befüllt werden kann. Diese Methode kann immer dann eingesetzt werden, wenn der Nutzer seine Daten in Dateiform aus seiner hausinternen Datenverarbeitung exportieren kann.

Bei der Dateneingabe können durch geeignete Mittel bereits Daten plausibilisiert werden. Hier sollte aber mit großer Sorgfalt über den Umfang der Prüfungen und die Gestaltung der entspre-

chenden Meldungen nachgedacht werden. Im Grunde stehen sich Einfachheit und damit Akzeptanz eines Formulars und seine subtile Plausibilisierung gegenseitig im Wege. So bleibt die Online-Datenerfassung letztlich doch nur ein neuer Meldeweg an das StLA und kann nicht die Übertragung zusätzlicher, im Grunde fachstatistischer Arbeiten auf den Berichtspflichtigen bedeuten. Damit sind auch den mancherorts erwarteten Rationalisierungseffekten an dieser Stelle natürliche Grenzen gesetzt.

Die Übertragung der Daten zwischen den Rechnern der Nutzer und dem Server im Amt ist mit den Mitteln des Internets heute technisch überhaupt kein Problem mehr. Allerdings sind die Fragen der Datensicherheit, nämlich wer sendet wem Daten und kann zwischendrin mit Sicherheit kein Dritter mitlesen oder gar schreiben, mit besonderer Aufmerksamkeit zu beachten.

Weiterverarbeitung der Daten

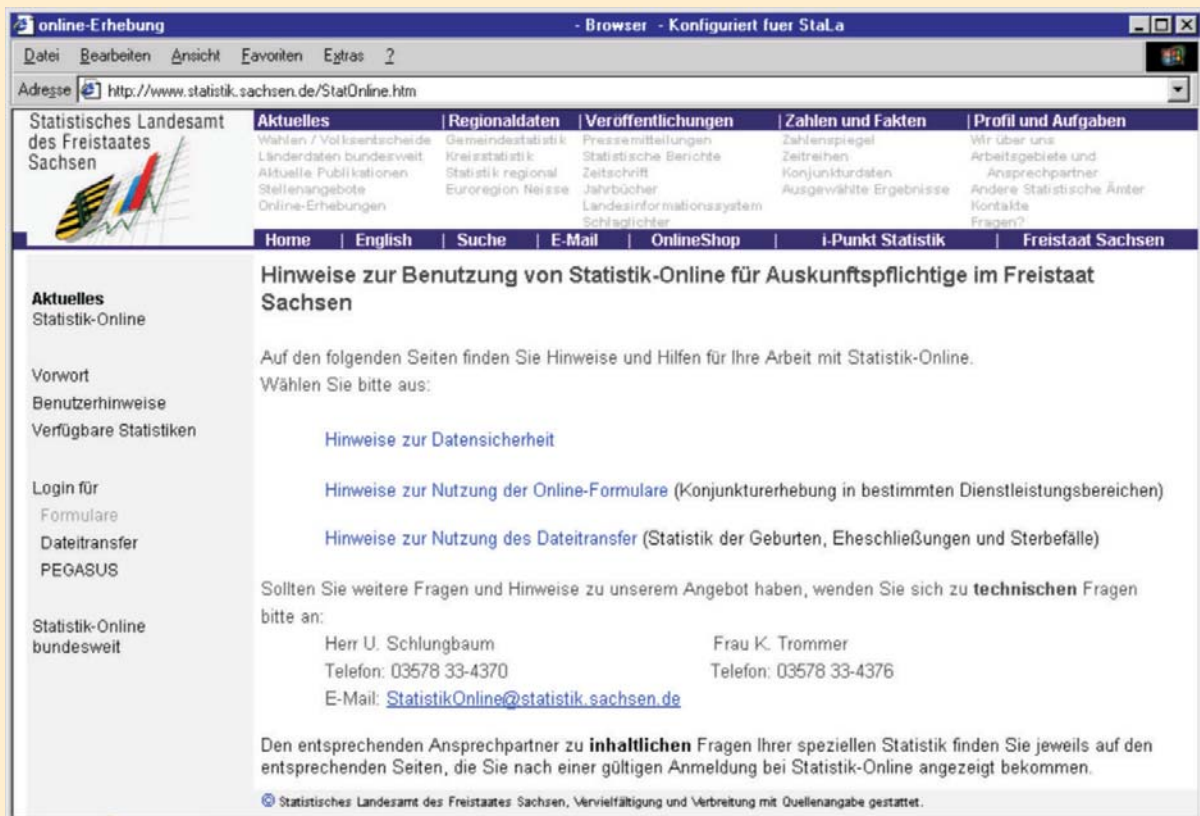
Dieser Komplex umfasst im Wesentlichen zwei Dinge. Daten, die online im StLA ankommen, müssen möglichst zeitgleich auch als Eingang erfasst werden. Das heißt noch nicht, die Dateninhalte

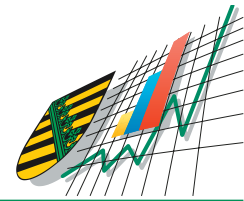
zu verarbeiten, aber die eigentliche Ankunft der Daten sollte dem Charakter des Mediums Internet entsprechend sofort signiert werden. Zu geeigneten Zeitpunkten sind dann die Dateninhalte in den bestehenden Statistikproduktionsprozess einzusteuern. Hierbei erwächst die eigentliche Schwierigkeit aus dem Bestandsschutz von EDV-Prozessen, die zu Zeiten konzipiert und programmiert wurden, als noch niemand an Online-Datenlieferung gedacht hat.

Vorschläge für die technische Umsetzung

Diese gewünschte Gesamtfunktionalität einer Online-Dateneinzugslösung für die amtliche Statistik war nun in Übereinstimmung mit den bestehenden technischen Möglichkeiten unter Berücksichtigung der vorhandenen IT-Infrastruktur in den einzelnen Ämtern zu bringen. Im Ergebnis dieses Diskussionsprozesses konnte durch die AG Internet ein Vorschlag entwickelt werden, wie eine Verbundlösung aussehen sollte. Die wesentlichen Eckpunkte, ohne auf Details einzugehen, seien hier genannt:

Abb. 1 Eingabemaske mit Login für Online-Erhebungen





- Formulare auf HTML-Basis,
- Formulare in Karteikartentechnologie,
- Datenübertragung über das HTTPS-Protokoll,
- serverseitige Verarbeitung,
- Nutzung von XML (DatML),
- Java-Programmierung,
- Nutzung von Open-Source-Produkten.

Interessanterweise sind alle diese Punkte auch Bestandteile des SAGA-Standards [1] der Koordinierungs- und Beratungsstelle der Bundesregierung für Informationstechnik (KBSt) und des Bundesamtes für Sicherheit in der Informationstechnik (BSI), der im Februar 2003 veröffentlicht wurde.

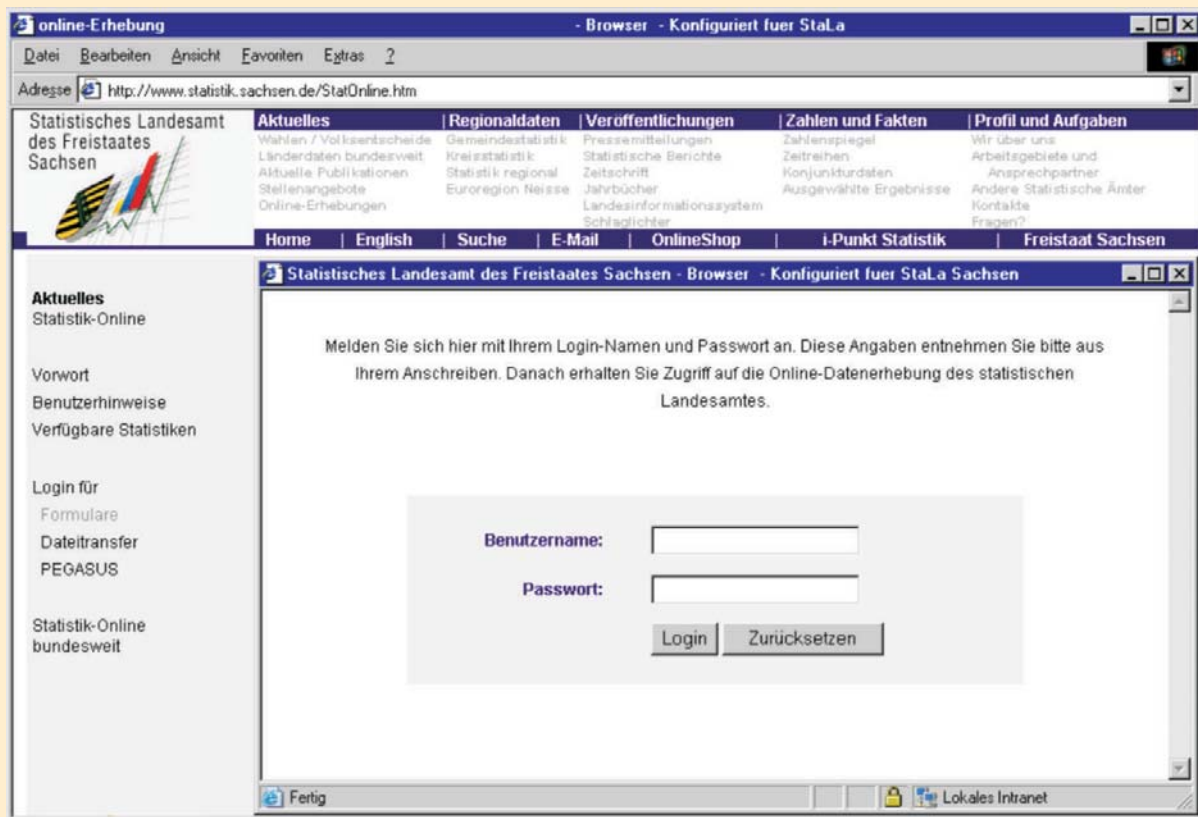
die Visualisierung der Formulare bereitzustellen. Dieser wurde dann recht kurzfristig über eine Lösung mit Perlscript zu einer einsatzfähigen Anwendung ausgebaut, befand sich im LDS NRW seit Januar 2001 zunächst im Testeinsatz und ist heute als Produktivsystem aktiv. Abweichend vom gerade entwickelten Verbundkonzept ist diese Lösung nicht ohne weiteres portabel, da sie nicht plattformunabhängig entwickelt wurde. Für notwendige Anpassungen zur Implementierung dieser Lösung in die bestehenden und teils recht unterschiedlichen IT-Infrastrukturen in den einzelnen Landesämtern sind spezielle IT-Kenntnisse nötig, die häufig nicht abgedeckt werden können. Wohl hauptsächlich aus diesem Grund beschloss ein Großteil der Statistischen Landesämter, das Web-Hosting-Angebot aus dem LDS NRW anzunehmen. Über die Einführung dieses Verfahrens wird ausführlich in den Statistischen Monatsheften Rheinland-Pfalz [2] berichtet.

Vorhandene Lösungen

Das Landesamt für Datenverarbeitung und Statistik Nordrhein-Westfalen (LDS NRW) erklärte sich bereit, zur Unterstützung der technischen Diskussion der AG Internet einen Prototyp für

Im Statistischen Landesamt in Sachsen wurde bereits sehr viel früher begonnen, mit kleineren Einzellösungen technologische Erfahrungen für die Entwicklung von Online-Dateneinzugslösungen unter Verwendung von Java zu sammeln. So entstand bereits im Sommer 2000 eine erste Online-Erhebung für eine

Abb. 2 Eingabemaske zur Nutzeranmeldung



Registerumfrage und in der Folgezeit eine Datei-Uploadlösung für die Pflegestatistik. Aufbauend auf diesen Anfangserfolgen gab es - begleitend zu den Anstrengungen in der Verbundarbeitsgruppe - erste Überlegungen zu einer komplexen Dateneinzugslösung.

Einerseits ergab die Prüfung des Erhebungs-Management-Systems EMSy [3] auf Tauglichkeit für den Aufbau von Nutzerkreisen für eine Online-Erhebung sehr schnell einen Erfolg. So ist es möglich, in EMSy zusammen mit dem herkömmlichen Postversand eine zusätzliche Datei zu generieren, die alle benötigten Nutzerangaben für eine Online-Erhebung bereitstellt. Damit war auf verblüffend einfache Art dieses eine Problem bereits gelöst, es musste nur noch ein verbindliches Format definiert und natürlich die entsprechende Schnittstelle programmiert werden.

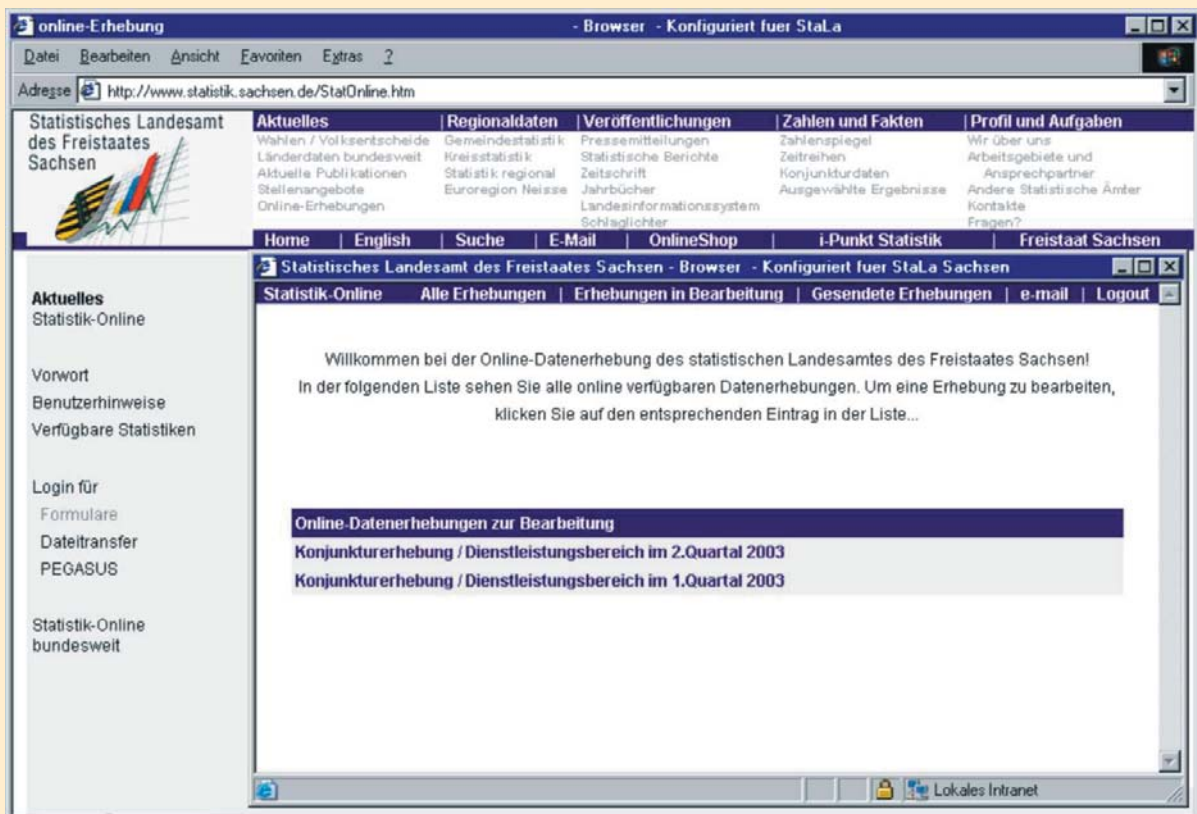
Bei den Überlegungen zum Trägermodul für die Online-Formulare wurde neben anderen auch eine gerade an der TU Dresden in Zusammenarbeit mit der Firma xima media GmbH entstehende Diplomarbeit in die Überlegungen mit einbezogen. Dabei wurde schnell deutlich, dass die theoretischen Ansätze dieser Arbeit

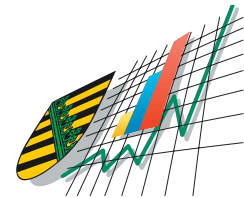
genau unserer angestrebten Funktionalität entsprachen. Der Diplomand seinerseits war an einem Gesprächspartner mit dem praktischen Anwendungshintergrund für solche Formularframeworks interessiert. Daraus ergab sich eine Kooperationsvereinbarung zwischen dem Statistischen Landesamt des Freistaates Sachsen und der xima media GmbH. Auf der Basis dieser Vereinbarung konnten die ursprünglich vorliegenden Ergebnisse der Diplomarbeit zu einer leistungsstarken Anwendung „xima-forms“ weiter entwickelt werden, die sich mittlerweile im Produktivbetrieb erfolgreich bewährt hat.

Beispiel Konjunkturstatistik in bestimmten Dienstleistungsbereichen

Das Statistische Landesamt startete Anfang April 2003 für die erstmalige vierteljährliche Erhebung zur Konjunkturstatistik in bestimmten Dienstleistungsbereichen die erste echte formularbasierte Dateneinzugseinzugslösung im Internet in Sachsen. Hierbei wurde der Gesamtheit der Berichtspflichtigen mit dem

Abb. 3 Eingabemaske mit Formularliste





regulären Aufforderungsschreiben zur Auskunftserteilung der personalisierte Zugang zur Erhebung bekannt gemacht, und ihnen wurde angeboten, diesen neuen Weg zu nutzen. Allen wurden vorsorglich ihre individuellen Zugangsdaten mitgeteilt. Der dafür notwendige Berichtskreis konnte auf einfache Weise aus dem Erhebungsmanagementsystem EMSy entladen werden, das diese Erhebung steuert. Dieses Vorgehen unterscheidet sich insofern von dem im LDS NRW entwickelten Verfahren, da dort zunächst eine gesonderte Ermittlung und Registrierung der Teilnahmewilligen vorangehen muss.

Das Verfahren aus der Sicht des Nutzers

Unter der im Anschreiben angegebenen URL (<http://www.statistik.sachsen.de/StatOnline.htm>) befindet sich neben den ausführlichen Nutzerhinweisen auch der Link zum Login für die Online-Erhebung (vgl. Abb. 1).

Hier müssen das persönliche Nutzerkennzeichen und das Passwort eingegeben werden (vgl. Abb. 2).

Bei erfolgreicher Anmeldung wird durch das Formularframework erkannt, zu welchen Statistiken der angemeldete Nutzer berichtspflichtig ist, und er bekommt eine Liste mit den entsprechenden Formularen angeboten, aus denen er jetzt wählen kann (vgl. Abb. 3).

Das Formular wird durch Tastatureingaben gefüllt, wobei einzelne Eingabefelder mit formularseitigen Prüfungen (Javascript) hinterlegt sein können. Diese Prüfungen beziehen sich zum Beispiel auf das Eingabeformat oder andere formale Merkmale. So wird für den Umsatz einer Firma beispielsweise eine rein numerische Eingabe ohne Nachkommastellen erwartet (vgl. Abb. 4).

Das Ausfüllen eines Formulars kann durch den Nutzer jederzeit unterbrochen werden. In diesem Fall besteht die Möglichkeit, den erreichten Arbeitsstand serverseitig zwischenspeichern. Nach Wiederanmeldung bei Statistik-Online kann die Arbeit mit diesem Stand dann fortgesetzt werden, bereits eingegebene Inhalte müssen nicht erneut eingegeben werden. Nach Abschluss der Bearbeitung wird ein Formular an das StLA gesendet, wobei auch hier die Möglichkeit einer Wiederholung bis zum Abschluss

Abb. 4 Eingabemaske für die Konjunkturerhebung in bestimmten Dienstleistungsbereichen

Konjunkturerhebung in bestimmten Dienstleistungsbereichen 2003
 Verkehr und Nachrichtenübermittlung
 Datenverarbeitung und Datenbanken
 Erbringung von wirtschaftlichen Dienstleistungen

Statistisches Landesamt des Freistaates Sachsen

Allgemein | Schwerpunkt | Umsatz, Personen | Niederlassungen | Erläuterungen

III Angaben zu Niederlassungen bei großen Unternehmen

Frage 1 Hatte Ihr Unternehmen/Einrichtung im Geschäftsjahr 2002 Umsatz / Einnahmen aus selbständiger Tätigkeit von 250 000 EUR und mehr ?
 NEIN -> Eine weitere Beantwortung folgender Fragen ist nicht erforderlich
 JA -> Bitte weiter mit Frage 2

Frage 2 Hat Ihr Unternehmen/Einrichtung Niederlassungen einschließlich der Hauptniederlassung in mehreren Bundesländern ?
 NEIN -> Eine weitere Beantwortung folgender Fragen ist nicht erforderlich
 JA -> Bitte teilen Sie die in II gemachten Angaben auf die Niederlassungen einschließlich der Hauptniederlassung in den Bundesländern auf.

Bundesländer	Umsatz/Einnahmen aus selbständiger Tätigkeit im aktuellen Berichtsquartal EUR	Tätige Personen am letzten Tag des Berichtsquartals Anzahl
08 Baden-Württemberg	12743,50	5
09 Bayern	12700	7
11 Berlin		

Bitte eine positive ganze Zahl eingeben.

der Erhebung besteht. Es wird immer die letzte Sendung als gültige Korrektur anerkannt und entsprechend weiterverarbeitet.

Jede Datensendung eines Nutzers wird mit einer Quittung beantwortet. Diese Seite beinhaltet alle originalen Karteikarten des Formulars mit den eingetragenen Daten. Über eine Zoomfunktion kann diese Seite auf Druckerbreite gebracht, bei Bedarf durch den Nutzer gedruckt und zu den Akten genommen werden. Als

zweite Möglichkeit kann ein Download dieser Seite durch den Nutzer für eine elektronische Speicherung dieser Quittung ausgelöst werden (vgl. Abb. 5).

Damit ist die eingangs erwähnte Forderung nach Einfachheit und Transparenz für den Nutzer soweit wie möglich erfüllt. Durch die Einstellung eines ordentlichen ssl-Zertifikates entstehen beim Nutzer auch keine Irritationen oder gar Zusatzaufwand für die

Abb. 5 Quittung des Formulars mit den gemeldeten Angaben

Das Formular wurde an die Behörde gesendet.
Ihre Angaben wurden unter der Bearbeitungsnummer 10625769619891014 abgelegt.
Drucken Sie sich diese Seite bitte aus.
Bei Rückfragen geben Sie bitte Ihre Bearbeitungsnummer an.
Der Ausdruck besteht aus 5 Seiten.

Konjunkturerhebung in bestimmten Dienstleistungsbereichen 2003
Verkehr und Nachrichtenübermittlung
Datenverarbeitung und Datenbanken
Erbringung von wirtschaftlichen Dienstleistungen

Statistisches Landesamt des Freistaates Sachsen

Algemein | Schwerpunkt | Umsatz, Personal | Niederlassungen | **Erhebungen**

Angaben zu Ihrer Anschrift

Ihre Kennnummer: 0000915000

Ihre Anschrift:	vorliegende Angaben	Korrektur
Firmenname	Oilbeantzer GmbH	
	Testbetrieb	
Strasse	Oilbeeweg 1	
PLZ	02345	
Ort	Ilberetha	

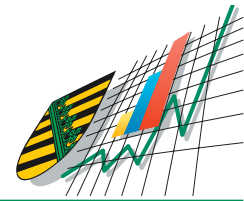
Ihre Fragen beantwortet:
Frank Kötzig
03578 33 - 3255
konjunktur-dig@statistik.sachsen.de

An wen können wir uns bei Rückfragen wenden:
Name:
Telefon:
e-mail:

Konjunkturerhebung in bestimmten Dienstleistungsbereichen 2003
Verkehr und Nachrichtenübermittlung
Datenverarbeitung und Datenbanken
Erbringung von wirtschaftlichen Dienstleistungen

Statistisches Landesamt des Freistaates Sachsen

Fertig Lokales Intranet



verschlüsselte Datenübertragung. Es erscheint lediglich der Standardhinweis des benutzten Webbrowsers, dass jetzt auf eine gesicherte Kommunikation gewechselt wird.

Das Verfahren aus der Sicht der Administration

Die Bereitstellung der Nutzerdatei zum Laden in das Formularframework xima-forms erfolgt gleichzeitig mit dem Postversand aus EMSy. In diesem Arbeitsschritt wird zusätzlich die benötigte Schnittstellendatei geschrieben. Für die Bereitstellung des benötigten Online-Formulars gibt es den „xima-forms designer“, ein Werkzeug, mit dem im Baukastenprinzip entsprechende Formulare einschließlich eventuell benötigter Plausibilisierungen entwickelt werden können. Aus diesem Designer werden die Formulare direkt auf den Webserver geladen.

In einem weiteren Arbeitsschritt müssen jetzt noch einige wenige Konfigurationen durch den xima-forms-Administrator vorgenommen werden. Einzustellen sind die Dauer der Veröffentlichung des Formulars, das Zielverzeichnis, und es sind die Nutzergruppe und das Formular zueinander zuzuordnen. - Damit ist die Erhebung für den Dateneinzug per Internet vorbereitet.

Die daraufhin eingehenden Datensätze werden automatisch zum Eingangszeitpunkt durch EMSy eingangsregistriert und vom Formularserver entfernt, indem sie in ein Arbeitsverzeichnis des zuständigen Fachbereiches verschoben werden. Diese Dateien liegen im DatML-Format vor. Für künftige Entwicklungen wird im statistischen Verbund angestrebt, dass alle Statistikdaten verarbeitende Programme dieses Format verstehen. Da das umfangreiche Umstellungsarbeiten erfordert, wird für eine Übergangszeit ein Programm benötigt, das die DatML-Dateien in das sequenzielle Datenformat für die herkömmlichen Eingangsschnittstellen konvertiert. Ein solches Programm wurde im Rahmen der Zusammenarbeit des Statistischen Landesamtes mit der Berufsakademie Bautzen als Ergebnis einer Diplomarbeit entwickelt und steht künftig zur Verfügung. [4]

Erste Ergebnisse

Nachdem die ersten beiden Erhebungszeiträume (1. und 2. Quartal 2003) beendet sind, zeigt sich folgendes Ergebnis:

	1. Quartal	2. Quartal
Anzahl der Berichtspflichtigen:	2 328	2 194
Anzahl der eingegangenen Meldungen:	1 897	1 792
darunter Internetmeldungen:	257	262
Anteil der Online-Melder (Prozent):	13	14

Die Beteiligung der Auskunftspflichtigen am Online-Verfahren liegt damit über den Erwartungen. Dabei dürfte sicherlich eine wesentliche Ursache im besonders einfachen Zugang zum Verfahren zu suchen sein.

Ulrich Schlungbaum,
Dipl.-Math., Sachgebietsleiter WebRedaktion

Quellen- und Literaturverzeichnis:

- [1] SAGA - Standard und Architekturen für E-Governmentanwendungen. Internet: <http://www.kbst.bund.de/saga>
- [2] Klug, S.: Einführung des Verbundverfahrens. In: Statistische Monatshefte Rheinland-Pfalz, 6/2003, S. 99 - 100.
- [3] Vgl. Lenk, K., J. Lesch: Einführung eines allgemeinen Erhebungs-Management-Systems. In: Statistik in Sachsen, 3-4/2001, S. 30 – 33.
- [4] Fiedler, S.: Entwicklung eines Konvertermoduls zur Transformation zwischen modernem Dateneinzugs- und bestehenden Produktionsverfahren in der amtlichen Statistik, Diplomarbeit, 2003, Berufsakademie Sachsen, Staatliche Studienakademie Bautzen