

Cena zł 12,00 (VAT 0%)

90  
lat  
1918-2008

Indeks 381306  
PL ISSN 0043-518X

# WIADOMOŚCI STATYSTYCZNE

GŁÓWNY  
URZĄD  
STATYSTYCZNY

POLSKIE  
TOWARZYSTWO  
STATYSTYCZNE

MIESIĘCZNIK  
ROK LIII  
WARSZAWA  
LISTOPAD 2008

11

w numerze m.in.:

TADEUSZ WALCZAK

Informatyka jako czynnik rozwoju statystyki w okresie 90 lat działalności GUS

KAMILA MIGDAŁ-NAJMAN, KRZYSZTOF NAJMAN

Wykorzystanie wskaźnika Dunna do ustalania optymalnej liczby skupień

JAROSŁAW KRAJEWSKI

Prognozowanie konsumpcji w Polsce na podstawie wybranych modeli

DOROTA PRUSKA

Wieloletowość na rynku pracy w Polsce i w niektórych krajach Unii Europejskiej





---

## KOLEGIUM REDAKCYJNE:

prof. dr hab. Tadeusz Walczak (redaktor naczelny, tel. 0-22 608-32-89, t.walczak@stat.gov.pl), dr Stanisław Paradysz (zastępca red. nacz.), prof. dr hab. Józef Zegar (zastępca red. nacz., tel. 0-22 826-14-28), inż. Alina Świdarska (sekretarz redakcji, tel. 0-22 608-32-25, a.swiderska@stat.gov.pl), mgr Jan Berger (tel. 0-22 608-32-63), dr Marek Cierpień-Wolan (tel. 0-17 853-26-35), mgr inż. Anatol Kula (tel. 0-668231489), mgr Wiesław Łagodziński (tel. 0-22 608-30-57), dr Grażyna Marciniak (tel. 0-22 608-33-54), prof. dr hab. Walenty Ostasiewicz (tel. 0-71 368-03-47), dr hab. Krystyna Pruska (tel. 0-42 635-51-76), mgr Lucyna Przybylska (tel. 0-22 461-36-11), prof. dr hab. Bogdan Stefanowicz (tel. 0-22 849-53-95), mgr Małgorzata Żyra (tel. 0-22 608-32-40)

---

## REDAKCJA

al. Niepodległości 208, 00-925 Warszawa, gmach GUS, pok. 347, tel. 0-22 608-32-25  
<http://www.stat.gov.pl/pts>

Elżbieta Grabowska (e.grabowska@stat.gov.pl)

---

## RADA PROGRAMOWA:

dr Halina Dmochowska (przewodnicząca, tel. 0-22 608-34-25), prof. dr hab. Czesław Domański, mgr Małgorzata Fronk, prof. dr hab. Jan Kordos, dr Tomasz Pawlak, mgr Stanisława Szwałek, dr Teresa Smiłowska, prof. dr hab. Kazimierz Zajac

---



## ZAKŁAD WYDAWNICTW STATYSTYCZNYCH

al. Niepodległości 208, 00-925 Warszawa, tel. 0-22 608-31-45.  
Informacje w sprawach nabywania czasopism tel. 0-22 608-32-10, 608-38-10.  
Zbigniew Karpiński (redaktor techniczny), Ewa Krawczyńska (skład i łamanie),  
Wydział Korekty pod kierunkiem Teresy Chmielewskiej, mgr Andrzej Kajkowski (wykresy).

## Indeks 381306

### WARUNKI PRENUMERATY REALIZOWANEJ PRZEZ RUCH S.A.

#### Prenumerata krajowa:

Wpłaty na prenumeratę przyjmują jednostki kolportażowe „RUCH” S.A. właściwe dla miejsca zamieszkania lub siedziby prenumerującego. Termin przyjmowania wpłat na prenumeratę krajową do 5 każdego miesiąca poprzedzającego okres rozpoczęcia prenumeraty.

W Internecie <http://www.prenumerata.ruch.com.pl>

#### Prenumerata opłacana w złotych ze zleceniem wysyłki za granicę:

Informacji o warunkach prenumeraty i sposobie zamawiania udziela „RUCH” S.A. Oddział Krajowej Dystrybucji Prasy, 01-248 Warszawa, ul. Jana Kazimierza 31/33.

Telefony: 0-22 5328-731, 5328-816, 5328-819, 5328-820.

Infolinia: 0-800-1200-29, wpłaty na konto w banku PEKAO S.A. IV O/Warszawa. Nr 12401053-40060347-2700-401112-005 lub w kasie Oddziału.

Dokonując wpłaty na prenumeratę w banku czy też w urzędzie pocztowym należy podać: nazwę naszej firmy, nazwę banku, numer konta, czytelny pełny adres odbiorcy za granicą, okres prenumeraty, rodzaj wysyłki (pocztą lotniczą czy zwykłą) oraz zamawiany tytuł.

Warunkiem rozpoczęcia wysyłki prenumeraty jest dokonanie wpłaty na nasze konto.

#### Terminy przyjmowania wpłat na prenumeratę „WIADOMOŚCI STATYSTYCZNYCH”:

do 05.12 — na I kwartał roku następnego lub na cały rok następny,  
do 05.03 — na II kwartał roku bieżącego,  
do 05.06 — na III kwartał roku bieżącego,  
do 05.09 — na IV kwartał roku bieżącego.

---

# WIADOMOŚCI STATYSTYCZNE

CZASOPISMO GŁÓWNEGO URZĘDU STATYSTYCZNEGO  
I POLSKIEGO TOWARZYSTWA STATYSTYCZNEGO

## DZIEWIĘDZIESIĄT LAT DZIAŁALNOŚCI GŁÓWNEGO URZĘDU STATYSTYCZNEGO

**Tadeusz WALCZAK**

### Informatyka jako czynnik rozwoju statystyki w okresie 90 lat działalności GUS

Przedstawiając od początku tego roku w kolejnych numerach „Wiadomości Statystycznych” doświadczenia i dorobek polskich statystyków w 90-letnim okresie ich służby dla statystyki, nie sposób pominąć, bodaj krótkiego, omówienia ogromnego wkładu w rozwój statystyki licznej kadry informatyków, którzy w niemałym stopniu przyczynili się do osiągnięcia obecnego jej poziomu.

Początkowo, celem zastosowania środków technicznych informatyki w statystyce było umożliwienie i skrócenie czasu opracowania materiałów badań statystycznych, zwłaszcza badań masowych, takich jak spisy ludności i mieszkań czy spisy gospodarstw rolnych. W miarę jednak ulepszania środków technicznych, a także podnoszenia kwalifikacji pracowników, powiększał się również wpływ informatyki na usprawnienie wielu innych elementów systemu informacyjnego statystyki — kontrolę kompletności i dokładności danych, metody przesyłania danych, pogłębienie analizy z wykorzystaniem metod matematycznych, usprawnienie udostępniania informacji itp.

Zwiększenie wpływu informatyki na rozwój statystyki było ściśle związane z rozwojem technologicznym środków informatyki i ich dostępnością. W porównaniu z innymi krajami rozwój ten dokonywał się w Polsce z dużymi trudnościami i opóźnieniami. Początkowo, po powołaniu do życia GUS w lipcu i rozpoczęciu jego właściwej działalności w listopadzie 1918 r., statystycy nie mieli żadnych środków technicznych wspomagających opracowanie danych. Jedynymi narzędziami, które mogły w jakimś stopniu usprawnić pracę statystyków były liczydła, w które organizujący się Urząd był w stanie wyposażyć tylko niektóre działy.



Przewidziane na 30 września 1921 r. przeprowadzenie pierwszego w odrodzonym kraju powszechnego spisu ludności spowodowało niezwykle pilne zadanie zakupu, uruchomienia i przygotowania do pracy niezbędnych środków technicznych koniecznych do opracowania wyników spisu.

Najbardziej liczne dane jednostkowe badane w spisach ludności dotyczą demograficznych i społeczno-ekonomicznych cech poszczególnych osób: płci, wieku, stanu cywilnego, wykształcenia, źródeł utrzymania. W spisie 1921 r. podobny charakter miały również takie badane cechy, jak wyznanie religijne, narodowość i język ojczysty. Do opracowania wyników spisu trzeba było zatem zakupić takie maszyny, które zapewniłyby automatyczne posortowanie (pogrupowanie) i policzenie osób odznaczających się poszczególnymi cechami i przedstawienie wyników obliczeń według tych cech.

Przy ówczesnym poziomie techniki obliczeniowej nie było urządzeń umożliwiających automatyczny odczyt danych bezpośrednio z papierowych kwestionariuszy spisowych, a bez automatycznego odczytu niemożliwe było również automatyczne porządkowanie (sortowanie) danych. Trzeba było najpierw przenieść dane z kwestionariuszy spisowych na specjalne nośniki pośrednie — karty maszynowe wykonane z elastycznego kartonu, na które nanoszono symbole badanych cech w formie prostokątnych lub okrągłych dziurek. Maszyny te przyjęto nazywać maszynami systemu kart dziurkowanych lub maszynami licząco-analitycznymi (MLA).

W latach 20. ub. wieku na światowym rynku były dostępne dwa typy tych maszyn: skonstruowane w 1888 r. według pomysłu H. Holleritha stosujące elektryczny odczyt z kart oraz maszyny konstrukcji J. Powersa stosujące odczyt mechaniczny. Oba wynalazcy byli ściśle związani z amerykańskim Biurem Spisów i swoje wynalazki oparli na dokładnej znajomości najbardziej masowych czynności występujących w toku opracowania materiałów badań masowych. Maszyny H. Holleritha zastosowano po raz pierwszy do opracowania statystyki zgonów i urodzeń, a następnie do opracowania materiałów spisu ludności w 1890 r.

W 1903 r. dyrektor Biura Spisów Stanów Zjednoczonych, poszukując tańszych urządzeń, zlecił współpracującemu z biurem wynalazcy J. Powersowi opracowanie konkurencyjnych maszyn stosujących odczyt z kart dziurkowanych. Przy użyciu maszyn Powersa opracowano materiały spisu ludności w 1910 r. To właśnie maszyny konstrukcji J. Powersa postanowiono wykorzystać do opracowania wyników przeprowadzonego w 1921 r. spisu ludności w Polsce. Zakupiono w tym celu w Czechosłowacji 57 maszyn do dziurkowania kart, 29 sorterów do porządkowania i grupowania kart według symboli poszczególnych cech oraz 3 tabulatory — maszyny, które na podstawie odczytu danych z kart sporządzały w sposób automatyczny zestawienia z możliwością sumowania liczbowych danych. Pierwszą partię maszyn zakupiono w lipcu 1922 r., a drugą w styczniu 1923 r. Na początku 1926 r. uruchomiono w GUS produkcję kart maszynowych, początkowo z kartonu importowanego z Czechosłowacji, a następnie z uruchomionej staraniem GUS produkcji krajowej w mirkowskiej fabryce papieru<sup>1</sup>.

Maszynowe opracowanie danych spisu rozpoczęto w październiku 1922 r. i zakończono 30 kwietnia 1926 r. (J. Miller, 1930). Uwzględniając obszerny program tego spisu, obejmujący spis ludności, spis zwierząt domowych oraz spis gospodarstw rolnych, ogrodowych i leśnych, a także spis budynków, co wymagało wydziurkowania, sprawdzenia i maszynowego opracowania ponad 41,4 mln kart, ogólny czas opracowania, trwający ok. 3,5 roku, należy uznać za sukces.

<sup>1</sup> W latach powojennych, mimo usilnych starań GUS, nie udało się uruchomić w Polsce produkcji kartonu na karty dziurkowane.



Dodatkowym pozytywnym efektem pierwszego zastosowania w statystyce maszyn licząco-analitycznych było zapoczątkowanie nowej organizacji ewidencji obiegu dokumentów, w warunkach mechanizacji przetwarzania danych, w sposób zapewniający niezbędny porządek i ochronę dokumentów statystycznych oraz zagwarantowanie niezbędnej jakości opracowania danych. Zaczęto także przywiązywać większą wagę do właściwego projektowania badań statystycznych. Doświadczenia te miały bardzo duże znaczenie dla usprawniania opracowań statystycznych w następnych latach, zwłaszcza wtedy, gdy na maszynach zaczęto opracowywać wiele różnych tematów badań, co wymagało szczególnej staranności w zapewnieniu właściwego porządku w postępowaniu z dokumentami.

W 1924 r., ze względu na opóźnienia w opracowaniu dokumentów z obrotów towarowych z zagranicą, równoległe z opracowaniem spisu, na maszynach liczących zaczęto opracowywać również dane statystyki handlu zagranicznego. W 1925 r. przystąpiono do opracowania danych statystyki ruchu urodzeń, małżeństw i zgonów oraz statystyki zatrudnienia. W kolejnych latach przyjęto do opracowywania na maszynach liczących materiałów statystyki przewozów kolejowych według powiatów, statystyki produkcji przemysłowej, statystyki poborowych oraz statystyki kryminalnej.

Na początku 1930 r. GUS zrobił kolejny krok na drodze unowocześnienia informatyki<sup>2</sup>. Korzystając z okazji, że do Polski sprowadzono maszyny licząco-analityczne konstrukcji H. Holleritha, stosujące elektryczną metodę odczytu danych z kart dziurkowanych i uznawane wówczas za bardziej nowoczesne w porównaniu z maszynami J. Powersa, głównie dzięki możliwości zastosowania bardziej elastycznych metod sterowania ich pracą, Urząd wydzierżawił, początkowo jeden, a następnie drugi zestaw tych maszyn. Maszyny te, razem z zakupionymi poprzednio maszynami Powersa, zostały wykorzystane do opracowania kolejnego spisu powszechnego ludności z 9 grudnia 1931 r.

Obok maszyn licząco-analitycznych, które stosowane były głównie do opracowania materiałów masowych badań, takich jak powszechne spisy, sprawozdania z licznych przedsiębiorstw, dokumenty ruchu naturalnego ludności itp., zaczęto stosować również maszyny rachunkowe zwane maszynami małej mechanizacji. Były to początkowo maszyny mechaniczne, ręczne, dwudziałaniowe, wykonujące wyłącznie operacje dodawania i odejmowania oraz maszyny czterodziałaniowe, wykonujące również mnożenie i dzielenie. Maszyny te były niezwykle pomocne do wykonywania ogromnej ilości opracowań drobnych oraz do rachunkowego opracowania tablic wynikowych — obliczania wskaźników, procentów itp. Trzeba bowiem pamiętać, że ówczesne maszyny licząco-analityczne nie były przystosowane do wykonywania operacji mnożenia i dzielenia.

Mimo znacznych ograniczeń finansowych na badania, GUS w latach 1918—1939 dokonał dużego postępu w rozwoju zastosowań informatyki. Przed wybuchem II wojny światowej oddział maszyn liczących Urzędu dysponował kilkoma zestawami maszyn licząco-analitycznych z ogólną liczbą ok. 200 dziurkarek, sprawdzarek, sorterów tabulatorów, własną maszyną do produkcji kart dziurkowanych, niezłym wyposażeniem warsztatu naprawczego oraz ok. 120 maszynami rachunkowymi. Prócz tego, w pewną liczbę maszyn rachunkowych były wyposażone wydziały Urzędu. Niestety, po wybuchu wojny, w pierwszych dniach września 1939 r., wszystkie maszyny licząco-analityczne zainstalowane w lokalu przy ul. Chałubińskiego 9 oraz w Biurze Spisów Ludności przy

<sup>2</sup> Pojęciem informatyki (ang. *computing science, information technology*) określamy dziedzinę nauki i techniki zajmującą się przetwarzaniem danych lub informacji. Wprowadzenie tego terminu do języka polskiego zaproponował w 1968 r. prof. Romuald Marczyński (1921—2000), współtwórca pierwszych polskich komputerów. Informatykę często kojarzy się z zastosowaniem komputerów. W szerszym rozumieniu pojęcie informatyki obejmuje wszelkie narzędzia mające na celu usprawnienie i zwiększenie efektywności przetwarzania danych (uwaga T. W.).



ul. Nowowiejskiej 43, w czasie bombardowania Warszawy, spłonęły wraz z materiałami spisu ludności z 1931 r. Spłonęło również całe wyposażenie warsztatu oraz maszyna do produkcji kart dziurkowanych. Z pożogi wojennej ocalały jedynie maszyny rachunkowe znajdujące się w lokalu Urzędu w Al. Jerozolimskich 32. Maszyny te, wraz z całym majątkiem GUS wywieziono do Krakowa, do tworzącego się tam Urzędu Statystycznego dla Generalnej Guberni (Łukasiak, 1979). Tylko nieznaczną część tych maszyn udało się odzyskać po wojnie.

12 marca 1945 r., kiedy trwały jeszcze działania wojenne na Ziemiach Zachodnich i Północnych, Rada Ministrów reaktywowała działalność GUS. W ramach tworzącej się struktury organizacyjnej już od początku sierpnia 1945 r. rozpoczął pracę, początkowo 18-osobowy, Wydział Opracowań Masowych<sup>3</sup>. Obowiązkiem tego Wydziału było opracowanie materiałów spisu ludności Warszawy przeprowadzonego według stanu na 15 maja 1945 r. Pracę tę wykonywano ręcznie w taki sposób, że zasymbolizowane cechy (ok. 378 tys. osób i 48 tys. mieszkań) przeniesiono na przecięte na pół karty maszynowe, które następnie porządkowano według kolejności i liczono ręcznie, a otrzymane wyniki wpisywano do tablic (tamże). Jedynym „narzędziem informatycznym”, jakie wykorzystywano w tym opracowaniu były liczydła oraz niewielka liczba sumatorów i arytmometrów do podliczania danych w tablicach powstałych z zapisów wynikających z ręcznego policzenia kart według badanych cech. Również metodą ręczną opracowano wyniki powszechnego sumarycznego spisu ludności z 14 lutego 1946 r.

### *Pierwsze próby mechanizacji statystyki po II wojnie światowej*

W czerwcu 1946 r. Wydział Opracowań Maszynowych wyposażono w pierwszy zestaw maszyn licząco-analitycznych. Maszyny systemu Hollerith, częściowo zdekompletowane, pochodzące z rewanentów powojennych, wydzierżawiono od Centrali Eksploatacji i Konserwacji Maszyn Biurowych. Ich uruchomienie i wykorzystanie do opracowań wymagało znacznego wkładu pracy zespołu technicznego Wydziału. Pierwszą pracą wykonaną na tych maszynach było opracowanie badania wyników szkolnictwa powszechnego za rok 1945/46.

Zwiększający się zakres bieżących badań statystycznych oraz perspektywa kolejnego powszechnego spisu ludności (3 grudnia 1950 r.) stworzyły konieczność wcześniejszego zakupu i uruchomienia nowych środków technicznych. Na opracowanie danych spisu z 1950 r. zakupiono (z terminem dostaw na lata 1949—1950) łącznie 15 zestawów maszyn licząco-analitycznych firmy Bull (Francja). Były to, na ówczesne czasy, maszyny bardzo nowoczesne, o rozszerzonym zestawie jednostek uwzględniających szczególne wymagania opracowań statystycznych (zwłaszcza sortery wyposażone w indywidualne liczniki pozwalające nie tylko na porządkowanie kart według badanych cech, ale także na liczenie i drukowanie wyników).

Ten park maszynowy, uzupełniany nieznacznie w kolejnych latach, stanowił podstawowe narzędzie do opracowania przede wszystkim badań masowych. Chodzi tu szczególnie o powszechne spisy ludności i mieszkań przeprowadzone w latach 50. i 60. ub. wieku, spisy rolnicze, ale również badania: ruchu naturalnego ludności, zatrudnienia i wynagrodzeń, produkcji przemysłowej, przychodów, rozchodów i zapasów materiałów, handlu zagranicznego, ruchu budowlanego, szkolnictwa, statystyki zdrowia, statystyki kryminalnej. Równocześnie unowocześniano i rozszerzano opracowania przy użyciu maszyn rachunkowych oraz wielolicznikowych maszyn księgujących, które aż do czasu pojawienia się mikrokomputerów stanowiły bardzo pożyteczne narzędzie opracowania materiałów badań o mniejszej pracochłonności.

<sup>3</sup> Jego pierwszym kierownikiem był Zygmunt Padowicz, późniejszy prezes GUS.



Począwszy od 1950 r. wszystkie dostarczane maszyny instalowano w nowym gmachu przy al. Niepodległości 208. Tu też przeniesiono w połowie 1950 r. maszyny eksploatowane dotąd w lokalu przy ul. Narbutta 33, a maszyny dzierżawione zwrócono.

W 1949 r. w nowym statucie organizacyjnym GUS przewidziano utworzenie Biura Techniki Statystycznej i Maszyn. Nałożono na nie obowiązek nie tylko opracowania wyników badań statystycznych, przy użyciu maszyn licząco-analitycznych i maszyn rachunkowych, ale także zadania stanowiące załączek projektowania badań i opracowań (w tym dotyczące określenia najbardziej racjonalnych metod wykorzystania środków technicznych, organizacji i techniki opracowań i zaopatrywania całego Urzędu w środki informatyki oraz zapewnienia ich racjonalnego wykorzystania<sup>4</sup>).

Decyzją Prezydium Rządu z lutego 1955 r., 1 kwietnia 1955 r. został utworzony Zakład Techniki Statystycznej (na miejsce Biura Techniki Statystycznej i Maszyn)<sup>5</sup>. Decyzja ta sprzyjała podejmowaniu zwiększonego zakresu zadań związanych z modernizacją metod opracowania kolejnych tematów opracowań. W okresach mniejszego obciążenia pracami Zakład wykonywał nieodpłatnie usługi m.in. dla Ministerstwa Finansów, Ministerstwa Handlu Zagranicznego, Ministerstwa Sprawiedliwości, Instytutu Ekonomiki Rolnictwa, Naczelnej Organizacji Technicznej (Łukasiak, 1979).

Zakład Techniki Statystycznej podjął się również realizacji szeregu innych — ważnych dla dalszego rozwoju systemu informatycznego statystyki — zadań. W szczególności należy tu wspomnieć realizację obowiązków organizacji wojewódzkiej sieci stacji maszyn liczących, stanowiących techniczne wsparcie rozszerzających się opracowań wojewódzkich urzędów statystycznych (WUS). Początkowo stacje te powstawały jako oddziały terenowe ZTS, a następnie przekazano je pod bezpośredni nadzór dyrektorów WUS.

Pierwszą wojewódzką stacją techniki statystycznej (WSTS) utworzono w IV kwartale 1961 r. w Katowicach. Wyposażono ją w 3 zestawy, importowanych z ZSRR, maszyn licząco-analitycznych (MLA). W rok później powstała stacja w Łodzi, a następnie w Krakowie. Do 1968 r. WSTS utworzono we wszystkich województwach. Do ich podstawowych zadań należało opracowanie materiałów badań prowadzonych przez wojewódzkie urzędy i powiatowe inspektoraty statystyczne na potrzeby władz i społeczeństw lokalnych oraz udział w opracowaniach ujmowanych w programach centralnych wynikających z rocznych programów badań GUS oraz masowych badań jednorazowych.

W początkowym okresie organizowania WSTS nie ustrzeżono się od popełnienia błędów. Polegały one na tym, że przed przystąpieniem do wyposażenia technicznego ośrodków nie przeprowadzono szczegółowej analizy rodzaju prac występujących w poszczególnych urzędach. Spowodowało to, że instalowane tam MLA nie były w pełni dostosowane do występującej, w większości WUS, dużej różnorodności opracowań.

W ówczesnych warunkach maszyny te były bardzo wydajne. Stosowały one automatyczny odczyt danych. Te informacje trzeba było jednak najpierw przenieść ze sprawozdań na karty dziurkowane. Dziurkowanie kart było czynnością bardzo pracochłonną i opłacało się ją wykonać tylko w tych rodzajach opracowań, które polegały na sporządzaniu złożonych zestawień w wielu przekrojach. Mało efektywne było natomiast stosowanie tych maszyn do sporządzania różnorodnych prostych zestawień i tablic wymagających następnie dodatkowej kontroli i obliczania wskaźników w taki sposób, aby nadawały się one do publikacji lub bezpośredniego udostępniania odbiorcom.

<sup>4</sup> Pierwszym kierownikiem nowego biura został inż. Juliusz Miller, a jego zastępcami — Antoni Łukasiak i Edward Otrębski.

<sup>5</sup> Od 1 czerwca 1955 r., wkrótce po przejściu inż. J. Millera na emeryturę, dyrektorem ZTS został Zygmunt Wojcieszak.



Mankamentem dostępnych wówczas MLA było także to, że nie mogły one drukować tekstu. Sporządzane zatem tablice mogły być wykorzystywane wyłącznie jako zestawienia robocze wymagające ręcznego uzupełniania i przepisywania. W celu usunięcia trudności wkrótce w pierwszych WSTS utworzono sekcje małej i średniej mechanizacji, wyposażone w elektryczne maszyny rachunkowe oraz w maszyny licząco-piszące (księgujące). Sekcje te stanowiły bardzo efektywne uzupełnienie maszyn licząco-analitycznych i przyczyniły się do znacznego usprawnienia opracowań dla urzędów.

Uruchomienie wojewódzkich ośrodków maszyn liczących wpłynęło również na wprowadzenie nowych zasad współdziałania urzędów statystycznych, będących głównymi użytkownikami usług informatycznych z pracownikami ośrodków obliczeniowych świadczącymi te usługi. Po początkowym okresie trudności — wynikających zarówno ze specyficznych wymagań i oczekiwań statystyków, jak i niedostatecznych doświadczeń informatyków — zaczęła się harmonijna praca obu tych grup zawodowych. Doświadczenia zdobyte w latach 1960—1970 okazały się szczególnie cenne po zastosowaniu w statystyce bardziej nowoczesnej i wydajnej technologii komputerowej, wymagającej ścisłego współdziałania statystyków z informatykami.

Na przełomie lat 1965/66 między GUS i WUS rozpoczęto uruchamianie łączności dalekopisowej. W połowie 1966 r. wszystkie województwa zostały włączone do sieci teleksowej. Wobec braku innych, bardziej nowoczesnych i szybkich środków transmisji danych, sieć dalekopisową od stycznia 1966 r. zaczęto stosować do przesyłania do GUS operatywnej informacji statystycznej. Pozwoliło to Urzędowi przyspieszyć opracowanie operatywnych informacji, zarówno dla kraju, jak i w układach wojewódzkich. Znaczenie tego systemu łączności znacznie wzrosło po utworzeniu w 1975 r. 49 województw oraz po zmianie obiegu informacji statystycznej po likwidacji zjednoczeń w 1981 r.

### *Pora pomyśleć o zastosowaniu komputerów w statystyce*

Lata 60. ub. wieku przyniosły bardzo poważne zmiany w zadaniach i usytuowaniu służb statystycznych. Na mocy ustawy z 15 lutego 1962 r. o organizacji statystyki państwowej terenowe służby statystyczne zostały wydzielone z rad narodowych i podporządkowane bezpośrednio prezesowi GUS. Towarzyszyło temu stopniowe przejmowanie przez te służby znacznej części sprawozdań opracowywanych dotąd przez poszczególne resorty i zjednoczenia. Znacznie wzrosło zapotrzebowanie na informacje statystyczne ze strony władz terenowych, zwłaszcza wojewódzkich. W tej sytuacji specjaliści GUS, odpowiadający za realizację programu przetwarzania danych, zaczęli zdawać sobie sprawę, że stosowane dotąd w statystyce techniczne środki w postaci maszyn licząco-analitycznych nie będą w stanie sprostać rosnącym wymaganiom użytkowników informacji. Podstawowe mankamenty MLA polegały na konieczności stosowania bardzo sztywnego układu danych, wynikającego z ograniczonej pojemności informacyjnej kart dziurkowanych oraz bardzo sztywnych wymagań maszyn odczytujących i opracowujących te dane, a także bardzo ograniczonych możliwości logicznej kontroli poprawności informacji.

Jednocześnie, z różnych informacji, w tym zwłaszcza uzyskiwanych przez specjalistów w ramach współpracy GUS ze specjalistycznymi grupami roboczymi, działającymi w Konferencji Statystyków Europejskich (KSE) wynikało, że w innych krajach do przetwarzania wyników badań statystycznych coraz to szerzej i z dużą efektywnością upowszechnia się zastosowanie komputerów. W tych warunkach podjęcie decyzji w sprawie dalszego rozwoju informatyki w statystyce polskiej było sprawą niezwykle ważną i pilną, zwłaszcza wobec zbliżającego się kolejnego powszechnego spisu ludności i mieszkań, planowanego na grudzień 1970 r. Opracowanie tego spisu według formuło-



wanych oczekiwań użytkowników informacji nie byłoby możliwe przy użyciu przestarzałych i mocno już wyeksploatowanych maszyn licząco-analitycznych. Inwestowanie w zakup maszyn tej klasy, wobec zarysowującego się gwałtownego rozwoju techniki elektronicznej, byłyby wielce nieefektywne.

Specjaliści GUS zdawali sobie jednocześnie sprawę, że racjonalne wykorzystanie zupełnie nieznaney w warunkach polskich elektronicznej techniki obliczeniowej będzie wymagać odpowiednich przygotowań. Nie było bowiem w tym czasie w Polsce żadnych doświadczeń w tej dziedzinie. Pojedyncze komputery zainstalowane wówczas w kraju stosowane były głównie do celów naukowych lub do obliczeń o charakterze inżyniersko-konstrukcyjnym.

Do najważniejszych spraw, wymagających odpowiednio wcześniejszego przygotowania, należały w szczególności:

- wytypowanie wspólnie ze statystykami tematów opracowań, które w pierwszej kolejności powinny być realizowane z użyciem komputerów. Było to niezbędne ze względu na konieczność wcześniejszego przygotowania niezbędnej dokumentacji opracowań, w tym zwłaszcza programów komputerowych;
- dokonanie wyboru typu i konfiguracji komputera spełniającego w maksymalnym stopniu wymagania opracowań statystycznych z uwzględnieniem możliwości jego zakupu w ramach posiadanych środków;
- dokonanie niezbędnej adaptacji pomieszczeń odpowiadających technicznym wymaganiom określonym przez producenta i zapewniających bezpieczną eksploatację komputera;
- przeszkolenie kadry projektantów, programistów oraz personelu obsługi technicznej i operatorskiej dla zapewnienia sprawnego funkcjonowania ośrodka obliczeniowego w systemie pracy trzymianowej.

Za całokształt przygotowania warunków zakupu i uruchomienia komputerów w GUS odpowiadał, utworzony w czerwcu 1963 r. w Zakładzie Techniki Statystycznej, Dział Studiów i Analiz (DSiA). Dział ten kierowany przez dra Stefana Semczuka, liczący początkowo 4 pracowników, był załączkiem utworzonego parę lat później Ośrodka Elektronicznego GUS.

Prowadzone przez DSiA prace studialno-badawcze od początku były ściśle łączone z praktycznym uruchamianiem opracowań statystycznych na dostępnym wówczas w Polsce komputerze ZAM2, zainstalowanym w Instytucie Maszyn Matematycznych oraz na komputerze ICT-1300 pracującym w Centralnym Ośrodku Kadry Kierowniczych. Na tych komputerach, mimo że żaden z nich nie odpowiadał wymogom prac statystycznych, przygotowano programy opracowania sprawozdawczości ze szkolnictwa podstawowego, zatrudnienia, spisu ludności oraz badań demograficznych. Opracowania te nie miały większego znaczenia praktycznego, ale pozwoliły zgromadzić wiele cennych doświadczeń na temat metod formułowania założeń do przetwarzania danych oraz na temat wymagań, jakim powinny odpowiadać komputery do prac statystycznych. Doświadczenia te okazały się szczególnie cenne na etapie prowadzenia negocjacji z potencjalnymi dostawcami komputerów.

Po szczegółowej analizie parametrów techniczno-użytkowych, oprogramowania, cen, warunków dostaw, obowiązków gwarancyjnych, warunków pomocy w szkoleniu specjalistów, warunków zaopatrzenia w części zamiennej itp., najważniejszą rolę pod uwagę komputery produkowane przez amerykańską firmę IBM, brytyjską firmę ICT (później ICL) oraz komputery amerykańsko-francuskiej firmy General-Electric-Bull.

Komputery IBM należały wówczas do najbardziej rozpowszechnionych. Miały dobrą opinię użytkowników w wielu krajach. Były także stosowane w większości urzędów



statystycznych w krajach zachodnich. W 1964 r. firma IBM rozpoczęła produkcję nowych maszyn serii 360, należących do maszyn III generacji zbudowanych w nowej technologii obwodów scalonych. Maszyny te uznano jednak za wyrób strategiczny i wydano zakaz ich sprzedaży do niektórych krajów, w tym do Polski. Firma IBM zaproponowała nam sprzedaż starszego modelu maszyn z serii 1400. Informatycy GUS zdawali sobie sprawę, że wobec uruchomienia produkcji nowej serii, produkcja maszyn serii 1400 zostanie wkrótce zaniechana, mimo przeciwnych zapewnień przedstawicieli IBM. W ostatniej fazie negocjacji wyszło na jaw, że firma IBM proponuje nam sprzedaż używanego komputera IBM 1410 po renowacji za cenę zbliżoną do ceny maszyny nowej. W tej sytuacji odstąpiono od dalszych rozmów z firmą IBM i sfinalizowano prowadzone równolegle rozmowy z brytyjską firmą ICL oraz podpisano kontrakt na dostawę dla GUS w 1967 r. komputera ICL 1905, mającego opinię najszybszego i najnowocześniejszego komputera w Europie.

Podpisanie kontraktu stanowiło impuls do znacznego przyśpieszenia i skonkretyzowania prac przygotowawczych do organizacji centrum obliczeniowego. Zintensyfikowano przygotowanie programów dla konkretnych tematów opracowań, których część sprawdzono na komputerach podobnego typu w ramach praktyk szkoleniowych programistów w W. Brytanii. Pierwszymi tematami przygotowanymi do opracowania były: statystyka handlu zagranicznego, ruch wędrowności ludności, statystyka zgonów, budżety rodzinne, statystyka pracy, sprawozdawczość z rocznej działalności przedsiębiorstw przemysłowych oraz ankietowe badanie warunków bytu ludności.

Dobra synchronizacja przygotowań z terminem dostawy komputera umożliwiła przystąpienie do uruchomienia i bardzo szczegółowego odbioru technicznego komputera niezwłocznie po jego dostawie. 16 października 1967 r. rozpoczęła się normalna eksploatacja komputera na 2 zmiany, a krótko potem w systemie pracy ciągłej.

Wybór dla GUS komputera ICL 1905 okazał się szczęśliwy również z tego względu, że kilka lat później polski przemysł komputerowy, na mocy porozumienia z firmą ICL, zastosował system oprogramowania maszyn rodziny ICL 1900 w polskich komputerach Odra 1300. Dzięki temu oba typy stosowały w pełni wymienne oprogramowanie, co znacznie ułatwiło stosowanie obu tych komputerów w służbach statystycznych w naszym kraju (Walczak, 1993).

### *Początek ery komputeryzacji statystyki*

Uruchomienie w Ośrodku Elektronicznym GUS pierwszego komputera zapoczątkowało nowy, niezwykle ważny okres w rozwoju badań statystycznych. Ocena wyników pracy komputera była od początku bardzo pozytywna, chociaż w okresie wdrażania tej nowej technologii nie brak było wielu napięć, typowych dla początkowego etapu komputeryzacji.

W grudniu 1970 r. przeprowadzono w Polsce kolejny Narodowy Spis Powszechny, obejmujący tematykę: ludność, mieszkania, budynki oraz indywidualne gospodarstwa rolne. W tym czasie komputer ICL 1905, pracujący od połowy 1968 r. na 3 zmiany, był całkowicie obciążony opracowaniami bieżącymi. Jednocześnie decyzjami rządu GUS został zobowiązany do opracowania wyników spisu w ciągu 2,5 roku, czyli w terminie niemal dwukrotnie krótszym w porównaniu ze spisem z 1960 r., mimo zwiększenia liczby ludności i znacznie obszerniejszego programu opracowań. Było to możliwe jedynie w przypadku użycia komputerów do opracowania wyników spisu.

Dzięki energicznym staraniom prezesa Wincentego Kawalca, GUS uzyskał dodatkowe pieniądze na opracowanie materiałów NSP 1970. Pozwoliło to m.in. zakupić nowy



komputer ICL 1902A<sup>6</sup> oraz komputer produkcji polskiej Odra 1304. Oba te komputery uruchomiono pod koniec 1970 r.

W następnych latach komputery Odra zainstalowano także w kilku ośrodkach wojewódzkich: w 1970 r. w Krakowie, w 1971 r. w Katowicach i Poznaniu, w 1973 r. w Radomiu, w 1975 r. we Wrocławiu, Zielonej Górze i Lublinie, w 1976 r. w Łodzi i Koszalinie, a w 1983 r. w Olsztynie.

W krótkim czasie po wprowadzeniu pierwszych tematów opracowań statystycznych do komputerów, w warunkach radykalnego zwiększenia prędkości obliczeniowej tych maszyn, szczególnie widoczna stała się dysproporcja pomiędzy bardzo dużą wydajnością czynności przetwarzania wynikowego (kontrola automatyczna, sortowanie i grupowanie danych, obliczenia, opracowanie i drukowanie tablic) a czynnościami przygotowania danych, głównie dziurkowania i kontroli kart, a także ograniczoną prędkością odczytu kart przez komputer.

Kiedy na przełomie lat 1889/90 H. Hollerith konstruował dla amerykańskiego Biura Spisów swoje pierwsze maszyny licząco-analityczne, mogące odczytywać informacje dziurkowane na kartach, wynalazek ten stanowił ogromny przełom w technologii przetwarzania danych. Miało to szczególne znaczenie w badaniach statystycznych, w których mamy do czynienia z masowo występującymi operacjami porządkowania i grupowania badanych cech oraz liczenia jednostek odznaczających się interesującymi badacza cechami. Uzyskanie wyników badań statystycznych wymaga wielokrotnego odczytywania informacji, z jednoczesnym liczeniem częstotliwości występowania badanych cech lub sumowania cech ilościowo-wartościowych.

Przykładowo, materiały ostatniego opracowywanego za pomocą MLA Powszechnego Spisu Ludności i Mieszkań 1960 r. zostały przeniesione na 42,3 mln kart dziurkowanych. Opracowanie zaś przewidzianych programem wyników wymagało 40-krotnego ich odczytu i grupowania przez maszyny sortujące oraz 4,5-krotnego odczytu przez tabulatory w celu sumowania wartości cech mających charakter ilościowo-wartościowy. Przy tej technologii ponoszenie znacznych nakładów na dziurkowanie i kontrolę kart znajdowało uzasadnienie w znacznym przyspieszeniu czynności wielokrotnego automatycznego odczytu informacji z kart.

W warunkach komputerowego przetwarzania danych karty dziurkowane wykorzystywano wyłącznie w celu jednokrotnego wprowadzenia danych do komputera, które po przeprowadzeniu automatycznej kontroli logicznej i arytmetycznej zapisywano na nośnikach (wówczas taśmach) magnetycznych. Dalsze przetwarzanie odbywało się wyłącznie z wykorzystaniem zapisu danych na nośnikach magnetycznych.

Karty dziurkowane wykorzystywano jedynie w celu wprowadzenia ewentualnych poprawek błędów stwierdzonych w czasie wczytywania i kontroli. To pozwoliło znacznie przyspieszyć przetwarzanie danych, ponieważ odczyt danych z nośników magnetycznych w czasie przetwarzania był co najmniej 100 razy szybszy w porównaniu z ich odczytem z kart dziurkowanych. Niemniej jednak, nawet konieczność jednokrotnego odczytu danych z kart poważnie ograniczała ogólną wydajność przetwarzania danych. Wystarczy uzmysłowić sobie, że w czasie odczytu jednej karty, trwającego ok. 140 milisekund, nawet stosunkowo wolny, jak na obecne czasy, komputer mógł wykonać kilkanaście tysięcy operacji obliczeniowych.

Ze stosowaniem kart dziurkowanych wiązało się także wiele niedogodności, w tym zwłaszcza bardzo duża pracochłonność dziurkowania i kontroli kart, zbyt mała pojem-

<sup>6</sup> Przebudowany następnie na miejscu w ośrodku elektronicznym na model większy — ICL 1903A.



ność informacyjna karty (80 znaków), łatwość zagubienia kart, konieczność uzyskiwania zgody na import papieru do produkcji kart, kłopoty z przechowywaniem i transportem wewnętrznym ogromnej liczby kart.

### **Koniec epoki kart dziurkowanych**

W latach 1974—1978 w informatyce GUS dokonał się kolejny, o dużym znaczeniu, przełom w metodach przygotowania danych do przetwarzania. W 1974 r. zastosowano w GUS po raz pierwszy urządzenia do rejestracji danych, pozwalające zapisywać dane z klawiatury stanowiska operatorskiego. Zapisem sterował minikomputer. Pozwoliło to nie tylko zwiększyć nieco wydajność przenoszenia danych na nośnik magnetyczny w porównaniu z dziurkowaniem kart, ale co istotniejsze — wprowadzić automatyczną kontrolę dokładności zapisu. Dzięki temu można było już w czasie zapisu danych wykryć ewentualne błędy i w ten sposób zmniejszyć ogólną pracochłonność kontroli i korekty danych.

W służbach statystycznych w Polsce do zapisu danych na nośnikach magnetycznych zastosowano produkowany w Polsce, na licencji brytyjskiej firmy Redifon, wielostanowiskowy (do 32 stanowisk operatorskich) system rejestracji danych Mera 9150. Urządzenia te sprawdziły się nie tylko jako bardzo sprawne narzędzie do przygotowania nośników danych, ale zostały wnet ocenione bardzo pozytywnie jako wielostanowiskowy minikomputer do przetwarzania niezbyt dużych zbiorów danych. Znalazł on bardzo efektywne zastosowanie, zwłaszcza w służbach terenowych.

Powszechny Spis Ludności i Mieszkań w 1978 r. został opracowany po raz pierwszy na podstawie danych zarejestrowanych na taśmach magnetycznych. Pozwoliło to zrezygnować z dziurkowania ponad 40 mln kart o wadze ok. 120 ton, zaoszczędzić znaczne środki finansowe na ich import<sup>7</sup> i zmniejszyć pracochłonność zapisu. Według ostrożnych szacunków, uwzględniając znacznie zwiększony program tamtego spisu w porównaniu ze spisem z roku 1970, zastosowanie taśm magnetycznych pozwoliło zaoszczędzić na etapie przygotowania danych ok. 450 tys. godzin. Dodatkowo, znaczne oszczędności uzyskano na etapie wynikowego przetwarzania danych (Walczak, 1985).

Oznaczało to naturalny koniec kart dziurkowanych jako masowego nośnika danych. Przez pewien czas karty dziurkowane stosowano jeszcze do zapisu i odczytu źródłowych programów komputerowych.

Dzięki inwencji informatyków Zarządu Mechanizacji i Automatykacji Opracowań Statystycznych (ZMiAOS) minikomputer Mera 9150 znalazł również dodatkowe oryginalne i niezwykle efektywne zastosowanie w realizacji załączka sieci transmisji danych w statystyce.

W wyniku znacznego zwiększenia zakresu informacji operatywnej przesyłanej w relacji WUS—GUS, zwłaszcza po likwidacji zjednoczeń i przejęciu przez GUS znacznej części informacji bezpośrednio od przedsiębiorstw, uruchomiona w 1966 r. sieć teleksowa nie była w stanie zapewnić bardzo krótkich terminów opracowania informacji miesięcznych w układach wojewódzkich. Zbierane przez WUS w wielkim pośpiechu informacje z przedsiębiorstw, przekazywane drogą teleksową do GUS w 2 dni po upływie miesiąca, często zawierały błędy. Dodatkowe błędy powstawały w trakcie transmisji danych liniami telegraficznymi.

<sup>7</sup> Taśmy magnetyczne również kupowane były za dewizy, ale trzeba pamiętać, że karty dziurkowane były nośnikiem jednokrotnego użytku, natomiast nośniki magnetyczne można było wykorzystywać do zapisu wielokrotnie.



Wykrywanie błędów, a zwłaszcza procedury wyjaśniania ich przyczyn i korygowania, były bardzo kłopotliwe i pracochłonne. Niezwykle cenną pomoc w rozwiązaniu tych trudności okazali informatycy ZMiAOS, którzy zaprojektowali, wykonali i wdrożyli oryginalny system, zwany w skrócie TELZIS (telegraficzne zbieranie informacji sprawozdawczych). System ten polegał na sprzężeniu sieci teleksowej ze standardowym minikomputerem Mera 9150 w taki sposób, że nadawane przez WUS za pomocą dalekopisów meldunki były automatycznie wprowadzane do minikomputera. Po przeprowadzeniu kontroli ich poprawności, komputer wysyłał sygnał akceptacji, a w przypadku wykrycia błędu w meldunku, generował komunikat informujący o liczbie i rodzaju błędów oraz o sposobie ich poprawienia (Rawski, 1981).

Poszukując innych alternatywnych metod usprawnienia metod przygotowania danych do przetwarzania na komputerach, w 1976 r. po raz pierwszy zastosowano w statystyce polskiej optyczny czytnik dokumentów (typu Scan-Data 2250). Czytnik ten umożliwiał odczyt danych bezpośrednio z kwestionariuszy papierowych wypełnianych ręcznie lub na maszynie do pisania wyposażonej w specjalną, nieco stylizowaną czcionkę.

Czytnik zastosowano początkowo do odczytu dokumentów statystyki małżeństw oraz wykazów wyjazdów za granicę, a następnie w reprezentacyjnym badaniu przyczyn zmian miejsca zamieszkania i dojazdów do pracy, przeprowadzonym w ramach NSP 1978 r. Należy zauważyć, że opinie na temat efektywności zastosowania tego czytnika były podzielone. Pozwolił on znacznie przyspieszyć opracowanie informacji w ośrodkach elektronicznych, stawiał natomiast zbyt szybyne wymagania wobec osób wypełniających dokumenty. Również oprogramowanie rozpoznawania znaków pozostawiało wiele do życzenia, gdyż powodowało konieczność dość częstych interwencji operatorów w czasie odczytu w każdym przypadku, gdy system nie był w stanie jednoznacznie zdefiniować odczytywanego znaku.

### *Mikrokomputery w statystyce*

W połowie lat 80. ub. wieku rozpoczął się nowy etap rozwoju zastosowań informatyki w statystyce, który wywarł i wywiera nadal wszechstronny wpływ na poziom rozwoju wielu dziedzin statystyki. W 1984 r., początkowo w Departamencie Statystyki Przemysłu, a następnie w wielu innych komórkach (w tym w Ośrodku Badawczo-Rozwojowym oraz w Zakładzie Badań Statystyczno-Ekonomicznych oraz w WUS) zaczęto wykorzystywać mikrokomputery do wielu bardzo różnorodnych czynności występujących na co dzień w pracach statystycznych. Było ich dużo począwszy od najprostszych, związanych z opracowaniem korespondencji, wytycznych, instrukcji, wykazów, tablic do publikacji itp.<sup>8</sup>, poprzez mniej lub bardziej zaawansowane próby tworzenia systemów, klasyfikacji, nomenklatur, informacji o zasobach informacyjnych do tworzenia i aktualizacji katalogów, zawierających informacje o badaniach i opracowaniach oraz o dostępnym oprogramowaniu potrzebnym do wtórnego przetwarzania danych statystycznych przez samego użytkownika. Mikrokomputery zaczęto także stosować do opracowania analiz statystycznych, projektowania kwestionariuszy — aż do samodzielnej realizacji pełnego cyklu przetwarzania danych, bez udziału zawodowych informatyków. Pojawiły się także próby podejmowania przez statystyków pełnego cyklu projektowania, opracowania, analizy i publikacji niektórych mniejszych tematów badań. Przykładowo wymienię wykorzystanie oprogramowa-

<sup>8</sup> Początkowo, można było spotkać się z uwagami i wątpliwościami na temat słuszności stosowania stosunkowo drogiej i bardzo szybkich mikrokomputerów do tak prostych prac, jak pisanie tekstów. Pamiętając jednak, jak kłopotliwe i pracochłonne było sporządzanie i nieraz wielokrotne poprawianie różnych tekstów przez halę maszyn, niedługo pracownicy przekonali się, że mimo wszystko wykorzystanie do tego celu mikrokomputerów było efektywne. Po pewnym czasie utrzymywanie hali maszyn w urzędzie okazało się zbędne.



nia standardowego w rodzaju systemu BLAISE, który stanowił stosunkowo przyjazne w użyciu narzędzie nie tylko opracowania i analizy danych, ale i projektowania układu ankiet, ich kontroli itp.

W bardzo krótkim czasie mikrokomputery, oferując ogromną moc obliczeniową, obszerną pojemność pamięci, prostą i łatwą obsługę oraz przyjazne i wszechstronne oprogramowanie, zyskały powszechne uznanie jako niezwykle sprawne narzędzie pracy statystyka. Jednocześnie, co jest bodaj ważniejsze z punktu widzenia przyszłego rozwoju statystyki, dzięki rozpowszechnieniu mikrokomputerów zostały stworzone warunki zmiany modelu współpracy pomiędzy statystykami i informatykami.

W warunkach tradycyjnej organizacji wykorzystania dużych komputerów, statystycy, podobnie jak i użytkownicy informacji w systemach informacyjnych przedsiębiorstw i instytucji poza statystyką, byli coraz silniej uzależnieni od systemów skomputeryzowanych, które na skutek niedoskonałych i bardzo skomplikowanych metod programowania, były mało zrozumiałe dla użytkownika. Jakakolwiek ingerencja użytkownika we własny system informacyjny była praktycznie niemożliwa bez pośrednictwa kwalifikowanego zespołu informatyków. Sytuacja ta rodziła poczucie izolacji użytkowników od własnych zasobów informacyjnych i utraty kontroli nad nimi. Stwarzała również wiele praktycznych trudności w korzystaniu z informacji, wydłużała drogę i czas uzyskiwania informacji na doraźne żądania oraz poważnie utrudniała wprowadzanie jakichkolwiek modyfikacji do systemów informacyjnych.

Wprowadzenie mikrokomputerów rozpoczęło przełamywanie monopolu profesjonalnych zespołów informatycznych na opracowanie informacji statystycznych i większe zaangażowanie się statystyków nie tylko w okresie przygotowania i organizacji badań, ale również na etapach opracowań, począwszy od gromadzenia, przetwarzania i kontroli danych oraz analizy i udostępniania informacji. Wpłynęło to bardzo pozytywnie na profesjonalną integrację wszystkich zespołów pracowników odpowiedzialnych za rozwój statystyki, zwiększyło atrakcyjność pracy statystyków, co podkreślali zwłaszcza pracownicy młodzi, wiążący swoje szanse życiowe z zawodem statystyka.

Pozytywne skutki ściślejszego współdziałania statystyków z informatykami stały się szczególnie widoczne, kiedy to trzeba było przystąpić do kolejnego etapu rozwoju statystyki z wykorzystaniem metod wspólnych baz danych, ale już przy wykorzystaniu nowej generacji komputerów i sieci komputerowych.

### *Od komputerów autonomicznych do sieci komputerowych*

Na przełomie lat 80. i 90. ub. wieku, wobec dokonujących się zmian w systemie politycznym i społeczno-gospodarczym kraju, zaszła konieczność zasadniczej przebudowy statystyki. Konieczność ta wynikała zarówno ze zmieniających się potrzeb krajowych użytkowników informacji, jak i wymogów dostosowania rozwiązań metodologicznych do standardów międzynarodowych, co pozwoliłoby zapewnić porównywalność naszej statystyki ze statystyką innych krajów. Wprowadzenie niezbędnych zmian w statystyce wynikało również z decyzji władz państwowych w sprawie podjęcia starań o przystąpienie Polski do Unii Europejskiej (UE).

W przeważającej części zmiany te miały charakter metodologiczno-prawny. Dotyczyły one zwłaszcza określenia nowej roli statystyki w warunkach gospodarki rynkowej i społeczeństwa demokratycznego. Chodziło o zapewnienie gwarancji niezależności statystyki w sprawach decydowania o tematyce i metodach badań, stworzenia przesłanek dla odzyskania i utrzymania społecznego zaufania do statystyki, opracowania i wdrożenia klasyfikacji i nomenklatur zgodnych ze standardami międzynarodowymi.



Pomyślne wprowadzenie niektórych zmian wymagało jednocześnie zastosowania bardziej nowoczesnych i wydajnych środków technicznych. Dotyczyło to zwłaszcza objęcia badaniami statystycznymi dużej liczby nowo powstających małych i średnich przedsiębiorstw prywatnych, uruchomienia wielu nowych badań (np. prywatyzacji, koniunktury gospodarczej, rozszerzenia i doskonalenia badań cen itp.). Bardzo pracochłonne były także czynności związane z utrzymaniem i aktualizacją znacznie rozszerzonego liczebnie rejestru jednostek gospodarczych REGON, zarządzaniem nowymi klasyfikacjami i nomenklaturami.

Dodatkowym argumentem za koniecznością unowocześnienia informatycznej bazy statystyki był fakt, iż zainstalowane w końcu lat 60. ub. wieku komputery były na początku lat 90. ub. wieku nie tylko przestarzałe technologicznie, ale również poważnie fizycznie zużyte w wyniku intensywnej eksploatacji.

Dzięki nawiązaniu współpracy z Urzędem Statystycznym Wspólnot Europejskich — Eurostatem, sformalizowanej podpisaniem 30 października 1990 r. przez prezesa GUS i dyrektora generalnego Eurostatu Wspólnej Deklaracji o Współpracy Statystycznej, możliwa była pomoc techniczna i finansowa UE w ramach programu PHARE<sup>9</sup>. Od 1993 r., kiedy to uruchomiono pierwsze, nowo zakupione superminikomputery HP serii 9000, rozpoczęto w statystyce polskiej realizację niezwykle potrzebnych inwestycji informatycznych. Miały one charakter kompleksowy; obejmowały bowiem zakup i instalację nowego sprzętu komputerowego, projektowanie i uruchomienie lokalnych sieci komputerowych (LAN) dla GUS i urzędów statystycznych oraz sieci rozległej (WAN), łączącej wszystkie sieci lokalne w jeden spójny system. Dzięki temu przygotowano kadry do bardziej efektywnego wykorzystania sprzętu, przez wdrożenie nowych metod przetwarzania, wykorzystanie zalet sieciowej eksploatacji sprzętu, wykorzystanie nowoczesnego oprogramowania i nowych metod gromadzenia i korzystania ze zbiorów danych czy zapewnienia bezpieczeństwa zbiorów danych przed uszkodzeniem i nieupoważnionym dostępem itp.

Uruchomiony pod koniec 1993 r., w ramach wspomnianych inwestycji informatycznych, Internet był początkowo wykorzystywany jako media informujące zainteresowanych użytkowników o możliwości uzyskania wyników badań statystycznych oraz o zasadach metodologicznych niektórych ważniejszych badań. Te funkcje informacyjne były stopniowo rozszerzane i obecnie Internet stanowi jedną z najważniejszych form udostępniania informacji opracowywanych przez GUS i urzędy statystyczne dla szerokiego kręgu użytkowników. O rosnącym zainteresowaniu witryną internetową GUS świadczy stale rosnąca liczba „odwiedzających” ją użytkowników. W 1999 r. liczbaostępów do witryny internetowej GUS wynosiła (w tys.) — 240, w 2000 r. — 355, w 2001 r. — 531, w 2002 r. — 776, a w 2007 r. — 4350.

Od pewnego czasu służby statystyczne wykorzystują Internet w charakterze narzędzia i metody uzyskiwania informacji bezpośrednio od respondentów, zastępując stosowane dotychczas sprawozdania i kwestionariusze papierowe przesyłane tradycyjną pocztą lub zbierane bezpośrednio przez ankieterów. Jeśli metoda ta znajdzie uznanie respondentów i zostanie konsekwentnie wdrożona, to przyczyni się ona do zasadniczego usprawnienia najbardziej kłopotliwego i pracochłonnego oraz najbardziej obciążonego błędami etapu opracowań statystycznych. Dzięki zapewnieniu automatycznej kontroli kwestionariuszy elektronicznych przyczyni się więc do poprawy jakości informacji. Ze względu na ogromne znaczenie tej metody badań dla usprawnienia całego systemu zbierania i kontroli danych, zasługuje ona na

<sup>9</sup> Program PHARE (*Poland and Hungary Assistance for Restructuring their Economies*), rozszerzony następnie na pozostałe 8 krajów, które wraz z Polską i Węgrami przystąpiły do Unii w 2004 r., a później dodatkowo na Bułgarię i Rumunię. Statystyka polska korzystała kilkakrotnie z różnych podprogramów tematycznych programu PHARE w latach 90. ub. wieku oraz w pierwszych latach tego tysiąclecia. Szerzej na ten temat patrz Piskorz (2001).



bardzo wnikliwą analizę, w celu uprzedzenia i ewentualnego usunięcia mogących wystąpić, zwłaszcza w początkowym okresie jej stosowania, trudności i zahamowań.

Ważną częścią składową unowocześnienia systemu informacji statystycznej, chociaż niewiążąca się bezpośrednio z przetwarzaniem i analizą danych, ale mająca bardzo duże znaczenie dla podnoszenia wiedzy i kwalifikacji kadr statystycznych oraz udostępniania informacji i popularyzacji statystyki, było rozpoczęcie w latach 1996—1998 i kontynuowanie w latach 2000—2005 komputeryzacji Centralnej Biblioteki Statystycznej i włączenie jej do krajowego i międzynarodowego systemu wymiany informacji.

Realizacja dużym wysiłkiem organizacyjnym i finansowym GUS oraz otrzymana pomoc z UE umożliwiły poważne unowocześnienie całej infrastruktury informatycznej polskiej statystyki. Pozwoliło to nie tylko usprawnić gromadzenie, przetwarzanie, przechowywanie, analizę i udostępnianie informacji statystycznej, ale, co bardziej istotne, wprowadzić zasadnicze pozytywne zmiany w całym systemie informacji statystycznej. Na ważniejsze elementy zmian zwrócimy uwagę w następnym punkcie.

### WPLYW METOD I NARZĘDZI INFORMATYCZNYCH NA ROZWÓJ STATYSTYKI

Wykorzystanie środków technicznych do opracowania materiałów badań statystycznych w początkowym okresie, aż do zastosowania komputerów, przyczyniło się w istotnym stopniu do usprawnienia i przyspieszenia realizacji bardzo pracochłonnych i żmudnych operacji przetwarzania danych. Wpłynęło również na porządkowanie wewnętrznego obiegu dokumentów statystycznych, poprawę dokładności operacji obliczeniowych oraz podniesienie organizacji i dyscypliny, wynikającej z wymogów technologicznego procesu zmechanizowanych opracowań. W tym okresie zaczęto również zwracać większą uwagę na konieczność bardziej starannego przygotowania badań, w wyniku czego w sposób bardziej systematyczny zaczęto rozwijać projektowanie badań statystycznych. Ograniczone możliwości technologiczne „przedkomputerowych” środków informatyki nie wpłynęły jednak na bardziej istotne przekształcenia w samej statystyce. Stało się to możliwe dopiero w warunkach zastosowania komputerów.

Jeśli spojrzeć na rozwój zastosowań komputerów w statystyce z punktu widzenia ich wpływu na kształt i poziom statystycznego systemu informacyjnego, to w tym procesie można wyróżnić wyraźnie trzy etapy. Ze względu na ich wpływ na kształtowanie stopnia dojrzałości tego systemu, można je nazwać **generacjami zastosowań** lub **generacjami statystycznego systemu informacyjnego**. Można tu wyróżnić **generację pierwszą** — systemy formularzowe (tematyczne), **generację drugą** — systemy informacyjne oparte na wspólnej bazie danych oraz **generację trzecią** — systemy informacyjno-analityczne, tworzone przy bezpośrednim udziale statystyków, z wykorzystaniem mikrokomputerów i sieci komputerowych.

Nie należy traktować tego podziału w sposób sztywny, podobnie jak trudno ustalić ściśle przedziały czasowe trwania poszczególnych generacji. Przebiegają one w różnych krajach w rozmaity sposób. Często poszczególne generacje istnieją równolegle w różnych podsystemach. Podział ten jest jednak bardzo pomocny dla ilustracji istotnych problemów związanych z zastosowaniem informatyki i zwiększeniem jej wpływu na rozwój statystycznego systemu informacyjnego.

#### **Systemy formularzowe**

Mniej więcej do połowy lat 60. ub. wieku, a w Polsce o kilkanaście lat dłużej, zastosowanie komputerów w statystyce sprowadzało się do autonomicznego przetwarzania materiałów poszczególnych badań. Każde badanie oparte, z zasady, na odrębnym kwestionariuszu



lub formularzu sprawozdawczym traktowano jako odrębny system przetwarzania danych. Opracowywano dla niego osobno założenia oraz przeważnie bardzo dużą liczbę programów komputerowych niezbędnych do tworzenia zbiorów, szczegółowej kontroli poprawności danych w zbiorze, korygowania danych oraz opracowania i drukowania tablic. Taki system organizacji przetwarzania panował powszechnie we wszystkich służbach statystycznych (Cox, Croot, 1991). W Polsce przyjęto go nazywać przetwarzaniem tematycznym lub **formularzowym**.

W gospodarczych zastosowaniach komputerów, poza statystyką, podobne systemy przetwarzania przyjęto nazywać systemami **transakcyjnymi** lub **dziedzicznymi** (Flakiewicz, 1990; Radzikowski, 1990). Istota tych systemów sprowadza się do automatyzacji występujących masowo czynności związanych z opracowaniem materiałów poszczególnych tematów badań dotyczących, przykładowo, zużycia materiałów, stanu zapasów, zatrudnienia i wynagrodzeń, inwestycji, środków trwałych, wyników finansowych itp. Zakres tematyczny informacji wynikowych ograniczony jest tu do tematyki określonego kwestionariusza badania lub formularza sprawozdawczego. W przypadku formularzy wielotematycznych, odrębnym systemem mogą być poszczególne ich działy tematyczne.

Główny wpływ zastosowania komputerów na podniesienie poziomu informacji statystycznej w systemach formularzowych polegał na:

- 1) znacznym zmniejszeniu pracochłonności opracowania danych i dzięki temu skróceniu terminów uzyskiwania wyników. Źródłem tego przyspieszenia był fakt, że dane jednostkowe raz wprowadzone do komputera (do początku lat 70. ub. wieku głównie z kart dziurkowanych) i zapisane na magnetycznych nośnikach informacji mogły być wykorzystywane wielokrotnie do uzyskiwania tablic statystycznych w wielu wymaganych przekrojach. Na tym etapie opracowań wynikowych prędkość przetwarzania stosowanych wtedy komputerów przewyższała 500-krotnie prędkość stosowanych poprzednio maszyn systemu kart dziurkowanych;
- 2) znacznym rozszerzeniu zakresu opracowań wynikowych, głównie poprzez zwiększenie liczby tablic sporządzanych w różnych układach, co stworzyło podstawę rozszerzenia i pogłębienia analiz wyników badań;
- 3) istotnym rozszerzeniu kontroli poprawności danych źródłowych. Obowiązkową częścią składową każdego systemu przetwarzania danych były szczegółowe procedury kontrolne, zarówno poszczególnych pozycji, jak i pewnych informacji grupowych odznaczających się logicznymi powiązaniem w ramach danej grupy informacji (powiązanie pomiędzy cechami poszczególnych osób w ramach rodziny lub gospodarstwa domowego, zgodność logiczna i rachunkowa między danymi w ramach przedsiębiorstwa itp.). W wielu krajach rozwinęły się różnorodne, często bardzo skomplikowane metody kontroli poprawności danych, oparte na porównaniach z analogicznymi informacjami za okresy ubiegłe, porównaniach zakresu wartości poszczególnych cech między podobnymi jednostkami badania, porównaniach granic wielkości poszczególnych wartości względnych, obliczonych na podstawie wybranych cech (średnie ceny, średnie wynagrodzenia, wartość środków trwałych na zatrudnionego itp.).

Szczególnie skuteczną formą kontroli danych z zawartością rejestrów lub katalogów. Przykładowo, istniała możliwość kontroli poprawności wiodących cech poszczególnych przedsiębiorstw metodą ich porównania z zawartością rejestru jednostek typu REGON lub porównania przyczyn zgonów w indywidualnych kartach zgonu z oficjalnym wykazem przyczyn zgonów, z możliwością różnicowania prawdopodobieństwa ich występowania w zależności od płci, wieku, wykonywanego zawodu itp. Kontrola indywidualnych danych z zawartością rejestrów, oprócz sprawdzenia poprawności poszczególnych symboli wiodących (rodzaj działalności, forma organizacyjno-prawna, symbol terytorialny, powiązanie



danej jednostki z jednostką macierzystą itp.) umożliwia również sprawdzenie kompletności zbioru danych oraz automatyczne przeniesienie wszystkich niezbędnych symboli informacyjnych i klasyfikacyjnych do pozycji jednostkowych danego zbioru. Pozwala to uniknąć ręcznej symbolizacji oraz wielu kłopotów związanych z niebezpieczeństwem popełnienia dodatkowych błędów, gdyby ta czynność była wykonywana jak poprzednio — ręcznie;

4) możliwości szerszego rozwoju zastosowania nowoczesnych metod matematycznych i ekonometrycznych, zastosowania metody reprezentacyjnej oraz ekonometrycznych metod analiz i opracowywania prognoz. Stworzenie tych możliwości, nawet w ograniczonym zakresie, sprzyjało jednocześnie rozwojowi teoretycznych badań w tej dziedzinie oraz realizacji praktycznych obliczeń i analiz (por. *Zastosowanie ...*, 1969). Pozwoliło to również opracować pierwsze programy wdrażania metod matematycznych w pracach GUS. Należy zwrócić tu uwagę, że organizacja przetwarzania, ukształtowana pod wpływem wieloletniej tradycji przetwarzania przy użyciu maszyn systemu kart dziurkowanych, dość szybko ujawniła cały szereg mankamentów. Chodzi tutaj zarówno o sposoby zaspokojenia informacyjnych potrzeb użytkowników, jak i racjonalność organizacji procesu eksploatacji systemów informatycznych.

Z punktu widzenia użytkowników, zwłaszcza spoza oficjalnych służb statystycznych, podstawowym mankamentem systemów formularzowych było rozdrobnienie i izolacja tematyczna wyników poszczególnych badań. Powstawały trudności w dostarczeniu odbiorcom całościowych, zintegrowanych informacji, charakteryzujących w sposób wszechstronny określone zjawiska oraz procesy gospodarcze i społeczne.

Typowymi procedurami opracowania informacji wynikowych w systemach formularzowych było sporządzanie i publikowanie obszernych informacji tematycznych o charakterze źródłowym. Tego rodzaju opracowania, w wielu przypadkach były przez użytkowników oceniane pozytywnie, zwłaszcza wtedy, gdy dotyczyły jednorodnych zjawisk masowych. Nie odpowiadały one jednak potrzebom prowadzenia pogłębionych analiz, w których konieczne jest porównanie szerokiego zakresu powiązanych ze sobą informacji. Aby zaspokoić rozszerzone potrzeby w warunkach systemów formularzowych niezbędne byłoby sortowanie i łączenie poszczególnych zbiorów i tworzenie nowych, połączonych zbiorów. Wymagałoby to opracowywania dodatkowych programów oraz manipulacji na zbiorach, powodując w konsekwencji zwiększenie pracochłonności i kosztów opracowań. Co gorsza, nie zawsze było to możliwe, ponieważ często różne systemy formularzowe nie były z sobą w pełni spójne i ich wyniki nie zawsze mogły być porównywane.

Innym mankamentem systemów formularzowych, którego skutki odczuwali zarówno użytkownicy informacji, jak i sami informatycy było zastosowanie przestarzałej technologii przetwarzania, opartej na indywidualnym oprogramowaniu opracowywanym dla każdego tematu (systemu formularzowego). To prowadziło do znacznego wzrostu pracochłonności programowania oraz do opóźnień w realizacji dodatkowych zamówień użytkowników. Sytuację tę scharakteryzował obrazowo szef wielkiego banku amerykańskiego stwierdzeniem: *...przedtem zwykle na moje zapytanie otrzymywałem odpowiedź w ciągu dwóch dni. Teraz muszę za każdym razem czekać aż napiszą program. Trwa to dwa tygodnie* (Withington, 1974). Mankament ten szczególnie silnie odczuwany był w początkowym okresie komputeryzacji. W kolejnych latach, w wielu urzędach statystycznych opracowano szereg programów uniwersalnych, zapewniających realizację typowych operacji (wprowadzanie danych, aktualizacja, obliczanie i drukowanie tablic). Znacznie usprawniło to procedury uzyskiwania doraźnych informacji w sposób bardziej operatywny.

Najbardziej rozpowszechnionym wariantem realizacji systemów formularzowych, zwłaszcza w początkowych latach ich stosowania, było powszechne stosowanie tzw. przetwarzania wsadowego z przewagą stosowania sekwencyjnej organizacji zbiorów. Brak



w owym czasie odpowiedniego sprzętu i oprogramowania, zapewniającego bezpośredni dostęp do zbiorów za pośrednictwem większej liczby urządzeń końcowych (terminali), uniemożliwiał nie tylko uzyskiwanie informacji na doraźne żądania użytkowników. Komplikował też i wydłużał procedury kontroli i korekty danych, co w przypadku większej liczby błędów w danych źródłowych powodowało konieczność wielokrotnego drukowania i poprawiania tzw. wykazów błędów, wydłużając nadmiernie ogólny czas przetwarzania. Często czas wielocyklicznej kontroli, zwłaszcza w przypadku konieczności udziału urzędów statystycznych w poprawianiu wykazów błędów, zajmował do 50% ogólnego czasu opracowania.

Brak łatwego, bezpośredniego dostępu do informacji gromadzonych w komputerze oraz trudności w operatywnym uzyskiwaniu dodatkowych, nieprzewidzianych w założeniach, informacji wynikowych prowadziły do tego, że w ramach każdego systemu formularzowego opracowywano zwykle bardzo dużą liczbę tablic wynikowych w wielu szczegółowych przekrojach. Służyły one do opracowania publikacji oraz stanowiły zasób informacyjny dla poszczególnych departamentów do prowadzenia prac analitycznych i udzielania informacji w przypadku dodatkowych zapytań. W ten sposób departamenty gromadziły u siebie swoje „bazy danych” w postaci obszernych tablic, co było rozwiązaniem nieracjonalnym. Było to jednak usprawiedliwione trudnościami w operatywnym uzyskiwaniu dodatkowych informacji z obszernych zbiorów tematycznych, przechowywanych w sposób sekwencyjny na taśmach magnetycznych.

Systemy formularzowe (a w zastosowaniach informatyki poza statystyką — systemy transakcyjne) oceniane są również krytycznie przez samych informatyków, z powodu szeregu cech prowadzących do obniżenia efektywności eksploatacji systemów. Pierwsza uwaga krytyczna dotyczyła tzw. redundancji (nadmiaru) informacji, czyli konieczności powtarzania niektórych informacji w wielu różnych zbiorach tematycznych. Przykładowo, w przypadku statystyki gospodarczej w każdym zbiorze tematycznym musiał wystąpić symbol przedsiębiorstwa, rodzaj działalności, symbol terytorialny, forma własności, forma organizacyjno-prawna itp. Powodowało to ogólne zwiększenie objętości przechowywanej informacji oraz zwiększenie pracochłonności jej przetwarzania. W rzeczywistości redundancja w zbiorach tematycznych była wielokrotnie większa z uwagi na konieczność przechowywania tych samych zbiorów w różnych układach, aby zmniejszyć czas wielokrotnego sortowania zbiorów przed opracowaniem tablic w pokrewnych układach. Zmora informatyków było również zapewnienie jednolitego systemu aktualizacji **wszystkich** zbiorów tematycznych w przypadkach korekty cech wspólnych dla rozmaitych zbiorów, zwłaszcza gdy dysponentami zbiorów były różne komórki organizacyjne służb statystycznych.

Drugi „informatyczny” zarzut wobec tradycyjnej organizacji przetwarzania formularzowego (transakcyjnego) dotyczył stosowanych metod programowania, które zakładały szczegółowy opis struktury rekordów zbioru w każdym programie użytkowym. Powodowało to konieczność zmiany znacznej liczby programów w każdym przypadku wprowadzenia nawet „drobnych” zmian w układzie dokumentów lub w układzie tablic wynikowych. Rozwiązanie tej trudności przyniosło dopiero przejście na zasadniczo różny system oprogramowania, zwany Systemem Zarządzania Bazą Danych (SZBD), w warunkach przejścia do drugiej generacji systemu informacyjnego.

### *Systemy informacyjne wykorzystujące wspólną bazę danych*

Pojawienie się na światowym rynku w połowie lat 60. ub. wieku komputerów trzeciej generacji, odznaczających się większą prędkością obliczeniową, urządzeniami masowej pamięci o bardzo krótkim (rzędu kilkunastu milisekund) czasie dostępu do informacji, możliwością bezpośredniego kontaktu użytkownika z komputerem z wielu stanowisk (terminali)



oraz znaczne udoskonalenie metod programowania, doprowadziło do sytuacji, w której dotychczasowa organizacja przetwarzania formularzowego stawała się coraz bardziej anachroniczna.

Nowe podejście do organizacji systemów informacyjnych, uwzględniające z jednej strony krytyczną ocenę osiągniętego poziomu zaspokojenia rzeczywistych potrzeb informacyjnych użytkowników, a z drugiej strony, pojawiające się znacznie większe możliwości nowych środków informatyki i oprogramowania, polega na zamianie przetwarzania opartego na systemach formularzowych, w nowy system przetwarzania i gromadzenia informacji, oparty na koncepcji tzw. **baz danych**. Podobny proces zmian w zastosowaniach informatyki w gospodarce przyjęto nazywać przechodzeniem od systemów transakcyjnych do systemów informacyjnych zarządzania.

Statystyk pod pojęciem bazy danych rozumie zorganizowany w logiczny sposób zasób informacji. Składa się on z połączonych zbiorów informacji charakteryzujących wszechstronnie określone zjawiska i procesy. Przechowywany jest ten zasób w taki sposób, aby użytkownik w każdej chwili miał zapewniony do niego bezpośredni dostęp (zarówno do poszczególnych elementów informacji zawartych w bazie, jak i do informacji pochodnych, możliwych do uzyskania na podstawie informacji zawartych w bazie po określeniu przezeń algorytmu uzyskiwania tych informacji). W systemie takim rolę podstawowego źródła informacji spełnia wspólna baza danych, tworzona w wyniku łączenia (integracji) zbiorów tematycznych (formularzowych) oraz wszelkich innych zbiorów danych. Funkcje obsługi użytkownika realizuje SZBD, z którym użytkownik kontaktuje się za pośrednictwem stosunkowo łatwego języka, niewymagającego kwalifikacji informatycznych, poprzez klawiaturę i monitor ekranowy urządzenia końcowego (stacji roboczej).

Zalety systemu, wykorzystującego wspólną bazę danych w porównaniu z systemami formularzowymi, są tak istotne, że z pełnym uzasadnieniem można tu mówić o nowej generacji systemu informacyjnego. Na podkreślenie zasługuje zwłaszcza:

- 1) umożliwienie tworzenia połączonego zbioru danych na podstawie informacji uzyskiwanych z różnych źródeł (własnych badań statystycznych, badań prowadzonych przez inne instytucje, dokumentów i rejestrów administracyjnych itp.);
- 2) umożliwienie uzyskiwania przez samych statystyków zestawień analitycznych z połączonych zasobów informacji rzeczowych i finansowych, umożliwiających bardziej wszechstronną analizę wyników działalności podmiotów gospodarczych oraz kształtowania się sytuacji gospodarczej kraju;
- 3) zapewnienie bezpośredniego dostępu do zintegrowanych zasobów informacyjnych gromadzonych w porównywalnych szeregach czasowych oraz uzyskiwania różnorodnych danych na doraźne żądania użytkowników;
- 4) dalsza poprawa jakości informacji poprzez usunięcie różnorodnych niespójności pomiędzy danymi zawartymi w różnych systemach formularzowych. Usunięcie tych niespójności w warunkach systemów formularzowych, mimo bardzo szczegółowej ich kontroli, było praktycznie niewykonalne, ponieważ dopiero porównanie danych zawartych w rozmaitych zbiorach formularzowych pozwala wykryć ewentualne dalsze niedopuszczalne różnice między nimi;
- 5) zapewnienie jednolitego systemu aktualizacji połączonego zbioru danych w przypadku zmian w cechach jakiegokolwiek zbioru formularzowego lub innego zbioru danych tworzącego bazę;
- 6) zapewnienie programowej ochrony informacji gromadzonej w bazie danych przed nieupoważnionym dostępem oraz przed świadomym lub mimowolnym zniekształceniem zawartości bazy. Cecha ta ma szczególne znaczenie z uwagi na to, że wspólna baza danych z samej swojej istoty ma charakter wieloużytkowy i wielodostępny.



W Polsce do pierwszych prac związanych z przejściem od systemów formularzowych do systemu opartego o bazy danych przystąpiono na początku lat 70. ub. wieku, a więc tylko kilka lat później w stosunku do krajów przodujących w rozwoju systemu informacji statystycznej. Z szerszym wdrożeniem do eksploatacji pierwszych baz danych było gorzej z uwagi na wieloletnie opóźnienia w rozwoju bazy sprzętowej i brak nowoczesnego oprogramowania. Co więcej, należy stwierdzić, że wdrożenie pierwszych baz danych w statystyce polskiej w latach 1975—1982 nie przyniosło rzeczywistego przełomu jakościowego w funkcjonowaniu systemu informacyjnego statystyki. Wypracowane w Ośrodku Badawczo-Rozwojowym Statystyki (OBRS) nowoczesne koncepcje baz i banków danych zetknęły się z twardymi realiami przestarzałej bazy informatycznej, dostosowywanej z wielkim trudem do rosnących rozmiarów opracowań w systemach formularzowych i praktycznie nie nadającej się do realizacji systemu opartego na bazie danych.

Eksploatowane wówczas w GUS komputery nie miały podstawowych cech niezbędnych do racjonalnej eksploatacji baz danych. Posiadały natomiast ograniczoną pojemność pamięci o bezpośrednim dostępie, a nie miały ani technicznych, ani programowych możliwości współpracy z większą liczbą terminali, zapewniających łatwy dla użytkownika dostęp do bazy, ani odpowiedniego oprogramowania typu SZBD. W tej sytuacji trzeba było z konieczności wprowadzić szereg rozwiązań połowicznych. Między innymi, zamiast jednej bazy danych o przedsiębiorstwach wprowadzono kilka baz, zamiast bezpośredniego, operatywnego dostępu do bazy przez samych użytkowników zastosowano system korzystania z danych na zamówienie. Było to chyba głównym powodem rozczarowania statystyków do propagowanych przez informatyków idei baz danych. Brak gotowego systemu zarządzania bazą danych zmusił również informatyków do opracowania własnego oprogramowania o ograniczonych możliwościach, co udało się zrealizować jedynie dzięki ogromnemu wysiłkowi utalentowanego zespołu programistów GUS.

Może tu rodzić się pytanie, czy słuszne było podejmowanie nieprostych przecież prac analityczno-projektowych w dziedzinie baz danych, skoro nie było niezbędnych warunków do ich efektywnego wdrożenia, a przyjęte z konieczności rozwiązania połowiczne obniżyły zaufanie do samej idei baz danych. Oceniając jednak te prace z perspektywy z górą 30 lat, z pełnym przekonaniem twierdę, że biorąc pod uwagę wszystkie okoliczności wybrana wówczas droga była słuszną i dobrze świadczy o kwalifikacjach i zaangażowaniu ówczesnego zespołu statystyków i informatyków. Mimo bowiem szeregu ograniczeń, uwarunkowanych poziomem dostępnej technologii, osiągnięto rzeczywistą poprawę jakości informacji. I co niemiernie ważne, proces wdrażania baz danych stworzył warunki praktycznego rozwiązywania wielu problemów, mających duży wpływ na doskonalenie metodologii badań i opracowań statystycznych. Dzięki temu kilkanaście lat później, w nowych warunkach technologicznych, prace w dziedzinie systemów opartych na bazach danych można było wykonywać o wiele pewniej i szybciej. Zdobyte doświadczenia pozwalają nam również obecnie, w ramach międzynarodowej współpracy, występować w charakterze bardziej doświadczonego partnera.

Z punktu widzenia potrzeb dnia dzisiejszego, uwzględniając zmieniające się potrzeby informacyjne w nowych warunkach ustrojowych oraz istniejące zasadniczo nowe możliwości technologiczne, szczególne znaczenie mają następujące aspekty dotychczasowych doświadczeń prac nad tworzeniem baz danych.

**Po pierwsze**, w warunkach systemu informacyjnego opartego na bazach danych inaczej rozkłada się odpowiedzialność za jakość, w tym zwłaszcza za spójność danych w bazie. Przy systemach formularzowych każdy departament, według najlepszej wiedzy i rozeznania potrzeb, definiował tematykę i zakres podmiotowy badań. Zakres podmiotowy ustala się zwykle na podstawie generowanego dla każdego badania podzbioru jednostek z REGON-u, które zobowiązane są złożyć określony kwestionariusz badania (sprawozdanie). Jeśli kryte-



ria doboru jednostek do badania dla różnych formularzy nie są identyczne, a sprawozdania mają tworzyć wspólną bazę, to z góry wiadomo, że w czasie łączenia zbiorów wystąpią przypadki, w których w części zbiorów brak będzie pewnej liczby sprawozdań. Czasem różnice te wynikają z różnej definicji jednostek badanej zbiorowości (inaczej rozumiane pojęcie przedsiębiorstwa lub zakładu w badaniach z zakresu inwestycji, inaczej w sprawozdawczości finansowej, inaczej w sprawozdawczości z zatrudnienia i wynagrodzeń itp.).

Čzęsto przyczyną różnic są niejednolite kryteria podziału jednostek na duże, średnie i małe (przykładowo, wielkość jednostki według stanu zatrudnienia lub liczby pracujących w różnych dniach, przyjęcie w niektórych badaniach jako kryterium określenia klasy wielkości jednostki formy prowadzonej ewidencji). Zdarzają się również bardziej prozaiczne przyczyny rozbieżności, jak np. generowanie katalogu jednostek dla rozmaitych badań na podstawie REGON-u według stanu aktualności w różnych terminach, zastosowanie indywidualnych zwolnień ze składania niektórych sprawozdań czy wreszcie niezłożenie niektórych sprawozdań przez część jednostek. Szereg trudności występuje w czasie łączenia zbiorów miesięcznych, półrocznych i rocznych, w przypadku gdy część jednostek w ciągu roku w sposób niejednolity zmienia kategorię wielkości itp.

Wszystkie wymienione, pozornie drobne i techniczne, sprawy są (obok występujących różnic w sformułowaniach i interpretacji niektórych pojęć) najczęstszą przyczyną braku spójności zbiorów danych i ogromnych trudności w toku opracowań rachunków zbiorczych.

Aby statystyka mogła dysponować spójnymi bazami danych, musi istnieć silny (niekoniecznie liczny) zespół spełniający funkcje administratora bazy danych. Ma on być odpowiedzialny w pełni za kompletność, integralność i spójność informacji w bazach, w tym również za inicjowanie opracowania procedur przywracających spójność informacji w przypadku jej naruszenia. Jednym z rodzajów takich procedur powinno być szacowanie brakujących danych lub korygowanie jawnych błędów.

**Po drugie**, w warunkach oparcia systemu informacyjnego na bazach danych, kiedy połączone zasoby informacyjne pochodzą z rozmaitych źródeł (różne sprawozdania, dokumenty administracyjne, rejestry, badania reprezentacyjne, szacunki ekspertów itp.), wykształca się pewna autonomizacja między czynnościami gromadzenia i bieżącego utrzymania danych (aktualizacja, przywracanie porównywalności) i czynnościami związanymi z wykorzystaniem informacji uzyskiwanych ze wspólnej bazy (opracowanie analiz, publikacji, wnioskowanie w sprawie modyfikacji i rozwoju bazy danych, ocena metod zapewniających spójność danych itp.).

Chodzi o to, aby z jednej strony zagwarantować w pełni jednolitość dyspozycji i odpowiedzialności za zachowanie integralności bazy, a z drugiej — o stworzenie odpowiednich warunków rozwoju prac badawczych, analitycznych i działalności promocyjnej. W urzędach statystycznych za granicą, w których technologia baz danych zajmuje bardziej znaczące miejsce, ten aspekt podziału pracy ma nawet odpowiednie konsekwencje organizacyjne.

Wynika to z tego, że bazy danych są wieloużytkowe. Z tego też względu musi istnieć jeden gestor określonej bazy, aby każdy użytkownik nie musiał się kontaktować z wieloma dysponentami poszczególnych części tematycznych bazy oraz aby w jednym miejscu mógł uzyskać informacje zarówno co do możliwości otrzymania określonych danych, jak i co do możliwości skorzystania z oprogramowania niezbędnego do wykorzystania tych danych zgodnie z jego potrzebami. Takie podejście i taka organizacja zakłada zarówno konieczność odejścia od wąskodzielzinowej (działowej) organizacji komórek statystycznych, jak i zerwania z tradycyjnie rozumianym podziałem na statystyków i informatyków.

**Po trzecie**, wdrożenie baz danych w statystyce znacznie utrudnia opanowanie szczegółowej wiedzy o zasobach informacyjnych statystyki. Chodzi tu o ekonomiczną treść i sposób liczenia poszczególnych wskaźników, o dopuszczalność ich porównywania, rodzaje i szczegółowość klasyfikacji i nomenklatur, w jakich można wyrazić wskaźniki i kategorie



ekonomiczne występujące w bazie, organizację, sposób i miejsce przechowywania określonej bazy czy rodzaje i dostępność programów komputerowych niezbędnych do opracowania danych itp.

Przez cały poprzedni okres stosowania informatyki w statystyce mieliśmy do czynienia z sytuacją, w której coraz konsekwentniej automatyzowano proces systemu informacyjnego statystyki w części dotyczącej informacji liczbowych. Jednocześnie utrzymywano tradycyjny „ręczny” opis wszystkich elementów tego systemu. W warunkach systemów formularzowych sytuację tę można było jeszcze tolerować, ponieważ statystyk odpowiedzialny za wąski tematycznie odcinek badań (gospodarz tematu), na ogół wszystkie wymienione informacje był w stanie zapamiętać. W warunkach wspólnej bazy danych powstaje konieczność zastosowania komputerowych narzędzi opracowania nie tylko danych liczbowych tworzących bazę, ale i szczegółowych opisów oraz komentarzy odnoszących się do informacji liczbowych zawartych w bazie, jak też do całej „infrastruktury informacyjnej” towarzyszącej bazie. Dlatego bardziej złożone, wielotematyczne bazy danych muszą się w rzeczywistości składać z bazy danych liczbowych oraz z bazy identyfikującej, opisującej i komentującej informacje zawarte w bazie danych liczbowych. Tę drugą część bazy informacyjnej statystyki przyjęto nazywać bazą metadanych lub metainformacji.

Pojęcie metadanych definiuje się zwykle jako wszelkie dane, za pomocą których można opisać dane liczbowe, ich pochodzenie, znaczenie, powiązanie z innymi danymi, sposób zbierania, stopień dokładności itp. Podobnie, jako metainformacje określa się zwykle wszelkie informacje, które służą do opisu informacji. Tak więc elementami metainformacji są podstawowe rejestry opisujące obiekty obserwacji, klasyfikacje i nomenklatury, katalogi wskaźników statystycznych, wykazy formularzy, instrukcje i opisy metodologiczne, słowniki obowiązujących pojęć i definicji, informacje o miejscu przechowywania danych, o zasadach korzystania z danych, o stopniu dokładności i wiarygodności danych, informacje o programach komputerowych niezbędnych do przetwarzania i korzystania z danych itp. (Oleński, 1978, 1997).

Problematyką tworzenia i wykorzystania systemów metainformacji urzędy statystyczne zajmują się od dawna. W latach 1980—1984, w ramach KSE i przy pomocy finansowej UNDP, realizowano wspólny program badawczy „Zastosowanie komputerów w statystyce”. W jego ramach pracowała wydzielona grupa badawcza Metis (*Meta Information System*), przy aktywnym udziale statystyków polskich<sup>10</sup>. Grupa Metis zainicjowała międzynarodową dyskusję na temat roli i zakresu pojęciowego metadanych i metainformacji. Niezwykle cennym dorobkiem grupy było uzgodnienie ważniejszych pojęć i wymiana poglądów na temat strategii projektowania systemu metainformacji. Konkretnym wynikiem pracy tej grupy było opracowanie i opublikowanie poradnika na temat Metis w statystyce (*Users Guide...*, 1984). Od tego czasu nie zmalało zainteresowanie statystyków tym systemem. KSE, wspólnie z OECD i Eurostatem organizują co 2 lata sesje robocze na ten temat.

Cały system wspólnego oprogramowania baz danych liczbowych i baz metainformacji zmierza do tego, aby zapewnić możliwość udostępnienia użytkownikom nie tylko wskaźników liczbowych, ale również wszelkich dodatkowych informacji. Chodzi o podanie, gdzie dane są publikowane, gdzie publikowane są opisy metodologii badań, co oznaczają poszczególne wskaźniki, według jakich algorytmów się je oblicza, jakie programy kompute-

<sup>10</sup> Grupa projektowa systemu metainformacji (Metis) była jedną z 4 grup badawczych utworzonych w ramach programu pod nazwą Statistical Computing Project (SCP). Oprócz grupy Metis pracowały grupy specjalistyczne zajmujące się: opracowaniem standardowego programu tworzenia tablic statystycznych, automatyczną kontrolą i korektą danych oraz analizą funkcjonowania SZBD RAPID. Przedstawiciele GUS uczestniczyli aktywnie w realizacji tego projektu, w tym zwłaszcza w grupie projektowania tworzenia tablic statystycznych — Stanisław Sielużycki i Krystyna Pruska a w grupie Metis Józef Oleński. Pracami całego projektu z ramienia EKG/UNDP kierował Komitet Sterujący, któremu przewodniczył autor artykułu.



rowe można wykorzystać do dalszej analizy danych itp. Taki system, oparty na wykorzystaniu wspólnych baz danych, zintegrowany z systemem metainformacji, stwarzający dla użytkowników warunki bezpośredniego korzystania z zasobów informacyjnych oraz informacji opisującej te zasoby i informacje pochodne, będzie miał zasadnicze znaczenie dla radykalnego podniesienia jakości i efektywności funkcjonowania statystyki. Można uznać go za **III generację statystycznego systemu informacyjnego**.

Mimo wieloletnich, trwających już ponad 30 lat, wysiłków krajowych i międzynarodowych zespołów z udziałem wybitnych specjalistów i mimo wielu cząstkowych prób wdrożeń, nadal jednak do pełnego urzeczywistnienia docelowego systemu informacji statystycznej III generacji, integrującego bazy danych liczbowych z systemem metainformacji, jest jeszcze droga daleka. Wydaje się, że najważniejszą przyczyną tego stanu jest ogromna rozległość tematyczna, złożoność i zróżnicowanie tego systemu, zapewne większe niż w przypadku baz danych liczbowych. Wynika to z bardzo zróżnicowanych wymagań jej użytkowników, dla których musi być projektowany system metainformacji. Do użytkowników tego systemu zaliczamy (*Common...*, 2006; Oleński, 1997):

- użytkowników zewnętrznych krajowych: pracowników organów administracji, polityków, ekonomistów, działaczy gospodarczych, analityków, pracowników naukowych, organizacji pozarządowych, bibliotek, dziennikarzy, studentów i innych osób zainteresowanych informacjami statystycznymi;
- użytkowników międzynarodowych: osoby indywidualne korzystające z danych do analiz międzynarodowej sytuacji gospodarczej i społecznej, przedsiębiorstwa wielonarodowe, organizacje międzynarodowe, w tym zwłaszcza takie, jak: ONZ, ORCD, UE, MOP, MFW, Europejski Bank Centralny. Pracownicy tych organizacji są szczególnie zainteresowani uzyskiwaniem informacji, na podstawie których mogliby się przekonać o zgodności metodologicznej statystyki krajowej ze standardami międzynarodowymi;
- pracowników krajowych służb statystycznych: personel kierowniczy, projektantów systemów statystycznych, metodologów, analityków, pracowników odpowiedzialnych za udostępnianie informacji, administratorów systemów informacyjnych, pracowników badawczo-rozwojowych, informatyków odpowiedzialnych za przetwarzanie i przechowywanie danych;
- respondentów dostarczających dane ujmowane w programach badań. Część z nich korzysta jednocześnie z informacji do prowadzenia własnych ocen i analiz i dlatego potrzebuje wielu wyjaśnień, komentarzy i uzupełnień. Prócz tego część respondentów potrzebuje informacji niezbędnych im do wyjaśnienia wątpliwości dotyczących poprawnego wypełnienia kwestionariuszy objętych programami badań.

W każdej z wymienionych grup użytkowników Metis potrzeby dotyczące zakresu i szczegółowości poszczególnych elementów informacji są bardzo zróżnicowane. Przykładowo, w zależności od kategorii pracowników służb statystycznych, mogą oni oczekiwać odpowiedzi na pytania: do jakich celów wykorzystywane są poszczególne informacje statystyczne, czy użytkownicy są zadowoleni z jakości danych i metadanych, ich treści, dokładności, terminowości, jakie uwagi krytyczne są zgłaszane przez użytkowników, jakie są koszty opracowania poszczególnych tematów badań, jaka jest ich struktura, z jakich źródeł danych korzysta statystyka, na podstawie jakich wzorów dokumentów, jaki jest system udostępniania informacji, jaka jest treść dokumentacji projektowania badań (wzory tablic, kwestionariuszy, klasyfikacje, rejestry, oprogramowanie), jaka jest treść poszczególnych badań zmiennych (słownik pojęć), kim są główni użytkownicy informacji, jakie są wymagania poszczególnych organizacji międzynarodowych itp.

Z przytoczonego wyliczenia potencjalnych potrzeb informacyjnych jednej tylko grupy użytkowników Metis można się zorientować, jak trudne zadania stoją przed projektantami tego systemu. Informacje zamieszczane w Metis nie mogą pełnić roli podręcznika ani ency-



klopedii statystyki. Jednocześnie znaczna ich część musi mieć charakter bardziej szczegółowy. Pamiętać także należy o konieczności bieżącej aktualizacji każdej zamieszczanej w Metis informacji, w przypadku każdej zmiany w przepisach, w programach badań, w podziale administracyjnym itp. Do zbudowania takiego systemu niezbędna jest aktywna rola w jego projektowaniu statystyków korzystających z bezpośredniego wsparcia najwyższego kierownictwa służb statystycznych. Bardziej szczegółowe zalecenia KSE, Eurostatu i OECD w sprawie metod projektowania systemu informacji statystycznej sterowanego systemem metainformacji można znaleźć w *Common...* (2006).

Obecnie w wielu urzędach statystycznych obserwuje się znaczny wzrost aktywności w kwestii projektowania systemów informacyjnych z wykorzystaniem wiodących funkcji Metis, jako elementu sterującego całym systemem. Prace w tej dziedzinie prowadzone są w wielu krajach, w tym zwłaszcza w Kanadzie, Nowej Zelandii, Australii, a także w krajach UE, w tym w nowych krajach Unii — w Czechach i na Łotwie (Zeila, 2005).

Z przyjemnością należy odnotować także wysiłki polskich statystyków i informatyków w tej dziedzinie. Zmierzają one w kierunku utworzenia systemu, w którym główną rolę odgrywać będzie system metainformacji i hurtowni danych. Hurtownia danych stanowi podstawowy zasób informacji liczbowych pochodzących z różnych badań i innych źródeł, po ich szczegółowej kontroli i integracji tematycznej i usunięciu ewentualnych błędów i niespójności między danymi pochodzącymi z różnych źródeł. Projekt przewiduje możliwość bezpośredniego korzystania z zasobów hurtowni danych, zarówno danych jednostkowych jak i zagregowanych, z zapewnieniem przestrzegania zasad pełnej ochrony przed nieupoważnionym dostępem. Rolę wiodącą w projektowanym systemie spełniać będzie podsystem metainformacji obejmujący wszystkie elementy metadanych odnoszące się zarówno do ogólnych zagadnień funkcjonowania statystyki, jak i do poszczególnych tematów badań. Będzie on zawierał słownik pojęć, klasyfikacje, informacje dotyczące poszczególnych badań oraz wszelkie inne informacje niezbędne dla różnych grup użytkowników (Kurkowski, Sielużycki, 2005).

W jednym artykule nie sposób omówić wszystkich aspektów nowego systemu oraz problemów czekających jeszcze na rozwiązanie. Wyrażam nadzieję, że moje młodsze koleżanki i koledzy, uczestniczący bezpośrednio w pracach projektowych nad doskonaleniem statystyki, z uwzględnieniem rosnących możliwości technologii informacyjnych, podejmą wysiłek zaznajomienia pracowników statystyki z najważniejszymi wyzwaniem stojącym przed statystyką, wynikającym z przechodzenia do III generacji systemu informacyjnego statystyki.

## Podsumowanie

W okresie 90 lat działalności GUS byliśmy świadkami ogromnych zmian w technice i metodach opracowania informacji w statystyce. Zmiany, zasługujące ze wszech miar na miano przełomowych, zachodziły niemal we wszystkich rodzajach środków technicznych. Wielu Czytelników pamięta zapewne stosowane w latach 1950—1970 hałaśliwe czterodziałaniowe elektryczne maszyny liczące, ważące 15—22 kg. Zajmowały one niemal połowę powierzchni biurka i wykonywały działania arytmetyczne wielokrotnie wolniej od ich następców — kalkulatorów elektronicznych, pracujących bezgłośnie, setki razy szybciej i ważących zaledwie kilkanaście do kilkuset gramów.

Maszyny licząco-analityczne, dzięki którym w latach 1922—1970 w naszej statystyce dokonała się prawdziwa rewolucja przez wprowadzenie automatycznego odczytu danych z kart dziurkowanych, zostały zastąpione przez komputery odczytujące dane wielokrotnie szybciej z nośników magnetycznych, pracujące kilka milionów razy szybciej od MLA i mające nieporównanie większe możliwości wykonywania działań logicznych. Pozwoliło



to zrewolucjonizować procesy kontroli i pamiętania danych oraz wzbogacić zakres analiz wyników badań z wykorzystaniem metod ekonometrycznych.

Wolne i obciążone błędami przesyłanie danych za pośrednictwem linii telegraficznych zainicjowane w połowie lat 60. ub. wieku zastąpiono elektronicznymi metodami transmisji z wykorzystaniem w latach 90. ub. wieku sieci komputerowych.

Zastosowanie mikrokomputerów zmieniło zasadniczo nie tylko metody pracy statystyków, ale i całą organizację przygotowania i obiegu informacji zarówno wewnątrz służb statystycznych, jak ich współdziałanie ze światem zewnętrznym.

Dzięki możliwości zapisu i „pamiętania” nieograniczonych niemal zasobów informacji zrealizowano idee wspólnych baz danych. W znacznym stopniu przyczyniło się to do bardziej wszechstronnego wykorzystania połączonych informacji i stworzyło warunki urzeczywistnienia marzeń wielu użytkowników w sprawie zapewnienia spójności między poszczególnymi informacjami.

Uruchomienie na początku lat 90. ub. wieku łączności internetowej i umożliwienie bezpośredniego korzystania z informacji statystycznych przez zewnętrznych użytkowników zapoczątkowało trwający proces przełamywania barier między informacyjnymi statystyki a ich użytkownikami. Obecnie stworzyły się również realne szanse bezpośrednio otrzymywania danych od respondentów z pominięciem szeregu czynności pośrednich (w tym sporządzania dokumentów papierowych, ich ponownego przenoszenia na nośniki elektroniczne oraz kontroli ich poprawności).

Informatycy w całym okresie dziewięćdziesięciolecia wspierali statystyków w wykonywaniu najbardziej pracochłonnych i żmudnych czynności. Jednocześnie inicjowali wprowadzanie zmian w samej statystyce, wyrażających się w bardziej racjonalnych metodach projektowania badań, organizacji ich przeprowadzania, w wysiłkach na rzecz poprawy jakości i pogłębianiu analiz, w metodach przechowywania i udostępniania informacji użytkownikom.

---

prof. dr hab. Tadeusz Walczak — GUS

#### LITERATURA

- Common Metadata Framework, Part A — Corporate Context* (2006), United Nations Economic Commission for Europe
- Cox N. W. P., Croot D. A. (1991), *Data editing in a mixed DBMS environment*, *Statistical Journal of the United Nations Economic Commission for Europe*, Vol. 8, No. 2, IOS Press
- Flakiewicz W. (1990), *Informacyjne systemy zarządzania. Podstawy budowy i funkcjonowania*, PWE, Warszawa
- Kurkowski K., Sielużycki S. (2005), *Statistical metainformation system and data warehouse as main subsystems of SIS — current situation and plans, supporting paper for Joint ECE/Eurostat/OECD Meeting on the Management of Statistical Information System (MSIS)*, Bratislava, Slovakia, 18—20 April 2005
- Lukasiak A. (1979), *Analiza ilościowego i jakościowego rozwoju techniki obliczeniowej w czasie 60-letniej działalności Głównego Urzędu Statystycznego*, wyd. powielaczowe, GUS-ZMiAOS
- Miller J. (1930), *Historia, rozwój i stan obecny maszynowego opracowania dat statystycznych w Głównym Urzędzie Statystycznym*, „Kwartalnik Statystyczny”, tom VII, zeszyt 2
- Oleński J. (1978), *W sprawie metodyki projektowania centralnych systemów informatycznych*, *Systemy informatyczne*, Seminarium SPIS'77, Wyd. GUS OBR, t. 1, Warszawa
- Oleński J. (1997), *Standardy informacyjne w gospodarce*, Wydawnictwa Uniwersytetu Warszawskiego, Warszawa
- Piskorz J. (2001), *Program rozwoju informatyki w statystyce publicznej*, „Wiadomości Statystyczne”, nr 9
- Radzikowski W. (1990), *Komputerowe systemy wspomagania decyzji*, PWE, Warszawa



- Rawski R. (1981), *Automatyzacja procesu pozyskiwania informacji dla sprawozdawczości operatywnej GUS*, „Wiadomości Statystyczne”, nr 4
- Users Guide to Metainformations Systems in Statistical Offices* (1984), United Nations, Printed by WUSEI-AR, Bratislava, October
- Walczak T. (1985), *Wprowadzanie masowych danych do komputerów*, wyd. II poprawione i uzupełnione, PWE, Warszawa
- Walczak T. (1993), *Informatyka w 75-leciu GUS*, „Wiadomości Statystyczne”, nr 1
- Withington F. G. (1974), *Five generations of Computer*, Harvard Review, No. 7
- Zastosowanie metod matematycznych w statystyce* (1969), Plenarna sesja Naukowej Rady Statystycznej, „Biblioteka Wiadomości Statystycznych” tom 7, GUS, Warszawa
- Zeila K. (2005), *Metadata Driven Integrated Statistical Data Processing and Dissemination System*, Central Statistical Bureau of Latvia

## SUMMARY

*The article presents a historical review of development of data processing's technical devices used in ninety years of official statistical services in Poland. In this period the significant changes have occurred — from the use of the simplest tools improving statisticians work, implementation of punched cards machines, through computers and micro-computers, computer networks and Internet. Modernisation of data preparation methods have been accomplished, starting from punched cards use, through the use of magnetic tape recording devices, the use of machine-readable data from paper documents, to currently used methods of direct data transmission via Internet.*

*Special attention was paid to the use of methods influenced on statistics transformation — from automation of individual survey data to the creation of common databases system controlled by meta-information subsystem.*

## РЕЗЮМЕ

*Статья является историческим обзором развития использования технических средств обработки данных в девятидесятилетней деятельности статистических служб в Польше. В этот период осуществились значительные перемены — с момента использования самых простых приборов (инструментов) улучшающих работу статистиков, использования счетно-перфорационных машин, а потом компьютеров и микрокомпьютеров, компьютерных сетей и Интернета. В статье описывается модернизация методов подготовки данных, которая началась с использования перфокарт, через устройство для непосредственной записи данных на магнитных носителях, автоматическое считывание бумажных документов, по используемый в настоящее время метод прямой передачи данных с помощью Интернета.*

*Особенное внимание статья обращает на влияние информатических методов на изменения в статистике — от автоматизации материалов отдельных обследований, по созданию общих баз данных управляемых подсистемой метаинформации.*