

91. DGINS-KONFERENZ

KOPENHAGEN, 26. + 27. MAI 2005

Die nationalen statistischen Systeme und die Forschung

Svante Öberg
Director General, Statistics Sweden

Die nationalen statistischen Systeme und die Forschung

Svante ÖBERG
Director General, Statistics Sweden

Im vorliegenden Papier befasse ich mich mit den Beziehungen zwischen den nationalen statistischen Systemen und der Forschung aus der Sicht des Statistikproduzenten. Ich spreche mich für den kostenlosen Internet-Zugang zu Veröffentlichungen und Datenbanken, die Entwicklung von Systemen für den Online-Zugang zu Mikrodaten für Forscher und die Verbesserung der Zusammenarbeit mit Forschungseinrichtungen aus. Als Beispiel ziehe ich dabei die Situation in Schweden heran. Ich stelle diese Überlegungen zur Diskussion, ohne den Anspruch zu erheben, zu wissen, wie gut sie für die einzelnen statistischen Ämter geeignet wären.

1. Einleitung

Die Forschungsgemeinschaft ist eine wichtige Gruppe von Nutzern statistischer Daten. Für uns in den nationalen statistischen Ämtern (NSÄ) ist es äußerst wichtig, dass wir den Forschern die Statistiken und Mikrodaten an die Hand geben, die sie als Grundlage für ihre Forschungsarbeit benötigen. Im vorliegenden Papier präsentiere ich einige Überlegungen, wie wir der Forschung künftig besser zuarbeiten können, indem wir über das Internet freien Zugang zu Datenbanken anbieten und insbesondere einen besseren Zugriff auf Mikrodaten, was ich als nächsten Schritt auf dem Weg zu einer noch flexibleren Art der Verbreitung statistischer Daten sehe. Ferner stelle ich Überlegungen darüber an, wie bei den NSÄ wie auch auf europäischer Ebene die Zusammenarbeit mit der Forschung weiterentwickelt werden kann.

2. Freier Zugang zu Veröffentlichungen und Datenbanken über das Internet

Vor fünfundzwanzig Jahren waren gedruckte Tabellen in Veröffentlichungen auf Papier die übliche Art der Verbreitung von Statistiken. Im Hinblick auf die Flexibilität der Produktion brachte dies wenig. In der Regel mussten die Statistiker entscheiden, welche Tabellen veröffentlicht werden sollten, und die Erhebungen entsprechend gestalten. Wenn Forscher die statistischen Daten für weitere Berechnungen nutzen wollten, mussten sie die Zahlen per Hand auf ihre eigenen Programme übertragen bzw. übertragen lassen.

Mitte der 70er Jahre begann Statistics Sweden, Datenbanken aufzubauen. Sie wurden auf gewerblicher Basis über an das Telefonnetz angeschlossene Terminals zur Verfügung gestellt. Wir

erhoben Gebühren bei den Nutzern, um die mit den Datenbanken verbundene Arbeit zu finanzieren. Im Januar 1997 wurden die Datenbanken über das Internet zur Verfügung gestellt.

Im Januar 2000 änderten wir unsere Verbreitungspolitik und stellten die Datenbanken kostenlos zur Verfügung. Davor hatten wir von jedem Nutzer eine jährliche Gebühr von 6 000 SEK (etwa 650 EUR) erhoben. Wir hatten rund 350 zahlende Nutzer und etwa 100 nicht zahlende Nutzer (Bibliotheken usw.). Der Einnahmefall wurde uns von der Regierung erstattet (rund 200 000 EUR).

Innerhalb von drei Monaten hatten wir mehr als 10 000 Nutzer. Ihre Zahl erhöhte sich kontinuierlich - bei 30 000 hörten wir auf zu zählen, denn die Nutzer mussten sich jetzt nicht länger auf unserer Website anmelden, um Zugang zu den Datenbanken zu erhalten. Diese letzte Verfahrensänderung führte dazu, dass sich die Zahl der Datenbankabfragen verdoppelte.

Im Januar 2000 beschlossen wir außerdem, alle unsere Veröffentlichungen in elektronischem Format auf unserer Website anzubieten und von den Nutzern keine Gebühren für das Herunterladen zu verlangen. Ein wichtiger Grund für den Verzicht auf Gebühren war der, dass wir finanziell dadurch nicht schlechter dastehen würden: Unsere Einnahmen aus Abonnements der Print-Publikationen würden zwar aller Wahrscheinlichkeit nach zurückgehen, aber gleichzeitig würden sich auch unsere Kosten in ungefähr dem gleichem Umfang verringern. Die Einnahmen deckten nur die Kosten für Papier, Druck, Vertrieb und Verwaltung.

Heute ist die Website unser wichtigstes Medium für die Verbreitung von Statistiken. Im vergangenen Jahr wurde sie 2,7 Millionen Mal besucht, die Abfragen über Suchmaschinen nicht mitgerechnet.

Diese Erfahrung verdeutlicht den Nutzen des freien Zugangs zu Veröffentlichungen und Datenbanken über das Internet. Auch stichhaltige theoretische Gründe sprechen dafür, Statistiken als ein öffentliches Gut zu betrachten, das aus allgemeinen Steuern finanziert und unentgeltlich zur Verfügung gestellt werden sollte. Der Hauptgrund ist, dass die Grenzkosten für die Verbreitung einer zusätzlichen Ausgabe einer Veröffentlichung oder Datenbank sehr gering sind, wesentlich geringer als für eine Druckausgabe mit demselben Inhalt und in der Regel auch geringer als der Nutzen für den Empfänger.

Gleichzeitig müssen die NSÄ aber auch Wege finden, ihre Entwicklungskosten zu finanzieren. Wenn es nicht möglich ist, finanzielle Unterstützung vom Staat zu bekommen, könnte es notwendig werden, Benutzungsgebühren zu erheben. Eine andere Möglichkeit, die Kosten für die unentgeltliche Bereitstellung von Veröffentlichungen und Datenbanken zu verringern, besteht darin,

das Produktionssystem zu rationalisieren, indem die Verbreitung über das Internet als Hauptvertriebsweg gewählt wird. Wenn Veröffentlichungen vorbereitet werden, sollten sie in einer Weise erstellt werden, dass sie sowohl gedruckt als auch ohne zusätzliche Kosten von der Website heruntergeladen werden können.

Ich bin ein entschiedener Verfechter des kostenlosen Zugangs zu Veröffentlichungen und Datenbanken über das Internet. Daher freut es mich besonders, dass Eurostat unlängst seine Veröffentlichungspolitik in dieser Hinsicht geändert hat. Statistische Daten über die Entwicklungen in der Europäischen Union und in der Eurozone erlangen in dem Maße zunehmende Bedeutung, wie die Mitgliedstaaten besser integriert werden und die europäische Politik weiter voranschreitet.

In diesem Zusammenhang möchte ich zwei Vorschläge machen: Als erstes schlage ich vor, dass das OECD-Sekretariat seine Politik im gleichen Sinne ändert. Das OECD-Sekretariat gibt Publikationen von sehr hoher Qualität heraus und unterhält ein breites Spektrum äußerst wertvoller Datenbanken, die besser genutzt werden könnten, wenn sie unentgeltlich bereitgestellt würden. In Europa hat man das Augenmerk in den letzten zehn bis fünfzehn Jahren sehr stark auf die Harmonisierung der Statistik gerichtet. Jetzt ist es an der Zeit, sich mehr auf die Vergleichbarkeit zwischen Europa und anderen Teilen der Welt, vor allem anderen OECD-Ländern, zu konzentrieren. Diesem Anliegen käme ein kostenloser Zugang zu dem umfangreichen Datenbestand der OECD zugute.

Zweitens möchte ich anregen, dass diejenigen NSÄ, die noch immer Gebühren für ihre Publikationen und Datenbanken auf ihren Websites erheben, ihre Politik überprüfen, die Argumente für und gegen den kostenlosen Zugang abwägen und über Möglichkeiten nachdenken, wie sie ihre Veröffentlichungen und Datenbanken im Internet gebührenfrei zur Verfügung stellen könnten.

3. Systeme für den Online-Zugang von Forschern zu Mikrodaten

Anonymisierte Mikrodatenfiles, also Dateien mit Datensätzen über Einzelpersonen oder einzelne Unternehmen, die keine Identifikationsnummern oder dergleichen enthalten, können als nächster Schritt für die NSÄ gesehen werden, ihren Nutzern ein flexibleres Produktangebot zur Verfügung zu stellen. Über das Internet zugängliche Datenbanken ermöglichen es dem Nutzer, Tabellen vielfältiger Art zu erstellen. Anonymisierte Mikrodatenfiles bieten eine noch größere Flexibilität. Zudem hat es der technologische Fortschritt den Forschern leichter gemacht, bei ihren Arbeiten Mikrodaten zu verwenden, und die Nachfrage nach Mikrodaten nimmt ständig zu.

Statistics Sweden blickt auf eine lange Tradition der Erhebung von Verwaltungsdaten und der Übertragung dieser Daten in Register für statistische Verwendungszwecke zurück. Die schwedischen Statistiken beruhen zu einem großen Teil auf Verwaltungsregistern – zu 85 bis 95 %,

je nachdem, wie gemessen wird. Zum Registersystem gehören auch einige erhebungsbasierte Register, zum Beispiel die Ergebnisse aus den Arbeitskräfteerhebungen. Über viele Jahre hinweg hat Statistics Sweden anonymisierte Mikrodaten an eine Vielzahl von Forschungseinrichtungen und andere Stellen verteilt, wobei als Träger Magnetbänder, CD-ROM, DVD und andere Formate verwendet wurden. Gleichzeitig mit der Zunahme des Volumens erhöhte sich auch die Zahl der Veröffentlichungen/Aufträge. Im vergangenen Jahr brachten wir 170 Veröffentlichungen heraus.

Die Geheimhaltung der Daten von Einzelpersonen und Unternehmen ist einer der wichtigsten Grundsätze der amtlichen Statistik und muss berücksichtigt werden, wenn es um Mikrodaten geht. Der Einzelne hat ein Recht darauf, von unzulässigen Eingriffen in seine Privatsphäre geschützt zu werden. Die Nutzung statistischer Informationen wird daher normalerweise durch Gesetze und/oder einen Verhaltenskodex geregelt.

In Schweden besagen die Rechtsvorschriften über die Verwendung statistischer Informationen, dass alle Daten einschließlich anonymer Daten, die für statistische Zwecke erhoben wurden, vertraulich sind, egal aus welcher Quelle sie stammen. Für statistische Zwecke gesammelte Daten dürfen nur für die Erstellung von Statistiken oder für wissenschaftliche Zwecke verwendet werden. Auf vertrauliche Daten dürfen nur hierzu befugte Personen zugreifen. Der Zugang kann dergestalt gewährt werden, dass keine direkte oder indirekte Identifizierung von Einzelpersonen oder anderen erfassten Einheiten wie beispielsweise Unternehmen möglich ist. In der Praxis gewährt Statistics Sweden nur Zugang zu Mikrodaten ohne Namen, Anschrift und Identifikationsnummer. Die Rechtsvorschriften in Schweden wie auch in anderen nordischen Ländern enthalten keine spezifischen Bestimmungen, die die Freigabe von Mikrodaten einschränken. Solange die allgemeinen gesetzlichen Anforderungen erfüllt werden, kann die jeweils am besten geeignete Methode gewählt werden.

Bei Statistics Sweden entwickeln wir derzeit Verfahren, die es einfacher, preiswerter und sicherer machen, in der Forschung mit Mikrodaten zu arbeiten. Dahinter steht die Grundidee, die Mikrodaten physisch bei Statistics Sweden zu behalten, anstatt sie auf CD-ROM usw. an Forschungseinrichtungen weiterzugeben, Techniken zu entwickeln, die es Forschern erleichtern, über das Internet Berechnungen anhand von Mikrodaten bei Statistics Sweden anzustellen, und Mittel für die Entwicklung und Verwaltung dieses Systems zu beschaffen, damit die Grenzkosten für den einzelnen Forscher möglichst gering sind.

Im vergangenen Jahr wurde ein Referat eingerichtet, das sich mit dieser Initiative befasst. Es heißt „Register Coordination and Microdata Access“ (Koordination der Register und Zugang zu Mikrodaten) und ist der Abteilung Forschung und Entwicklung zugeordnet. Eine seiner Aufgaben

ist der Aufbau eines umfassenden Registersystems, des „Data Warehouse“ von Statistics Sweden, in dem Variablen und Grundgesamtheiten aus verschiedenen Registern standardisiert und miteinander verknüpft werden, wobei die Zuständigkeiten für die einzelnen Register und für unterschiedliche Phasen der Koordinierungsarbeit genau abgegrenzt sind.

Für den Online-Zugang zu den Mikrodaten von Statistics Sweden ist ein System entwickelt worden, das sich weitgehend an einem ähnlichen System bei Statistics Denmark orientiert und MONA (Microdata Online Access at Statistics Sweden) genannt wurde. Es ermöglicht Forschern den Online-Zugang zu bestimmten Servern bei Statistics Sweden. Ein Desktop mit der entsprechenden Software (z. B. SAS oder SPSS) und mit Zugang zu den Mikrodaten wird für den Forscher eingerichtet, der dann völlig frei arbeiten, seine eigenen Datensätze erstellen, Berechnungen durchführen kann usw. Die gesamte Datenverarbeitung wird jedoch auf dem Server bei Statistics Sweden durchgeführt, Downloads sind nicht gestattet. Die Ergebnisse werden dem Forscher häufig in Tabellenform per E-Mail übermittelt.

Ein wichtiger Aspekt bei der künftigen Entwicklung ist die Erstellung mehrerer neuer thematischer Register, die so konzipiert sind, dass sie dem Bedarf der Forschung besser gerecht werden (z. B. LISA – ein Integrationsregister über Krankheitsurlaub und Beschäftigung). Um dieses Ziel zu erreichen, muss noch beträchtliche Arbeit geleistet werden, und zwar sowohl seitens der Sachverständigen für Methodikfragen als auch seitens der Experten auf dem jeweiligen Fachgebiet. Ein weiterer Zukunftstrend ist die Entwicklung von Verfahren, die eine Verknüpfung von Daten aus verschiedenen Quellen, innerhalb wie auch außerhalb von Statistics Sweden, ermöglichen.

Der bessere Zugang zu Mikrodaten ist allerdings mit relativ hohen Kosten verbunden. Ohne Finanzierung durch den Staat, die wir nicht erhalten, müssen die Kosten von den Forschern selbst getragen werden. Daher hat sich Statistics Sweden an den Schwedischen Forschungsbeirat gewandt und angeregt, der Rat möge ein System für den Zugang zu Mikrodaten finanzieren. Durch eine solche Finanzierung wären die Grenzkosten für die Forscher bei der Verwendung von Mikrodaten geringer. Dies würde eine verstärkte Nutzung von Mikrodaten in der Forschung begünstigen. Erfahrungen in anderen Bereichen, wo man sich auf eine Grundfinanzierung einigen konnte und die Forscher nur für geringe Grenzkosten aufkommen müssen, waren äußerst positiv. Eine solche Lösung würde eine umfassende IT-Unterstützung für den Online-Zugang über das Internet, einen Kundendienst („Front Office“) zur Betreuung und Beratung der Forscher sowie thematische Datenbanken beinhalten.

Statistics Sweden erhält zahlreiche Anfragen nach Mikrodaten von Hochschulen und Forschungseinrichtungen in den Vereinigten Staaten, die ihre eigenen Datenbanken zum Beispiel

um Volkszählungsdaten und Registerdaten von uns bereichern möchten. Aus mehreren Gründen, darunter gesetzlichen Beschränkungen, sind wir nicht in der Lage, solchen Anfragen zu entsprechen, wenn mit ihnen die physische Herausgabe von Daten verbunden ist. Es scheint indessen technisch möglich zu sein, das vorstehend beschriebene dänisch-schwedische Modell zu einer allgemeineren, internationalen Netzlösung weiterzuentwickeln, bei der alle Daten physisch dort verbleiben könnten, wo sie hingehören, während sie gleichzeitig von Forschern in der ganzen Welt abgerufen werden können. Auf nationaler Ebene wird eine solche Netzlösung gegenwärtig in Australien verwirklicht: Beim so genannten National Data Network (NDN) ist das Australian Bureau of Statistics einer der Knotenpunkte in dem Netz, dem auch andere Einrichtungen und Forschungsinstitute angeschlossen sind.

In diesem Zusammenhang habe ich drei Vorschläge: Erstens möchte ich anregen, dass die nationalen statistischen Ämter, die bereit sind, Systeme für den Online-Zugang zu Mikrodaten zu entwickeln, sich zu einem Netzwerk zusammenschließen, um Gedanken und Erfahrungen auszutauschen. Ich bin überzeugt, dass die Forschungsgemeinschaft ganz wesentlich vom Zugang zu solchen Strukturen profitieren würde. Statistics Sweden hat diesen Weg bereits eingeschlagen und würde natürlich gerne eine aktive Rolle in einem solchen Netz übernehmen.

Zweitens möchte ich vorschlagen, dass wir uns auf internationaler Ebene für ein schrittweises Vorgehen entscheiden und uns zunächst auf einige Leitlinien über Geheimhaltung und Mikrodaten einigen. Solche Leitlinien werden derzeit im Rahmen der Konferenz Europäischer Statistiker von einer Taskforce unter der Leitung des australischen Chefstatistikers Dennis Trewin erarbeitet. Als im vergangenen Jahr sowohl bei Eurostat im Ausschuss für das Statistische Programm als auch bei der OECD im Statistikausschuss Versuche unternommen wurden, diese Entwicklungen zu beschleunigen, gab es hierfür keine nennenswerte Unterstützung.

Drittens möchte ich anregen, dass wir die Möglichkeiten zur Einrichtung internationaler Mikrodatennetze untersuchen, über die Forscher aus anderen Ländern als den unseren auf sichere Weise Zugang zu unseren Daten erhalten könnten.

4. Organisation der Zusammenarbeit mit der Forschung

In den meisten, wenn nicht gar in allen nationalen statistischen Ämtern gibt es irgendeine Art der organisierten Zusammenarbeit mit der Forschung. Diese Zusammenarbeit kann unterschiedliche Formen annehmen, je nach Größe des Landes, Traditionen usw.

Bei Statistics Sweden gibt es zwölf Nutzergruppen für verschiedene Themenbereiche: Wirtschaftsstatistik, Sozialschutzstatistik, Bevölkerungsstatistik, Regionalstatistik usw. Diese

Nutzergruppen bestehen aus externen Sachverständigen einschließlich Forschern und beraten uns über die Entwicklung der Statistik auf ihren jeweiligen Fachgebieten. Ferner haben wir einen Wissenschaftlichen Beirat, einen Ausschuss für den Verbraucherpreisindex und einen Ausschuss für den Gebäudeindex mit Vertretern aus der Forschung. In diesem Jahr haben wir einen Ausschuss für Volkswirtschaftliche Gesamtrechnungen eingerichtet, der uns Ratschläge in Fragen der Methodik erteilt, natürlich im Rahmen des Europäischen Systems Volkswirtschaftlicher Gesamtrechnungen.

Darüber hinaus gibt es eine organisierte Zusammenarbeit insbesondere mit den Universitäten Stockholm und Örebro, also den Städten, in denen wir unsere Büros haben. Eine Kooperationsvereinbarung mit der Universität Örebro wurde 2001 unterzeichnet. Sie beinhaltet einen gemeinsamen Statistikprofessor, die gemeinsame Finanzierung von Dissertationen sowie regelmäßige Lehrgänge, Seminare und Sommerkurse. Ich selbst bin Mitglied des Verwaltungsrats der Universität Örebro, und mehrere Professoren der Universität Örebro sind Mitglieder in den verschiedenen oben erwähnten Gruppen. Ein Lehrstuhl an der Universität Stockholm ist speziell der amtlichen Statistik gewidmet, und der Professor verbringt einen Teil seiner Zeit bei Statistics Sweden.

Zusätzlich zu dieser recht formalisierten und umfassenden Kooperation gibt es auch eine Zusammenarbeit auf breiter Ebene zwischen den verschiedenen Teilen von Statistics Sweden und Forschern auf zahlreichen Fachgebieten. Einige unserer Mitarbeiter sind gleichzeitig Hochschullehrer, und einige Hochschulforscher arbeiten auf Teilzeitbasis bei uns. Und schließlich führen wir regelmäßig Erhebungen über die Kundenzufriedenheit durch, bei denen unter anderem auch Forscher befragt werden.

Beim Vergleich von Statistics Sweden mit anderen NSÄ habe ich festgestellt, dass wir nicht den Vorteil einer so umfassenden Zusammenarbeit mit den Hochschulen genießen, wie es in den Vereinigten Staaten der Fall ist. Auch die NSÄ einiger anderer europäischer Länder, zum Beispiel die des Vereinigten Königreichs und Frankreichs, sind in dieser Hinsicht in einer weitaus besseren Position als Statistics Sweden. Auch führen wir nicht so viele interne Forschungen und Analysen durch wie die NSÄ in anderen Ländern, beispielsweise in Kanada, Australien und Norwegen. Und wir bieten auch keine interne Statistikausbildung auf Hochschulniveau an, wie das INSEE in Frankreich.

Auch hierzu möchte ich zwei Vorschläge unterbreiten: Erstens bin ich der Meinung, dass wir in den meisten NSÄ die Zusammenarbeit mit der Forschung wahrscheinlich verstärken könnten und dies auch tun sollten. Ich glaube an Benchmarking und daran, dass wir von den guten Beispielen anderer

Länder lernen können. In der Tat gehen viele der aktuellen Veränderungen bei Statistics Sweden auf Besuche in anderen Ländern zurück. Ich meine nicht, dass wir die Organisation der Zusammenarbeit einheitlich gestalten sollten, aber wir könnten durchaus Ideen aufgreifen, die für unsere jeweiligen nationalen Systeme geeignet wären.

Im übrigen hat es mich sehr erstaunt, wie unterschiedlich die Beziehungen zu den Nutzern in den NSÄ und bei Eurostat gehandhabt werden. In den meisten nationalen statistischen Ämtern haben wir Systeme für den Dialog mit unseren Nutzern entwickelt. Sie sind von Land zu Land unterschiedlich. Einige NSÄ haben nationale Statistische Beiräte, die das gesamte Spektrum der amtlichen Statistik abdecken. Andere, wie Statistics Sweden, haben einzelne Nutzergruppen für die verschiedenen Statistikbereiche. Die meisten NSÄ pflegen rege Kontakte mit Branchen- und Fachverbänden usw.

Auf europäischer Ebene gibt es dies zum größten Teil nicht. Gewiss, wir haben den CEIES, der gegenwärtig einer Überprüfung unterzogen wird. Es ist sicher nicht unfair zu sagen, dass der CEIES als Nutzergruppe keine wesentliche Rolle gespielt hat. Die Ansichten der Nutzer wurden vielmehr durch die verschiedenen Generaldirektionen der Kommission und die Arbeitsgruppen und Ausschüsse von Eurostat hindurchgeleitet. Die Seminare, die der CEIES organisiert hat, waren allerdings von hoher Qualität und fanden sehr viel Anklang.

Mein zweiter Vorschlag auf diesem Gebiet ist daher, dass Eurostat Überlegungen darüber einleitet, wie es seine Beziehungen zu seinen Nutzern einschließlich der Forschung unter Berücksichtigung der Erfahrungen anderer Mitgliedstaaten weiterentwickeln könnte. Wären Nutzerbeiräte für die europäische Statistik insgesamt oder für einzelne Statistikbereiche eine sinnvolle Einrichtung für das Europäische Statistische System? Wären ein Wissenschaftlicher Beirat oder besondere Beratungsgremien für Volkswirtschaftliche Gesamtrechnungen, Verbraucherpreise oder andere Bereiche der Statistik für das ESS von Vorteil? Würden ihm Erhebungen über die Nutzerzufriedenheit in einzelnen Bereichen der Statistik nützen?

5. Schlussfolgerung

In der vorliegenden Unterlage habe ich die Beziehungen zwischen den nationalen statistischen Systemen und der Forschung aus der Sicht des Statistikproduzenten dargelegt. Ich sehe eindeutig Möglichkeiten für uns, unsere Arbeitsergebnisse zu verbessern und für Forscher wertvoller zu machen. Insbesondere befürworte ich den kostenlosen Internet-Zugang zu Veröffentlichungen und Datenbanken und die Entwicklung von Systemen für den Online-Zugang zu Mikrodaten für Forscher. Darüber hinaus sehe ich Möglichkeiten zur Verbesserung der Zusammenarbeit mit der

Forschung. Ich stelle diese Überlegungen zur Diskussion, wobei ich mir durchaus bewusst bin, dass für einzelne nationale statistische Ämter andere Lösungen geeigneter sein könnten.