

Informatique et démographie

Chronique d'un dialogue obligé

Frans J. WILLEKENS et Michel ORIS

Compte rendu de «*Expert Meeting on Demographic Software and Computing: Strategies for the Future*», The Hague, June 29–July 3, 1992.

Le NIDI (Netherlands Interdisciplinary Demographic Institute) et le groupe de travail sur les logiciels démographiques et la micro-informatique, créé au sein de l'Union internationale pour l'étude scientifique de la Population (U.I.E.S.P.), se sont associés pour organiser à La Haye, du 29 juin au 3 juillet 1992, une réunion d'experts sur « Les logiciels démographiques et les stratégies informatiques pour le futur ». Leurs réflexions présentent un intérêt général dans la mesure où la démographie est à la fois la plus et la moins spécifique des sciences humaines. Dans la découverte et la description des structures et des mouvements des populations, elle est certainement la discipline la plus quantitative de toutes. Quand il s'agit de passer de la description à la compréhension, elle intègre des dimensions économiques, sociologiques, de psychologie sociale, d'ethnologie, etc., pour former un ensemble composite qualifié « d'études de population ». En outre, si l'analyse des sociétés contemporaines est privilégiée, tant l'histoire que la prospective font l'objet de nombreux travaux.

Au départ, les objectifs du colloque ont été exprimés clairement : « *Technological change is having a significant impact on data collection, data processing, teaching and research. In order to effectively utilise technological innovations in computer hardware and software, a strategy is required that outlines where we want population studies and demography to be at some time in the future and how technology may be used to help us getting there. The aim of the workshop is to outline a strategy and, by doing so, to lay the basis for a computer-based system for the acquisition, management and transfer of demographic information and knowledge* »¹.

¹ *Bulletin de Liaison de l'UIESP*, n° 41, janvier–avril 1991, p. 11.

✉ F.J. WILLEKENS, Professeur; Netherlands Interdisciplinary Demographic Institute; Box 11650; NL-2506 The Hague (Netherlands).

Fax : + 31 70 3647 187

✉ M. ORIS, Chargé de recherches FNRS; Université de Liège; Histoire Économique et Sociale Contemporaine; 32, Place du 20-Août; B-4000 Liège (Belgique).

Fax : + 32 41 2325 45

Un tel énoncé ne peut que susciter l'approbation, puisqu'il ne s'agit plus de gérer au jour le jour l'intégration des progrès techniques, mais bien de réfléchir sur une démarche scientifique, indépendamment des apports de la micro-informatique, dont l'intérêt de la contribution est mesuré à cette aune. Il faut toutefois constater que cette excellente méthode a été peu suivie. Les interventions se sont fondées beaucoup plus classiquement sur des réflexions rétrospectives — en quoi l'usage de l'ordinateur a-t-il transformé la démographie? —, sur les problèmes du présent et leurs perspectives de résolution.

Outre une séance introductive, sept sessions ont été distinguées :

- 1) la collecte et la mise en forme des données;
- 2) les bases de données et les systèmes d'information;
- 3) l'intégration de nouveaux produits dans un environnement informatique pré-existant;
- 4) l'analyse des données;
- 5) le micro-ordinateur dans l'enseignement de la démographie ou des relations entre population et développement;
- 6) population et développement;
- 7) les systèmes experts et les diverses perspectives dans le domaine technique.

L'ensemble des participants a souligné l'importance des apports du micro-ordinateur. Les améliorations purement techniques ont joué un rôle indéniable : extension des mémoires, processeurs plus rapides, disques durs plus vastes, cassettes ou CD-ROM, etc. Mais le succès a été dû plus aux logiciels, tableurs, traitements de texte, gestionnaires de bases de données, qu'au matériel lui-même. C'est la possibilité d'un usage sans compétence particulière qui a réellement popularisé l'informatique, qui a révolutionné au quotidien le travail scientifique (GERLAND). Dans la pratique, il en a résulté un éclatement des centres de collecte et de traitement, une plus grande autonomie des acteurs de la recherche, la constitution d'échantillons plus vastes pour mieux asseoir nos connaissances, et des investigations plus approfondies.

L'impact sur le traitement des données a été le plus immédiat dans une science du chiffre. Des calculs rebutants et fastidieux sont devenus monnaie courante, des méthodes sophistiquées ont été mises à la portée du premier venu, comme l'analyse multivariée des relations causales (NOUMBISSI & SALHI). Dans le domaine de la saisie, les améliorations ont été plus progressives. Les procédures de tests et d'évaluation, le codage automatique, la gestion ergonomique des masques, la saisie directe, même lors d'enquêtes sur le terrain, la lecture optique des formulaires de recensement, ont rapidement additionné leurs effets en l'espace d'une quinzaine d'années.

C'est au niveau de l'organisation et de la gestion des données que subsistent les plus fortes carences. Les logiciels commerciaux sont peu satisfaisants lorsqu'ils s'appliquent à des enquêtes complexes, à questions croisées ou ouvertes, ou quand il faut organiser un encodage massif sans vouloir ou pouvoir payer autant de licences que de postes de travail. C'est pourquoi différents produits spécifiques ont été conçus : *PC-EDIT* par les Nations-Unies (PAGEAU), *ISSA (Integrated Systems for Survey Analysis)* [CROFT], ou encore *BLAISE*, qu'utilise le Bureau central de Statistique des Pays-Bas (HUNDEPOOL).

Ils résolvent avec plus ou moins de bonheur le problème de l'hétérogénéité des fichiers de population. En effet, comme l'a fort justement souligné VU DU Y MAN, les bases de données démographiques contiennent soit des enregistrements nominatifs, d'individus ou de ménages, soit des tableaux de chiffres. Leur gestion optimale exige

un matériel et surtout des logiciels distincts. L'O.N.U. joint pour ce faire *XTABLE* à *PC-EDIT*. La plupart des concepteurs de produits récents prévoient des possibilités d'exportation vers les tableurs ou programmes graphiques commerciaux, qu'ils n'estiment pas pouvoir concurrencer sérieusement.

De nombreux intervenants ont d'ailleurs dénoncé l'amateurisme de la plupart des logiciels spécifiques à la démographie. Le premier reproche est l'absence de standards, donc de compatibilité. Le second est le manque de souplesse, tant lorsque l'utilisateur veut modifier en profondeur les paramètres des modèles utilisés que lorsqu'il veut intégrer un nouveau modèle. Le troisième porte sur leur caractère souvent peu explicite quant à la méthodologie de calcul utilisée, aux algorithmes mis en œuvre, aux intervalles de confiance dans lesquels se situent les résultats (LEETE). À partir de ces critiques, les participants ont dessiné le portrait du logiciel idéal. Il serait souple, convivial, doté d'un didacticiel clair exposant la méthode, privilégiant l'interaction et la communication.

Implicitement ou explicitement, ces interventions condamnent une informatique « presse-bouton », un usage irréfléchi de l'ordinateur qui est aussi celui de techniques statistiques délicates, complexes, qui ne s'appliquent utilement qu'à des situations délicates et complexes. Le risque de dérive vers un technicisme gratuit est grand dans une science aussi quantitative que la démographie quand le potentiel de calcul explose. Aussi des réflexions et contributions critiques comme celles de Gérard CALOT, qui propose un procédé d'analyse de tendances, ou de S. GRIBBLE, qui décrit l'intégration des modèles du type Monte Carlo dans les logiciels de micro-simulation longitudinale à Statistique Canada, sont-elles toujours précieuses. C'est d'autant plus vrai que selon HAKKERT, la méthodologie n'a pas toujours été rénovée autant que l'apport technique des micros le permettait. Des logiciels proposent des procédés d'estimation « simples », conçus au temps de la statistique « pré-informatique », alors qu'existe maintenant la possibilité de les raffiner, de construire des modèles plus complexes, etc.

Lutter contre les excès est d'abord une question de pédagogie. ARRIAGA rappelle de ce point de vue l'apport essentiel, à un apprentissage efficace de la démographie, des feuilles pré-programmées dans un tableur quelconque comme *Lotus 1-2-3*. Elles installent l'utilisateur dans une relation interactive qui lui permet de visualiser immédiatement le résultat de ses choix, l'influence d'un changement de paramètre, l'impact d'une méthode d'approche. Ce type de formation évite de placer directement le novice devant des logiciels performants avant que la méthodologie ait été correctement assimilée.

Le souci d'une informatique transparente et d'un usage responsable se retrouve dans de nombreuses interventions, ce qui témoigne de l'ampleur du problème. BURCH vante les mérites de *MATHCAD*, programme mathématique qui permet de composer et d'enchaîner la plupart des algorithmes nécessaires à l'analyse démographique. Son utilisation implique nécessairement une bonne domination des outils statistiques et de l'ensemble de la démarche. Aussi BURCH en arrive-t-il à poser une question provocante : est-il encore nécessaire de produire des logiciels à l'usage des démographes ? On peut répondre affirmativement, mais uniquement pour des programmes plus clairs, plus explicites, dotés de véritables manuels consultables en ligne, offrant à l'utilisateur un contrôle complet des opérations.

Se pose alors le problème de la création et plus encore de la maintenance de produits aussi performants. Seules des firmes privées ou des institutions internationales semblent

capables d'un tel effort (MORELAND; GERLAND). Il faut surtout compter sur l'O.N.U., le *U.S. Bureau of Census*, etc., car le marché est limité et pauvre. Alors que l'on connaît les problèmes démographiques dramatiques dans lesquels se débat l'Afrique sub-saharienne, il est significatif que, malgré les efforts consentis par les organisateurs, elle n'ait délégué que trois ressortissants parmi les 75 participants au colloque, et encore deux d'entre eux sont-ils assistants dans une université européenne... La démographie n'est pas que la science du vieillissement des riches, mais aussi celle de la survie des pauvres. On ne peut dissenter sur les apports de l'informatique en négligeant de s'interroger sur leur coût, sur leur transportabilité dans des contextes socio-économiques situés aux extrêmes de l'échelle planétaire.

De ce point de vue, M. VLASSOFF a dressé un constat dénué de complaisance de l'action de son organisation, le *United Nations Fund for Population Activities*. Il en résulte que les logiciels dont le succès a été réel avaient des buts plus spécifiques que généraux, qu'ils étaient une composante d'une assistance plus vaste, et qu'un dialogue de qualité entre producteurs et utilisateurs avait présidé à leur mise au point. Il constate également le peu de réussite des programmes qui intègrent les variables démographiques dans une planification de développement global (Voir cependant les contributions de SEGAL, MORELAND ou de LBETE sur *WORKERS*). On peut en dire autant des outils d'aide à la recherche dans un domaine en pleine expansion ces dernières années, celui des rapports entre population et environnement.

C'est une question de concepts, mais aussi de données. Des ensembles considérables et hétérogènes sont nécessaires pour passer de la démographie pure à la planification, l'assistance aux prises de décision, la définition de politiques de population (GERLAND). De véritables banques, sous forme de Système d'Information Géographique, sont encore rares. À un niveau général, ce n'est que depuis peu que l'O.N.U. propose *PopMap*, un logiciel de cartographie automatique doté de quelques options de tabulation ou de gestion et, surtout, de possibilités d'importation-exportation qui en permettent l'usage dans un environnement informatique préexistant. Se pose cependant le problème de la difficile constitution des fonds de carte. À un niveau local, de nombreux développements sont en cours, y compris dans des pays où la collecte est malaisée (KARL).

Dans les contrées riches, on peut espérer une croissance parallèle à la technologie du CD-ROM, et se féliciter d'une manière générale de la diffusion de plus en plus fréquente des données sur support électronique. Cela ne suffit cependant pas à permettre une réelle intégration des variables. Les formats et les structures divergent, ce qui impose de nombreuses opérations techniques et de réaménagement. Plus encore, la comparabilité à l'échelon supra-national est souvent illusoire, l'uniformité des notions utilisées inexistante (HEILIG).

Cette situation pénalise les chercheurs, mais c'est aussi un handicap certain pour les programmes d'instruction et d'information démographiques dont HEILIG, à nouveau, souligne tout l'intérêt. En exploitant au maximum les possibilités graphiques des ordinateurs personnels, ils ont pour cibles moins les démographes professionnels que tous ceux, politiciens, intellectuels, journalistes, qui ne s'intéressent pas aux calculs mais bien aux résultats. Récemment, lors de la Chaire Quetelet '92, les participants se sont accordés à souligner le décalage entre la qualité et la quantité des travaux consacrés au Tiers-Monde d'une part, les lieux communs véhiculés dans la presse d'autre part. Il est facile d'accuser

les néophytes, mais il plus sain de rappeler aux scientifiques leur devoir de communiquer avec la société.

La communication est d'ailleurs au centre de toutes les discussions sur les solutions aux problèmes du présent, sur les « stratégies du futur ». Sans même revenir sur la question importante de l'unification des notions statistiques, il est évident qu'une standardisation des fichiers d'entrée ou de sortie est souhaitable. D'autre part, la large autonomie de traitement accordée aux chercheurs grâce aux micro-ordinateurs n'a de sens que si les organismes producteurs des données de base les diffusent sur des supports appropriés dans des formats compatibles. Il y a de toute évidence encore beaucoup à faire puisque, parmi les différents thèmes qui étaient proposés aux participants au colloque, un des rares à ne pas avoir été traité était le suivant : « Vers un langage d'interrogation pour les données de population »². Les historiens démographes, qui étaient peu présents, semblent être allés plus loin en ce domaine que les démographes contemporains (ORIS).

La communication doit aussi être nettement améliorée entre producteurs et « consommateurs » de logiciels. Ces derniers doivent être conçus en fonction des attentes des utilisateurs, qui ne peuvent être déterminées dans un quartier général à New York ou à Genève. Afin de les identifier, la Division Technique et Évaluation de la U.N.F.P.A. réalise une enquête dans les pays en voie de développement (VLASSOFF)³. La communication doit aussi porter sur les programmes existants.

Les évaluations objectives sont rares. BOGUE fait œuvre de pionnier en analysant six logiciels d'étude des relations entre population et développement. Sa contribution s'inscrit dans le cadre des activités du Groupe de Travail sur les Logiciels démographiques et la Micro-Informatique, créé au sein de l'U.I.E.S.P. en 1990. D'autres thèmes devraient être couverts et il serait souhaitable que, désormais, tous les nouveaux produits soient soumis à une recension critique, qu'accueillerait une des principales revues de démographie (LEETE).

Dans l'attente, sous la direction de F. WILLEKENS, le Groupe prépare la publication d'un inventaire détaillé des logiciels disponibles. Il s'adresse d'abord aux producteurs, qui sont priés d'arrêter le gaspillage. Alors qu'il existe plus de trente programmes différents pour calculer des projections de population (NOUMBISSI & SALHI), il est inutile d'en ajouter un de plus à la liste. Les ajouts doivent désormais être spécifiques et combler des vides, comme par exemple en matière de projection des petits groupes.

Cela ne signifie pas que la recherche d'une certaine cohérence soit abandonnée. Toutefois, à partir du moment où il apparaît illusoire de vouloir concurrencer certains produits commerciaux, il devient clair qu'il faut privilégier la cohabitation des logiciels plutôt qu'une intégration en un vaste programme généraliste. L'expérience du Bureau de Statistique du Bangladesh est à cet égard tout à fait significative (BHUIYAN). L'environnement *Windows* suscite beaucoup d'enthousiasme, mais sa diffusion dans le Tiers-Monde n'est pas pour demain. Un appel est donc lancé pour rencontrer une attente

² *Bulletin de Liaison de l'UIESP*, n° 41, janvier-avril 1991.

³ Voir déjà *Asia-Pacific POPIN Directory : databases, services, products and institutions/Economic and Social Commission for Asia and the Pacific*, New York, United Nations, 1990, XII + 234 pp.; *Report of the Asia-Pacific Popin consultative workshop, 29 september-4 october 1990*, New Delhi - New York, 1991, 44 pp.

unanime. Au moins serait-il possible de réunir un grand nombre de procédures, comme dans *POPSYN* ou *MortPak-Lite*.

L'inventaire et les évaluations s'adressent bien évidemment aussi aux simples utilisateurs, qui ont souvent du mal à réaliser une sélection, tant le foisonnement est grand. RAO propose de compléter ce service en créant une banque des logiciels gratuits ou tombés dans le domaine public, en un centre unique auquel tous pourraient s'adresser.

En guise de conclusion, il n'est sans doute pas inutile de résumer les neuf recommandations à l'usage des créateurs de logiciels démographiques dans les années 1990, émises par le groupe d'expertise. Quatre ont trait à la gestion des ressources informatiques, tant en matériel qu'en logiciels. Le gaspillage actuel devrait être limité par une approche plus coordonnée. L'effort de coordination doit naturellement venir des bailleurs de fonds que sont les Nations Unies, l'Union internationale pour l'étude scientifique de la Population, le *U.S. Bureau of the Census*, etc., dont les équipes ont tendance à travailler de manière pratiquement indépendante les unes des autres. Leur coopération peut s'exercer tant au travers de la constitution d'une banque de logiciels démographiques accessibles à bas prix, que par des dons d'équipement aux pays du Tiers-Monde, dons qui impliquent une assistance en matière de développement, de maintenance, de formation, etc.

Au niveau des programmes, le groupe d'expertise insiste sur le soin avec lequel devraient être conçues les interfaces utilisateurs, d'une part dans une optique de clarté, avec des didacticiels qui exposent la méthodologie mise en œuvre, d'autre part dans une optique de « portabilité ». Un produit puissant mais très spécifique quant à sa technologie s'intégrera difficilement dans un environnement de travail pré-existant. Il indique en outre que des développements sont particulièrement attendus dans deux directions : la création de systèmes d'informations démographiques et de systèmes de gestion et de prévision incluant tant les variables de population que de santé et de planning familial.

Liste des contributions

- G. ALTER & M.P. GUTMANN, *Casting spells: database concepts for event history analysis.*
- E.M. ANDREEV, *Using of demographic software in the Soviet Union.*
- E.A. ARRIAGA, *Demographic software as a didactic and analytical tool.*
- J. BERGMAN, *Current use of software for teaching demographic methods course at the University of Michigan.*
- M.H.H. BHUTYAN, *The programme for the development of a population information system at the Bangladesh Bureau of Statistics.*
- D.J. BOGUE, *An evaluation of microcomputer software for population and development.*
- T.K. BURCH, *Do we need special demographic software?*
- G. CALOT, *Monitoring first marriage and fertility change.*
- T. CROFT, *The ISSA system.*
- P. GERLAND, *Software development : past, present and future trends and tools.*
- S.F. GRIBBLE, *Design issues for longitudinal microsimulation software.*
- R. HAKKERT, *Computing in demographic analysis: beyond paper and pencil algorithms.*

- G.K. HEILIG, *Strategies for the development of PC-based demographic information and instruction systems.*
- A. HUNDEPOOL, *Computer assisted survey processing at the Netherlands Central Bureau of Statistics.*
- Z. JIANG, *The use of information technology in family planning policies and programs in China.*
- J. JOZWIAK & I. KOTOWSKA, *Brief information on the use of computers in the Institute of Statistics and Demography.*
- G. KARL, *Statistical data bases and GIS application at unchs (Habitat).*
- F. KOSCHIN, *Demography and computers at the Prague University of Economics.*
- M. LACKNER & VU DUY MAN, *PopMap: a GIS for population activities.*
- R. LEETE, *Recent developments in software for demographic analysis: inventory of micro-computer packages.*
- I.W. MOLENAAR, *Expert systems and prospects: the role of uncertainty.*
- R.S. MORELAND, *Computer software requirements for sectoral planning models: recent experience with host-based models.*
- A. NOUMBISSI & M. SALHI, *L'informatique en démographie : bilan et perspectives.*
- J.A. NELDER, *Intelligent software and expert systems for statistical analysis.*
- M. ORIS, *Liège's databases and history of populations between the industrial revolution and demographic transition.*
- M. PAGEAU, P. GERLAND, N. MCGIRR, *Data processing strategies: hardware and software for data entry editing and tabulating.*
- K.V. RAO, *Software for survival and transition data analysis: a comparison and evaluation of packages.*
- J. SEGAL, *Issues and priorities in developing/improving integrated databases and software for population policies and programmes.*
- K. SRINIVASAN, *Note on use of software packages in the teaching of demography.*
- F.N. STOKMAN & E.P.H. ZEGGELINK, *Object oriented modeling of social networks.*
- T. UNALAN & A. HANCIOGLU, *Computer use in teaching and research at the Institute of Population Studies.*
- VU DUY MAN, *Demographic database: design and implementation.*
- F. WILLEKENS, *The IUSSP software project.*
- M. VLASSOFF, *Software development for population policy and programmes: experiences and future demands.*