

Günter Krombholz/Volker Blank

## Moderne Techniken in der Statistik

*Die Statistik setzte schon frühzeitig für die Erstellung statistischer Ergebnisse die Mittel der Datenverarbeitung ein. In der Zeit der ehemaligen DDR wurde insbesondere auf den Einsatz von Großrechnern orientiert. Nach der Wende kam es mit dem Aufbau des Amtes zu einer stürmischen Entwicklung in der arbeitsplatzbezogenen Computertechnik. Zunehmende Vernetzung und Verteilung der Statistikerarbeiten auf unterschiedliche Rechnerebenen prägen den letzten Abschnitt dieser Entwicklung.*

### Einleitung

Die amtliche Statistik hat seit ihrem Bestehen die jeweils vorhandenen modernen Aufbereitungstechniken für die Bewältigung ihrer Massenarbeiten eingesetzt.

Schon frühzeitig war zu erkennen, daß allein mit manuellem Aufwand die gestellten Ziele, schnell und aktuell Daten für Entscheidungsprozesse und Planungszwecke bereitzustellen, nicht erreicht werden konnten. Bereits um die Jahrhundertwende wurden zur Bewältigung von großen Zählungen Sortier- und Tabelliermaschinen von der Deutschen Hollerith-Gesellschaft im damaligen Kaiserlichen Statistischen Amt angemietet.

Aufgaben und Umfang in der amtlichen Statistik haben sich im Laufe der Zeit grundlegend gewandelt. Geblieben ist der Anspruch, neue und zusätzliche Aufgaben mit modernen Mitteln schnell einer Lösung zuzuführen und die Qualität der Ergebnisse der statistischen Arbeit immer weiter zu verbessern. Dies zeigt sich gerade in den vergangenen Jahren, wo neben den bereits über einen längeren Zeitraum laufenden Erhebungen zunehmend Aufgaben im Rahmen der statistischen Arbeiten der Europäischen Union zu realisieren sind und die Komplexität der damit verbundenen Aufgabenstellungen ständig zunimmt.

Gleichzeitig werden die Anforderungen an die amtliche Statistik immer größer und differenzierter. Und das vor dem Hintergrund einer komplexen, wirtschaftlichen, konjunkturellen und gesellschaftlichen Entwicklung.

Der technische Fortschritt in den letzten Jahren hat dazu geführt, daß ganze Berufe und Berufsbilder von der Bühne der Arbeitswelt verschwunden sind, jedoch sind auch neue entstanden, die zunehmend die Flexibilität und Kreativität des Einzelnen fordern.

Gerade das Aufkommen der Mikroelektronik hat fast alle Berufsbilder der Gesellschaft entscheidend verändert. Durch den breiten Einsatz von Personalcomputern gibt es keinen Zweig der Wirtschaft mehr, der davon nicht betroffen ist. Selbst in der Freizeitwelt der Menschen hat er seinen Siegeszug angetreten. Damit werden aber auch Chancen einer Arbeiterleichterung bei veränderten Arbeitsabläufen sowie einer Befreiung von lästigen und zeitaufwendigen Routinearbeiten eröffnet.

### Rückblick auf die Zeiten der Staatlichen Zentralverwaltung für Statistik (SZS)

Mit der „Verordnung über die Aufgaben und die Organisation der Statistik in der DDR“ im Jahre 1956 wurde als Grundstein der Anwendung der Rechentechnik bei der Staatlichen Zentralverwaltung für Statistik ein Rechenbetrieb gebildet. Dieser Rechenbetrieb sollte für bessere Voraussetzungen bei der zukünftigen Anwendung von mechanisierten und automatisierten Methoden in der Statistik sorgen.

In den einzelnen Bezirksstellen wurden die gebildeten maschinellen Abteilungen in den Folgejahren ausgegliedert und ein Rechenbetrieb, der VEB Maschinelles Rechnen, in den damaligen Verwaltungsgliederungen, den Bezirken gegründet.

Die Rechenbetriebe aller Bezirke wurden als Vereinigung Volkseigener Betriebe (VVB) der Leitung der Staatlichen Zentralverwaltung für Statistik unterstellt. Daran änderte auch die spätere Zusammenfassung der Datenverarbeitungszentren zum Kombinat Datenverarbeitung nichts. Die amtliche Statistik nutzte von Anfang an die Rechenkapazität dieser geschaffenen Einheiten, hatte mit dem Rechenzentrum für Statistik bei der Staatlichen Zentralverwaltung für Statistik sogar einen eigenen Rechenbetrieb.

Mit Beginn der sechziger Jahre kam die erste Technik, Robotron 100, ein Lochkartenrechner zum Einsatz. Bei diesem Gerät wurde die Schaltung noch „gestöpselt“. Ende der sechziger Jahre kam der Robotron R300 zum Einsatz, ein Rechner, der bereits einen Hauptspeicher von 30.000 Zeichen besaß und maschinenorientiert programmierbar war. Bereits bei der Nutzung dieser Rechner spielte die Statistik eine Vorreiterrolle, ihre Verfahren und Projekte waren die ersten, die auf diesen Rechnern liefen.

1975 wurde der Robotron R21 in Betrieb genommen, welcher bereits einen Hauptspeicher von 65.535 Zeichen besaß und eine, wenn auch geringe Multiverarbeitung zuließ. Mit dem Jahr 1982 erfolgte der Einsatz eines ES1040 mit einem Hauptspeicher über 1 Million Zeichen, welcher 1987 vom ES1055 und ES1057 mit über 4 Millionen Zeichen abgelöst wurde. Diese Anlagen ließen schon eine Mehrfachprogrammabarbeitung zu und entsprachen in ihrer Leistungsklasse der 360-iger und 370-iger Serie der IBM. Die Programme wurden zentral erstellt und meistens auf der Basis der Programmiersprachen ASSEMBLER oder PL/1 implementiert.

Die Statistikverfahren lasteten die vorhandenen Rechner in den Datenverarbeitungszentren vor allem in den ersten Arbeitstagen mit der monatlichen Berichterstattung fast vollständig aus. Der damalige Ablauf sah eine Fertigstellung der monatlichen Berichterstattungen innerhalb der jeweils ersten drei Arbeitstage des Monats vor. Die Daten wurden anschließend zur Staatlichen Zentralverwaltung für Statistik geliefert. Dies erfolgte am Anfang in Magnetbandform über einen Kurierdienst mit Pkw, später auch in einer Datenübermittlung über das Telefonnetz.

Mit den zuletzt eingesetzten Anlagen war auch ein Übergang zur Terminalarbeit, zur dialoggestützten Arbeit möglich. Und die Frage einer statistischen Datenbank wurde gelöst. Diese Datenbank (SPAZ) nahm alle statistischen Daten der Erhebungen auf und gestattete einfach zu parametrierende Abfragen und ein standardmäßig erzeugtes Tabellenprogramm.

Jedoch zeigte sich, daß bestimmte Kennziffern und Merkmale auch vor Ort in den vorhandenen Kreisstellen gebraucht wurden. Deshalb kamen ab 1987/1988 auch die ersten Bürocomputer auf 8-Bitbasis zum Einsatz. Jetzt wurden Daten aus der vorhandenen Datenbank des Bezirkes am Großrechner auf Magnetband abgezogen und mittels eines Bürocomputers A5130 auf Diskette konvertiert und an die damals vorhandenen Kreisstellen versandt. Zum

Einsatz kamen Bürocomputer des Typs A5120, A5130 und PC1715, alles Rechner der 286-Klasse. Im letzten Jahr vor der Wende 1989 wurden mit dem A7100 erstmalig Personalcomputer der 16-Bit-Klasse eingesetzt. Mit dieser Arbeitsplatztechnik wurden in hohem Maße Aufgaben der Analyse- und Informationstätigkeit im damaligen statistischen Berichtswesen realisiert.

### Die Zeit nach der Wende

Ab 1990 lösten sich zunehmend die Bindungen der Betriebe mit der Staatlichen Zentralverwaltung. Mit dem wiedergeschaffenen Land Thüringen bildete sich nach und nach auch der föderative Gedanke in der Statistik wieder heraus.

Als wichtiges technologisches Mittel zur Bewältigung der ersten freien Volkskammerwahl wurden zum damaligen Zeitpunkt für die 3 Bezirke Erfurt, Gera und Suhl jeweils Midrangerechner des Typs IBM AS/400 angeschafft. Diese Geräte waren jeweils mittels Modem mit den in den Kreisstellen ebenfalls neu installierten PS/2-Systemen (286-Prozessor) verbunden und ermöglichten damit eine Ferneingabe der Wahlergebnisse in den vorhandenen Kreisstellen. Diese Technik wurde auch noch zu den im Herbst 1990 stattfindenden Bundestagswahlen genutzt. Dabei bestanden in beiden Wahlen noch Fernverbindungen nach Berlin und zwischen den Lokationen Erfurt, Gera und Suhl.

1991 wurde zu einem Umbruchjahr in der Statistik in Thüringen.

Die drei Bezirke gaben ihre Eigenständigkeit auf. Das Thüringer Landesamt für Statistik befand sich in Gründung. Neben einer Verwaltungsabteilung und drei statistischen Fachabteilungen entstand eine Abteilung Maschinelle Datenverarbeitung. Verbunden mit der Schaffung dieser Abteilung, die am Anfang aus nur 3 Mitarbeitern bestand, erfolgte eine Orientierung der Abarbeitung der zukünftig zu bewältigenden statistischen Arbeiten und eine Konzentration der datenverarbeitungstechnischen Leistungen auf einen Standort. Als technische Hinterlassenschaft waren lediglich einige überalterte Bildschirmterminals, die drei AS/400 und Personalcomputer aus DDR-Produktion vorhanden.

Nach der Wende wurden alle statistischen Verfahren der alten DDR nach und nach eingestellt. Um schnell wieder statistische Daten bereitzustellen, stand Anfang 1991 nun die Aufgabe, die Verfahren der Bundesstatistik auch im

neuen Bundesland Thüringen zu übernehmen. Nach einer personellen Erweiterung der Abteilung durch erfahrene Datenverarbeitungskräfte wurde dies mit dem Ziel einer weitestgehenden Eigenleistung in Angriff genommen. Die vorhandenen drei AS/400 in den Lokationen Erfurt, Gera und Suhl wurden mittels Datenfernverarbeitung vernetzt und damit auch der Zugriff von allen drei Standorten auf den Großrechner IBM 4381, einer Anlage mit über 32 Millionen Zeichen und einer Rechnerleistung von 6 Millionen Operationen in der Sekunde, im damaligen Rechenzentrum und Softwarehaus Thüringen ermöglicht. Parallel dazu erfolgte in Erfurt durch die Installation eines Terminalsteuergerätes und 16 Terminals die Schaffung der Voraussetzungen für eine Einführung der statistischen Verfahren am Großrechner.

Auch wurden Ende 1991 die Arbeiten begonnen, sich an das Datenübermittlungssystem der Statistischen Landesämter anzuschließen. Der Datentransfer über DATEX-P wurde Anfang 1992 als Testbetrieb aufgenommen und im Laufe des Jahres voll in Betrieb genommen, so daß der aufwendige Transport von magnetischen Datenträgern in andere Statistische Landesämter und in das Statistische Bundesamt entfielen.

Damit war erst einmal die Arbeitsfähigkeit der Datenverarbeitung im Thüringer Landesamt für Statistik hergestellt und die Datenfernverarbeitung ermöglichte einen problemlosen Zugriff und Datentransfer an die drei Standorte. Es zeigte sich jedoch sehr bald, daß diese Infrastruktur nicht ausreichend für alle zukünftigen Arbeiten der Statistik war. Deshalb begannen schon 1991 Planungen für eine Inhouse-Vernetzung jeden Arbeitsplatzes, die Ende 1991 bereits konzeptionell abgeschlossen wurden.

Das Herz einer jeden Informationstechnikstruktur bildet das darunterliegende Netzkonzept. Es bestimmt in entscheidendem Maße, wie effizient man in einer neuen Umgebung mit den einfließenden Informationen arbeiten kann. Die saubere Definition von Schnittstellen im Rahmen der aufzubauenden Architektur, eine strukturierte Verkabelung bis an jeden Arbeitsplatz bietet den Vorteil, daß teure Neuverkabelungen bei Umzügen von Mitarbeitern entfallen. Deshalb wurden gerade in die konzeptionelle Gestaltung viel Kraft und Ideen gesetzt.

Auch wurden sich bereits Gedanken einer Speicherung der Daten in eine eigene Datenbank gemacht, die mit einer Untersuchung der in anderen Landesämtern betriebenen Lösungen begann. Dabei zeigte es sich, daß hier unter-

schiedliche Lösungswege gegangen wurden. Die Entscheidung fiel auf das Datenbankbetriebssystem ADABAS als für die zu diesem Zeitpunkt für das Amt günstigste Lösung. Diese Datenbanklösung, das Statistische Informationssystem, wurde von der Software AG für einen Verbund von Statistischen Ämtern der Städte und Länder entwickelt. Sie läßt eine Gliederung der Daten nach räumlichen, zeitlichen und fachlichen Aspekten zu und ermöglicht mit ihren beschreibenden Daten einen relativ einfachen Zugriff.

### **Zeit ab Gründung des Thüringer Landesamtes für Statistik**

Im Jahre 1992 wurde der Grundstein für eine moderne, modular aufgebaute Inhouse-Vernetzung gelegt. Dieser besteht aus einem Glasfaserbackbone, welcher 2 Gebäude und innerhalb der Gebäude die Etagen miteinander verbindet. Die einzelnen Etagen wurden mittels Kupferkabel vernetzt. Jeder Raum, jeder Arbeitsplatz war somit erreichbar. Begonnen wurde 1992 in 2 Etagen des Gebäudes, in den folgenden 2 Jahren folgten jeweils eine weitere Etage. Ebenfalls erfolgte in den Standorten Gera und Suhl eine Verkabelung. Integriert in diese Vernetzung wurde ein Zugang zum Großrechner, welcher 64 Terminaldialoge zuläßt. Traditionell sind fast alle statistischen Erhebungen ursprünglich für den Großrechnereinsatz konzipiert worden. Mit dem zunehmenden Einsatz von Personalcomputern, den damit verbundenen verbesserten Arbeitsmöglichkeiten und Effektivitätsverbesserungen verlagern sich jedoch zunehmend immer mehr Arbeitsprozesse und komplette statistische Erhebungen auf die Ebene der Arbeitsplatzcomputer. Das entstandene Netz läßt eine verteilte Verarbeitung zu, so daß Arbeitsstationen auf die im Netz vorhandenen Dienste der Datenhaltung, Kommunikation und Archivierung zugreifen können.

Das gesamte Netz wurde so segmentiert, daß die vorhandenen Datenserver immer eine Abteilung bedienen und dadurch der Verkehr auf dem Netz minimiert wird. Auch erfolgte über das Landesdatennetz eine Kopplung der Netze zwischen den Standorten. So ist es heute möglich, problemlos Dateien und Listen zwischen den Standorten auszutauschen, von den Standorten Gera und Suhl Dienste der Datenverarbeitung in Erfurt zu nutzen und damit eine schnelle Kommunikation zu ermöglichen. Im Laufe der vergangenen drei Jahre stand im Vordergrund, das Netz auf Grund der rasanten Entwicklung gerade auf dem Personalcomputersektor an die neuen Bedingungen anzupassen. So erfolgte der Einsatz von neuen Softwarepaketen und verbesserter Netzwerktechnologie.

Um die Veröffentlichungen des Thüringer Landesamtes für Statistik in einer hohen Qualität und aktuell erstellen zu können, wurden zwei Grafikarbeitsplätze bereits 1992 auf Basis von Apple Macintosh-Gerätetechnik installiert. Damit waren neben der layoutmäßigen Gestaltung von Veröffentlichungen auch die Neukonzipierung von Erhebungsunterlagen und Vordrucken schnell und mit geringem Aufwand zu lösen. Mittels hochauflösender Belichtungs-technik erfolgt dann in der hauseigenen Druckerei eine entsprechende Vervielfältigung.

Auch Möglichkeiten der Öffentlichkeitsarbeit wie die Teilnahme am Bildschirmtext (Btx) innerhalb des gemeinsamen Programms aller Statistischen Landesämter und des Bundesamtes wurden realisiert. Die Möglichkeit einer Faxsendung aus dem Inhousesnetz wurde integriert und wird heute insbesondere für den Versand der Pressemitteilungen des Amtes genutzt.

Im Jahre 1993 erfolgte der Einsatz eines größeren Rechners im Thüringer Landesrechenzentrum. Es wurde eine IBM 9121 mit über 128 Millionen Zeichen installiert, die Rechengeschwindigkeiten von über 15 Mio. Operationen in der Sekunde erlaubte. Dadurch verbesserten sich die Arbeitsbedingungen der betreffenden Mitarbeiter des TLS bei der maschinellen Aufbereitung der statistischen Erhebungen entscheidend. Wert gelegt wurde in allen technischen Weiterentwicklungen im Amt immer auf eine Verbesserung der Arbeitsumgebung für die Mitarbeiter, auf zu erzielende Verbesserungseffekte im Arbeitsablauf und auf eine Integration aller technischen Möglichkeiten an den Arbeitsplatz des Mitarbeiters.

Ende 1992 war die Übernahme der mehr als 240 statistischen Erhebungen weitestgehend abgeschlossen. Nunmehr lagen die Anstrengungen im Großrechnerbereich - neben der normalen Abarbeitung der statistischen Erhebungen - vorwiegend an der Beteiligung in der Programmierung von Verfahren, die im Verbund aller statistischen Landesämter gemeinsam entwickelt werden. Hier konnte auf die Qualifikation der Mitarbeiter des TLS aufgebaut werden, die über eine zum Teil langjährige Erfahrung in der Programmierung verfügen.

Auch hier war schnell der Anschluß an das Niveau der alten Bundesländer erreicht. Ab Ende 1993 erreichte man eine neue Qualität, als Sprache der 4. Generation wurde nun vorwiegend NATURAL eingesetzt. Damit wurde auch nahtlos ein Übergang zu Datenbankanwendungen geschaffen.

In der Datenbankarbeit wurde der beschrittene Weg weiter gegangen und das Statistische Informationssystem mit Daten gefüllt. So ist es jetzt möglich, als externer Nutzer auf dieses Informationssystem zuzugreifen. Zunehmend erfolgt für die öffentlichen Verwaltungen der Anschluß an das Statistische Informationssystem über das Landesdatennetz.

Mit dem Internet, einem Kommunikationsweg für derzeit weltweit mehr als 32 Millionen Menschen, eröffnen sich hier wahrscheinlich weitere Spektren einer noch effizienteren Öffentlichkeitsarbeit. Nachdem bereits erste Anfänge einer Nutzung in der Thüringer Landesverwaltung zu verzeichnen sind, könnte diese Art der Präsentation gerade im Hinblick auf eine weitere Verbesserung der Aktualität und Präsenz von Daten des Thüringer Landesamtes für Statistik künftig eine wichtige Rolle spielen.