

CEZARY KUKLO
(Białystok)

TECHNIKI KOMPUTEROWE W DEMOGRAFII HISTORYCZNEJ

Już prawie od pół wieku rozwojowi naszej cywilizacji towarzyszą elektroniczne maszyny cyfrowe służące do automatycznego przetwarzania danych¹. Podobnie jak i wiele innych wynalazków również komputery mają swoją historię, w której arytmometr zbudowany w 1642 r. przez Blaise Pascala (działał na zasadzie przekładni zębatej, dodając i odejmując liczby kilkucyfrowe) zajmuje miejsce znaczące². Chociaż pierwsze maszyny cyfrowe były konstruowane z myślą o obliczeniach naukowo-technicznych, to prawie natychmiast – w początkach lat pięćdziesiątych – znalazły zastosowanie w badaniach nauk ścisłych i przyrodniczych, a nieco później w naukach społecznych, zwłaszcza w demografii historycznej. Przerodnictwo demografii przeszłości na polu zastosowania komputerów było uwarunkowane jej bazą źródłową, która jak wiadomo w badaniach nad stanem zaludnienia wykorzystuje spisy ludności różnorodnej proveniencji, zaś w badaniach nad ruchem naturalnym ludności metryki parafialne i akta stanu cywilnego. Zarówno spisy, jak i księgi metrykalne są źródłami o charakterze masowym i co ważniejsze o ujednoliconym schemacie, gdzie pewien określony zespół cech występuje systematycznie niemal w każdym zapisie³. Pozwalają więc one na daleko szersze stosowanie metody ETO niż jest to w wypadku innych źródeł – o mniej zunifikowanej postaci.

Trudno dokładnie zaprezentować wszystkie problemy badawcze demografii historycznej, przy których rozwiązywaniu stosowane są techniki komputerowe. Wydaje się jednak, że od samego początku badacz oczekiwał od maszyny cyfrowej możliwości operowania przetwarzaniem statystycznym coraz większych rozmiarów kartotek numerycznych celem obliczenia najprzeróż-

¹ Pracę wykonano w ramach projektu badawczego Nr 1 1662 91 02 finansowanego przez Komitet Badań Naukowych w latach 1992–1994.

² Zob. szerzej H. Kaufmann, *Dzieje komputerów*, Warszawa 1980.

³ Por. opinie W. Kuli, *Problemy i metody historii gospodarczej*, Warszawa 1983, s. 358–360, 465–470, który w opracowaniu źródeł dotyczących faktów indywidualnych a występujących w skali masowej widział przyszłość badań demograficzno-społecznych. Zob. też S. Kowalska-Glikmin, *Akta masowe w badaniach historii społecznej*, w: *Metody i wyniki. Z warsztatu historyka źródeł społeczeństwa polskiego*, pod red. S. Kalabińskiego, przy współudziale J. Hensel i I. Rychlikowej, Warszawa 1980, s. 179–206.

niejszych współczynników demograficznych i wykonania innych analiz. Z drugiej strony, to właśnie komputer mógł podjąć zadanie próby modelowania zjawisk demografii społecznej zarówno w przeszłości⁴, np. jaką szansę miały dzieci w społeczeństwie przedprzemysłowym, aby poznać swoich dziadków⁵, jak i w przyszłości – usiłując zrekonstruować podział rodzin według typów i wielkości w zależności od zmian wybranych parametrów demograficznych⁶. Wreszcie chyba największe oczekiwania historyka demografa były i są związane do chwili obecnej z automatyczną rekonstrukcją rodzin.

Ożywienie, a nawet dynamiczny rozwój badań społeczno-demograficznych spowodowane formowaniem się dwu wielkich szkół naukowych: francuskiej i angielskiej, wybijało się jako najbardziej charakterystyczne zjawisko od lat sześćdziesiątych w historiografii światowej. Przypomnijmy, że pierwsza z nich – uznawszy księgi parafialne za kamień węgielny demografii prestatystycznej opracowuje je metodą L. Henry’ego – rekonstrukcji rodzin. Druga – wobec jakościowej słabości metryk angielskich – rozwinęła do perfekcji zbiorczą metodę ich wykorzystywania, ale równocześnie szeroko włączyła analizę wczesnych spisów ludności (parafialnych, list imiennych itp.). Swoje sukcesy światowe obie zawdzięczają w dużej mierze zastosowaniu technik komputerowych, choć ich kontynuatorzy w wielu innych krajach za bardziej miarodajne uznają prowadzenie na szeroką skalę badań nad najmniejszą komórką demograficzną – rodziną, najczęściej w zróżnicowaniu socjoekonomicznym.

Metoda rekonstrukcji rodzin autorstwa L. Henry’ego, zaproponowana w połowie lat pięćdziesiątych, w ciągu ostatnich lat nie została ulepszona. Polega ona na wyodrębnieniu z rejestrów parafialnych elementów dotyczących każdej rodziny, tj. łączeniu chrzcin, ślubu, zgonu współmałżonków i ich dzieci, a następnie na umieszczaniu całości na tej samej karcie⁷. Jej pomyslnie zakończenie umożliwia policzenie dla dawnych populacji na przestrzeni nieraz kilku wieków współczynników zawierania związków małżeńskich, płodności i umieralności. Przy czym nie chodzi tutaj o wyniki liczbowe interesujące same w sobie, ale również o poznanie ściślejszych związków istniejących w przeszłości pomiędzy ekonomiką, strukturą społeczną i ruchem ludności, które działały na poziomie pojedynczego człowieka i jego najbliższego otoczenia. Wokół kart rodziny historyk może grupować inne dokumenty, np.: testamenty, spisy, inwentarze. W ten sposób staje się możliwe socjologiczne badanie rodzin w zależności od charakterystyk demograficznych.

⁴ Zob. E. Van der Walle, *Analyse non nominative des sources écrites de tout ordre. De l'emploi des modèles en démographie historique*, „Annales de Démographie Historique” (dalej: „Annales DH”), 1972, s. 154–177; tenże, *Démographie historique et modèles*, w: *Démographie historique*, pod red. M.L. Marcilio i H. Charbonneau, Paris 1979, s. 131–151; J. Dupâquier, *Méthodes d'interprétation des sources non nominatives pour la construction modèles*, „Annales DH” 1972, s. 179–191.

⁵ Por. H. Le Bras, *Parents, grand-parents, bisaïeux*, „Population” 1, 1973, s. 9–37.

⁶ Zob. próbę J. Dupâquiera i M. Demoneta, *Ce qui fait les familles nombreuses*, „Annales – Economies, Sociétés, Civilisations” (dalej: „Annales ESC”), 4–5, 1972, s. 1025–1045.

⁷ Zob. M. Fleury, L. Henry, *Des registres paroissiaux à l'histoire de la population. Manuel de dépouillement et d'exploitation de l'état civil ancien*, Paris 1956; szerzej o niej – C. Kukło, *Rodzina w osiemnastowiecznej Warszawie*, Białystok 1991, s. 39 n.

Pierwsze próby zastosowania informatyki w badaniach historycznych, w tym także demografii przeszłości, przyniosły pierwsze korzyści, ale i rozczarowania⁸. Do pierwszych należałoby zaliczyć: automatyzację obliczeń, szybkość i niezawodność wykonania, wreszcie możliwość operowania ogromną liczbą danych źródłowych. Do drugich: konieczność formalizowania kartotek (ujednoczenia zapisu, co powodowało odejście od źródła oryginalnego, niekiedy wręcz okaleczenie informacji) i ich przetwarzania w zależności od systemu wprowadzania danych, języka programowania, w końcu samej struktury maszyny cyfrowej. Należy podkreślić, że pierwsze języki programowania, jak FORTRAN czy ALGOL, powstały dla potrzeb matematyków i fizyków⁹. Nieodzowna zatem stawała się pomoc zawodowego programisty, który często narzucał historykowi swój punkt widzenia, m.in. wprowadzanie danych w ustalony sposób, czy kodowanie numeryczne wszystkiego co było możliwe. Na tym etapie dominowało bardziej stosowanie metod numerycznych (tj. operacji arytmetycznych) niż nienumerycznych (łączenie, dobieranie, wybieranie, klasyfikowanie, sortowanie).

Trzeba ponadto zwrócić uwagę, że rozmiary pamięci pierwszych generacji komputerów były niewystarczające na przetwarzanie ogromnej masy danych za każdym razem, a koszty tych operacji były wysokie, zważywszy na ich usytuowanie w wielkich centrach obliczeniowych. Dochodziły do tego różnorodne błędy – najczęściej na etapie wprowadzania danych, których usunięcie pochłaniało sporo czasu. Przypomnijmy, że w latach sześćdziesiątych rejestrowano kodowane dane za pomocą dyktowania do magnetofonu, które było z korzyścią dla szybkości wprowadzania danych, ale uciążliwe dla osób zajmujących się transkrypcją¹⁰. Jedyny postęp, o jakim ówczesnie można mówić, był związany z porzuceniem przez użytkownika taśmy magnetofonowej (łatwo ulegającej zniszczeniu) i posługiwanie się kartą perforowaną (mniej więcej od połowy lat sześćdziesiątych). W praktyce oznaczało to postęp na jednym z istotniejszych etapów współpracy historyka-demografa z komputerem, a mianowicie wprowadzania danych przy pomocy czytelnej klawiatury, która od tej pory pojawiła się na wyposażeniu wszystkich komputerów. W tym systemie maszyna cyfrowa sama dokonywała kodo-

⁸ Zakres tematyczny i chronologiczny prac zagranicznych powstałych z zastosowaniem ETO do początku lat siedemdziesiątych omawia M. Wolański, *Zastosowanie komputerów w badaniach historycznych*, „Acta Universitatis Wratislaviensis” 279, Historia, 27, Wrocław 1976, s. 315–335.

⁹ Por. J. Hawryluk, *Maszyna cyfrowa narzędzie współczesnego człowieka*, wyd. 2, Warszawa 1976, s. 145 nn. M. in. dlatego M. Couturier pracując przez 15 lat nad strukturami społecznymi miasta Châteaudun w XVI–XVIII w. zdecydował się na opracowanie innego języka – FORCOD A i FORCOD B bardziej uproszczonego i tym samym przystępniejszego w badaniach historycznych. Podobnie A. Chamoux, *La reconstitution des familles: espoirs et réalités*, „Annales ESC” 4–5, 1972, s. 1083–90, na podstawie własnych doświadczeń rekonstrukcji rodzin w Reims (XVII–XVIII w.) podziela opinię o dużej użyteczności wspomnianego języka (i jego nowej odmiany FORCOD C). Odmienne stanowisko prezentuje M. Wolański, op. cit., s. 327.

¹⁰ Por. relację A. Wyczańskiego z pracy wykonanej w 1974 r. w EHESS w Paryżu: *Współpraca historyka z komputerem*, „Kwart. Hist.” 1, 1976, s. 65–73.

wania danych kwalifikowanych. Nadal jednak pozostawały wysokie ceny przetwarzania i koszty programowania. Większość programów, które powstały w latach siedemdziesiątych były pisane najczęściej dla konkretnego, nieraz pojedynczego badacza i nie były zbyt powszechnie wykorzystywane przez innych zainteresowanych.

Dopiero rozwój mikroinformatyki jaki dokonał się w latach osiemdziesiątych zmienił zasadniczo problem danych i pozwolił na znaczne zredukowanie kosztów finansowych. Od tej chwili historyk-demograf mając do dyspozycji mikrokomputer (zaopatrzony w ekran, klawiaturę i drukarkę) mógł sam osobiście wprowadzać dane i tworzyć programy. Dysponując z kolei przenośnym notebookiem może dokonywać rejestracji danych nawet bezpośrednio w miejscu przechowywania źródeł.

Wydaje się że najwcześniej, bo już w końcu lat pięćdziesiątych próbę łączenia danych nominatywnych zaczerpniętych z wielu źródeł, m.in. akt stanu cywilnego, rejestrów imigracyjnych, kartotek sanitarnych, podjął H.B. Newcombe z kanadyjskiego Centrum Atomowego w Chalk River¹¹. We Francji dr J. Sutter z Narodowego Instytutu Badań Demograficznych (INED) wprowadził w 1955 r. mechaniczny zapis w celu rekonstrukcji stopnia pokrewieństwa Eskimosów z Thule¹², a następnie podjął się przetwarzania danych dotyczących ludności wsi bretońskiej Galien z lat 1827–1962. Z drugiej strony na początku lat sześćdziesiątych J.– C. Perrot, późniejszy autorytet francuskiej gospodarki preindustrialnej, opracował program automatycznej rekonstrukcji rodzin parafii Saint-Gilles w Caen¹³. W jego ślady podążyli wkrótce R.S. Schofield, realizując ambitne projekty dotyczące ludności wsi angielskich¹⁴, i L.L. Cavalli-Sforza, obejmujące mieszkańców doliny Parmy we Włoszech¹⁵.

Rekonstrukcja rodzin przez komputer jest zasadniczo szczególnym przypadkiem łączenia danych. Jak już wspomniano chodzi o przybliżenie danych dotyczących tej samej osoby wydobytych z aktów urodzeń (chrztów), małżeństw (ślubów) i zgonów (pogrzebów) rozproszonych w długim okresie czasu. W praktyce nie stanowiłaby ona większych problemów, gdyby osoby we wszystkich przypadkach były określone w ten sam sposób.

¹¹ H.B. Newcombe, et al., *Automatic Linkage of Vital Records*, „Science” 130, 1959, s. 954–959.

¹² M. Mugnier, J. Sutter, J.– M. Goux, *Organigrammes pour l'étude mécanographique de la parenté et de la fécondité dans une population*, „Population” 1, 1966, s. 75–98.

¹³ J.– C. Perrot, *Observation sur l'emploi de l'ordinateur pour l'étude démographique d'une paroisse de Caen au XVIIIe siècle*, „Annales DH” 1967, s. 29–41; Y. Daubeze, J.–C. Perrot, *Un programme d'étude démographique sur ordinateur*, „Annales ESC” 4–5, 1972, s. 1047–1070.

¹⁴ Badania prowadzono przy współpracy z Centrum Informatyki Uniwersytetu Newcastle-upon-Tyne rejestrując dane na taśmie perforowanej i posługując się Newcastle File Handling System. Serię programów opracowano w jęz. ALGOL – szerzej na ten temat zob. R.S. Schofield, *La reconstitution de la famille par ordinateur*, „Annales ESC” 4–5, 1972, s. 1071–1082.

¹⁵ M.H. Skolnick i in., *A Reconstruction of Historical Persons from the Parish Registers of Parma Valley, Italy*, „Genus” 3–4, 1973, s. 103–155.

Podobnego zdania jest spora grupa badaczy zainteresowanych (np. R.S. Schofield¹⁶, czy J. Dupâquier¹⁷, skłonna twierdzić, iż nie jest ona zadaniem trudnym, czy nazbyt skomplikowanym, to niemal zaraz dalej dodają – pod warunkiem, że dokumenty identyfikujące jednostkę zawierają wystarczającą ilość szczegółów.

Wystarczyłyby wówczas nieskomplikowany program do przeprowadzania połączeń. Niestety – w okresie wczesnonowożytnym nazewnictwo osobowe w całej Europie rozwijało się bardzo intensywnie¹⁸, choć niektóre z nazwisk (przezwoisk) mogły nosić już postać niezmienną¹⁹. Dla historyka, którego celem jest odtworzenie losów demograficznych rodzin w przeszłości, występowanie w aktach metrykalnych homonimów, istotnych zmian w pisowni patronimów, używania przez osoby dwóch, a niekiedy i więcej mian – stanowi największą trudność²⁰. Dlatego podstawowym problemem było uproszczenie pisowni nazwisk rodowych. W początkowych programach starano się zastępować nazwiska grupami liter i cyfr zgodnie z całością reguł fonetycznych. I tak np. w systemie kanadyjskim (Russel Soundex Code) zachowywano pierwszą literę nazwiska, zastępując charakterystyczne spółgłoski cyframi (np. b, p, v=1; d, t=2, itp.), pomijano samogłoski i spółgłoski (h, w). Eliminowano ponadto podwójne litery i litery poczynając od czwartej; wreszcie dorzucano „O” aby zawsze była grupa 3 cyfr po literze pierwszej. Jednakże w opinii R.S. Schofielda system ten nie jest w stanie ogarnąć ogromu różnorodności pisowni, np. nazwisk angielskich XVI–XVII w.²¹. Dlatego też zaproponował on, aby sam badacz wybierał typową pisownię nazwiska²² i rejestrował je w pamięci maszyny cyfrowej w postaci „szkieletu” – tylko litera początkowa wraz ze spółgłoskami (z pominięciem spółgłosek niemych)²³.

¹⁶ R.S. Schofield, op. cit., s. 1081.

¹⁷ J. Dupâquier, *Pour la démographie historique*, Paris 1983, s. 131.

¹⁸ Zob. zbiorowy tom – *Noms et prénoms: aperçu historique sur la dénomination des personnes en divers pays*, pod. red. L. Henry'ego, Liège 1974, w którym S. Songer omawia nazewnictwo w krajach skandynawskich, A. von Nell w Niemczech, P. Deprez we Flandrii, E. Hélin w Belgii frankońskiej, J. Nadal w Katalonii, T.H. Hollingsworth w Anglii, J. Kovacsics i S. Mikery na Węgrzech, M. Livi Bacci i L. Del Panta we Włoszech, H. Palli w Estonii, H. Charbonneau w Kanadzie.

¹⁹ Z ostatnich prac poświęconych nazwom osobowym mieszkańców Rzeczypospolitej szlacheckiej należy odnotować rozprawę M. Górnego, *Przezwoiska i nazwiska chłopów pałaczkich w XVII wieku. Pochodzenie i budowa*, Wrocław 1990 (z obszerną literaturą wcześniejszą). Z kolei bogaty zestaw prac światowych czytelnik znajdzie w książce L. Perouas i in., *Léonard, Marie, Jean et les autres. Les prénoms en Limousin depuis un millénaire*, Paris 1984, s. 193–211.

²⁰ Zwracają na to uwagę m.in. J.– C. Perrot, op. cit., s. 32 n, R.S. Schofield, op. cit., s. 1071 n; J. Dupâquier, op. cit., s. 132; H. Charbonneau, J. Légare, *Utilisation des ordinateurs en démographie historique*, w: *Démographie historique*, pod. red. M.L. Marcilio i H. Charbonneau, Paris 1979, s. 115 n; I. Gieysztorowa, *Wstęp do demografii staropolskiej*, Warszawa 1976, s. 211.

²¹ R.S. Schofield, op. cit., s. 1074.

²² Dotyczyło to również najczęściej występującej pisowni zawodów i miejscowości co prowadziło do powstania nowych (ujednoliconych) kartotek. Przy czym kartoteki oryginalne (z różną grafiką nazwisk, zawodów i miejscowości) zachowywano w całości, co umożliwiałało w razie potrzeby powrót do decyzji wcześniejszych.

²³ R.S. Schofield, *The Standardization of Names and the Automatic Linking of Historical Records*, „Annales DH” 1972, s. 379 nn.

Z kolei w systemie zaproponowanym przez L. Henry'ego zachowywano generalnie pierwszą literę spółgłoski nazwiska i dwie znaczące spółgłoski następujące po sobie²⁴.

Oczywiście zarówno we wspomnianych systemach, jak i nawet w późniejszych bardziej przygotowanych, znajdujemy pewną część arbitralności, co objawiało się wykonywaniem przez komputer połączeń błędnych, a zwłaszcza często chybionych. W sumie obok ogromnych nakładów finansowych i utworzenia całych ekip interdyscyplinarnych trzeba było dokonywać setek interwencji ręcznych – w konsekwencji takie systemy trudno było nazwać automatyczną rekonstrukcją rodzin.

Dlatego też od połowy lat siedemdziesiątych w niektórych ośrodkach badawczych, jak np. w Laboratoire de Démographie Historique w EHESS w Paryżu, nie rezygnując z celu zasadniczego podjęto prace nad daleko szerszym wykorzystaniem ETO do przetworzenia setek tysięcy informacji zawartych na kartach rodzin już zrekonstruowanych ręcznie. Chodziło m.in. o przetworzenie 6000 rodzin tzw. zamkniętych²⁵ z Rouen (1650–1819), odtworzonych przez J.– P. Bardeta, 4500 z Vexin (1660–1850) – przez J. Dupâquiera i 2500 z Reims (1660–1800) – przez A. Fauve-Chamoux. W rezultacie blisko trzyletnich prac kiluosobowego zespołu historyków i informatyka, zaangażowania paru ośrodków badawczych²⁶, a przede wszystkim znacznych subwencji CNRS (100 tys. franków) powstał „Logiciel CASOAR” – zintegrowany system programów obliczeń i analiz demograficznych dla rodzin zrekonstruowanych²⁷. W części użytkowo-badawczej składa się nań 5 programów generalnych, 10 dotyczących małżeństw, 30 płodności w rodzinie oraz 9 umieralności. Jednakże tak zbudowany system „CASOAR” zawiera pewne niedogodności wynikające z przyjętych przez jego twórców założeń. Ograniczenie go do operowania wyłącznie rodzinami zrekonstruowanymi uniemożliwia w tym systemie oddzielne przetwarzanie pierwotnych źródeł informacji (tj. aktów ślubu, chrztów i pogrzebów) i uzyskanie w rezultacie końcowym wielu istotnych analiz czy zestawień opracowywanych do tej pory przez samych badaczy. Dodajmy, że cały system mógł być eksploatowany w większych ośrodkach badawczych dysponujących ówczesznie dużymi komputerami IBM 360/370²⁸.

²⁴ Por. L. Henry, *Variations de noms de famille et changements de prénom. Problèmes qui en résultent pour le couplage automatique des données. Introduction*, „Annales DH” 1972; tenże, *Projet de transcription phonétique des noms de famille*, „Annales DH” 1976, s. 200–214.

²⁵ Są to rodziny, dla których badacz dysponuje obok daty ślubu także datą zgonu pierwszego ze współmałżonków.

²⁶ Laboratoire d'Informatique pour les Sciences de l'Homme, EHESS, Ecole Normale Supérieure, Centre de Recherche d'Histoire Quantitative de l'Université de Caen.

²⁷ M. Hainsworth, J.– P. Bardet, *Logiciel CASOAR. Calculs et Analyses Sur Ordinateur Appliqués aux Reconstitutions*, Paris 1981. 1 er Cahier des Annales de démographie historique.

²⁸ Zbliżone oprogramowanie do „Logiciel CASOAR” opracowano także w 1983 r. na Uniwersytecie Teksaskim w Austin. Jego autorzy użyli uniwersalnego Statistical Analysis System (SAS), który pozwolił na dokonanie na Uniwersytecie w Liège analiz demograficznych ponad 3600 rodzin zrekonstruowanych z XIX w. we wschodniej Belgii, zob. R. Leboutte, G. Alter, M. Gutmann, *Analysis of Reconstituted Families. A Package of SAS Programs*, „Historical Methods” 1, 1987, s. 29–34.

Dlatego też na początku lat osiemdziesiątych rozpoczęto prace nad nowym systemem umożliwiającym łączne i rozdzielne operowanie danymi z rejestrów parafialnych – tzw. półautomatyczną rekonstrukcją rodzin i, co ważniejsze, przy użyciu coraz powszechniejszych i tańszych mikrokomputerów. W maju 1982 r. inż. J.– P. Péliissier z Narodowego Instytutu Badań Agronomicznych zaprezentował system rejestrowania danych na mikrokomputerze APPLE IIe napisany w języku BASIC, który pozwalał historykowi osobiście wprowadzać dane z klawiatury w bardzo szybkim tempie, przy małych stosunkowo kosztach i bez pośrednictwa transkrypcji²⁹. Kilkanaście programów pozwala kontrolować spójność danych, formalizować i analizować fiszki według rodzajów aktów, a przede wszystkim tworzyć tzw. fiszkę odwróconą. Ponieważ pierwsze mikrokomputery nie dysponowały zbyt wielką mocą obliczeniową, autor systemu dalszy ciąg operacji łączenia danych (budowa genealogii wstępnych i zstępnych, rekonstrukcji rodzin) widział w wykonaniu komputerów o dużej mocy przetwarzania, po uprzednim przeniesieniu do nich danych. W tym 3 etapie operacji wszystkie informacje, które można było zgrupować bez dwuznaczności, są przekształcane w fiszki biograficzne i uzupełniane automatycznie z pomocą informacji zawartych w innych fiszkach. Całość tych operacji dokonuje komputer bez bezpośredniej interwencji badacza.

W rezultacie do swojej dyspozycji otrzymuje on dwa rodzaje fiszek: z jednej strony fiszki biograficzne, które zostały utworzone samodzielnie przez elektroniczną maszynę cyfrową, z drugiej fiszki „odwrócone” tzn. takie, które nie mogły być połączone w sposób pewny. A zatem historyk będzie dysponował częściową rekonstrukcją rodzin i listą wypadków spornych, wątpliwych. I to on sam, dysponując maksimum informacji, będzie podejmował decyzje o dalszych ewentualnych połączeniach.

Trzeba wyraźnie powiedzieć, że system opracowany przez J.– P. Péliissiera jest zintegrowanym pakietem programów mającym zasadniczo pomóc w rekonstrukcji rodzin, nie jest natomiast systemem całkowitej automatycznej rekonstrukcji rodzin.

Pierwszy projekt, a zarazem jeden z najciekawszych, który kwalifikujemy jako makro-rekonstrukcję rodzin, powstał w wyniku współpracy dwóch ośrodków naukowych w Kanadzie i we Francji. W Departamencie Demografii Uniwersytetu w Montrealu pod kierunkiem H. Charbonneau i J. Légare podjęto systematyczne wykorzystanie rejestrów parafialnych mieszkańców Québec od czasów ich powstania, tj. od początku XVII w. aż do 1851 r. Do chwili obecnej wspomniana ekipa z informatyzowała w pierwszym etapie (dane do 1766 r.) 300 tys. kanadyjskich aktów stanu cywilnego, które były przedmiotem różnych już prezentowanych analiz, i 400 tys. aktów

²⁹ J.– P. Péliissier, *Démographie, généalogie, micro-informatique*, Paris 1983. 2 er Cahier des Annales de démographie historique, wyd. 2, Paris 1985.

dotatkowo, które były konieczne ażeby rozszerzyć obserwację do początku XIX w.³⁰

Z kolei drugi uczestnik ekipy badawczej – Grupa Badań nad Chorobami Rendu-Osler z Lyonu, rekonstruuje ludność 5 wsi doliny Valserine w departamencie Ain (płd.-wsch. Francja). Do pamięci maszyny cyfrowej wprowadzono ponad 60 tys. aktów stanu cywilnego od XVII w. do naszych czasów i 35 tys. wzmianek indywidualnych zaczerpniętych ze spisów XIX–XX w., które obecnie są w trakcie łączenia. Dodajmy, że grupę badawczą z Lyonu interesował przede wszystkim charakter przenoszenia rzadkich chorób dziedzicznych, co w praktyce zaowocowało nie tylko potrzebą budowy klasycznej rekonstrukcji rodzin, ale także łańcuchów genealogicznych osobników występujących w populacji³¹.

Objętość badań grupy z Montrealu i optyka genealogiczna ekipy z Lyonu doprowadziła do rozszerzenia przyjętej wcześniej struktury rejestrowania informatycznego ludności w kierunku jak najprostszego użytkowania i dużej szybkości realizacji zadań³².

W rezultacie blisko dwudziestoletnich prac (!) powstała najnowsza wersja systemu informatycznego, noszącego nazwę – SYGAP, który jest systemem modułowym analiz i zarządzania rejestrem ludności na mikrokomputerze³³. Posiada on otwartą architekturę, w której zarządzanie danymi jest zapewnione przez moduł centralny, ale który jednocześnie służy do powoływania modułów satelitarnych realizujących analizy demograficzne i genetyczne. SYGAP redukuje przedstawianie ludności do dwóch fiszek podstawowych: zapisu pojedynczej osoby i związku. Tak skonstruowana baza danych składa się z następujących 4 stałych kartotek: dwóch zawierających wykazy osobników i par małżeńskich (obie z indeksami); słownika zmiennych bazy danych oraz kartoteki parametrów funkcjonowania systemu.

Zasadą główną jest, że wszystkie informacje znajdujące się w bazie danych są łączone dowolnie, ażeby dać odpowiedzi na pytania badacza. System jest

³⁰ Zob. J. Légaré, A. La Rose, R. Roy, *Reconstitution de la population canadienne au XVIIe siècle: méthodes et bilan d'une recherche*, „Recherches sociologiques” 3, 1973; P. Beauchamp, H. Charbonneau, B. Desjardins, J. Légaré, *La reconstitution automatique des familles: un fait acquis*, „Population” numéro spécial, Hommage à Louis Henry, 1977, s. 375–396; H. Charbonneau, J. Légaré, op. cit. Wprawdzie projekt tylko w ciągu 5 lat pochłonął 5 mln dolarów kanadyjskich, to zaowocował monumentalnym 47-tomowym, jak dotąd, wydawnictwem – *Répertoire des actes de baptême, mariage, sépulture et des recensements du Québec ancien*, pod red. H. Charbonneau i J. Légaré, Montreal 1980–1990, obejmującym okres XVII w. – 1765 r., oraz cenny studium zbiorowym – H. Charbonneau i in., *Naissance d'une population. Les Français établis au Canada au XVIIe siècle*, Paris–Montreal 1987.

³¹ A. Bideau, H. Plauchu, G. Brunet, *Etude épidémiologique de la maladie de Rendu-Osler en France: répartition géographique et prévalence*, „Population” 1, 1989, s. 9–28 (tamże wcześniejsza literatura).

³² Por. A. Bideau, B. Desjardins, G. Brunet, J. Légaré, *Démographie historique et génétique de population: collaboration franco-québécoise autour de la constitution de registres de population*, w: „Population et Cultures”. *Etudes réunies en l'honneur de François Lebrun*, Rennes 1989, s. 9–20.

³³ S. Poulard, E. Heyer, A. Guillemette, G. Brunet, „SYGAP”, *Un logiciel de gestion et d'analyse de population*, „Annales DH” 1990, s. 455–467.

skonstruowany w sposób interakcyjny, tzn. dialogowy. W SYGAP „menu” zarządzające rejestrem proponuje 5 możliwości, które z kolei same uruchamiają „pod-menu”: dwa pierwsze umożliwiają wyszukiwanie fiszki osobnika lub pary małżeńskiej, jej modyfikowanie, usunięcie lub tworzenie nowej, itd.; trzecie – tworzenie genealogii wstępnych i zstępnych oraz biografii osobnika; w końcu dwa ostatnie pozwalają na wyszukiwanie dwóch wykazów dotyczących wszystkich osobników lub par małżeńskich występujących w rejestrze ludności w porządku numerycznym bądź alfabetycznym i według płci.

W końcu dwa podstawowe moduły: analizy demograficznej proponuje wykonanie szeregu obliczeń w zakresie płodności, związków małżeńskich, umieralności, sporządzania hipotetycznych spisów ludności o wyznaczonej dacie i ruchu naturalnego ludności; natomiast analizy genetycznej jest bazą dla obliczeń genealogii wstępnych zarówno co do pojedynczego człowieka, jak i całej jego populacji (parametry przywiązania do niej, ocena otwartości populacji, miary intensywności zmian między różnymi grupami osobników, itd.).

Dodajmy, że pomimo wdrożenia już do pracy systemu SYGAP w 1990 r.³⁴, jego twórcy zamierzają rozbudowywać go o kolejne moduły, które pozwalałyby np. na dokonywanie analiz poprzecznych ludności czy automatycznego rozpoznawania struktury gospodarstw domowych według typologii opracowanej przez P. Lasletta³⁵.

Drugi wielki system, o nazwie SOREP, na który chcemy zwrócić uwagę czytelnika, powstał również w Kanadzie. Jego budowę G. Bouchard, R. Roy, B. Casgrain rozpoczęli w 1974 r. i w ciągu ostatnich dziesięciu lat stworzyli zupełnie nowy system automatycznej rekonstrukcji rodzin, który już w chwili obecnej przetworzył 660 tys. aktów rejestracji stanu cywilnego społeczeństwa Saguenay z lat 1842–1971³⁶.

W opinii twórców ich system działa zarówno w sytuacjach bardzo prostych, jak i bardzo trudnych. Za rzecz najważniejszą uznali oni zabezpieczenie się przed tym, co nazwalibyśmy tutaj związkiem sprzecznym łączenia automatycznego (problem więzów podwójnych lub konkurencyjnych). Opracowana procedura łączenia danych nominatywnych działa czterokrotnie progresywnie, akumulując informacje z kroku na krok w fiszkach rodzin. System działa w możliwie najbardziej bezpieczny sposób, poczynając od najpewniej-

³⁴ Napisany w jęz. CLIPPER pracuje on w systemie operacyjnym PC DOS wersji 3.0 (ewentualnie późniejszej) na mikrokomputerze zgodnym z IBM PC XT/AT lub PS przy użyciu dBase III lub dBase IV.

³⁵ Zob. P. Laslett, *La famille et le ménage: approches historiques*, „Annales ESC” 4–5, 1972, s. 847–872.

³⁶ Podstawowe informacje o systemie zaczerpnięto z prac: G. Bouchard, P. Brard, *Le programme de reconstitution automatique des familles saguenayennes: données de base et résultats provisoires*, „Histoire Sociale” 12, 1979, s. 170–185; G. Bouchard, Ch. Pouyez, *Name Variations and Computerized Record Linkage*, „Historical Methods” 2, 1980, s. 119–125; a zwłaszcza z G. Bouchard, R. Roy, B. Casgrain, *De la micro à la macro-reconstitution des familles. Le système „SOREP”*, „Genus” 3–4, 1986, s. 33–54. Nie miałem niestety dostępu do podstawowej pracy ze szczegółowym opisem, por. G. Bouchard, B. Casgrain, R. Roy, *Reconstitution automatique des familles: le système SOREP*, t. 1–2, Chicoutimi 1985, ss. 763.

szych połączeń nominatywnych i zostawiając najbardziej skomplikowane sytuacje na koniec³⁷.

Tak skonstruowany system SOREP składa się z kilku podsystemów. Pierwszy grupuje i wykonuje wszystkie operacje wstępne przeznaczone do przygotowania fiszki wzmianek o parze w celu połączenia. Polegają one na usunięciu akt uznanych jako niedopuszczalne (są to wzmianki, które nie zawierają 4 podstawowych elementów nominatywnych tzn. nazwisk i imion współmałżonków) i eliminacji zmian nominatywnych najbardziej powierzchownych³⁸.

Warto w tym miejscu zauważyć, że ta standaryzacja danych nominatywnych przed łączeniem nie jest sądem arbitralnym badacza, ale jest zrealizowana przez komputer na podstawie reguł jednakowo sprecyzowanych i zastosowanych. Pozwala to na uniknięcie tworzenia sztucznych nowych zmienionych nominatywnych.

Drugi podsystem przystępuje do konstrukcji biografii rodzinnych lub fiszek par małżeńskich poprzez 7 sortowań³⁹. Pierwsze pozwala na znalezienie par małżeńskich o 100% zgodności informacji. Pozostałe 6 pozwala komputerowi na łączenie par gdzie są zgodne tylko 2 lub 3 elementy z czterech. Dla analizowanych przez autorów danych ten krok dał 98% wyników. Ponadto w 12 dodatkowych sortowaniach brane są pod uwagę coraz bardziej liberalne kryteria. Tworzą one podzbiory, gdzie identyczny jest jeden element i tylko część drugiego, tzn. 3 pierwsze litery.

W końcu trzeci podsystem opierając się na biografiach rodzinnych kreuje biografie indywidualne, tzn. łączy ślub rodziców z ich chrztem i wykonuje połączenia dotyczące dzieci (więzi chrzest – pogrzeb i chrzest – ślub).

Należy jednak zaznaczyć, że niewątpliwy sukces badawczy komputerowej rekonstrukcji rodzin był też możliwy w jakimś stopniu dzięki spożytkowaniu przez historyków źródeł ze stosunkowo bliskiego nam okresu, w którym aż 99,6% małżonków zapisano w aktach z nazwiska i imienia (!)⁴⁰. Cały system SOREP, i to warto zaakcentować, wprawdzie oparty na podstawie technicznej – kartotece ludności – ma służyć badaniom interdyscyplinarnym, np. powstawaniu społeczeństwa regionalnego, mobilności społecznej, kształtowaniu się klasy robotniczej, systemu rodzinnego, reprodukcji demograficznej, zapobieganiu chorobom dziedzicznym przez rozpoznanie populacji z ryzykiem⁴¹.

³⁷ Por. G. Bouchard, *The Processing of Ambiguous Links in Computerized Family Reconstruction*, „Historical Methods” 1, 1986, s. 9–19.

³⁸ W tym celu autorzy posługują się programem FONEM – kodem fonetycznym, który w odróżnieniu od systemu Russel Soundex Code anuluje wyłącznie zmiany nominatywne, nie zmieniając struktury fonetycznej nazwisk i imion, zob. szerzej G. Bouchard, P. Brard, Y. Lovoie, *FONEM: un code de transcription phonétique pour la reconstitution automatique des familles saguenayennes*, „Population” 6, 1981, s. 1085–1103.

³⁹ G. Bouchard, R. Roy, B. Casgrin, op. cit., s. 44–46.

⁴⁰ Por. B. Bouchard, op. cit., s. 10 n.

⁴¹ Szczegółowe omówienia dorobku i planów naukowych grupy badaczy Uniwersytetu Québec w Chicoutimi są zawarte w raportach rocznych programu SOREP. Duży rozgłos zyskało

Z ciekawszych projektów zagranicznych warto odnotować też przepracowaną wersję systemu CASOAR, który z dawnego systemu obliczeń i analiz przekształcił się za sprawą J.P. Bardeta i informatyka J. Renarda w nowoczesny system z tzw. półautomatyczną rekonstrukcją rodzin. Wprawdzie do tej pory nie ukazały się publikacje omawiające szczegółowo takich zagadnień, jak chociażby filozofii łączenia danych nominatywnych, to kilkumiesięczny pobyt w paryskim Centre de Recherches sur la Civilisation de L'Europe Moderne (Paris IV) i rozmowy z jego autorami dowiodły, że nowy CASOAR, działający na mikrokomputerach, przekroczył etap testowania i został wdrożony do normalnej pracy. W chwili obecnej baza danych mikrokomputerowego CAOSARA zawiera blisko 77 tys. zrekonstruowanych rodzin zamkniętych, głównie z okolic Normandii⁴². Jest ona dziełem studentów i młodszych pracowników naukowych, skupionych wokół prof. J.– P. Bardeta⁴³.

Wreszcie trzeba wspomnieć, że w ostatnim dziesięcioleciu zorganizowane zostały inne równie ważne i interesujące projekty. Wśród nich warto odnotować m.in. zakrojone na najszerszą skalę badania prowadzone na Uniwersytecie Utah w Salt Lake City (USA), mające na celu stworzenie bazy danych genealogicznych kościoła Mormonów (gigantyczna kolekcja 1 234 000 mikrofilmów, każdy o długości 30 m, i 8 mln fiszek rodzin!)⁴⁴, prace S. Ackermana oraz G. Sandströma i J. Sundina w Szwecji⁴⁵, na Uniwersytecie Massachussets w Amhrst mający za cel ludność 12 wsi doliny Connecticut⁴⁶, czy dotyczący epidemiologii genetycznej, kierowane przez K.M. Weissa na przykładzie miasta Laredo w Teksasie⁴⁷.

Wreszcie próby rekonstrukcji rodzin i budowy ciągów genealogicznych, już na mniejszą skalę, podjęto także w Hiszpanii, gdzie D.– S. Reher na podstawie

także niedawno wydane przez nią opracowanie zbiorowe – G. Bouchard, M. de Braekeller, et al., *Histoire d'un Génome: population et génétique dans l'est Québec*, Québec 1991 (praca znana mi jedynie z recenzji J.– P. Poussou, „Annales DH” 1992, s. 350–353).

⁴² I tak Crceem zrekonstruował ok. 26 tys. rodzin w Vernon-les-Andelys; J. Renard ok. 23 tys. w Pont l'Eveque (1690–1836) i 9 tys. w Cherbourgu (te ostatnie na podstawie spisów); J.–P. Bardet ok. 6,5 tys. w Rouen (XVII–XVIII w.); A. Mauger ok. 6,5 tys. w Cherbourgu (1720–1794); C. Renard ok. 4 tys. w Barentin-Pavilly (1740–1880); S. Beauvalet ok. 2 tys. w Verdun z XVIII w.

⁴³ Pragnę w tym miejscu podziękować prof. J.– P. Bardetowi oraz J. Renardowi za umożliwienie mnie i mgr. W. Gruszeckiemu praktycznego zapoznania się tak w Paryżu, jak i w Białymstoku z funkcjonowaniem mikrokomputerowej wersji CAOSARA.

⁴⁴ Szerzej o tym bezprecedensowym przedsięwzięciu finansowanym przez 3 duże organizacje amerykańskie – zob. M.H. Skolnik, L.L. Bean, G.P. Mineau, L.B. Jorde, *Etudes génétiques et démographiques et base de données généalogiques de l'Utah*, „Annales DH” 1984, s. 103–113 i przytoczoną w nim literaturę.

⁴⁵ Por. S. Akerman, *An Evaluation of the Family Reconstitution Technique*, „The Scandinavia Economic History Review” 2, 1977, s. 160–170; G. Sandström, J. Sundin, *Computer Analysis of Life Histories from Swedish Church Records: a Case Study from the Demographic Data Base at Umea University*, w: *Historical Social Research. The Use of Histor and Process-Produced Data*, Stuttgart 1980, s. 192–202.

⁴⁶ Zob. A.C. Swedlung, R.S. Meinde, M. Gradie, *Family Reconstitution in the Connecticut Valley: progress on Record Linkage and the Mortality Survey*, w: *Genealogical Demography*, eds. B. Dyke, W.T. Morrill, New York 1980, s. 139–155.

⁴⁷ A.V. Buchanan, K. Weiss et al., *Reconstruction of Genealogies from Vital Records: The Laredo Epidemiology Project*, „Computers and Biomedical Research” 17, 1984, s. 326–351; R.J. Schwartz, K.M. Weiss, A.V. Buchanan, *Looking into the Black Box: Practical Approaches to Record Linkage*, „Annales DH” 1984, s. 119–128.

37 tys. aktów metrykalnych zrekonstruował rodziny w mieście Cuenca (poł. XVI – poł. XIX w.)⁴⁸, a we Włoszech E. Lucchetti i L. Soliani dla miasteczka Bellino na przestrzeni dwóch wieków (1770–1970) i na podstawie ponad 10 tys. metryk⁴⁹.

Ograniczone ramy artykułu sprawiają, że nie o wszystkich obszarach zastosowania elektronicznej maszyny cyfrowej jesteśmy w stanie dokładnie poinformować czytelnika. Niemniej jednak choć w największym skrócie należy odnotować, iż w Instytucie Max-Plancka w Getyndze w latach 1986–1987 pod kierunkiem M. Thallera zbudowano KLEIO – system specjalnie stworzony do pracy w naukach humanistycznych, a w szczególności w dyscyplinach historycznych⁵⁰. Pozwala on na stworzenie bazy danych historycznych, literackich, archiwalnych, a od niedawna malarskich i kartograficznych. Dzięki jego zastosowaniu zrekonstruowano: rodziny gminy alpejskiej St. Lamprecht (1600–1850), gminy westfalskiej Spenge (1600–1850), utworzono bazy danych demograficznych o strukturach rodzinnych dużej liczby gmin europejskich z XVI–XX w. (250 tys. osób w przestrzeni geograficznej między Bazyleą, Wiedniem, Agram i Rzymem) oraz 3 wsi reńskich (1700–1840) i 5 wsi langwedockich (1715–1840)⁵¹.

W Polsce prace nad zastosowaniem technik komputerowych w badaniach demografii przeszłości najwcześniej – w połowie lat sześćdziesiątych rozpoczęto w Zakładzie Antropologii PAN we Wrocławiu pod kierunkiem prof. E. Piaseckiego. Głównym celem prac zespołu wrocławskiego była analiza prawie 33 tys. kart osobników parafii bejskiej (woj. kieleckie) z XVIII–XX w. pod kątem odtworzenia modelu przemienności generacji⁵². Dlatego też główny autor projektu odrzucił metodę „rekonstrukcji rodzin”, każda zaś osoba miała oddzielną kartę, na którą nanoszono wszystkie dotyczące jej informacje z trzech rodzajów ksiąg metrykalnych⁵³. Do pamięci maszyny cyfrowej nie wprowadzono ani imienia i nazwiska, ani miejsca urodzenia czy zamieszkania, ani też zawodu czy przyczyny zgonu. Funkcję identyfikacyjną zamiast nazwiska i imienia spełniał numer karty danej osoby. Trudności w procesie identyfikacji osobników zmusiły jej autorów do całkowitego zrezygnowania z prac nad materiałem metrykalnym z okresu 1607–1745. Mniejszych trudności w tym zakresie dostarczyły księgi parafialne od połowy XVIII w., w których

⁴⁸ D.– S. Reher, *Introduction à la étude de l'information nominative à partir de la mise sur ordinateur des archives paroissiales espagnoles*, „Annales DH”, 1984, s. 137–146.

⁴⁹ E. Lucchetti, L. Soliani, *L'utilisation de l'ordinateur dans la reconstitution des familles et des généalogies à Bellino*, „Annales DH” 1984, s. 129–135.

⁵⁰ Por. J. Smets, *Créer une base de données historiques avec KLEIO*, Göttingen 1990, s. 7–16.

⁵¹ *Ibid.*, s. 9.

⁵² Zob. E. Piasecki, R. Wrona, *Przeniesienie do pamięci maszyny cyfrowej danych za lata 1746–1968 parafii bejskiej w województwie kieleckim*, „Przeszłość Demograficzna Polski” (dalej: „Przesz. Dem. Pol.”), 10, 1977, s. 17–37. Podsumowanie ponad dwudziestoletnich badań z zastosowaniem ETO przynosi książka E. Piaseckiego, *Ludność parafii bejskiej (woj. kieleckie) w świetle ksiąg metrykalnych z XVIII–XX w. Studium demograficzne*, Warszawa–Wrocław 1990 (z powołaniem wszystkich wcześniejszych prac autora), która wywołała krytyczną dyskusję publikowaną na łamach „Przesz. Dem. Pol.” 19, 1993 (w druku).

⁵³ E. Piasecki, *op. cit.*, s. 22–38.

wprawdzie nie było już osób określonych wyłącznie imieniem, ale byli tacy, którzy używali jednocześnie dwóch lub trzech nazwisk o różnym brzmieniu. Cały materiał został zapisany początkowo na maszynie eliott 803, prymitywnym komputerze jeszcze lampowym; dwukrotnie był przenoszony na inne maszyny, końcowe zaś analizy zrealizowano na odrze 1304.

Doceniając nowatorstwo metodyczne na gruncie polskim uczonych wrocławskich trzeba stwierdzić krytycznie, że zbudowany w ten sposób system, oparty na kartach pojedynczych osób, ma mniejszą przydatność w badaniach demograficzno-historycznych, posługujących się szerszym (społeczno-demograficznym) kwestionariuszem badawczym⁵⁴, poza oczywiście genetyką populacyjną. Jego podstawowym mankamentem jest niemożliwość podjęcia badań nad rozmiarami i strukturą rodzin oraz płodnością i umieralnością w stratyfikacji społecznej i zawodowej. System wrocławski autorstwa E. Piaseckiego nie uwzględnia także badania problemów trwałości barier społecznych między poszczególnymi stanami, warstwami lub grupami zawodowymi, zasięgu kontaktów społeczno-towarzyskich i zawodowo-rodziny.

W połowie 1979 r. pod kierunkiem prof. A. Wyczańskiego rozpoczął pracę zespół stawiający sobie za zadanie informatyzację badań społeczno-demograficznych, na podstawie rejestrów parafialnych, opracowywanych metodą rekonstrukcji rodzin⁵⁵. Organizację oparto na porozumieniu Instytutów Historii PAN i Filii UW w Białymstoku, co gwarantowało do połowy lat osiemdziesiątych obsługę badań i środki finansowe. Generalnie chciano badać dawną rodzinę polską, znaną dotąd tylko fragmentarycznie, w odniesieniu do szczytów społecznych, a której obraz w świadomości społecznej rysował się błędnie. Jednocześnie chciano uruchomić ogromne pokłady źródeł zalegających w archiwach kościelnych, tysiące ksiąg metrykalnych chrztów, ślubów i pogrzebów, obejmujących wszystkie warstwy dawnego społeczeństwa. Liczono przy tym na zebranie szerszych doświadczeń w zakresie zastosowania informatyki i komputera do badań historycznych, szczególnie w odniesieniu do źródeł masowych⁵⁶.

W tym wypadku przetwarzanie danych miało być stosowane w zakresach:

- 1) analizy materiałów dotyczących chrztów, ślubów i pogrzebów;
- 2) analizy materiału przygotowanego przez historyków w postaci tzw. rodzin zrekonstruowanych ręcznie;
- 3) ewentualnego wykorzystania komputera do automatycznej rekonstrukcji rodzin, zgodnie z opracowanym systemem informatycznym.

Szczegółowymi badaniami objęto cztery społeczności parafialne. I tak dla parafii Łomża (XVIII– 1 poł. XIX w.) i św. Jana w Warszawie (XVII w.)

⁵⁴ Por. prace poświęcone problematyce rodzinnej I. Gieysztorowej, M. Boguckiej, C. Kukli, A. Wyczańskiego i A. Wyrobisza.

⁵⁵ A. Wyczański, *Informacja o pracach zespołu nad systemem informatycznym dla badań społeczno-demograficznych*, „Przesz. Dem. Pol.” 15, 1984, s. 151–153.

⁵⁶ Zob. też C. Kuklo, A. Wyczański, *Badania nad dawną rodziną w Polsce XVI–XVII wieku za pomocą elektronicznego przetwarzania danych prowadzone w Instytucie Historii PAN w Warszawie oraz w Instytucie Historii FUW w Białymstoku*, „Studia Podlaskie” 1, 1990, s. 85–90.

przetwarzano tylko rodziny już uprzednio zrekonstruowane odpowiednio przez I. Gieysztorową i M. Sierocką-Pośpiech. Natomiast dla parafii św. Krzyża w Warszawie (XVIII w.) i Ostrowi Mazowieckiej (XVII–XVIII w.) ich autorzy, C. Kuklo i A. Siłuch-Błaszczec, przetworzyli w pierwszej kolejności metryki chrztów, ślubów i pogrzebów, uzyskując obok różnorodnych analiz bardzo pomocne w dalszych pracach skorowidze alfabetyczne nowożeńców, urodzonych dzieci, osób zmarłych, świadków ślubów i chrztów, w następnej – rodziny zrekonstruowane ręcznie.

Przechowywanie akt metrykalnych w archiwach kościelnych czy państwowych uniemożliwiało ich wypożyczenie do ośrodków informatycznych celem zapisywania danych bezpośrednio z ksiąg parafialnych wprost na kartę lub taśmę perforowaną. Ponieważ zespół warszawsko-białostocki kładł większy nacisk na zagadnienia społeczno-demograficzne, nieprzydatne okazały się znane powszechnie francuskie fiszki-formularze autorstwa M. Fleury'ego i L. Henry'ego⁵⁷ czy polskie zaproponowane przed 20 laty przez S. Hoszowskiego i Z. Sułowskiego⁵⁸, jak również E. Piaseckiego i R. Wrony⁵⁹.

Opracowano zatem wspólnie z informatykami własne znormalizowane formaty (formularze) wejściowe (aktu ślubu, chrztu i zgonu), do których badacze przenosili dane z poszczególnych serii ksiąg metrykalnych i w końcowym etapie formularz – rodziny zrekonstruowanej ręcznie. Tak wypełnione formularze wejściowe początkowo perforatorki wprowadzały za pomocą taśmy magnetycznej do pamięci komputera. W dwa lata później mogli to już robić sami historycy (nawet z pominięciem formularzy, jeżeli dysponowali oczywiście źródłem w ośrodku obliczeniowym), korzystając z ekranowego urządzenia typu Mera 9150.

Podjęto ponadto szereg czynności upraszczających zapisywanie poszczególnych aktów metrykalnych celem zaoszczędzenia czasu pracy maszyny cyfrowej oraz związanych z tym kosztów i opracowano dodatkowo trzy kartoteki (zawodów i stanowisk, tytułów społecznych wraz ze skrótami, imion w wersji polskiej i łacińskiej).

W sumie opracowano i przetestowano pozytywnie kilkanaście programów merytorycznych analiz dotyczących ślubów, chrztów i pogrzebów, a także w części dotyczących rodzin zrekonstruowanych⁶⁰. Zasadniczo bowiem rodziny zrekonstruowane ręcznie zostały przetworzone z pomocą programów francuskich systemu CASOAR po uprzedniej ich konwersji na użytek polski. Zostały one nam udostępnione w ramach współpracy nawiązanej z EHESS w Paryżu.

⁵⁷ M. Fleury, L. Henry, *Nouveau manuel de dépouillement et d'exploitation de l'état civil ancien*, wyd. 2, Paris 1976, s. 44–53.

⁵⁸ W dodatku mało czytelne w bieżącej pracy historyka ze źródłem masowym, gdyż wszystkie pola do wypełnienia oznaczono wyłącznie cyframi, por. S. Hoszowski, Z. Sułowski, *Ewidencja ruchu naturalnego ludności oparta na dawnych metrykach parafialnych. Projekt ujednolicenia systemu gromadzenia danych*, „Przesz. Dem. Pol.” 4, 1971, s. 3–20.

⁵⁹ E. Piasecki, R. Wrona, op. cit., s. 19.

⁶⁰Pod koniec 1984 r. zespół zorganizował kolokwium, podczas którego przedstawił dotychczasowe rezultaty prac, doświadczenia i napotkane trudności – zob. A. Fauve-Chamoux,

Programy polskie napisano w języku PL/I, a dane przetwarzano na maszynach cyfrowych IS RIAD i IBM dostępnych w dużych ośrodkach obliczeniowych.

Poważną część pracy zespołu pochłonęło przedyskutowanie i wstępne opracowanie koncepcji budowy Informatycznego Systemu Rekonstrukcji Rodzin (ISRR), pracującego na podstawie tzw. relacyjnej bazy danych. Miał on umożliwić w przyszłości:

- automatyczną rekonstrukcję rodzin;
- tworzenie ciągów genealogicznych;
- dostęp do informacji biograficznych dotyczących poszczególnych osób objętych systemem;
- przetwarzanie wymienionych danych oparte na już posiadanych i przyszłych programach⁶¹.

Niestety malejące z roku na rok nakłady finansowe, a potem praktyczne zaprzestanie finansowania badań w połowie lat osiemdziesiątych, spowodowało bardzo znaczne osłabienie tempa prac zespołu (niemożność opłacenia wysokokwalifikowanych informatyków, przetworzenia danych). Pomimo wstępnego przetestowania ISRR nie wdrożono go do stałej pracy, albowiem wymagało to nie tyle nawet czasu na usunięcie ewentualnych błędów czy wprowadzenie poprawek, co znacznych kosztów finansowych, np. za dzierżawę i przechowywanie dysków magnetycznych, nie wspominając już o lawinowo rosnących po 1985 r. kosztach przetwarzania danych w wyspecjalizowanych centrach informatycznych.

Końcowym rezultatem prac wspomnianego zespołu było opublikowanie studium C. Kukli, poświęcone demografii rodzin warszawskich w czasach saskich i stanisławowskich⁶². Dwa dalsze – I. Gieysztorowej odtwarzające demograficzne dzieje rodzin w ziemi łomżyńskiej w XVIII–XIX w. i M. Sierockiej-Pośpiech w stołecznej parafii św. Jana w XVII w. – są na ukończeniu.

W odmiennym, ale jakże interesującym badawczo kierunku zmierzają prace zespołu kierowanego przez prof. A. Mączaka. Mianowicie magistranci i doktoranci w ramach programu badawczego „Społeczeństwo Korony w drugiej połowie XVII wieku” podjęli zadanie budowy baz danych dotyczących struktury społecznej i zaludnienia w świetle rejestrów pogłównego generalnego z lat 1662–1676. Początkowo tworzone bazy danych dla poszczególnych ziem i województw przy pomocy specjalnego programu i formularza opracowanego przez Jędrzeja Mączaka na bazie programu ISIS. Jednakże ostatnio rozpo-

Informatique et démographie historique en Pologne: Colloque du 3 décembre 1984, Varsovie, „Annales DH” 1985, s. 397–405; I. Gieysztorowa, Sprawozdanie z konferencji poświęconej zastosowaniu technik komputerowych w badaniach historyczno-demograficznych XVII i XVIII w., „Przesz. Dem. Pol.” 17, 1987, s. 265–275.

⁶¹ Por. P. Cegięła, J. Prokuratorowski, S. Senkowski, *ISRR – Informatyczny System Rekonstrukcji Rodzin na podstawie rejestrów parafialnych. Założenia ogólne*, „Przesz. Dem. Pol.” 14, 1983, s. 123–136.

⁶² Zob. C. Kuklo, *Rodzina...*

często prace nad nowym programem, dzięki któremu wszystkie bazy danych mają być dostępne za pośrednictwem sieci EARN⁶³.

Dynamiczny rozwój techniki mikrokomputerowej w ostatnich latach przyczynił się z jednej strony do nieznanego wcześniej wzrostu produkcji komputerów osobistych, z drugiej do rozwoju ich oprogramowania. Obecnie nawet w uboższych placówkach humanistycznych możemy spotkać nierzadko kilka lub kilkanaście mikrokomputerów i coraz powszechniej znajdujemy je w gabinetach historyków zawodowych. Istnieje na świecie kilkanaście tysięcy różnego typu pakietów oprogramowania i cały czas powstają nowe programy i systemy, poczynając od gier i programów edukacyjnych poprzez programy narzędziowe, po poważne aplikacje stosowane w różnorodnych sferach życia i działalności człowieka.

Warte odnotowania są programy narzędziowe, pozwalające na budowę aplikacji użytkowych na przykład baz danych. Istnieje szereg tego typu standardowych programów dostępnych na rynku polskim np. dBase, Rbase, Nutshell, Foxbase, oraz tzw. programy zintegrowane (wielofunkcyjne), jak: Lotus, Exel, Symphony⁶⁴ itp. Ze względu na swoje źródło pochodzenia prawie wszystkie są w angielskiej wersji językowej, choć ostatnio pojawiają się ich tłumaczenia polskie.

Mając przed sobą problem użycia mikrokomputera do celów badawczych można próbować go rozwiązać różnymi metodami. Wydaje się celowe wyróżnienie dwóch podejść:

- dostarczenie badaczowi narzędzi w postaci komputera i standardowego oprogramowania;
- przygotowanie specjalistycznego systemu realizującego określone funkcje badawcze.

Jak można się domyślać każde z tych rozwiązań ma swoje wady i zalety. W pierwszym przypadku niezbędne jest nauczenie użytkownika, jak ma samodzielnie budować swoją aplikację wykorzystując oprogramowanie wzięte z „półki”, z ewentualną koniecznością nauczenia programowania komputerów dla zastosowań, których nie można rozwiązać narzędziami nieprogramowanymi. Zaletą natomiast może być koszt takiego rozwiązania (może, ale nie musi ze względu na konieczność zakupu odpowiedniej do ilości użytkowników liczby programów narzędziowych) oraz czas realizacji.

Drugie podejście powinno dać w efekcie produkt rozwiązujący postawiony problem i dający użytkownikowi, co chcemy podkreślić, łatwe w wykorzystaniu narzędzie przy następujących założeniach:

- gruntownej analizy merytorycznej problemu przed przystąpieniem do jego oprogramowania;

⁶³ Zob. M. Kopczyński, *Archiwum na dysku*, „Przegl. Hist.” 1, 1992, s. 115–118 oraz na podstawie informacji pani mgr A. Laszuk.

⁶⁴ Ten właśnie program pozwolił K. Makowskiemu, *Rodzina poznańska w I połowie XIX wieku*, Poznań 1992, na utworzenie bazy danych dotyczących rodziny zaczerpniętych z ksiąg metrykalnych oraz dokonanie obliczeń i analiz, których zakres demograficzny jest jednak bardzo skromny – por. moją recenzję w „Przesz. Dem. Pol.” 19, 1993 (w druku).

- wykonania oprogramowania tak, aby zrealizować postawione zadanie;
- wykonanie odpowiedniej dokumentacji czytelnej dla nieprzygotowanego informatycznie badacza.

Wychodząc z powyższych założeń i dzieląc powszechną opinię, że badania metrykalne przynoszą wyniki nie tylko demograficzne i społeczne, ale szersze – interdyscyplinarne, dwaj członkowie zespołu prof. A. Wyczańskiego – historyk C. Kuklo i informatyk W. Gruszecki podjęli zadanie opracowania Informatycznego Systemu Rekonstrukcji Rodzin, Gospodarstw Domowych i Społeczności Lokalnych w Polsce przedrozbiorowej, finansowanego przez Komitet Badań Naukowych.

Celem powyższego projektu badawczego jest opracowanie i oddanie do dyspozycji historyków, studentów i innych osób zainteresowanych, wielostronnie rozbudowanego systemu informatycznego, który pozwoliłby nawet osobom bez specjalistycznego przeszkolenia na przetwarzanie i analizy demograficzno-społeczne masowych źródeł z XVI–XVIII w. Cały system będzie składał się z 4 większych podsystemów, z których każdy będzie zawierał po kilkanaście szczegółowych programów analitycznych, a także programy pomocniczo-techniczne do obsługi zbiorów danych.

Poszczególne podsystemy będą realizowały:

- 1) przetwarzanie list imiennych i spisów ludności pod kątem struktury ludności według płci, wieku, stanu cywilnego oraz struktury gospodarstw domowych;
- 2) przetwarzanie i analizy aktów metrykalnych ślubów, chrztów i pogrzebów;
- 3) półautomatyczną rekonstrukcję rodzin i analizy demograficzno-społeczne rodzin zrekonstruowanych;
- 4) przetwarzanie i częściowe analizy wiejskich i miejskich ksiąg sądowych pod kątem funkcjonowania społeczności lokalnych.

Wydaje się, że na gruncie polskim zaproponowany system ma charakter nowatorski, a jedną z jego zalet jest fakt, że będzie on eksploatowany na ogólnie dostępnych mikrokomputerach kompatybilnych z IBM PC⁶⁵.

Kilkumiesięczne bardzo intensywne prace zaowocowały (do końca 1992 r.) przygotowaniem oprogramowania do wprowadzania danych z list imiennych i prestatystycznych spisów ludności oraz z aktów metrykalnych trzech serii. Podstawą budowy oprogramowania pierwszego z podsystemów stały się informacje zawarte w spisach parafialnych sporządzonych przez plebanów w latach 1790–1792 na użytek Komisji Porządkowych Cywilno-Wojskowych⁶⁶. Przedmiotem rejestracji przez historyka w pamięci mikrokomputera obok adresu (nazwa miasta-wsi, ulicy, nr domu) jest osoba fizyczna: jej imię i nazwisko, stan cywilny, płeć, wiek, zawód, status społeczny, stopień pokrewieństwa

⁶⁵ Jego oprogramowanie jest przygotowywane przy użyciu języka CLIPPER.

⁶⁶ Chcąc ustalić ich zakres tematyczny i wewnętrzny układ oraz określić stopień szczegółowości, przeprowadzono liczne kwerendy w archiwach państwowych i kościelnych. Dokładniej rozpoznano spisy parafialne m.in.: z ziemi wieluńskiej, powiatu ostrzeszowskiego i radziejowskiego, parafii św. Krzyża w Warszawie (AGAD), parafii miasta Krakowa i innych z województwa krakowskiego (WAP w Krakowie).

w stosunku do głowy rodziny/gospodarstwa. Natomiast przy budowie oprogramowania drugiego z podsystemów odwołano się szeroko zarówno do własnych kilkuletnich doświadczeń w pracy z tysiącami osiemnastowiecznych metryk stołecznej parafii świętokrzyskiej⁶⁷, jak i innych badaczy (np. I. Gieysztorowej⁶⁸, W. Dworzaczka⁶⁹).

W chwili obecnej, po roku pracy, dwa pierwsze podsystemy przeszły pomyślnie próby testowe i po wprowadzeniu niezbędnych drobnych poprawek i uzupełnień mogą zostać udostępnione innym potencjalnym użytkownikom. Już teraz historyk–demograf może nie tylko rejestrować dane (według opracowanych prostych instrukcji), ale także dokonywać ich analiz według przygotowanych grup tematycznych⁷⁰.

Przed autorami pozostaje do realizacji najtrudniejszy – w świetle doświadczeń zagranicznych ośrodków badawczych – z podsystemów, a mianowicie – automatycznej czy półautomatycznej rekonstrukcji rodzin. Dodatkowym utrudnieniem jest świadomy wybór obiektu badawczego. Chodzi w tym wypadku o przeprowadzenie łączenia danych nominatywnych wziętych z ksiąg metrykalnych ślubów, chrztu i pogrzebów warszawskiej parafii św. Krzyża z lat 1670–1709. Migracyjny charakter parafii, okresowe pobyty w stolicy szlacheckich posiadaczy pałaców, dworów i dworków oraz przepływ ich obsługi nie sprzyjają kwerendzie genealogicznej. Nie ulega wątpliwości, iż budowa podsystemu rekonstrukcji rodzin przez komputer jest wielce ryzykowną próbą na tak wczesnym materiale. Być może właśnie przez swoje liczne trudności rokuje ona większe doświadczenia badawcze tak na polu zastosowania informatycznej techniki obliczeniowej, jak i samej warsawianistyki⁷¹.

Warto jednak pamiętać, że tego typu prace prowadzone są zaledwie w kilku, może w kilkunastu największych światowych ośrodkach naukowych przez wiele lat oraz z zaangażowaniem wieloosobowych i interdyscyplinarnych zespołów badawczych. I nawet jeżeli rodzime próby wykorzystania ETO w badaniach historycznych nie zawsze swoim zasięgiem czy rozmachem są porównywalne ze światowymi, to sądzimy, iż należy je podejmować. Są bowiem one jednym ze sposobów nawiązania kontaktu z czołówką światową, który w zakresie badań społecznych był w ostatnich latach luźny i nieregularny.

⁶⁷ Por. C. Kukło, *Rodzina...*, s. 31–49.

⁶⁸ I. Gieysztorowa, op. cit., s. 210–217.

⁶⁹ W. Dworzaczek, *O badaniach genealogicznych nad dawną rodziną*, w: „Społeczeństwo staropolskie”, pod. red. A. Wyczańskiego, t. 2, 1979, s. 177–191.

⁷⁰ Ich szczegółowy wykaz zamieszczam w artykule – *Prace nad budową Informatycznego Systemu Rekonstrukcji Rodzin, Gospodarstw Domowych i Społeczności Lokalnych w Polsce przed-rozbiorowej*, „Studia Podlaskie” 5 (w druku).

⁷¹ Ręczna rekonstrukcja rodzin (przepisanie akt po akcie ok. 15 tys. zapisów metrykalnych na pojedynczych fiszkach, klasyfikacja alfabetyczna celem ich połączenia na kartach rodzin, w końcu wykonanie obliczeń statystycznych) zajęłaby kilka lat pracy historyka.