

Note de lecture

Réflexion sur "Informatique et Histoire"

Informatique et prosopographie : Actes de la table ronde du CNRS (25-26 octobre 1984), Paris, CNRS, 1985, 358 p.

La table ronde organisée à Paris, les 25 et 26 octobre 1984, sur "informatique et prosopographie" regroupait des chercheurs historiens engagés sur des projets d'études prosopographiques à l'aide de l'ordinateur.

La prosopographie est une discipline historique qui "consiste à établir des notices biographiques individuelles et à les confronter. A la différence de la biographie, elle ne s'applique jamais à un seul homme [...] mais à un échantillon [...] qu'on suppose représentatif du groupe étudié [...] ; il s'agit, à partir de personnalités, de préparer des types, de faire jaillir les traits communs et les différences" (1). L'usage de l'ordinateur se prête volontiers à cette discipline qui nécessite la constitution d'échantillons représentatifs et une exploitation statistique des données à partir d'une documentation sérielle et éparpillée.

Le mariage entre la démarche prosopographique et l'informatique a fait ses preuves notamment en histoire sociale de l'art et des artistes : en 1984, Gerhard FRIES a recensé 160 projets d'application de l'informatique en histoire de l'art et en archéologie (2).

Toutes périodes confondues, les 24 contributions de notre table ronde s'attachent à présenter les travaux historiques et les méthodes d'approche qui se sont confrontés aux exigences de l'informatique. Au-delà de l'exposé des résultats abondamment recueillis et de la présentation des matériels et des logiciels mis en oeuvre, les participants ont insisté sur les problèmes de méthode propres à chaque projet, à chaque recherche, à chaque instrument informatique. La lecture de ces *Actes* est particulièrement enrichissante du fait que chaque étude devient l'occasion de dessiner, à travers un matériau que l'informatique oblige à définir et à délimiter, les voies méthodologiques empruntées, les impasses techniques (du fait du matériel) ou scientifiques (du

1. ANDREAU J., article "Prosopographie" dans *Dictionnaire des Sciences historiques*, p. 546-548.

2. FRIES Gerhard, "Projets de banques de données biographiques en histoire de l'art", dans *Actes...*, p. 179.

fait de l'invalidité des questions ou de la caducité des réponses) rencontrées, les étapes tracées et parcourues, les résultats obtenus ou à venir. Ainsi apparaissent, à travers chaque présentation, les caractéristiques de chaque recherche, ses difficultés, son apport.

La constitution du corpus

Car - que l'on parte de l'étude de textes cunéiformes d'Ebla (III^e millénaire av. J.-C.), de manuscrits médiévaux occidentaux ou orientaux, de registres d'état civil du XVIII^e siècle, de dictionnaires d'artistes, de listes de notables, de catalogues d'élèves boursiers, d'archives nobiliaires (Angleterre, XIV^e siècle) - ce qui importe d'abord, c'est de constituer un corpus.

Parfois donné d'avance - dans le cas de documentation sérielle ou suivie, telle les registres pontificaux avignonnais -, le constituant de l'enquête prosopographique exige le plus souvent d'être sélectionné dès le départ, soit par le groupement de données très dispersées - tels les sceaux byzantins du IV^e au XIV^e siècle -, soit par le recoupement de deux ou plusieurs sources : J. PETIT et A. ZYSBERG ont confronté, par exemple, un dictionnaire biographique de militaires français sous la Révolution et l'Empire avec des dossiers personnels d'officiers pour établir un échantillon de militaires et en étudier les carrières.

Cette opération de constitution du matériau de recherche, non seulement implique une pratique des archives, une familiarité avec la période historique concernée, une connaissance des sources d'information utiles, mais surtout exige un travail considérable de conceptualisation de la recherche : seules, en effet, la définition des hypothèses et la limitation des questions permettent, d'une part d'affronter la masse des archives et, d'autre part, de résister à la grisaille de l'outil informatique et au vertige des "listings".

La démarche prosopographique consistant à établir un corpus, il est bien évident qu'on n'a pas attendu la révolution informatique pour établir des fichiers. L'ordinateur vient ici à point, comme instrument puissant de calcul et de classement offrant la possibilité de traiter des fichiers complexes, de manipuler de variables multiples, d'opérer toutes sortes de recoupements. Toutes les communications s'attachent à démontrer les vertus du va-et-vient qu'autorise l'usage de l'ordinateur, ainsi que les possibilités d'ajustement ou

de figelage des hypothèses grâce à la multiplication des recoupements et la rapidité des comparaisons.

Même lorsqu'une lecture traditionnelle peut permettre d'aboutir à d'excellents résultats, la voie informatique, même testée sur un programme simple, aura l'avantage de transformer celui-ci en un utile point de départ pour une grille plus complexe (3). Il reste que l'informatique a surtout pour mission de réaliser des programmes plus amples et plus sophistiqués : l'expérience de l'*Onomasticon Arabicum* (4) prévoit, non seulement de classer les grands personnages déjà connus à travers la littérature biographique arabe médiévale (100.000 notices du I/VI^e au XI/XVII^e siècle), mais aussi de prolonger l'étude onomastique jusqu'à la période contemporaine (Algérie) (5). L'ordinateur, facilitant le tri instantané de un ou plusieurs éléments choisis, permet de travailler sur des groupes très étendus en tirant parti de renseignements très minces (6).

La métasource

Bien conscient qu'on ne peut sortir de l'ordinateur que ce qu'on y a mis (7), les intervenants insistent sur l'importance de la formulation des questions pour une élaboration pertinente des données. Analyser des profils de carrière de militaires, raconter la vie et la mort d'un vaisseau de guerre, établir l'hérédité de fonction ou de rang social parmi certaines élites au début du XX^e siècle, reconstituer la population de Pise au cours du XIV^e siècle, identifier les artistes de l'Angleterre du XIX^e siècle sont autant de voies d'accès aux listes, dossiers, recensements, dictionnaires, catalogues divers utilisés par les chercheurs et pris comme sources premières. Par la formulation des questions, par l'établissement du projet de recherche, la source pre-

3. LASSUS F., "Une base de données biographiques : les provisions d'offices en Franche-Comté aux XVII^e et XVIII^e siècles", dans *Actes...*, p. 71-90.

4. Entreprise gigantesque de l'Institut de Recherche sur l'Histoire des Textes, qui projette de mettre sur fichier informatisé, à partir de la littérature biographique arabe, les noms des savants, des hommes politiques et religieux évoqués par les *tarāḡim* et les *ṭabaqāt*.

5. SIBLET J., "Autour de l'*Onomasticon Arabicum*", dans *Actes...*, p. 125-129.

6. C. HAYEZ A.-M., dans *Actes...*, p. 140.

7. C. CHARLE C., *Actes...*, p. 241. Voir aussi E. LEROY-LADURIE, "L'historien et l'ordinateur", dans *Le territoire de l'historien*, Paris, Gallimard, 1973.

mière va donner naissance à un fichier, qui constitue une source au second degré, une *métasource*, selon un terme maintenant consacré (8).

La distance qui sépare la source de la métasource varie selon les questions posées; enregistrer intégralement les données est difficile à réaliser, même dans les cas de bases de données documentaires, d'où l'importance de la sélection des informations et de leur codage.

Le bordereau de saisie : sélection et contraintes techniques

L'énonciation des hypothèses est donc la phase maîtresse de la constitution de la métasource, le premier moment important de la rencontre entre le questionnement prosopographique (historique) et l'ordinateur. Cette étape est concrétisée par la confection d'un bordereau de saisie, oeuvre conjointe d'historien(s) et d'informaticien(s), rencontre intime entre les questions entrevues et les moyens de parvenir au meilleur usage de l'outil informatique pour les poser : sans questions, l'ordinateur est paralysé devant la source, autrement dit sans accès "éclairé" à l'ordinateur, les obstacles techniques peuvent devenir des impasses.

Puisqu'il s'agissait de travaux prosopographiques, les bordereaux de saisie présentés au cours de notre table ronde étaient une transposition d'informations biographiques données ou reconstituées. Par exemple, pour un dictionnaire biographique de l'Islam classique, la section arabe de l'Institut de Recherche et d'Histoire des Textes a conçu un bordereau divisé en deux parties : la première, à but proprement onomastique, s'attache à retrouver tous les éléments composant le nom du personnage (*ism, laqab, kunyâ, šuhrâ...*); la seconde, plus biographique, retrace les étapes de l'acquisition de son savoir et de sa transmission (maîtres, élèves, voyages, écrits ...) (9).

Les détails donnés par les communications au sujet de la structure des bordereaux de saisie sont en fait un exposé sur le choix des informations sélectionnées pour être confiées à l'ordinateur : quelles questions l'historien posera-t-il à la source et comment les formulera-t-il, compte tenu des possibilités du logiciel et du programme choisi ?

8. GENET J.-P., article "Informatique", dans *Dictionnaire des Sciences Historiques*, p. 367.

9. SUBLET J., "Autour de l'Onomasticon Arabicum", dans *Actes...*, p. 126.

Lecture très instructive, donc, de ces communications sur les logiciels et les programmes mis en oeuvre, et sur leur plus ou moins grande fiabilité (10). Chaque communication expose problématique et moyens, donne les raisons des choix techniques et présente les résultats en fonction du matériel utilisé. On apprend beaucoup de choses, par exemple, sur le logiciel CLIO, mis au point par l'Institut Max-Planck d'Histoire de Göttingen, à travers une étude sur deux communautés rurales aux XVIIIème et XIXème siècles (11) et grâce à une réflexion générale du concepteur de ce système intégré de banques de données permettant le traitement de sources historiques (12). Certaines recherches visent l'élaboration de logiciels spécifiques, et il est intéressant de retrouver les soucis communs à toute recherche prosopographique informatisée et d'y discerner les questionnements propres à chaque problématique (13).

Le codage et ses problèmes

Pour confier des données à l'ordinateur, il faut les coder, c'est-à-dire transposer en langage informatique les variables considérées, les informations sélectionnées : inspirées des schémas classiques élaborés en démographie historique (14) ou reprenant les articulations des notices biographiques qui constituent les données de base (15), les structures des fiches de saisie sont en fait une schématisation d'individus dont la représentation abstraite permettra de dessiner la carrière, la vie de nombreux personnages, les données historiques de leur ascension ou au contraire de leur régression sociale.

10. OSIRIS, S.P.S.S., S.A.S., CLIO, MYTOP, P.F.S., A.D.D.A.D.

11. SMETS Josef, "Comparaison de deux sociétés paysannes et informatique. Le logiciel CLIO", dans *Actes...*, p. 299.

12. THALLER Manfred, dans *Actes...*, p. 339-352.

13. LECAM J.-L., "Prosopographies d'élèves boursiers en Allemagne protestante au XVIème siècle", dans *Actes...*, p. 19-38. - Voir également B. DESJARDINS, "Quelques éléments de l'expérience informatique du Programme de Recherche en Démographie historique", dans *Actes...*, p. 159-178.

14. Cf. *Nouveau manuel de dépouillement et d'exploitation de l'état civil ancien*, cité par B. DESJARDINS, p. 160.

15. PETIT J. et ZYSBERG A., dans *Actes...*, p. 41.

Dans leur diversité, les travaux de notre table ronde évoquent cette étape "délicieuse, perverse et démoralisante" (16) au cours de laquelle il faut décider du traitement des données. Le dilemme devant lequel se trouvent souvent les historiens est celui de la rigueur qu'exige le codage confronté à la nécessité d'exploiter la richesse des documents. L'exhaustivité étant l'ennemie de l'informatique comme de toute méthode, les programmes sont finalement affaire d'équilibre entre la systématisation du codage et les nécessités de l'analyse concrète des données : d'une part les documents historiques sont souvent peu homogènes et comportent des ambiguïtés que le codage peut gommer en appauvrissant le contenu; d'autre part, le codage, écran entre la source et la métasource, peut mener à la dérive si les questions ne sont pas rigoureusement cernées et si le matériau n'est pas maîtrisé. "L'éloge de l'artisanat" se justifie alors car, sans compter le temps perdu, la sophistication des moyens, la tyrannie du matériel informatique... et des informaticiens, la recherche doit affronter encore la surabondance des recoupements, des "listings" (17) et des possibilités quasi innombrables offertes au chercheur. Toutes les communications de notre table ronde insistent sur la quantité de temps à investir dans la satisfaction des besoins préliminaires de chaque recherche informatisée, mais elles reconnaissent en même temps que c'est précisément grâce au tâtonnement dans les phases de sélection et de codage qu'on peut parvenir à maîtriser la documentation et à donner toute sa rigueur à la méthode.

Une série d'expériences menées en vue de remédier aux lourdeurs de l'informatique a donné naissance à PROSOP. Par le moyen de ce système de traitement automatique des données prosopographiques (18), J.-P. GENET et M. HAINSWORTH ont tenté d'uniformiser les étapes de la constitution d'un fichier de notices biographiques en élaborant un modèle de fiche bio-bibliographique pour un matériel individuel et facile d'accès (19), dont le traitement est fonction des bases déjà existantes. Ils visent ainsi à gagner du temps, tout en rendant plus efficace l'usage de l'ordinateur grâce à la systématisa-

16. CHARLE C., dans *Actes...*, p. 235.

17. "Le syndrome de la masse de papier"; cf. B. DESJARDINS, dans *Actes...*, p. 165.

18. HAINSWORTH M. et GENET J.-P., "PROSOP, un système de traitement automatique des données prosopographiques", dans *Actes...*, p. 279-298.

19. Micro-ordinateur Apple II, logiciel Applewriter qui existe au Laboratoire d'Informatique pour les Sciences de l'Homme (CNRS) et dont les données sont transmissibles aux gros ordinateurs.

tion dans PROSOP d'une série d'opérations communes à toute recherche prosopographique informatisée.

Rédaction, saisie, corrections

Une fois les programmes conçus ou tout au moins choisis, une fois formulées les hypothèses et confectionné le bordereau de saisie, il s'agit de passer à l'étape de la rédaction de la fiche en vue de l'entrée en ordinateur.

Cette phase - que les progrès actuels de l'informatique tendent à réduire par la suppression du bordereau - est la raison d'être concrète des étapes précédentes et représente une mise à l'épreuve de la structure et de la syntaxe du bordereau. Elle est suivie par l'opération de saisie que toutes les communications s'accordent à qualifier de fastidieuse et d'ingrate du fait que, en apparence, elle soustrait du temps à l'étape plus noble d'interprétation des résultats. Les chercheurs jugent cependant cette épreuve nécessaire pour l'appréciation de la grille de lecture qu'est le bordereau de saisie.

Plus ingrate encore est la phase de correction des erreurs : erreurs de programmation, d'écriture ou de saisie qu'il s'agit de détecter par l'épluchage des "listings" ou parfois par les tests de vraisemblance; la saisie en conversationnel permet de contrôler instantanément chaque entrée et évite de reporter toutes les rectifications en fin de phase. Il reste cependant que les révisions et les retouches des entrées sont assurément les opérations les plus mutilantes et les plus ingrates du travail informatique.

Vient enfin l'analyse des résultats, étape en soi familière à tout historien qu'il soit ou non usager de l'informatique. L'ordinateur en tout cas aura contraint le chercheur à la rigueur puisqu'il ne supporte ni les exceptions ni le flou; il aura permis également des corrélations difficiles à réaliser par les procédés manuels : "utiliser l'informatique équivaut à avoir de nouvelles idées. Il s'en dégage un supplément de réflexion entraînant l'approfondissement et la nuance des résultats" (20).

En historien familiarisé avec l'usage de l'ordinateur, J.-P. GENET souligne, en conclusion, l'enrichissement que procure l'usage de l'ordinateur à la démarche prosopographique. Celle-ci a déjà profité de l'apport de disciplines

20. ACERRA M., dans *Actes...*, p. 232.

NOTE DE LECTURE

telles que l'onomastique, la démographie et l'anthropologie. L'ordinateur la situe désormais à la croisée des apports de l'informatique et des sciences humaines.

Le métier d'historien, fort exigeant déjà en temps et en argent, se voit donc convoqué à un nouvel effort : c'est bien l'historien lui-même en tant que tel qui joue le rôle fondamental dans l'enrichissement même que lui apporte l'informatique; c'est lui en effet qui constitue la métasource et la structure de base qu'il fournit à l'informatique pour le seconder. Les travaux présentés à notre table ronde ont bien démontré "que l'informatique n'est pas un gadget extérieur à la pratique historique : elle est au contraire l'outil grâce auquel cette pratique évolue et s'adapte à son objet" (21).

Ces mêmes communications ont également bien souligné la nécessité du travail en équipe pour accumuler les acquis des bases de données, autrement dit elles ont montré que "le recours aux techniques modernes ne dispense pas du travail de réflexion sur les données". Mais les conclusions de fond concernent, à notre avis, les rapports entre l'historien et ses sources. Comme l'a noté M. LUZZATI, "notre idée du rapport ordinateur-prosopographie est pour l'instant étroitement instrumentale" et ne permet pas "d'aller plus loin que la technique d'édition des sources" (22). De leur côté, J. PETIT et A. ZYSBERG nous rappellent "la fragilité du terrain de départ : une documentation nécessairement lacunaire sur une société du passé" (23).

Kmar BENDANA-MECHRI
août 1988

21. GENET J.-P., "Conclusion", dans *Actes...*, p. 356.

22. LUZZATI M., dans *Actes...*, p. 151.

23. PETIT J. et ZYSBERG, dans *Actes...*, p. 40.