

Le courrier du CNRS 65

Auteur(s) : CNRS

Les folios

En passant la souris sur une vignette, le titre de l'image apparaît.

52 Fichier(s)

Les relations du document

Ce document n'a pas de relation indiquée avec un autre document du projet.□

Citer cette page

CNRS, Le courrier du CNRS 65, 1986-07

Valérie Burgos, Comité pour l'histoire du CNRS & Projet EMAN (UMR Thalim, CNRS-Sorbonne Nouvelle-ENS)

Consulté le 03/02/2026 sur la plate-forme EMAN :

<https://eman-archives.org/ComiteHistoireCNRS/items/show/157>

Copier

Présentation

Date(s)1986-07

Mentions légalesFiche : Comité pour l'histoire du CNRS ; projet EMAN Thalim (CNRS-ENS-Sorbonne nouvelle). Licence Creative Commons Attribution – Partage à l'Identique 3.0 (CC BY-SA 3.0 FR).

Editeur de la ficheValérie Burgos, Comité pour l'histoire du CNRS & Projet EMAN (UMR Thalim, CNRS-Sorbonne Nouvelle-ENS)

Information générales

LangueFrançais

CollationA4

Informations éditoriales

N° ISSN0153-985x

Description & Analyse

Nombre de pages52

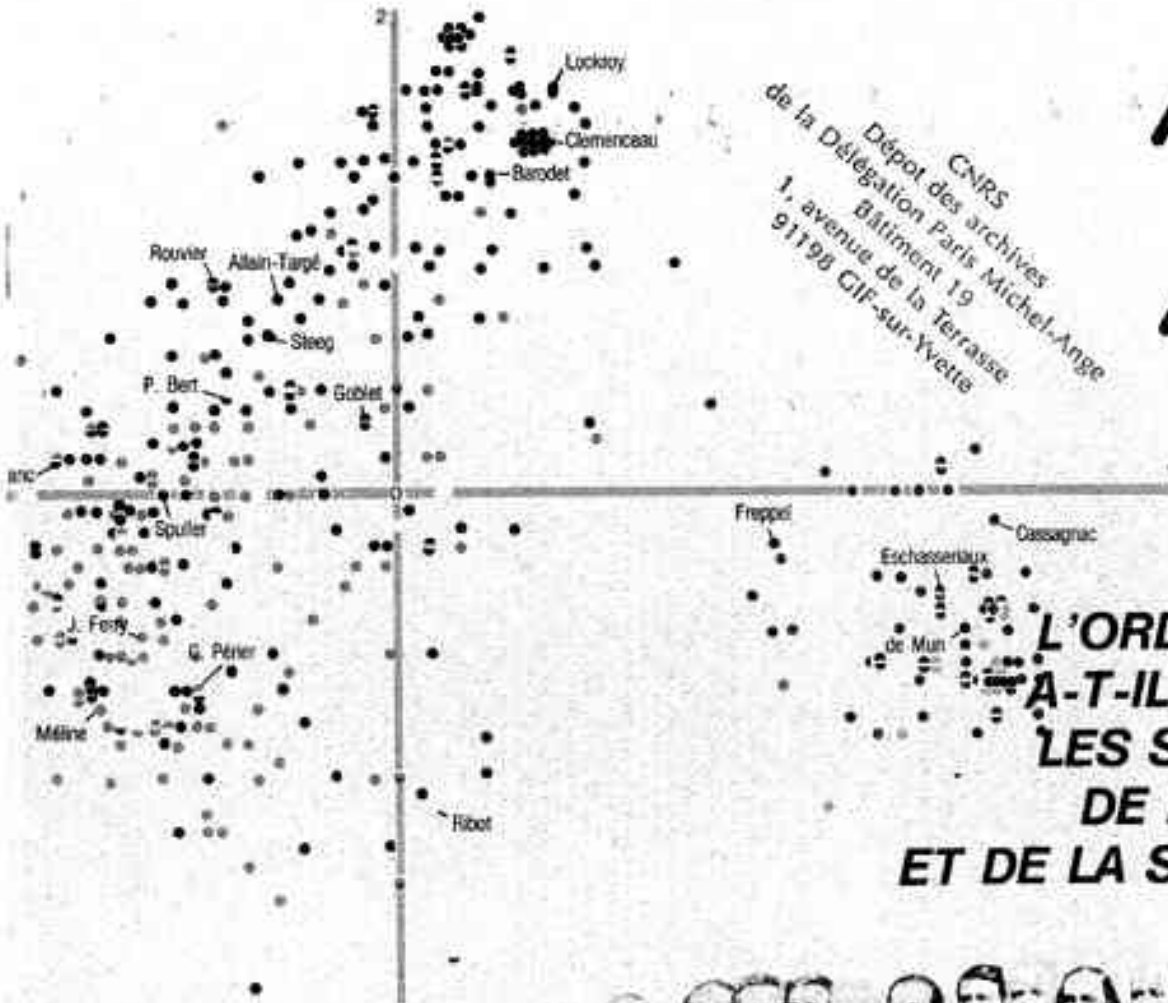
Notice créée par [Valérie Burgos](#) Notice créée le 05/10/2023 Dernière modification le 11/12/2024

LE COURRIER DU CNRS



65

CNRS
Dépôt des archives
de la Délégation Paris Michel-Ange
Bâtiment 19
1, avenue de la Terrasse
91198 GIF-sur-Yvette



DOSSIER :

**L'ORDINATEUR
A-T-IL CHANGÉ
LES SCIENCES
DE L'HOMME
ET DE LA SOCIÉTÉ ?**



GROUPE DE DÉPUTÉS - DÉPUTÉ DE LA SEINE - 21 MAI 1901

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
Mai-Juillet 1966 - 20F



Le *Courrier du CNRS* s'est vu obligé, pour des raisons budgétaires, de modifier le rythme de parution de ses numéros pour l'année 1986.

En conséquence, le numéro 65 sera diffusé au mois de septembre et non en juillet comme prévu.

Le numéro suivant, numéro triple en couleurs, paraîtra au tout début de l'année 1987.

Pour l'année 1987 nous retrouverons la périodicité habituelle de quatre numéros dont un numéro double.

La Rédaction.

Directeur de la publication : Geéry Delacôte.

Rédaction : Veronique Brossolet-Condé (rédacteur en chef), Régine Farré (rédacteur en chef-adjoint, responsable des sciences de l'homme et de la société), Françoise Bescond (secrétaire de rédaction, responsable de *La vie des laboratoires*).

Secrétariat : Martine Roche.

Entretiens : Monique Mounier-Kuhn.

Comité de rédaction : Robert Barbauld, Jean Bourdon, Georges Chapouthier, Bernard Dorny, Claire Dupas, Catherine Fuchs, Jean-Philippe Genet, Françoise Harrois-Monin, James Hiébiot, Louis Jaunou, Claudine Laurent, Jean Leca, Jacqueline Mirabel, Jean Montuelle, Janine Rondesi, Hervé Thiry, Maryvonne Tissier, Guy Vidal Naquet.

Membres correspondants : Nadine Chalem-Gouarin, Gérard Lilland, Jean Choufy.

Abonnement et vente au numéro. Le numéro : 20 F. Abonnement annuel : 74 F - 84 F pour l'étranger (voir bulletin d'abonnement pages 27-30). Tout changement d'adresse doit être signalé à la rédaction. Revue trimestrielle comportant cinq numéros par an. Nous remercions les auteurs et les organismes qui ont participé à la rédaction. Les interviews et les chapitres introduits ont été rédigés par la rédaction. Les textes et illustrations peuvent être reproduits sous réserve de l'autorisation écrite du directeur de la publication.

Direction artistique : Brilant et Associés/G. Moine, 56, rue Boissière, 75014 Paris. Réalisation : Bolo-France Impression, boulevard de Beaubourg, Emeraillville, 77200 Torcy.

C.P.A.D. 300-ISBN-2-222-02917-7-ISSN 0 153-945 X

© Centre national de la recherche scientifique.

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, 15, quai Anatole-France, 75700 Paris. Tél.: (1) 45.55.82.25. Téléc.: 260.034.

LE COURRIER DU CNRS



Page 1 de couverture

L'analyse factorielle des scrutins permet de reconstituer la structure de la Chambre des députés (1891-1894).

— Configuration de la Chambre d'après 15 scrutins législatifs (1811-1854): extrait de « La Chambre des députés (1811-1854). Analyse factuelle des scrutins », A. Prost, Chr. Remy, *Revue française de science politique*, février 1971, vol. XXI, n° 1, pp. 3-36.

ROUGE	Radicaux et extrême-gauche	JAUNE	Républicains mo- dérés
NOIR	Gambettistes	BLEU	Bonapartistes
ORANGE	Férentes	VERT CLAIR	Conservateurs
		VERT FONCÉ	Legitimistes

— Groupe des députés de la Seine (Élection du 21 août 1881). Musée Carnavalet, Paris. © Louvre-Limbourg.

Figure 3-4 de conversion :

Connaissance des milieux agricoles du Nord-Est de la Thaïlande (Sakon-Nakhon, Udon-Thani) à partir des données satellitaires LANDSAT
© IGN 1985

Sommaire N° 65

- Page 1 de couverture**
Analyse factorielle des scrutins permet de reconstituer la structure de la Chambre des députés (1881-1884).
Configuration de la Chambre d'après 13 scrutins (legislature 1881-1884), extrait de « La Chambre des députés (1881-1885). Analyse factorielle des scrutins », A. Prost, Chr. Rousselle, Revue française de science politique, février 1971, vol. XXI, n° 1, pp. 3-30.

DOUGE	Radiciaux et extrême-gauche	JAUNE	Républicains modérés
OR	Gambettistes	BLEU	Bonapartistes
RANGE	Férentistes	VERT CLAIR	Conservateurs
		VERT FONCÉ	Légitimistes

Groupes des députés de la Seine (Élection du 21 août 1881), Musée Carnavalet, Paris. © Lucien-Géraudon.

Pages 3-4 de couverture :
Reconnaissance des rizières agricoles du Nord-Est de la Thaïlande (Sakon Nakhon, Udon-Thani) à partir des données satellitaires LANDSAT-1, IGN, 1985.

Sommaire N° 65

 - Editorial**
Serge Feneuille
 - Flash actualités**
 - Débats et positions**
L'ordinateur a-t-il changé les sciences de l'homme et de la société ?
 - Présentation**
Hervé Théry
 - Des ordinateurs, pour quoi faire ?**
Alain Degenne
• Une application de l'informatique en musicologie
Hélène Charnassé
 - Impact de l'informatique sur la recherche historique**
André Zysberg
• Un who's who en histoire
Régine Ferré
 - Le sociologue et ses données**
Philippe Cibois
 - Ordinateur et terrain en Thaïlande**
Michel Bruneau
 - Géographie : un renouveau assisté par ordinateur**
Denise Pumain
 - La lexicométrie socio-politique**
Maurice Tournier
 - La linguistique informatique**
Jean-Pierre Desclés
 - Démographie : l'âge de raison**
Hervé Le Bras
 - Sciences économiques : quand l'outil valide le parole**
Jean Bourdon
 - Les nouveaux outils de l'archéologie**
Henri Ducasse
 - Archéologie et calculateurs : vers une archéologie théorique**
Jean-Claude Gardin
 - Coopération Internationale**
Aspects de la recherche scientifique et technique au Japon
Georges Rieu
 - Chercheurs-jeunes à la rencontre**
Initiation à l'astronomie
Monique Mounier-Kuhn - Jean-Pierre Sivan
 - Itinéraires**
Nicolas Franceschini ou l'exploration des petits cerveaux
Régine Ferré
 - Enquête**
Les réseaux informatiques
Max Fontet - Guy Pujolle

LE COURRIER DU CNRS

Contents N° 65



- 5 Editorial**
Serge Feneuille
- 6 Flash news**
- 7 Debates and points of view**
Has the computer changed the human and social sciences?
Presentation
Hervé Théry
- 8 Computers, to do what?**
Alain Degenne
The computerisation of research in the human and social sciences: an irreversible process which we must learn to control.
● An application of data processing to musicology
Hélène Charnassé
- 12 The impact of data processing on historical research**
André Zysberg
An historical breakthrough? A machine for filing records? or just a precious methodological tool?
● A who's who in history
Régine Ferré
- 16 The sociologist and his data**
Philippe Cibols
Data analysis, and in particular factorial analysis, has been the main contribution of the computer to sociology. It helps « visualize » data, though the interpretations that it suggests remain to be tested.
- 18 The computer and the terrain in Thailand**
Michel Bruneau
An exemplary case of the integration of the computer in a research practice: factorial analysis and remote sensing in Thailand.
- 22 Geography: a computer-assisted revival**
Denise Pumain
The computer, which is only partly responsible for the renewed interest in geography, has facilitated this revival by making it possible to test models and theories.
- 24 Sociopolitical lexicostatistics**
Maurice Tournier
Although it is not possible to directly process the contents of political speeches, the analysis of the words composing them clearly demonstrates choices, habits, and prejudices.
- 33 Computer linguistics**
Jean-Pierre Desclés
The computer, which helps us understand our own language, also adapts itself to the dialogue which we attempt to institute with machines.
- 35 Demography: the age of reason**
Hervé Le Bras
A tradition hostile to the computer explains the late conversion of the demographers, however, since 1975, this tool has gradually acquired respectability.
- 37 Economic sciences: when the tool validates the word**
Jean Bourdieu
The economic sciences may expect a lot from the computer. The data provided, however, must be reliable, the simulations must validate theories and in particular, in practice, they must assist decision.
- 40 The new tools of archeology**
Henri Ducasse
From the era of management to that of exploitation.
- 42 Archeology and computers. Towards a theoretical archeology**
Jean-Claude Gardin
Where archeological knowledge, with help from the computer, links up with the archeology of knowledge propounded by Michel Foucault in the Human and Social Sciences.
- 44 International cooperation**
Aspects of scientific and technical research in Japan
Georges Rieu
Japan, with a surface area of approximately four french departments in which to develop its industry and agriculture and to house its population of more hundred millions and a total lack of natural resources, has chosen to rely on scientific and technological development to compensate for these handicaps, insure economic growth and face international competition.
- 50 Scientists encounter the young**
Initiation into Astronomy
Monique Mounier-Kuhn - Jean-Pierre Sivan
A CNRS scientist helps fifty sixth form pupils studying science at the Joliot-Curie Lycée at Aubagne discover astronomical research.
- 51 Itineraries**
Nicolas Franceschini or the exploration of small brains
Régine Ferré
- 52 Enquiry**
Computer networks
Max Fontet - Guy Pujolle
The computer, which has become a means of communication, is an indispensable research tool. Scientists must be equipped with data processing equipment connected in networks in order to be integrated with European research institutions.

Le CNRS navigue au milieu des tempêtes, disent les journalistes. Il est vrai que depuis plusieurs mois, le CNRS connaît quelques turbulences, notamment avec l'annulation par le Conseil d'Etat de l'article concernant les modalités d'élection du Comité National.

La conséquence la plus grave de cette annulation a été l'arrêt des concours de recrutements pour 1986. Pour combler ce vide juridique, il était urgent de mettre en place des mesures transitoires permettant aux candidats d'attendre la reprise de ces concours.

C'est ainsi que deux cent quatre-vingts candidats se sont vu proposer des contrats temporaires d'un an, et que soixante autres candidats à un recrutement ont reçu une offre de détachement — notamment de l'Education Nationale —.

La presque totalité d'entre eux a répondu positivement. Par ailleurs, plus de cinquante candidats qui ne pouvaient bénéficier de ces mesures, ont reçu des propositions variées émanant de tous horizons. Ce qui leur permettra d'attendre la reprise des concours — qui devraient se tenir au début du printemps prochain.

Mais la solution provisoire trouvée à ce problème, ne saurait faire oublier que le CNRS fait actuellement l'objet de nombreuses critiques dans le monde politique, économique et universitaire. Son existence même est parfois remise en question.

Pourtant le contexte n'est pas aussi défavorable que ces attaques pourraient le laisser imaginer. Le CNRS a le soutien plein et entier de son ministre de tutelle. Monsieur Alain Devaquet ne cesse, en effet, de réaffirmer la nécessité de maintenir l'unité du CNRS et de conforter son rôle dans la recherche française, en particulier dans la recherche de base.

Pour ne pas décevoir cette confiance, il faut apprécier les forces du CNRS, nous appuyer sur elles puis détecter ses faiblesses réelles et les corriger. Durant ces trois derniers mois, j'ai pu par de multiples consultations identifier, parmi les critiques faites au CNRS, celles qui sont fondées. Cette analyse et ses conclusions sont déterminantes dans l'élaboration des réformes envisagées.

On reproche beaucoup au CNRS sa situation hégémonique, sa trop grande technocratie, la complexité et le manque de transparence de son organisation. L'esprit des réformes à entreprendre sera donc tourné vers une simplification, une déréglementation et une responsabilisation à tous les niveaux de la structure et notamment de la direction générale.

J'ai actuellement un nombre de collaborateurs directs trop important. Une réduction du nombre de ces collaborateurs directs s'avère nécessaire pour assurer une gestion et une politique de recherche plus efficaces. Un certain nombre de directions actuellement existantes seront vraisemblablement regroupées. Cette équipe de direction restreinte prendrait le nom de comité exécutif. Elle comprendrait un petit nombre de directeurs scientifiques du CNRS choisis pour leur compétence et leur personnalité et sans définition préalable des domaines de recherche qui leur seraient attachés.

Les relations avec les laboratoires et les unités de recherche seront également simplifiées afin d'assurer une véritable délégation des pouvoirs et une responsabilité plus affirmée de leurs directeurs.

Dans le même esprit, le nombre des départements scientifiques — intermédiaire essentiel entre la direction générale et les laboratoires — sera augmenté. Des structures réduites per-

mettent toujours de mieux cerner les problèmes existants et d'y répondre rapidement.

Par ailleurs, la communauté scientifique exercera pleinement son pouvoir de proposition en matière de politique scientifique. La prospective doit redevenir une de ses préoccupations majeures au détriment peut-être d'une programmation qui a été parfois excessive. L'élargissement envisagé de la base électorale du Comité National trouve son origine dans le souci d'assurer une meilleure représentation au sein de cette instance.

Certains thèmes de recherche seront peut-être moins soutenus que par le passé, au profit de nouvelles voies qu'il nous faudra développer. De nouveaux laboratoires, de nouvelles équipes devront être créés. De jeunes chercheurs particulièrement talentueux devraient pouvoir en prendre la responsabilité; évidemment ces créations supposeront le repli et parfois même la restructuration d'autres unités de recherche.

Cette réorganisation interne ira ainsi vers un renforcement au niveau stratégique des structures du CNRS et une meilleure adaptation de celles-ci aux besoins du terrain — deux faiblesses actuelles de l'organisation — qu'il fallait pallier.

Quant à la réputation d'hégémonie et de dirigisme qu'a le CNRS, nous y répondrons par un nouveau type de relations avec nos associés: le partenariat, c'est-à-dire la responsabilité partagée, la concertation et la prise de décisions en commun.

La conduite de cette réforme est essentielle pour l'avenir du CNRS. Il appartient à nous tous de participer activement à sa mise en place. Sa réussite prouvera à l'extérieur la fiabilité, la transparence, le dynamisme et l'ouverture de notre organisme.

Le rôle important qu'il joue dans le progrès de la recherche française et notamment des connaissances scientifiques de base n'en sera alors que plus reconnu et apprécié.

Serge Feneuille
Directeur général du CNRS



Serge Feneuille, Directeur général du CNRS depuis le 4 juin 1986.
© CNRS. Photo: OROP.

* De plus, près de soixante-dix candidats à un détachement de l'enseignement supérieur ont vu leur demande acceptée.

DES INSTITUTS FEDERATIFS VOIENT LE JOUR

Au printemps 1998, ont été inaugurées deux Instituts fédéraux du CNRS, l'Institut National de recherche sur les sociétés contemporaines (IRESCO) à Paris, et l'Institut de recherches et d'études sur le monde arabe et musulman (IREMAM) à Aix-en-Provence. Un troisième à Lille est en cours d'installation. L'Institut fédéral de recherches sur les sociétés industrielles (IRFSI) est...

Un institut fédéral est un regroupement, on un même bâtiment ou sur un même campus, de plusieurs laboratoires (groupes ou associations au CNRS) travaillant sur des problématiques communes ou complémentaires. Chaque équipe garde son indépendance, mais donne à l'ensemble une plus-value scientifique en contribuant au programme propre de l'institut afin de développer des recherches dans des domaines d'importance nationale. Les équipes profitent d'équipements collectifs nouveaux et lourds et de laboratoires qui seraient isolés, car un institut fédéral est aussi un projet d'utiliser au maximum de leur capacité des équipements qui demandent des investissements toujours croissants et qui finissent à l'échelle d'un seul laboratoire. Les instituts fédéraux sont des structures très locales et peuvent relever de plusieurs départements; ils se créent souvent à partir d'initiatives locales, plusieurs équipes demandant à travailler ensemble. Les programmes doivent évoluer et avec eux les équipes qui les composent. Les équipes d'ont peu un droit à vie à être hébergées, elles ont tout parti pour avant que leur programme soit une composante d'un programme plus vaste qui est celui de l'institut.

L'HEMAM et l'IRESCO illustrent bien cette notion d'intégrité fédérale. Dans le domaine de l'hémalogie et de l'étude du monde arabe, la France a une tradition intellectuelle fort marquée par des noms tels que Louis Massignon, Louis Massignon, Jacques Berque, mais aussi, et plus particulièrement, avec la mise au point de la Méditerranée, avec des responsables tels que par exemple l'arrivée de nombreux étudiants arabes, les origines de ces pays dont nous devons faire les chercheurs de demain au monde arabe. A Aix-en-Provence, des forces aussi nombreuses étaient réunies depuis plus de deux siècles et avaient déjà constitué un GSI Méditerranéen. Deux autres institutions, un comité de Maugras, l'autre au Maroc les ont rejointes, ce sont des chercheurs, une certaine œuvre, appartenant au GSI et à son université, que regroupent l'HEMAM. Ses statuts et ses représentations, coloniales du préhistoire, ethnologie, géographie, histoire, sociologie, sciences humaines, sciences.

LE PESCO s'intéresse certes à la France, mais aussi à d'autres sociétés prises comme exemples typiques de la mondialisation. Cet institut installé dans le nord de Paris, devant d'abord offrir un cadre de travail à une vingtaine d'étrangers peints, sculpteurs, musiciens, mais également d'artistes, qui ont en commun une certaine diversité de daccapores, non seulement un objet, le contemporain, mais aussi une démarche intellectuelle. L'analyse de données de terroirs contribue par les chercheurs empruntées à des observations réalisées d'origine administrative.

A life, sociological, geographical, historical, economic, political, scientific and cultural study of the multi-

tion industrielle dans différentes régions du monde. L'IFREI a une importance majeure dans cette région du Nord, d'industries comme on en trouve, et qu'il est à présent au cœur des bouleversements industriels, technologiques, économiques et sociaux actuels.

Mais les sciences sociales et humaines ne sont pas les seules disciplines qui organisent leurs recherches autour d'instituts fédéraux. A Montréal, un institut de chimie, qui regroupe plusieurs équipes travaillant notamment sur l'électrochimie et les polymères. A Nantes, un Institut des matériaux, en cours de constitution, regroupe chimistes et physiciens. Ces équipes travaillent sur des méthodes d'observation des matériaux, le plus souvent de basse dimensionnalité : cristallins minces, composés lamellaires, mono-cristaux. Par ces procédés, ils mettent au point des matériaux performants : composés d'interaction pour catalyse et électrochimie, matériaux pour l'optique, le micro-électronique, etc.

D'autres projets, moins avancés, se dessinent également à Orsay, à Strasbourg, à Montpellier et à Lyon.

LE PRIX ROUSSEL

Le prix Roussel qui couronne. Mais les deux ans une contribution majeure dans la détermination des mécanismes et auteurs, considérés vont d'être décerné en 1986 conjointement à Pierre-Alfred, directeur de recherche au CNRS, directeur de l'UJA 31 et à l'université Louis Pasteur de Strasbourg, et au professeur Guy Duranton, ex-directeur de l'UJA 31 et actuellement directeur de l'Institut de chimie des substances naturelles à Gif-sur-Yvette pour leurs travaux sur la décomposition et la polymérisation structurale de squamides océaniques et méditerranéens.

Les mixtes de Pierre-Albrecht et Ouy-Ouassan ont montré que tous les sédiments récents (quaternaires), les phytolites des rhizomes, contiennent, quels que soient leur âge et leur origine, des quantités notables d'unités linéaires particulières de tétracènes pentacycliques : les *hémipodes*. La décomposition structurale des géoponolides à mièvre que leur azéthane cumulé est général en C_{30} (36 atomes de carbone) et possède une chaîne latérale prolongée, contrairement à ceux en C_{28} (30 atomes de carbone) connus précédemment, mais n'ont été rencontrés dans les plantes. L'usage des géoponolides ne pouvait donc pas se limiter à l'étude de producteurs biologiques comme

l'identification dans les bactéries par un groupe de chercheurs américains du tétracyclinepéptidyle premier tétracycline un C_{12} possédant le même squelette carboné que les tétracyclines, permettant de formuler l'hypothèse, fondée sur des résultats pharmacologiques, d'une répartition beaucoup plus large de ce type de composés dans les micro-organismes, particulièrement celle dans la transformation de la matière organique du sous-sol. Cette hypothèse a rapidement trouvé sa confirmation dans les travaux de M. Rohmer qui a montré que le tétracyclinepéptidyle est bien le prototype d'une famille de tétracyclines, celles-ci présentant diverses variantes structurales sur le squelette de la chaîne latérale.

est donc la présence très générale des bioturbateurs dans les procaryotes (bactéries, algues bleues)

Qui explique la distribution universelle de leurs fossiles moléculaires dans les sédiments et les rochers.

Les travaux de biochimie et de biophysique du laboratoire strasbourgeois ont montré, d'une part, que les biophospholipides sont des constituants essentiels de la membrane cellulaire de nombreux procaryotes, qui en contiennent en général pas les stéréols, et jouent dans ces micro-organismes un rôle d'agent régulateur analogue à celui des stérols dans les membranes d'eucaryotes. Les architectures et encombrements moléculaires de ces deux types de molécules sont, en effet, très voisins. Le caractère amphipatique et la faible solubilité des biophospholipides expliquent les raisons pour lesquelles ces molécules, qui constituent probablement une des familles les plus répandues, dans le sous-sol et dans le biosphère, aient si longtemps échappé à l'investigation des chercheurs et n'ont été découvertes que grâce à leurs isolats microbiens. Quant on arrive à les séparer, on les perd, en effet, facilement sous forme de mousses à l'interface eau-solvant. Dans les sédiments et les sols, on les caractérise aisément de ces molécules est largement perdue par le jeu de transformations biochimiques qui éliminent les groupements hydroxyles. Les méthodes biochimiques donnent facilement satisfaisantes, ce qui a permis de les collecter, puis de les séparer.

Tous les procaryotes ne connaissent cependant pas de biotransformations. Dans certains micro-organismes, la stabilité mécanique des membranes est assurée par d'autres polymérisés : des lamelles de lipides ou des lamelles du chitinyle (dans les archaebactéries) qui traversent complètement la membrane, parfois des fibres de encore peut-être des terpènes incluses à chaîne zéroïque dont seuls les hydrogènes finaux sont ou à instant connus. Les fibres moléculaires de ces acides microbiens s'accumulent dans la même organelle sédimentaire et les périoles, dont le fluide est pur, mureux, conséquente ou l'inférence active microbienne qui transforme la matière organique dans les premiers stades de la sédimentation. Leur caractérisation et l'évaluation continue à la reconnaissance du rôle le plus général des polymérisés dans les organismes vivants, qui met d'assurer la stabilité biomécanique des membranes cellulaires.

La plupart des copolymères ne sont d'ailleurs pas seulement des analogues structuraux et fonctionnels, mais ils biomimètent sur des voies très similaires. Ceci a permis de formuler l'hypothèse, fondée à la fois sur des résultats géochimiques, biochimiques et biophysiques, d'une séquence phylogénétique évolutive des néoorganismes de biomimètiques conduisant de polymères simples à travers les biopolymères jusqu'aux protéines.

Les transformations des molécules biologiques en cours d'épuration fossiles sont largement régies par les conditions du milieu de dépôt, leur altération a été suivie conduisant à la découverte de nouvelles relations microbiennes. Les recherches fossiles elles-mêmes évoluent en fonction de l'enrichissement des séquences et de l'augmentation de température au l'accompagnement, elles subissent notamment des changements de stéréochimie conduisant à des isomères des matières, ce qui permet de les utiliser comme marqueurs moléculaires pour la recherche du pétrole. C'est l'équipe pour laquelle une grande partie du travail de l'équipe strasbourgeoise a été réalisée en collaboration avec la Société nationale De Aquatane production et l'Institut français du pétrole.

L'ORDINATEUR A-T-IL CHANGÉ LES SCIENCES DE L'HOMME ET DE LA SOCIÉTÉ ?

L'usage de l'ordinateur est devenu banal pour les chercheurs en sciences de l'homme et de la société : plus question d'asseoir prestige et pouvoir sur sa seule utilisation, presque plus de refus obstiné d'y recourir sous prétexte qu'il n'apporte rien ou n'a que des effets pervers. Le temps de la maturité est venu, et avec lui le moment de faire le bilan de ce qu'il a apporté, pour le pire et le meilleur, à ces disciplines.

Ce dossier le fait de deux façons différentes, en présentant un certain nombre de pratiques de recherche qui lui sont dues et en analysant son effet épistémologique : que permet-il, que l'on ne pouvait faire auparavant, et en quoi a-t-il changé la façon de penser, d'aborder et de résoudre un problème scientifique ?

Les articles ici rassemblés constituent un échantillon des diverses utilisations de l'ordinateur, dans des disciplines différentes, mais surtout sur des objets différents (données statistiques, textes, images, corpus d'objets), et nous avons demandé à leurs auteurs de répondre aux deux questions à la fois, ou à l'une ou l'autre (ce qui explique que figurent tantôt un, tantôt deux articles pour une même discipline). Tous ont utilisé et utilisent encore l'ordinateur dans leur recherche, car l'objet du débat n'était pas de faire dialoguer partisans et adversaires de son usage, mais de juger de son impact sur une pratique scientifique, et nul ne

conteste que l'on puisse parfaitement faire de la bonne science sans se servir d'un ordinateur.

Au-delà des éléments de réponses apportés pour l'histoire, la sociologie, la linguistique, la géographie, la démographie, l'économie et l'archéologie, ce dossier fait surgir de nouveaux débats. Ceux-ci recoupent les frontières disciplinaires et dépassent les problèmes de pure technique informatique, bien qu'ils aient été posés à propos de l'utilisation de l'ordinateur : comment apprécier la validité des données que l'on manipule, surtout quand les nouveaux moyens permettent d'en traiter de grandes quantités, d'aborder de nouveaux corpus ou d'entreprendre enfin l'exploitation de ceux que l'on croyait ne jamais pouvoir traiter ? Quelle place doit-on faire, dans la communication des résultats, à tout le traitement auquel on a soumis le

matériau de départ, est-ce lui qui est l'objet de l'exposé final, ou n'est-ce qu'un échafaudage que l'on retire avant de livrer un produit fini ? Quelles sont au fond les hypothèses majeures, les outils conceptuels que l'on utilisait, sans bien les expliciter parfois, et comment peut-on les formaliser pour les transposer sur ordinateur ?

On voit que des questions fondamentales, celles des méthodes et des objets, affleurent immédiatement quand on réfléchit sur l'effet de l'introduction de l'ordinateur dans ce groupe de disciplines. Même si cet effort de réflexion et de remise en question devait être le seul mérite des pionniers qui ont les premiers bricolé sur ces étranges machines – ce qui ne semble pas être l'avis des auteurs – il serait déjà à porter à leur crédit.

Hervé THÉRY



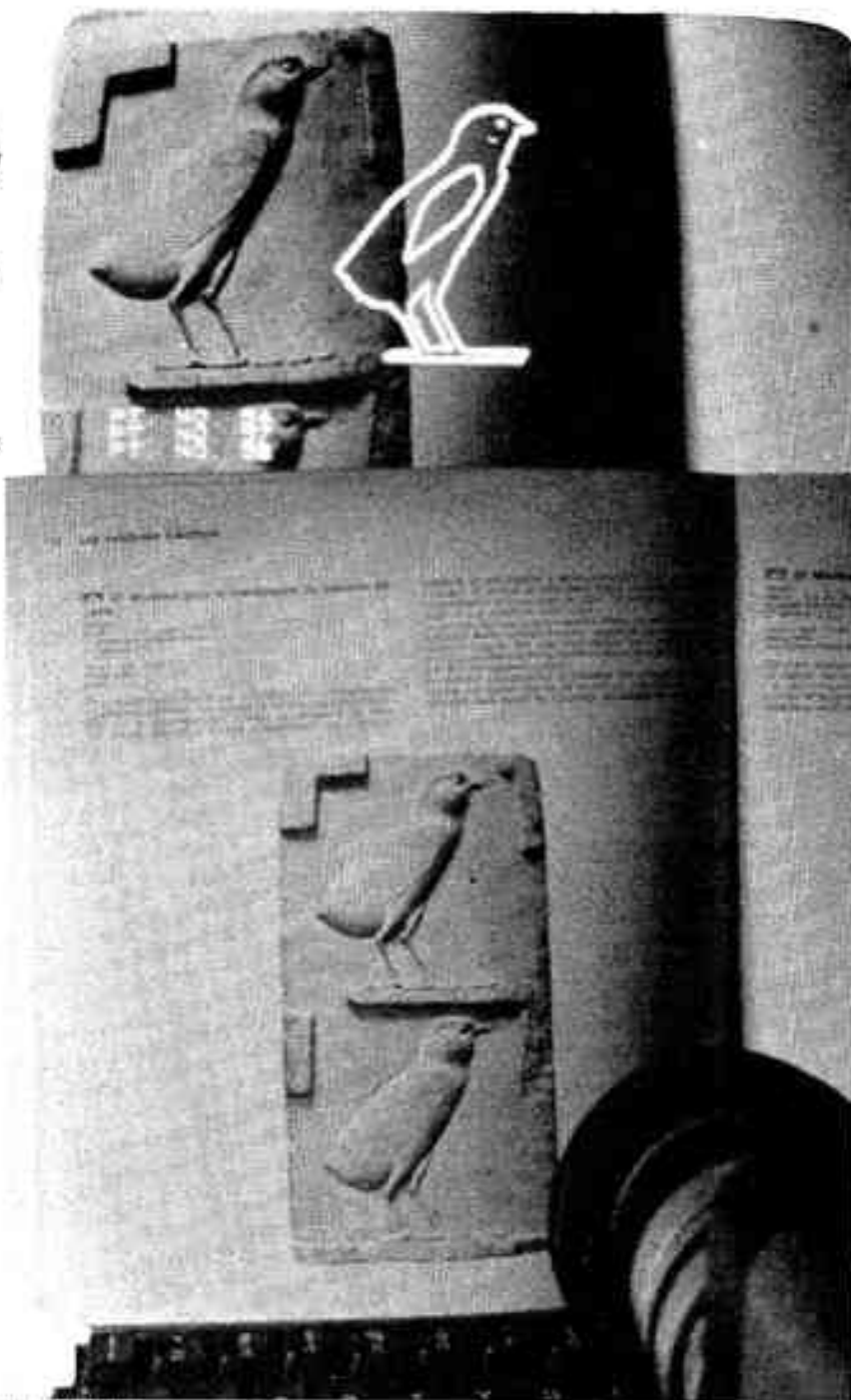
Salle des consoles du Laboratoire d'informatique pour les sciences de l'homme (LISH). © CNRS/ LISH. Photo Ph. Plailly.

■ Hervé Théry, chargé de recherche au CNRS, Centre de recherche et de documentation sur l'Amérique latine (CREDAL), UA 111, et Equipe de développement et d'enseignement de la cartographie informatisée des données (EDECID), Ecole normale supérieure, 45, rue d'Ulm, 75005 Paris.

Des ordinateurs, pour quoi faire ?

L'informatisation des recherches dans les sciences de l'homme et de la société : un mouvement sans doute irréversible, mais qu'il faut apprendre à diriger.

Alain DEGENNE



Informatique et égyptologie.

Détail du système de saisie : l'image (hiéroglyphes) filmée par la caméra-vidéo apparaît sur l'écran au fond et les données du micro-ordinateur sont superposées à la 1^{re} image. Au moyen de manettes graphiques, l'égyptologue enregistre chaque caractère point par point ; l'ensemble des coordonnées ainsi stockées est transmis grâce à un modem acoustique sur un gros calculateur du CIRCE (Centre interrégional de calcul électronique du CNRS à Orsay). À l'aide d'un langage évolué ces données sont manipulées de façon à restituer le texte imprimé.

© CNRS/LISIL. Photo Ph. Plailly.

L'impression qui prévaut dès que l'on s'intéresse à la gestion de la recherche dans les sciences de l'homme et de la société et que l'on tente de mesurer les besoins en matière d'équipement informatique est celle d'un retard accumulé depuis plusieurs années. Le

rapport sur les besoins de la recherche en matière d'équipements informatiques, présenté en décembre 1983 par Jean Yoccoz, notait que l'équipement et le budget informatiques dans les sciences de l'homme et de la société étaient « dramatiquement insuffisants ». Depuis cette date, des efforts ont été faits ; les budgets de 1985 et surtout celui de 1986 marquent une nette volonté de redresser cette situation et de mettre en œuvre une politique d'équipement.

Les sciences de l'homme et de la société auraient-elles négligé cet indispensable instrument de travail ? Se seraient-elles tenues à l'écart de la révolution intellectuelle et tech-

■ Alain Degenne, directeur de recherche au CNRS, chargé de mission au département des sciences de l'homme et de la société, responsable du Laboratoire d'analyse secondaire et des méthodes appliquées en sociologie (LASMAS), 59, rue Pouchet, 75017 Paris.

Les sciences de l'homme et de la société ont utilisé très tôt et massivement l'ordinateur.

En fait, un grand nombre d'applications sont passées dans la vie courante.

L'informatique s'est transformée. De plus en plus elle se décentralise.

L'informatique n'est plus l'affaire de quelques spécialistes. Ses applications concernent chaque chercheur.

nique que constitue l'informatique? Le public les classe en effet assez naturellement dans les activités à dominante littéraire et les imagine peu préoccupées par les nouvelles technologies. Il n'en est rien pourtant. Au contraire même, on peut dire que ces disciplines ont suivi de très près le développement de l'informatique, et ce, dès le début de l'apparition des ordinateurs dans les universités et les centres de recherche.

La première phase

Il ne faut pas oublier que les économistes, les sociologues se sont très vite appropriés l'énorme puissance de traitement et de calcul qui était mise à leur disposition pour dépouiller des enquêtes extensives et lancer des calculs qui auparavant étaient tout à fait inimaginables. Les linguistes ont, eux aussi, saisi l'ordinateur comme moyen privilégié d'expérimenter leurs théories sur la traduction automatique. Dans toutes les disciplines, l'ordinateur, comme support de traitements statistiques, c'est-à-dire comme outil de recherche, a été immédiatement et passionnément adopté. Il suffit d'ailleurs de regarder les domaines dans lesquels les sciences de l'homme ont débouché sur des applications diffusées sur le mode industriel et qui sont du domaine public, comme par exemple les sondages d'opinion. Nous sommes habitués aujourd'hui à ce que les instituts de sondages sachent prendre la température de l'opinion publique sur tous les problèmes du moment et mesurent avec une grande précision les intentions de vote. Nous sommes habitués à ce que, pour chaque consultation électorale nationale, ces mêmes instituts diffusent dès la fermeture des bureaux de vote, des estimations extrêmement fiables des résultats vrais qui ne sont connus que beaucoup plus tard. C'est là le produit du développement des recherches sur l'opinion publique et de la mise en œuvre de moyens informatiques puissants.

Les actions politiques dans le domaine économique et social, que nous sommes habitués à trouver commentées de façon très technique dans nos journaux quotidiens, sont, elles aussi, le fruit d'une maturation théorique et d'une meilleure connaissance des faits, mises en œuvre au moyen de puissants moyens de calcul.

Il convient donc d'être prudent. Les sciences de l'homme et de la société ne sont pas des spéculations hors du temps. Elles ont depuis bien longtemps débouché sur des applications courantes, qui transforment notre vie quotidienne et qui supposent la maîtrise de l'outil informatique.

Paradoxalement, c'est peut-être l'engouement initial pour les applications informatiques dans ces disciplines qui est à l'origine d'une partie du retard que l'on observe aujourd'hui dans l'équipement des laboratoires. Pendant de nombreuses années et jusqu'à une période récente, les travaux informatiques s'effectuaient pour la plupart de façon centralisée, sur des machines que l'on considérait comme relativement puissantes à l'époque, même si certaines pourraient être comparées à des micro-ordinateurs actuels, pour la taille de leur mémoire et la rapidité du calcul. Dans ce contexte, les linguistes ont mis à l'épreuve leurs hypothèses sur la traduction automatique. Ils en ont vu les points forts mais aussi les limites. Les économistes ont mis au point les premiers grands modèles à équations simultanées. Ils ont pris la mesure des difficultés que soulève leur mise au point et des limites de ce qu'ils apportent. Les sociologues ont mis en œuvre de grands programmes d'enquête. Ils ont mis des espoirs, probablement excessifs, dans l'analyse automatique des données mais ils

en ont aujourd'hui acquis la maîtrise. Les géographes ont construit la cartographie automatique et ils ont pu mesurer l'ampleur des perspectives qu'elle leur ouvrait et des problèmes qu'elle leur posait. De la même manière, on a traité par des moyens statistiques les œuvres d'écrivains ou de philosophes, Malebranche, par exemple, extrayant des concordances de termes, faisant ressortir des contextes.

On pourrait allonger cette liste, quelle que soit la discipline; tous ces chercheurs se sont trouvés confrontés simultanément à la nécessité de tirer les leçons d'une expérience qui amenait à remettre en cause certains des présupposés qui avaient fondé les travaux précédents et à l'apparition d'une autre informatique, légère, susceptible d'une plus grande décentralisation, qui permet de poser des problématiques nouvelles et qui devient tout autant un instrument de communication qu'un instrument de calcul.

La banalisation de l'outil informatique

La première phase avait structuré les tâches de calcul dans les équipes autour de quelques collègues mieux formés, ou plus motivés, ou autour de l'informaticien, pour celles qui avaient eu la chance de pouvoir en recruter. L'utilisation que l'on peut envisager aujourd'hui concerne l'ensemble des chercheurs, qui sont tous appelés à devenir responsables de leurs travaux informatiques qui du même coup prennent une place différente dans la stratégie de recherche. L'informatique que nous sommes en train de mettre en place en ce moment n'est donc plus l'apanage de quelques fers de méthodologie raffinée ou formés aux disciplines quantitatives et formelles. Cela devient un outil de travail de masse. Le rapport que les chercheurs entretiennent avec lui en est changé. Il n'est plus question de tirer avantage, dans la communauté scientifique, d'une attitude méthodologique d'avant-garde, comme certains ont pu le faire naguère. Bien au contraire, il n'est plus guère possible d'ignorer l'informatique sous peine d'être tout à fait marginalisé. Comme tout phénomène de masse, la généralisation de la micro-informatique dans la recherche pose des problèmes spécifiques. Laissons de côté la bureautique et le traitement de texte qui ne posent pas vraiment de question particulière aux chercheurs. Ils sont dans ce domaine des utilisateurs ordinaires. Au plus doit-on remarquer que, nécessité faisant loi, puisque la mise en forme dactylographique des documents a toujours constitué un goulet d'étranglement dans la production des publications qui sont relativement abondantes dans les sciences de l'homme et de la société, le traitement de texte représente une voie d'initiation à la micro-informatique qui peut avoir un impact pédagogique considérable. N'oublions pas, aussi, qu'il existe une grande variété de problèmes de représentation et de saisie graphique pour lesquels la micro-informatique apporte des solutions nouvelles: saisie et traitement de textes non latins (hébreu, chinois, arabe, par exemple), en musicologie, transcription de tablatures anciennes (encadré page suivante).

Une première classe de problèmes provient de la quantité des informations disponibles, des coûts de leur acquisition et de leur traitement. Songeons, par exemple, à ce que représente la saisie de l'œuvre de grands écrivains, comme cela se fait à l'Institut national de la langue française, pour l'élaboration du *Trésor de la langue française* et du *Dictionnaire* qui en découle. De la même manière, la télédétection par satellite fournit aux géographes et prochainement sans

Une application de l'informatique en musicologie

Hélène CHARNASSÉ *

L'Équipe de recherches sur l'analyse et la transcription des tablatures par ordinateur (ERATTO) a repris avec succès ses travaux de transcription automatique de notations musicales du XVI^e siècle (tablatures). Le stade de la transcription « élaborée » est désormais atteint. En effet, des analyses de la notation considérée en tant que langage ont mis en évidence des éléments d'interprétation implicites perçus comme tels et utilisés par le luthiste. Une grande part de logique et de tradition orale intervient dans sa mise en œuvre. Partant de ces données, des membres de l'équipe s'attachent à trouver les règles de transcription d'un type de tablature adopté par un auteur allemand, Hans Newwider, en partie inédite sur le plan de l'organisation musicale de la pièce à exécuter. De premiers essais, réalisés grâce à des algorithmes qui relèvent de l'intelligence artificielle et de la reconnaissance des formes, ont abouti à des résultats intéressants. Ils vont donc être poursuivis. En outre, un programme général interactif de transcription de tablatures, qui fonctionnera sur micro-ordinateur, est en cours de réalisation. Il facilitera l'accès à un vaste répertoire encore inconnu à ce jour.

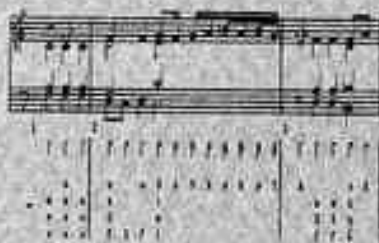
Exemple 1. Première étape des travaux : transcription élaborée de Sebastian Ochsenkun, *Tabulaturbuch*, Heidelberg, 1558. Tablature de type « structure ».



Notation originale. Dans cet ouvrage exceptionnel, la tablature reproduit la structure de la pièce à exécuter. Chaque caractère est présenté à la place qu'occupe le son dans le tissu musical. La disposition typographique de l'auteur est prise en compte lors de la saisie des données; la présence ou l'absence de caractère étant considérée comme un élément également significatif.



Exemple 2. Etape actuelle : transcription élaborée du recueil de Hans Newwider, *Der ander theil des Lautenbuchs*, Nuremberg, 1538. Tablature de type « note ».



Transcription exacte de l'original.



Transcription calculée.



Traduction de la tablature allemande, difficile à déchiffrer, en tablature française lue par tous les luthistes.

Réalisation informatique : B. Stepien.

* Hélène Charnassé, directeur de recherche au CNRS, responsable de l'Équipe de recherches sur l'analyse et la transcription des tablatures par ordinateur (ERATTO), ER 152, Collège de France, 11, rue de la Harpe, 75005 Paris. ER 1629 et B. Stepien (Université d'Orléans), Centre de recherches pluridisciplinaires, 27, rue Paul-Bert, 45000 Orléans-sur-Seine.

Une difficulté à dépasser dans l'avenir : la prolifération des bases de données et leur archivage.

doute à beaucoup d'autres utilisateurs, des quantités de données tout à fait considérables qui leur posent des problèmes nouveaux. Le traitement d'enquêtes sociologiques de plusieurs dizaines de milliers d'individus soulève des questions du même type. Plus que jamais, le choix des hypothèses de travail devient crucial.

Un aspect intéressant du développement de l'informatisation des sciences de l'homme et de la société, mais qui dépasse le secteur de la recherche, est celui de l'archivage. *A priori*, on se dit que l'on dispose d'un moyen extraordinairement efficace de se créer des archives très concentrées et facilement exploitables sur support magnétique. Or, la conservation des documents informatisés suppose une régénération

régulière. Elle a un coût non négligeable. Les supports sur lesquels on les enregistre sont fragiles, surtout s'il s'agit de disquettes. Les machines évoluent très vite, y compris les lecteurs et il n'est pas rare d'être embarrassé pour relire un document qui a quelques années et que l'on n'a pas pris soin de régénérer. Paradoxalement, l'informatique qui fait miroiter à nos yeux une très grande disponibilité de documents et la possibilité de créer de véritables banques de données archivées, risque fort de nous décevoir plus que nous ne l'imaginons. Créer de grandes banques de données informatisées est d'un coût prohibitif et ne se fera donc que sous forme d'entreprises concertées à l'échelle internationale. Des lors, les autres bases constituent une forme

de capital extrêmement fragile et éphémère. Les chercheurs qui voient dans l'informatique une plus grande facilité d'accès à certaines informations, risquent fort de se trouver rapidement confrontés au problème de la disparition de leurs archives. Il est vrai qu'il est difficile de prévoir, aujourd'hui, les conséquences à terme de l'innovation technologique très rapide que nous observons.

Une autre conséquence de la décentralisation des moyens est que tout chercheur est tenté de constituer sur l'objet de sa recherche une base de données informatisée. Qu'il s'agisse de textes, de données d'enquêtes, de données biographiques, etc., le micro-ordinateur donne à chacun le moyen de saisir, souvent au prix d'un effort considérable mais à sa portée, des données sur lesquelles il se propose de travailler. Ce foisonnement incite à la réflexion. On se trouve, en effet, en face d'une contradiction. D'un côté, la micro-informatique multiplie les sources d'information et fait entrevoir des perspectives de valorisation par la circulation des fichiers, de l'autre, elle favorise des attitudes de recherche individualistes, chacun entreprenant de créer sa propre base sans souci du voisin. La moindre des choses est de réfléchir à la mise en commun des données informatisées. Cela n'est pas toujours facile et se heurte à la pluralité des matériels, des standards de gestion.

Au moins en ce qui concerne la micro-informatique, le CNRS a décidé d'encourager la standardisation des équipements. Il encourage aussi la concertation avec les autres chercheurs intéressés, chaque fois qu'il est question de créer une base de données. Reste que dans toutes ces questions il convient d'être prudent. Les pratiques évoluent vite : les matériels aussi. Les politiques incitatives ou de contrôle doivent en tenir compte.

Il faut bien voir que dans cette période intermédiaire, nous n'avons pas encore pris la mesure des changements que l'informatique introduit dans les pratiques des chercheurs en sciences humaines. Chaque fois qu'elle permet un traitement de masse, il y a un effet en retour, pervers si l'on veut, dans la mesure où l'aspect descriptif peut prendre le pas sur la réflexion approfondie. On l'a bien vu avec la prolifération des traitements auxquels ont été soumis certains corpus tels que des textes ou des enquêtes. Avec le recul, l'on se demande parfois, aujourd'hui, si l'on avait bien mis en rapport les bénéfices attendus de ces traitements avec l'investissement qu'ils supposent. On l'a vu, également, avec la généralisation de traitements statistiques complexes comme les analyses de données. Ces outils descriptifs, qui peuvent être très précieux pour faciliter une première lecture d'un matériau, ont souvent fasciné les chercheurs au point qu'ils en ont oublié de formuler leurs hypothèses. Mais mes propos relèvent davantage de la constatation de faits que d'un jugement. Ces expériences étaient sans doute nécessaires pour que l'on puisse aujourd'hui en tirer les leçons. Finalement, bon nombre de travaux quantitatifs, qui font autorité dans notre pays, n'ont été possibles que grâce à la puissance de calcul dont nous disposons. La facilité des traitements n'interdit pas l'intuition, pas plus que leur lourdeur ne la favorisait hier. Le chercheur est un peu dans la situation du médecin, toujours tenté de demander un examen pour conforter ou affiner un diagnostic. Peut-être se fait-il, naguère, plus souvent à son intuition. Peut-être, aussi, se trompait-il plus souvent. Le sociologue, l'économiste, le géographe disposent maintenant d'un très grand nombre d'informations quantifiées ou documentaires. Ils ont des moyens de traite-

ment plus puissants. Les meilleurs de leurs travaux doivent plus à l'expertise, pas nécessairement moins à l'intuition, qu'il y a vingt ans.

Quand l'informatique remet les chercheurs en question

Je voudrais enfin évoquer un aspect de l'informatique dans les sciences de l'homme et de la société qui est encore très neuf et paraît prometteur. Il s'agit de retombées de la recherche en intelligence artificielle. Depuis quinze ans au moins, des archéologues, à la suite de Jean-Claude Gardin (voir article p. 42), ont réfléchi aux problèmes que leur posait la description des objets extraits au cours des fouilles. Afin de rendre ces descriptions pertinentes, ils ont été conduits à décomposer de très près et à simuler les raisonnements qui étudiaient les leurs. C'est ainsi que s'est développée dans ce milieu toute une réflexion sur la représentation et le traitement des connaissances.

Il ne faut donc pas s'étonner si les archéologues sont aujourd'hui parmi les premiers à s'intéresser aux systèmes experts. Ils ne sont d'ailleurs pas les seuls. On trouve à leurs côtés, par exemple, des juristes, des linguistes. Lorsqu'on laisse libre cours à son imagination, on peut attendre un effet en retour de ces sortes de travaux sur les pratiques de recherche elles-mêmes.

Construire un système expert, en effet, suppose que l'on élabore une base de connaissances, c'est-à-dire que l'on exprime, sous forme de propositions logiques simples, les principaux résultats sur lesquels on veut s'appuyer. Ensuite, il faut élaborer un moteur d'inférence qui, lui, peut être plus ou moins standardisé. Cette opération est toujours un exercice très exigeant pour les disciplines dans lesquelles les raisonnements demeurent souvent analogiques et la forme très littéraire. Nous n'en sommes qu'aux premiers balbutiements dans ce domaine, mais il faut rêver un peu. Il y a, là, un courant de recherche où la logique de l'instrument, qu'est le système expert, impose au chercheur un retour sur sa propre pratique et sur la nature de son discours. Même si l'on ne s'attend pas à des bouleversements très profonds, ces situations-là sont toujours stimulantes.

En conclusion, il me semble que l'informatique est actuellement dans les sciences de l'homme et de la société un enjeu de développement impossible à remettre en question. Le micro-ordinateur s'impose là comme il s'impose ailleurs. Le sous-équipement n'est plus acceptable. Par ailleurs, si le fétichisme du calcul et des outils statistiques tend à faire place à des pratiques mieux dominées, la facilité avec laquelle on peut disposer de mémoires de grosse capacité et la décentralisation entraînée par la micro-informatique encouragent l'attitude qui consiste à dire « créons d'abord notre base de données et nous verrons ensuite ce que l'on peut en faire ». Il y a là un risque qui commande un encouragement à la réflexion préalable et éventuellement une politique assez rigoureuse. Cela étant dit, l'ordinateur d'aujourd'hui est un outil merveilleux qui donne aux chercheurs en sciences de l'homme et de la société, la faculté, inconnue jusqu'alors, de manipuler une information toujours plus abondante et plus variée. Il nous apporte donc un moyen supplémentaire de fonder une réflexion mieux documentée. En ce sens, il est un atout pour que puisse se développer cette conscience critique dont Raymond Aron disait qu'elle est en sociologie ou en économie politique constitutive de la conscience scientifique elle-même.

L'informatique en sciences humaines : des méthodes qui se cherchent.

Mais l'ordinateur demeure un outil. Sa mise en œuvre suppose compétence et esprit critique.

Les systèmes experts constituent un domaine de recherche privilégié où l'informatique réagit sur la discipline qu'il utilise.

Impact de l'informatique sur la recherche historique

Une rupture historique ? Une machine à engranger des fiches ? Ou, plus simplement, un précieux apport méthodologique ?

André ZYSBERG

Les objets de l'histoire se sont peu à peu étendus à toutes les activités des sociétés humaines.

Ceux que l'histoire intéresse imagineront mal, *a priori*, le bénéfice qu'un historien peut retirer de l'emploi d'un ordinateur, sauf éventuellement pour écrire ses articles et ses livres avec un programme de traitement de texte. Or, qu'est-ce que le travail quotidien d'un historien, sinon la recherche minutieuse, la confrontation et le « traitement » d'informations provenant de diverses sources : archives, imprimés, images, objets ? Faire de l'histoire, c'est toujours écrire un récit ou un texte clair et intelligible (du moins si l'on veut être lu par d'autres personnes que les quelques spécialistes de telle ou telle question). C'est également, en amont, entreprendre une enquête, dont les sujets se sont peu à peu étendus à toutes les activités des sociétés humaines : celle des gens ordinaires, que Henri Marrou baptisait les « fantasmas » de l'histoire, comme les citoyens d'une colonie romaine, les marchands et clercs médiévaux, les paysans du XVII^e siècle, ou encore les banquiers et les chefs d'entreprise de la Monarchie de Juillet. Leur travail, leurs conflits, leurs relations familiales et professionnelles, leurs maladies, ainsi que leurs croyances nous intéressent désormais plus que les biographies des hommes illustres. « L'histoire bataille » elle-même, s'est parallèlement transformée et enrichie, devenant une histoire de l'État, de ses agents et de ses partenaires.

Un travail de défrichage

Cet élargissement des champs de l'histoire (car je préfère le mot champ au mot « territoire », qui sous-entend une idée de conquête aux dépens d'autres disciplines, alors qu'il s'agit plutôt d'un travail de défrichage) a induit, depuis au moins une quarantaine d'années, un bouleversement du métier d'historien. Si je m'en tiens aux époques moderne et contemporaine (mais on pourrait prendre d'autres exemples, aussi significatifs, pour le Moyen Âge et l'Antiquité) des documents méconnus ou utilisés ponctuellement, comme les registres d'état civil, les cadastres, les contrats notariés, les papiers des entreprises, les archives fiscales et judiciaires, et combien d'autres encore, ont fait l'objet de travaux dont les résultats parviennent maintenant au « grand public » sous la forme d'ouvrages de vulgarisation ou de « vie quotidienne ».

La manipulation et le dépouillement de ces sources ont nécessité l'acquisition de techniques complètement étrangères au savoir-faire traditionnel de l'historien. L'historien des populations et de la famille est ainsi devenu démographe, ou celui de la banque s'est initié aux méca-

nismes financiers et monétaires. De la même façon, la statistique est entrée dans la panoplie des historiens : trois d'entre eux n'ont-ils pas récemment publié un livre qui s'intitule *Outils statistiques pour les historiens* ?

Certains comptent et d'autres pas. La reconstitution et l'étude de séries numériques semblent aller de soi, lorsque l'on a affaire à des documents qui comportent déjà des mesures, même si celles-ci recèlent la plupart du temps bien des pièges et des ambiguïtés... Toutefois, les méthodes de l'histoire quantitative s'appliquent aussi fructueusement à des sources où l'information chiffrée est complètement absente : ce sont les ex-voto peints de Provence, témoignages de la religion populaire, tels qu'ils ont été regardés par Bernard Cousin (voir encadré p. 13), les comcristis languedociens du siècle dernier, chers à Jules Maurin, ou bien encore les proclamations électorales des débuts de la Troisième République « factorialisées » par Antoine Prost (cf. documents de 1^{er} page de couverture et ci-dessous).

Ces trois recherches, qui intéressent des domaines aussi divers que l'histoire des mentalités, l'histoire sociale et l'histoire politique, possèdent en commun l'utilisation de moyens informatiques. Leurs auteurs ont travaillé avec des logiciels de dépouillement d'enquête, autrefois assez rigides, qui peuvent maintenant s'appliquer à tous les types de données, alphabétique ou numérique, peu importe. La confection de la « grille » ou du questionnaire, ainsi que la définition de catégories ou de classes pour chaque question ou champ constituent la clef de cette démarche. Il s'agit bien d'opérations conceptuelles, que chacun d'ailleurs pratique implicite-

Afin d'exploiter les nouvelles sources, les historiens ont dû se familiariser avec des techniques complètement étrangères à leur savoir-faire traditionnel.

Avant même d'aborder le traitement informatique, la confection de la grille ou questionnaire est l'étape essentielle, qui conditionne le succès de l'enquête.



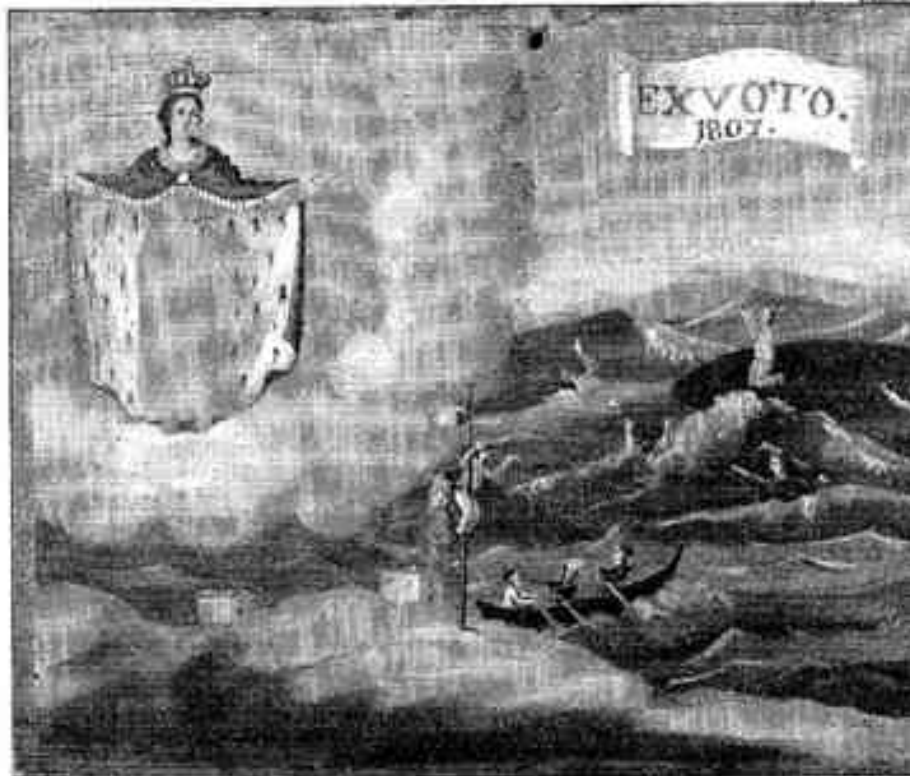
Extrait de la profession de foi de M. de Baudry-d'Arton, Recueil des textes authentiques des professions de foi et engagements électoraux des députés élus en 1881 (le Barodet), Chambre des députés, session extraordinaire de 1882, n° 808. Paris, Imprimerie nationale, 1882, t. 2, p. 534.

■ André Zysberg, chargé de recherche au CNRS, Institut d'histoire moderne et contemporaine, 45, rue d'Ulm, 75005 Paris.

Les ex-voto peints de Provence, l'historien et l'ordinateur

L'ex-voto peint est un petit tableau qui a été commandé, puis déposé dans une chapelle, une église, un sanctuaire, par un donateur unique ou collectif, en exécution d'un vœu ou d'une grâce obtenue. La surface de l'ex-voto peint est partagée en deux espaces. L'un, toujours situé dans la partie supérieure, représente celui qui est intervenu providentiellement en faveur du donateur : le Christ, la Vierge, un saint. L'autre espace, en dessous, figure le péril auquel a échappé l'auteur de l'offrande : une épidémie, un naufrage, un accident de la route... Si l'ex-voto témoigne de la dévotion, encore faut-il savoir interpréter cette « image codée ». Bernard Cousin a réuni un corpus de cinq mille ex-voto peints de Provence, dont les plus anciens remontent au XVII^e siècle, auxquels il a « posé » cent vingt-cinq « questions ». Celles-ci portent aussi bien sur les acteurs, sur les décors, le type de scène, que sur la surface respective occupée dans le tableau par l'espace céleste et l'espace terrestre. Les réponses ont été triées et analysées à l'aide d'un ordinateur. Le traitement révèle l'évolution dans le temps long des structures mentales : acte de foi inspiré à l'origine par les élites de la Contre-Réforme catholique, l'ex-voto peint, chemin faisant, devient l'expression à la fois naïve et roublarde de la religiosité populaire.

Documents de Bernard Cousin, *Le miracle et le quotidien, les ex-voto provinciaux, images d'une société*, université de Provence, Société, Mentalités, Cultures, 1983.



Tempête sur l'étang de Berre, 1897. (Notre-Dame de Miséricorde à Martigues). © R. Cousin.

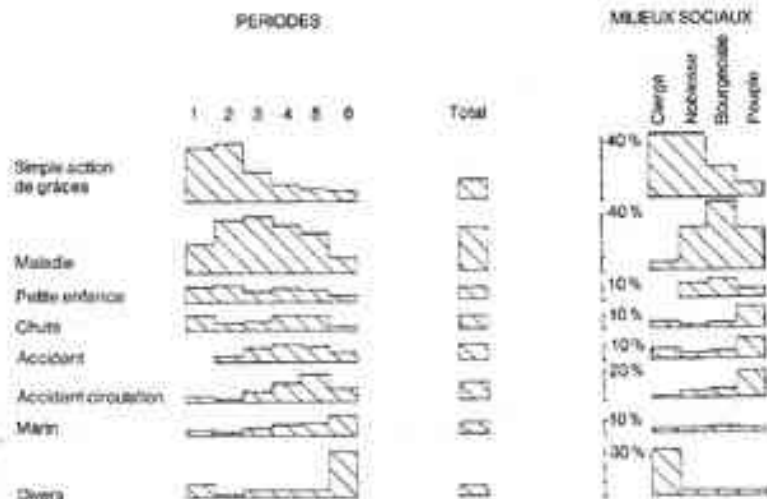
EXEMPLE DE FICHE CODÉE DE TRAITEMENT D'UN EX-VOTO

La description de l'image est formalisée grâce à une fiche codée, dont le traitement statistique et graphique est réalisé sur ordinateur.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

TYPES DE SCÈNES HUMAINES

On a représenté, pour chaque type de scène humaine, son pourcentage d'ex-voto par rapport au total des ex-voto de chaque période, à l'ensemble du corpus, et au total des ex-voto de chaque milieu social.



ment, avec ou sans ordinateur, dont les choix conditionnent, à l'évidence, la réussite de ce genre de démarche.

Cette moissonneuse-batteuse

Lorsque l'ordinateur paraît à l'horizon des sciences sociales, vers les années soixante-dix, l'on se représente avant tout celui-ci comme une sorte de mastodonte pas très finaud, juste bon à trier d'énormes quantités de fiches. Pour reprendre la narquoise expression d'Hervé Le Bras, les chercheurs à la page cessent de recueillir des fiches : ils engrangent des données. De part et d'autre de l'Atlantique, on lance alors de grandes enquêtes pour alimenter cette moissonneuse-batteuse : on perfore des dizaines de milliers de cartes et des kilomètres de ruban. Une fois ces opérations achevées, leurs promoteurs se trouvent souvent bien embarrassés, car si la machine a imprimé une pile impressionnante de papier, les résultats ne semblent pas toujours à la mesure des investissements... Des historiens sont également tombés dans ce piège, en supposant que la « masse des données » justifiait à elle seule l'emploi de la dite machine. On ira plus vite, disaient les « convertis », qui n'imaginaient pas tous les problèmes que posent la préparation des grands fichiers sur un support informatique. Nous avons des excuses, car les historiens sont souvent tentés par le vieux démon de l'exhaustivité : tout prendre, tout saisir, ne rien négliger... Chez certains même, cette thésaurisation représente une fin en soi, presque indépendamment des objectifs, des buts. N'ai-je pas entendu un universitaire affirmer, non sans gourmandise : *et le corpus ne sera jamais achevé...*

Je parle de tout cela en connaissance de cause, ayant moi-même entrepris (et heureusement achevé), pour les besoins d'une thèse, un fichier informatique réunissant les quelques 60 000 sujets de Louis XIV et Louis XV qui furent condamnés aux galères entre 1680 et 1748. Connaître la destinée de tous ces braves gens, savoir d'où ils venaient, qui les avait condamnés et pourquoi, apportera, je l'espère, une contribution originale sur le fonctionnement d'une institution militaire et punitive dans la France des XVII^e-XVIII^e siècles : et je ne crois pas avoir abouti à des évidences ou des tautologies. Toutefois, dans ce genre d'opération, il existe une fantastique disproportion entre les dix années qui me furent nécessaires pour la collecte et la mise au point des informations (compte tenu de quelques sérieux accidents de parcours comme la disparition de la machine sur laquelle j'avais commencée l'exploitation) et les quelques minutes, à peine, que demande la sortie des statistiques... D'autres historiens, particulièrement ceux qui travaillent sur la période contemporaine, ont construit une logique de sondage, qui a notamment permis d'exploiter des sources défiant l'exhaustivité du dépouillement, telles que les recensements et les listes nominatives. L'expérience acquise montre bien que c'est la finesse et la complexité d'un jeu de données, plus que le nombre de fiches, qui justifie l'emploi de l'informatique dans notre domaine. En irait-il ailleurs autrement ?

Un gain méthodologique

La recherche historique, à l'aide de l'informatique, n'implique pas nécessairement des calculs complexes et la production d'un appareil statistique très raffiné. Je me suis récemment servi d'un micro-ordinateur pour étudier

les activités d'une bande de brigands à la fin du XVIII^e siècle. Toutes les indications que j'ai retrouvées dans diverses sources à leur sujet ont été entrées en « clair » et, exception faite de quelques comptages (que j'aurais pu réaliser « manuellement »), le programme m'a surtout permis, au moyen de comparaisons successives sur des chaînes de caractères, de saisir les liens entre les membres de la troupe et d'analyser l'évolution de celle-ci au cours d'une dizaine d'années. A une échelle plus vaste, des travaux similaires sont en cours pour suivre les itinéraires des réfugiés huguenots après la Révocation (Michelle Magdelaine), ou bien pour reconstituer les biographies et les réseaux d'influence de groupes socio-professionnels.

L'éventail des possibilités apparaît donc très vaste, à la mesure des centres d'intérêt des chercheurs et de la souplesse croissante de l'outil informatique. Néanmoins, il serait excessif de prétendre qu'avec l'emploi de l'ordinateur nous avons aperçu ou découvert de nouveaux objets. L'évolution de la science historique, dont j'ai esquissé les grandes lignes, nous montre que le grand tournant de cette discipline, qui intervient entre les années 1930 et 1950, est pris et négocié bien avant l'introduction de l'informatique dans les sciences humaines. Cette introduction a accentué ou facilité des directions de recherche, mais elle n'a pas provoqué de rupture épistémologique. Toutefois, l'usage de l'ordinateur, quelle que soit sa taille, ne représente pas un phénomène aussi neutre que la manipulation d'un photocopieur ou d'un lecteur de microfilm. Il exerce une influence sur nos méthodes de travail, parce qu'il transforme notre approche de la documentation. En changeant de support, l'information programée acquiert une plasticité à la fois formelle et sémantique, qui permet d'expérimenter son traitement ou son exploitation. Le gain apparaît donc essentiellement d'ordre méthodologique, et il est partagé avec toutes les autres disciplines des sciences de l'homme et de la société.

BIBLIOGRAPHIE

- Andréani (R.), Heffer (J.), *La nouvelle histoire économique*, Paris, Gallimard, 1977.
Bardet (J.-P.) *Rouen aux XVII^e et XVIII^e siècles. Les mutations d'un espace social*, Paris, SEDES, 1983, 2 vol.
Béaur (G.), *Le marché foncier à la veille de la Révolution*, Paris, Editions de l'Ecole des hautes études en sciences sociales, 1984.
Cousin (H.), *Le miracle et le quotidien. Les ex-voto provençaux. Images d'une société*, Aix-en-Provence, université de Provence, 1983.
Couturier (M.), *Recherches sur les structures sociales de Châteaudun*, Paris, SEVPEN, 1969.
Fogel (R.W.), Engerman (S.L.), *Time on the cross*, Boston-Toronto, 1974, 2 vol.
Heffer (J.), Robert (J.-L.), Saly (P.), *Outils statistiques pour les historiens*, Paris, Publications de la Sorbonne, 1981.
Herlihy (D.), Klapisch (C.), *Les tisserands et leurs familles*, Paris, Presses de la Fondation nationale des sciences politiques, 1978.
Maurin (J.), *Armée, guerre, société : soldats laqueudois. 1889-1919*, Paris, Publications de la Sorbonne, 1982.
Pind (J.-L.), *Espace social et espace politique. Lyon à l'époque du Front populaire*, Lyon, Presses universitaires de Lyon, 1980.
Prost (A.), *Le vocabulaire des proclamations électorales de 1881, 1885, 1889*, Paris, PUF, 1974.
Zysberg (A.), « Galères et galériens en France à la fin du XVII^e siècle », in *Criminal Justice History*, New York, vol. 1, 1980.

Un choix difficile :
dépouillement exhaustif ou
sondage ?

On peut aussi utiliser
l'ordinateur pour reconstituer
des biographies ou étudier des
réseaux socio-professionnels.

Le recours à l'informatique
induit avant tout une remise
en cause des méthodes de
travail.

Un who's who en histoire

Régine FERRE

Qui est qui? Qui fait quoi? La communauté scientifique se pose souvent ces questions devant le nombre croissant de spécialistes et la complexité de leurs travaux. Pour y répondre, la direction scientifique des sciences de l'homme du CNRS a décidé, en 1981, de confier à l'Institut d'histoire moderne et contemporaine (IHMC) le soin de réaliser un instrument de consultation permettant à la fois de repérer les historiens français et de connaître l'état de leurs recherches, notamment par leurs publications. Un double objectif fut dès lors poursuivi: constitution d'un fichier traité par informatique et publication du *Repertoire des historiens français pour la période moderne et contemporaine*. Cette enquête supposait la collecte d'informations homogènes à partir d'un questionnaire adressé à une population aussi large que possible: historiens - y compris ceux appartenant à d'autres disciplines, droit, littérature, sciences, art, etc. - en activité au CNRS, dans les universités, les Grands établissements littéraires et scientifiques, aux archives, dans les bibliothèques et les musées. Sur deux mille questionnaires envoyés, mille deux cent soixante-sept réponses sont parvenues à l'IHMC, soit 63%, pourcentage encourageant car nombreux sont ceux qui reculent devant le fait d'avoir à remplir - encore - un questionnaire. Après le dépouillement systématique des questionnaires (opération très longue en raison des inexactitudes et imprécisions dans

les réponses à la rubrique "publications", les informations recueillies ont été traitées sur micro-ordinateur, avec l'agrément de la Commission nationale de l'informatique et des libertés (CNIL). Le fichier, ainsi constitué, est devenu une base de données interrogeable au CDSH. Une liste des notices individuelles et trois index (thématique, chronologique et géographique) ont été extraits de cette base pour être intégrés au *Repertoire* (voir exemple de notice individuelle). Ainsi sortis, parallèlement, deux éléments complémentaires d'un même outil de travail, chacun d'eux ayant sa propre spécificité: la base de données permettant des recherches plus « ciblées », le livre restant d'un accès traditionnel.

Sous ses deux aspects, ce reflet de la recherche historique en France, à une période donnée, a reçu un accueil très favorable de la part des chercheurs, des universitaires et des étudiants. Une remise à jour, qui étendra cette première enquête aux enseignants-chercheurs de l'enseignement secondaire, est prévue pour 1986.

* Faugères (A.), Ferré (R.), *Repertoire des historiens français pour la période moderne et contemporaine*, Paris, Editions du CNRS, 1983, 360 p.

Les géographes avaient déjà réalisé leur propre repertoire: *Repertoire des géographes français*, Paris, Laboratoire Interpéo du CNRS, 1984, 5^e édition, 313 p. (1070 notices).

CROUZET François (1922)

6 rue

ML

Professeur d'histoire de l'Europe du Nord

Université de Paris IV, 1 rue Victor Cochin, 75230 Paris Cedex 05

Centre de recherches sur la civilisation de l'Europe moderne.

Comité de rédaction de la Revue historique; Bureau de l'Association française des historiens économistes.

TE: L'économie britannique et le Blocus continental, 1806-1813 (Paris, 1956) Paris, 1956.

RC: Les industriels britanniques à l'époque de la révolution industrielle.

The sources of England's wealth: some French views in the 18th century, in Shipping, Trade and Commerce, Essays in Memory of Ralph Davis, Leicester, Leicester University Press, 1981, p.61-79. Sur les hiérarchies sociales et Grande-Bretagne au XVIII^e siècle, in Mélanges R. Gaxotte, Lyon, P.U.L., vol.1, 1980, p.119-129.

The Social Background of Industrialization during the Industrial Revolution in Britain, History Journal Middlesex Polytechnic, juin 1980, no.1, p.1-26.

Towards an export economy: British Exports during the Industrial Revolution, Explorations in Economic History, 1981, no.17.1, p.48-93.

■ Régine Ferré, ingénieur-d'études au CNRS, documentaliste à l'Institut d'histoire moderne et contemporaine (IHMC), LP 071, du 2 mai 1983 au 21 mai 1985.

Exemple de notice individuelle.

L'informatique, objet de recherche en sciences de l'homme et de la société

Le Centre de documentation sciences humaines (CDSH), sous la responsabilité du département sciences de l'homme et de la société et de la direction de l'information scientifique et technique, a constitué et tient à jour annuellement la base CNRS-SHS qui décrit les unités de recherche financées par le CNRS dans le domaine des sciences de l'homme et de la société (SHS).

Question: quelles sont les unités SHS liées à la section 08 (informatique, automatique, signaux et systèmes)?

Réponse: 17 unités.

Langage et systèmes informatiques - J. Vignolle (UA 347)
Laboratoire de la communication parlée - R. Carro (UA 388)
Centre de recherche de mathématique de la décision (CEREMADE) - L. Eliezer (UA 749)

Laboratoire d'automatique documentaire et linguistique - M. Gros (UA 819)

Cybernétique des entreprises, reconnaissance des formes, intelligence artificielle (CERFA) - S. Clastan (UA 824)

Laboratoire d'analyse et modélisation de systèmes pour l'aide à la décision (LAMSAD) - B. Roy (UA 825)

Groupe d'études pour la traduction automatique (GETA) - C. Bollet (UA 826)

Groupe de recherche en analyse de système et essai économique (GRASCE) - B. Munier (UA 933)

Institut de recherches et d'études pour le traitement de l'information juridique (IRETJ) - M. Vivant (UA 954)

Laboratoire de linguistique formelle - S. Cukier (UA 1026)

Centre d'analyse et de traitement automatique de la Bible et des traditions écrites (CATB) - G.-E. Weil (UA 1071)

Groupe représentation et traitement des connaissances (GRTCI) - E. Chouard (IER 270)

Théorie générale du rythme et ses applications (PRMT) - J. Rouleau, P. Lussion (RCP 700)

ELECTRE: un système expert d'enseignement basé sur le fonctionnement cognitif de l'élève - E. Courrière-Marmèche, J.-L. Laurène (RCP 768)

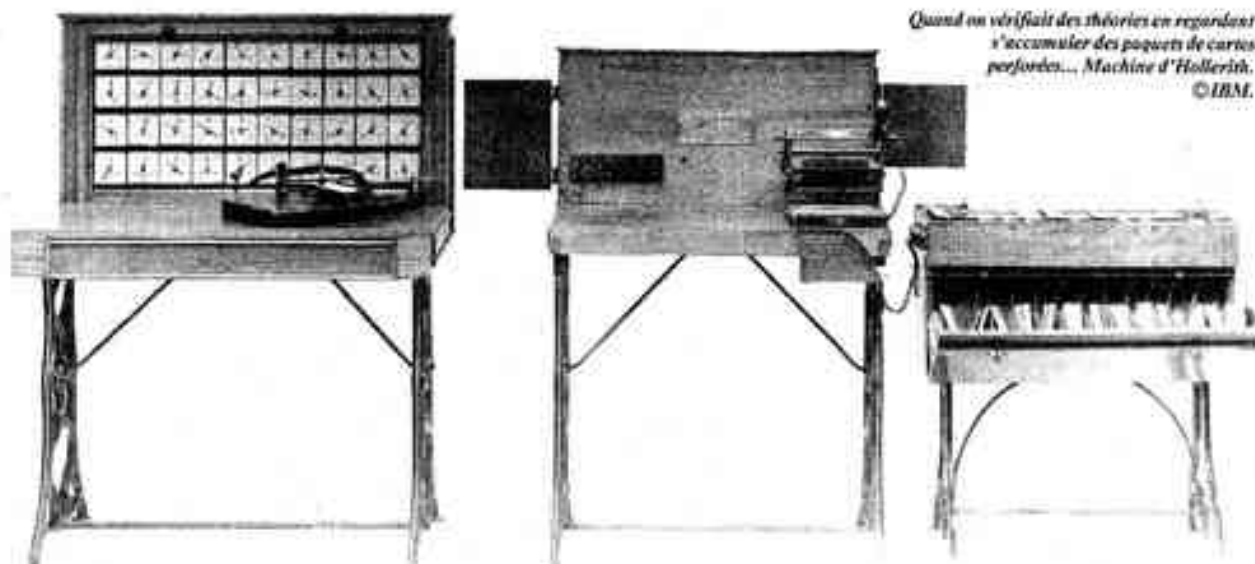
Didactique et acquisition des connaissances scientifiques - G. Vergnaud, G. Brousseau, M. Hulin (GRECO 71)

Centre d'analyse et de mathématique sociale (CAMS) - M. Barbut (IAM 17)

Action de recherche intégrée Sciences de la communication - D. Walton (ARI)

Pour toute précision, veuillez contacter le Centre de documentation sciences humaines (CDSH), service recherches en cours, 54, boulevard Raspail, 75006 Paris. Tél.: (1) 45.44.38.49, p. 357

* Les laboratoires utilisant ordinateur comme outil ne sont pas pris en compte.



Quand on vérifiait des théories en regardant s'accumuler des paquets de cartes perforées... Machine d'Hollerith. © IBM.

Le sociologue et ses données

L'analyse des données, et notamment l'analyse factorielle, a été l'acquis principal de l'ordinateur en sociologie. Elle permet de mieux « voir » les données, reste à vérifier les interprétations qu'elle suggère.

Philippe CIBOIS

L'informatique n'a pas révolutionné la sociologie.

Le dépouillement d'enquête avait déjà connu un développement important du fait de la mécanographie.

L'analyse des données est une manière de faire en réaction contre une pratique ritualisée de la statistique.

Disons-le franchement : l'informatique n'a pas révolutionné la sociologie. Mais a-t-elle modifié profondément la zoologie ? Pourquoi cette attente d'un *deus ex machina* sinon parce qu'on espère qu'il viendra donner de la rigueur à un ensemble de savoirs que l'on estime, sans trop oser le dire ouvertement, trop « mou », trop lié à la mode, aux chapelles.

Si l'on ne peut pas parler de révolution apportée par l'ordinateur, on doit cependant noter comme partout son impact qui a modifié certaines pratiques, comme en particulier le dépouillement d'enquête. Ce n'est pas la première fois que l'évolution technique modifie ce champ de recherches : ce fut le cas en particulier du fait de la mécanographie depuis le dernier recensement du XIX^e siècle aux États-Unis grâce à un certain M. Hollerith, futur fondateur de l'International Business Machine Corp., destinée à la croissance que l'on sait. Sans cette technique, la sociologie quantitative américaine n'aurait pas vu le jour. Dans les années soixante encore, la trieuse était l'instrument quantitatif par excellence et certains conservent encore la nostalgie de l'époque où ils voyaient leurs théories se vérifier, simplement en considérant la croissance rapide de paquets de cartes perforées dans les cases prévues.

■ Philippe Cibois, ingénieur d'études au CNRS, Laboratoire d'informatique pour les sciences de l'homme (LIH), LP 813, 54, boulevard Raspail, 75006 Paris.

En France, au moins, qui dit informatique en sociologie pense enquête et « analyse des données ». C'est sur ce point que nous voudrions insister.

« Voir » ses données

Le terme d'« analyse des données » nous vient de l'anglais où *data analysis* désigne, depuis une vingtaine d'années, une certaine manière de pratiquer la statistique qui consiste à regarder les données pour elles-mêmes, en mettant – au moins un moment – entre parenthèses les théories qui ont conduit à leur recueil. Cette attitude s'est opposée à la pratique statistique, alors dominante, qui voulait que les données soient soumises à des tests de « signification » qui engendrent une décision statistique (rejet ou acceptation d'une hypothèse d'indépendance, par exemple).

Cette manière de voir les données a été très facilitée par l'emploi de l'ordinateur. Celui-ci, en effet, a libéré le chercheur de calculs longs et fastidieux, mais lui a apporté surtout des possibilités de « voir » ses données par le biais d'une représentation graphique.

Ce désir de « voir » ses données conduira aux techniques anglo-saxonnes de *Multidimensional scaling* et, en France, à la technique de l'analyse factorielle des correspondances mise au point par J.-P. Benzécri dans les années

L'analyse des données cherche à observer sans a priori les données rassemblées.

Mais disposer d'un outil pour voir ne suffit pas : il faut être disposé à voir.

Une théorie bien ancrée peut empêcher de voir.

Pour qu'un observateur remette en cause une théorie bien admise, il faut qu'il ait confiance dans son instrument d'observation.

Dépouiller une enquête, c'est prendre en compte l'ensemble des données pour en discerner les oppositions fondamentales.

L'analyse factorielle est une méthode qui sert à explorer les données. Les résultats obtenus doivent ensuite être vérifiés.

soixante. Ce dernier, ainsi que les « analystes de données » qui emploient ses méthodes, prennent le terme d'analyse des données dans un sens beaucoup plus radical et même polémique. Pour J.-P. Benzecri, en particulier, l'analyse des données s'oppose aux pratiques des sciences humaines dans le sens où l'on refuse les *a priori* de ces disciplines, c'est-à-dire où l'on refuse de considérer que les « données » ne sont qu'une théorie mise en actes mais où l'on considère au contraire que les données nous sont, en quelque sorte, préexistantes, que nous n'avons qu'à les recueillir et non à les construire.

Il va de soi qu'un sociologue ne peut partager ce point de vue. Il sait que les données qu'il utilise sont le fruit d'une production sociale, consciente ou non. Les catégories employées dans les recueils statistiques produits par l'administration sont le reflet d'une vision sociale des rapports sociaux, on doit en tenir compte dans leur usage. Cependant, même s'il produit lui-même ses données, le sociologue sait bien qu'il introduit sa propre théorie dans les « faits » qu'il observe.

Cela dit, il n'empêche que le sociologue dispose maintenant d'un outil qui lui permet de voir ses données et qu'il se retrouve dans la même situation que l'astronome après l'invention de la lunette de Galilée ou le biologiste après celle du microscope.

L'instrument d'observation ne suffit pas : pour « voir », il faut être « disposé » à voir, c'est-à-dire qu'il ne faut pas que les préoccupations théoriques bloquent la perception du message. Nous nous retrouvons souvent dans la même situation que ces vingt-quatre professeurs de toutes les universités qui, du 24 au 25 avril 1610, firent des essais de la lunette de Galilée en présence de celui-ci. Le rapport final note que sur des objectifs terrestres l'instrument fonctionne admirablement bien et permet de lire des inscriptions sur des frontons d'églises à des lieues de distance. En revanche, braqué vers le ciel, il nous abuse. Ainsi, dirigé vers une étoile comme celle de la Vierge, celle-ci paraît double, ce qui est une preuve manifeste de l'inadéquation de l'instrument, puisque l'âme de la Vierge ne peut être double.

Comme nous ne partageons plus la théorie platonicienne des étoiles comme âmes du monde, une telle mise en cause de l'instrument nous paraît incroyable. Il n'est pas certain que d'ici quelques décennies, nous ne portions pas le même regard sur les usages actuels de l'analyse factorielle qui nous paraîtront viciés par nos théories du social.

Si les outils d'observation liés à l'utilisation de l'ordinateur sont disponibles, ils ne sont pas forcément bien utilisés : pour qu'ils le soient correctement, il importe de bien connaître l'instrument, de l'avoir en quelque sorte « étalonné » sur des données connues pour voir comment il traduit la réalité en langage factoriel. Ayant acquis expérience et confiance dans l'instrument, l'observateur acceptera ensuite de se faire remettre en cause par l'observation qu'il ne se sentira plus le droit de nier.

Les possibilités de l'outil

Dépouiller une enquête aujourd'hui, c'est sortir de ce genre littéraire particulièrement sophistique qu'est le « commentaire de tableaux croisés » et qui consiste à « croiser » le sexe, puis la catégorie socio-professionnelle, puis la tendance politique avec les différentes questions qui sont l'objet propre de l'enquête. De tels commentaires, par leur répétition, sont liés à une technologie, celle de la mécanographie qui ne

Laboratoire d'informatique pour les sciences de l'homme (LISH)

Ce laboratoire propre du CNRS, installé à la Maison des sciences de l'homme du boulevard Raspail à Paris, a pour fonction de permettre aux chercheurs de toutes les disciplines de sciences humaines l'utilisation de l'informatique dans leurs recherches.

Pour arriver à ce résultat, le LISH met à la disposition des chercheurs divers matériels : des consoles reliées au centre de calcul du CNRS, le CIRCE; des micro-ordinateurs, un lecteur optique de textes imprimés, des imprimantes et un traceur. Ces matériels sont utilisables par les chercheurs en libre-service.

Comme l'utilisation de ces matériels et des méthodes qu'ils impliquent suppose des connaissances dont ne disposent pas tous les chercheurs en sciences humaines, une grande partie des activités du LISH consiste en formation par le biais de nombreux stages et en information par l'intermédiaire de bulletins spécialisés dans diverses disciplines.

Contacts : directeur, Michael Hainsworth; secrétaire scientifique, Ugo Moonior.
Tél. (1) 42.22.43.25.

permettait de faire que des tableaux croisés, et non à un objectif qui doit être de discerner dans une population des types distincts, en fonction des réponses à l'ensemble des questions. Pour arriver à cette vision globale, il faut pouvoir prendre en compte l'ensemble de l'information, ce que peut faire l'analyse factorielle qui va examiner cet ensemble et en déterminer les oppositions les plus fondamentales.

L'analyse factorielle doit donc se situer en début de recherche : c'est une technique d'observation, heuristique en ce sens, qu'avec elle, on ne cherche pas à prouver quoi que ce soit, mais que l'on observe le matériel réuni. En revanche, quand des réponses ont été repérées comme étant associées à des caractéristiques du répondant, il est nécessaire de revenir aux données et de prouver la réalité des types sociaux que l'on croit pouvoir repérer. En effet, l'analyse factorielle étant une technique qui recherche les lignes d'ensemble des données, les généralisations auxquelles elle procède doivent être vérifiées quand on passe du général au particulier. En résumé, l'analyse factorielle sert d'échafaudage à la recherche : quand celle-ci est appuyée sur des résultats complémentaires, il n'est plus nécessaire d'exhiber les préliminaires. On démonte l'échafaudage et on n'est même pas forcé d'en faire état car l'exposition des résultats ne doit pas avoir grand chose à voir avec le journal de bord d'une recherche.

L'analyse factorielle est, pour le dépouillement d'enquête, un outil puissant qui permet au chercheur de prendre en compte la totalité du matériel qu'il a recueilli. Acceptera-t-il de remettre en cause ses hypothèses de départ au vu de l'observation ? Pour que cela soit possible, il faut qu'il dépasse la mode, les mythes et qu'il accepte d'apprendre le fonctionnement de la méthode, qu'il en comprenne les concepts fondamentaux, que finalement, bon artisan, il s'approprie son outil.

Ordinateur et terrain en Thaïlande

Un cas exemplaire de l'intégration de l'ordinateur dans une pratique de recherche : analyse factorielle et télédétection en Thaïlande.

Michel BRUNEAU

A l'inverse de leurs homologues anglophones, les géographes « tropicalistes » français ont jusqu'à présent peu utilisé les analyses multivariées appliquées à des enquêtes statistiques ou à des images de télédétection.

Les géographes « tropicalistes » français se sont toujours beaucoup méfiés des statistiques, arguant de la faiblesse et du caractère peu fiable de celles collectées par les États des pays en voie de développement. Ils plaçaient le travail de terrain au centre de leur démarche, ce qui supposait une lente imprégnation par les paysages, le contact avec les hommes au cours d'un séjour prolongé (plusieurs années). Cela impliquait un grand nombre d'enquêtes qualitatives et de parcours. L'objet d'étude, un pays ou une région assez vaste (plusieurs milliers voire centaines de milliers de km²), ne favorisait pas, comme en ethnologie, l'insertion dans une communauté et un contact aussi étroit avec une société paysanne. Dans les années soixante, le besoin de données chiffrées s'est fait sentir pour asseoir la crédibilité scientifique de la géographie auprès des aménageurs et planificateurs du développement. Cette géographie, qui était alors surtout africaniste, s'était lancée dans une cartographie à grande échelle des terroirs villageois, élaborée à l'aide de photographies aériennes et de levés sur le terrain. Le traitement graphique des données (fichiers-images) mis au point par l'équipe de J. Bortin était également utilisé, mais les possibilités offertes par l'analyse des données (analyses multivariées factorielles ou classification automatique), impliquant le recours à l'ordinateur, n'avaient pas été explorées par les « tropicalistes » français à l'inverse des chercheurs anglophones qui y avaient

eu largement recours pour ce que certains ont appelé la géographie du développement.

Enquêtes en milieu rural et analyse des correspondances des systèmes agraires à l'échelle villageoise

Ne disposant pas encore en 1967 de données statistiques systématiques à un niveau de divisions administratives suffisamment fines (canton ou district) pour l'étude géographique d'une région de taille moyenne (le Nord de la Thaïlande), j'ai pris le parti de recueillir moi-même mes données sur le terrain. J'ai donc procédé au choix raisonné d'une dizaine de villages en fonction du milieu écologique dans lequel ils se trouvaient et de leur niveau de développement (ayant déjà acquis depuis trois ans une connaissance qualitative de ce terrain). Cet échantillon était à la mesure des moyens dont je disposais : trois à quatre enquêteurs, étudiants de l'université de Chiang Mai et un véhicule.

La perspective d'utiliser un traitement de type analyse multivariée me permettait de choisir un grand nombre de variables (89 en tout) dans un questionnaire visant à recueillir des données démographiques mais également concernant les techniques de production, la terre, les activités non-agricoles, le bétail, les équipements, la commercialisation de la production, l'endettement, le niveau d'instruction et les obligations sociales. Ces données ont été recueillies auprès de chacune des unités socio-économiques de base, (les maisonnées), constituant un village, dont le nombre dans chacun de ces villages se situait autour de la centaine. L'outil informatique, l'analyse factorielle des correspondances, a été utilisé en 1975-76 pour le traitement de ces données villageoises du Nord de la Thaïlande grâce à une collaboration avec le Laboratoire de statistique de l'université Paris VI et avec la Maison des sciences de l'homme d'Aquitaine (O. Bouchta).

Cette démarche statistique utilisant l'ordinateur a eu des répercussions sur mon travail de terrain. Le recueil systématique de données sur questionnaire dans l'ensemble d'un village m'obligeait à un séjour prolongé (une dizaine de jours au moins et des retours périodiques pour vérifications et compléments) et à vivre pendant la durée de l'enquête dans le village en contact permanent avec ses habitants. Cela a permis ainsi, parallèlement à l'acquisition de données traitables statistiquement, une meilleure observation qualitative du fait même de l'immersion dans la société villageoise. Le travail de terrain a gagné en efficacité et en profondeur, à cause précisément des contraintes matérielles imposées par le recueil des données et à cause de la variété de celles-ci.

Ces analyses des correspondances effectuées pour chacun des villages ont fait ressortir les 3,

■ Michel Bruneau, directeur de recherche au CNRS, Centre d'études de géographie tropicale (CEGET), LP 8691, Domaine universitaire de Bordeaux, 33405 Talence cedex.



Enquête sur une maisonnée dans le village de Thap Pung (province de Sakon Nakhon).

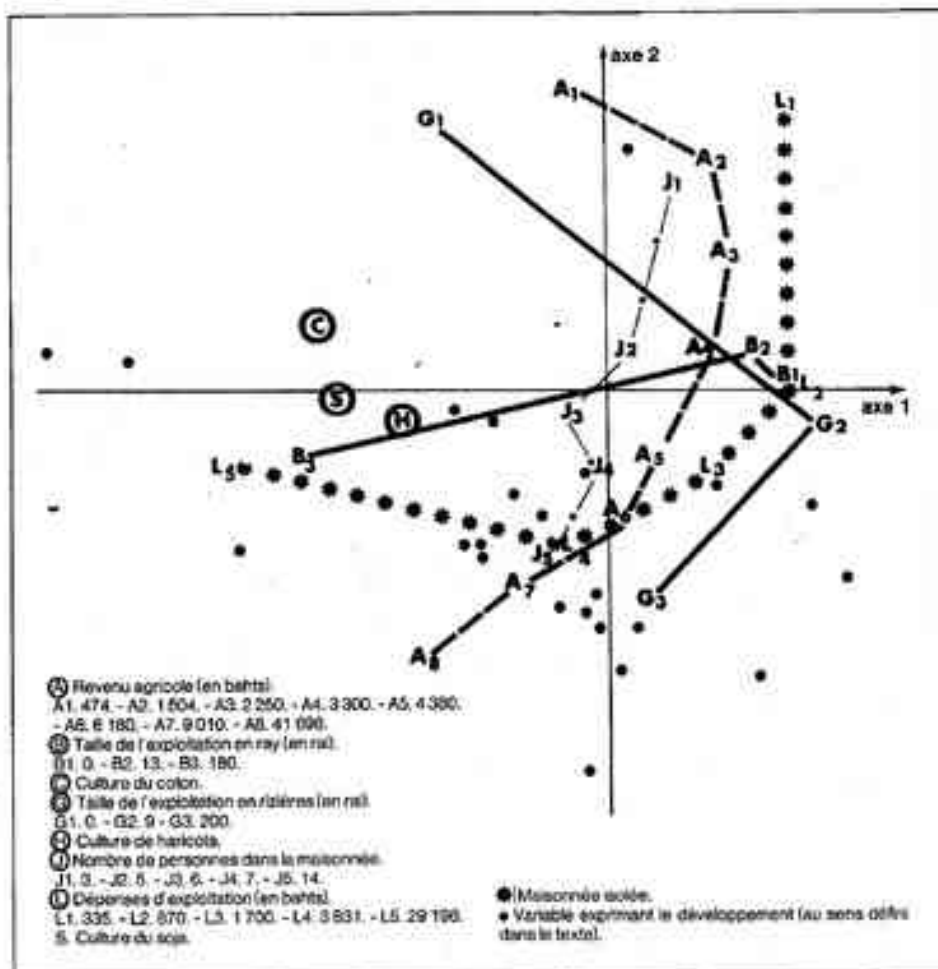
L'analyse des correspondances de dix systèmes agraires villageois a permis de vérifier statistiquement une hypothèse (la liaison très forte entre un développement basé sur la modernisation technique et une différenciation sociale de plus en plus accentuée) et de conforter le travail de terrain.

4 ou 5 variables qui différencient le mieux la population des maisonnettes étudiées. Les revenus agricoles et/ou non-agricoles, la taille des exploitations et/ou des propriétés en rizières et en cultures pluviales commerciales, l'âge du chef de famille, le nombre de buffles possédés, le nombre de personnes dans la maisonnette, les dépenses d'exploitation, les superficies consacrées à telle ou telle culture commerciale sont, sur les 89 variables traitées, celles qui se sont révélées les plus structurantes des populations étudiées.

J'ai pu ainsi vérifier statistiquement une hypothèse de travail qui pouvait être formulée à l'aide d'une connaissance qualitative: la liaison très forte qui existe entre un type de développement basé sur la modernisation technique et une différenciation sociale de plus en plus accentuée au sein des communautés villageoises. De même, l'importance des deux variables liées, âge du chef de maisonnette et nombre de personnes qui en font partie, apparaissait surtout dans

les milieux de densité de population plus faible (plaine de Sukhothai) et d'occupation plus récente, alors qu'elle était beaucoup moins grande dans les milieux très densément peuplés et d'occupation très ancienne (bassin de Chiang Mai par exemple).

L'analyse des correspondances portant sur l'ensemble des maisonnettes (854), appartenant aux dix villages, a montré une cohésion de ces villages: les maisonnettes de l'un s'y regroupaient dans un espace distinct de celui des autres au long de l'axe 1 (facteurs 1, degré de commercialisation de l'agriculture et 2, différenciation sociale) comme sur celui des axes 1 et 3 (types d'activités) voir plan factoriel ci-dessous. De même, la typologie des systèmes agraires à riziculture dominante par opposition à ceux à cultures pluviales commerciales dominantes (fronts pionniers) a été vérifiée avec la distinction très nette entre un bloc de quatre villages de riziculteurs et un autre bloc de quatre villages où les cultures pluviales commerciales l'empor-



Analyses des correspondances sur les variables dans les dix villages réunis.

Ce traitement a porté sur 854 maisonnettes caractérisées par 89 variables et appartenant à dix villages différents. Le premier axe représente le degré de commercialisation de l'agriculture. Il oppose une riziculture de petites exploitations avec un fort secteur d'autosubsistance à une agriculture commerciale s'étendant de plus grosses exploitations totalement dépendantes du marché. Le second axe est celui du niveau de vie, les variables revenus, dépenses, investissements ayant fortement contribué à sa formation. Une autre variable, le nombre de personnes par maisonnette, est bien corrélée avec cet axe. A ses deux extrémités s'opposent les plus grosses exploitations en rizières qui sont aussi des maisonnettes nombreuses, disposant d'une main-d'œuvre familiale importante, et les maisonnettes les moins nombreuses, sans terre. Le type d'activité et la différenciation sociale sont donc les deux critères dominants pour établir une typologie des maisonnettes dans les dix villages.

On peut également prendre une vue d'ensemble du plan constitué par ces deux premiers facteurs en opposant les quadrants deux à deux. En haut et à droite, figurent les maisonnettes n'ayant que peu de membres (1 à 3 personnes), pas ou peu de terres et vivant surtout de salaires de journaliers. A ce groupe s'opposent, en bas et à gauche, les maisonnettes nombreuses (8 à 14 personnes) ayant une grosse exploitation en faire-valoir direct sur brûlis (cultures pluviales commerciales), des revenus agricoles élevés et des dépenses importantes. Dans les deux autres quadrants s'opposent de grosses exploitations et propriétés en rizières bien équipées, rendant du riz en bas et à droite, et de très petites exploitations sur brûlis (cultures pluviales) sans rizières, devant acheter leur propre consommation de riz en haut et à gauche.

taient presque exclusivement. Entre ces deux types bien contrastés apparaissait un type mixte intermédiaire (2 villages).

Les techniques d'analyse des données sur les communautés rurales ont donc conforté mon travail de terrain dans la phase de recueil de ces données. Elles m'ont permis de préciser mes hypothèses de travail et de les étayer au cours de la phase de traitement. On a pu ainsi dégager les variables les plus significatives pour l'interprétation des inégalités du développement dans l'espace géographique. Une connaissance qualitative préalable et approfondie du terrain était nécessaire tant pour l'établissement des questionnaires et le choix des variables pertinentes, que pour le choix des villages étudiés. L'analyse des données est donc venue confirmer et préciser une démarche qu'elle n'a pas fondamentalement bouleversée. En revanche, la télédétection, utilisant des images satellitaires (Landsat), a plus profondément transformé mes rapports avec le terrain, avec les autres disciplines et avec la coopération scientifique et technique.

Les images de télédétection spatiale ou l'approche cartographique des systèmes agraires (en coopération)

Dans mes travaux sur l'organisation de l'espace en Thaïlande, j'avais largement utilisé la photo-interprétation à partir de séries chronologiques de photographies aériennes à diverses échelles en vue d'une cartographie de la dynamique des paysages ruraux. J'étais donc dès le lancement du premier satellite d'observation de la terre, le satellite américain Landsat 1 en 1972, tout à fait intéressé par la télédétection spatiale. La Thaïlande a été avec les Philippines, puis l'Indonésie, l'un des premiers pays d'Asie du Sud-Est à s'associer au programme de la NASA. Elle a conservé, depuis, son avance relative dans ce domaine, en construisant notamment la première station de réception de ces satellites en Asie du Sud-Est (Bangkok, 1981).

Dans un premier temps (1973-1976), j'ai utilisé les images Landsat sous leur forme analogique et me suis livré à des photo-interprétations sur des agrandissements au 1/500000 des canaux 5 et 7 et de couleurs composites (4, 5, 7). Leur échelle relativement petite me permettait de couvrir des superficies assez vastes, à l'échelle régionale, et de compléter et généraliser la cartographie à plus grande échelle que j'avais faite à l'aide de photographies aériennes. Des essais au Centre d'études spatiales des rayonnements (Toulon) et au Centre national d'études spatiales, CNES (Service traitement de l'image) m'avaient initié à l'utilisation des traitements informatiques sur des portions d'image encore limitées (500 points ou pixels en ligne sur 500 en colonnes, alors qu'une image Landsat comprend en tout $2\,983 \times 3\,240$ pixels).

Alors qu'il n'était possible de disposer que de rares couvertures photographiques aériennes à des dates distantes de plusieurs années (voire dizaines d'années), la télédétection permettait une couverture périodique régulière rendant possible des comparaisons entre saisons différentes et entre années plus ou moins éloignées. Le recours au traitement informatique d'images devenait nécessaire face à l'abondance et à la répétitivité des données. Dans le même temps, les systèmes de traitements interactifs d'images se développaient et il devenait possible de travailler sur des superficies de plus en plus vastes, pouvant aller jusqu'à la totalité d'une image Landsat ($185\text{ km} \times 185\text{ km}$).

La France prévoyant alors le lancement du satellite SPOT d'observation de la terre lancé en

février 1986) qui était le premier après celui des États-Unis (Landsat), une politique de coopération dans le domaine de la télédétection avec la Thaïlande, pays leader en Asie du Sud-Est, devenait nécessaire dans les années quatre-vingts. Avec le soutien de l'action thématique programmée CNES-CNRS, et celui du ministère des Relations extérieures, un premier programme de coopération avec le National Research Council (Division télédétection) et le département de géographie de l'université de Chiang Mai avait été lancé en 1979 sur une cartographie multitemporelle de l'occupation du sol du bassin de Chiang Mai. Le but était de tester, sur un terrain que nous connaissions bien, les traitements numériques d'images Landsat en vue d'une cartographie thématique et de former des enseignants et des chercheurs thaïlandais à ces traitements qui n'étaient pas encore pratiqués en Thaïlande.

Les responsables thaïlandais de la recherche m'ont ensuite orienté vers une demande plus précise de cartographie agro-écologique formulée par la faculté d'agronomie (département de sciences des sols) de l'université de Khon Kaen, située au cœur de la région Nord-Est, la plus pauvre du royaume. Il ne s'agissait plus d'une recherche uniquement fondamentale, sans application précise comme précédemment, mais il fallait répondre au besoin d'une cartographie extensive, couvrant une vaste région, cartographie qui devait être utile au développement agricole de cette région particulièrement sous-développée. Le recours à un chercheur, morpho-pédologue expérimenté, de l'Institut de recherche en agronomie tropicale de Montpellier (J. Kilián) et au département de télédétection de l'Institut géographique national (IGN) de Paris (H. Le Men), s'est alors révélé la meilleure solution en France pour répondre à ce type de demande. L'IGN avait en effet développé un système interactif de traitement d'images (TRIAS) et de restitution cartographique des résultats (système SEMIO) extrêmement performants et il était le seul en France à pouvoir couvrir de grandes surfaces (une ou plusieurs images Landsat).

Les données de la télédétection spatiale Landsat reflètent la superficie terrestre dans le spectre du visible et du proche infrarouge. Elles fournissent donc des informations sur tout ce qui s'y trouve et réfléchit ou absorbe les rayons du soleil. Face à la complexité et à la densité de ces informations qui sont données par les valeurs affectées aux points élémentaires de résolution ou pixels, il faut effectuer une première simplification ou classement des groupes de points de valeurs radiométriques voisines sur l'ensemble de l'image. Des classes ou états de surface sont ainsi définis par référence au terrain. Le travail de préparation sur le terrain doit être fait de la façon la plus rigoureuse possible à partir d'une première visualisation de l'image sur laquelle sont choisies des parcelles d'entraînement ou zones échantillons, observées ensuite directement sur le terrain.

Les états de surface, qu'on est ainsi amené à définir, reflètent à la fois le milieu naturel plus ou moins modifié par les hommes et les installations humaines proprement dites (aménagements, parcelles, habitats...). Cette réalité complexe, constituée par la totalité de ce qui est observable à la surface de la terre à différentes saisons, c'est-à-dire dans différentes conditions d'humidité et de végétation, nécessite pour être appréhendée dans sa globalité une approche interdisciplinaire. Il est impossible de séparer rigoureusement ce qui est d'origine humaine et naturelle. On ne peut pas maintenir une sépara-

Des traitements d'images satellitaires en collaboration avec l'Institut géographique national (IGN) et le Centre de coopération internationale en recherche agronomique pour le développement (CIRAD) ont été effectués à la demande de la Thaïlande en vue de l'aménagement d'une région pauvre, le Nord-Est.

La notion de paysage agro-écologique, qui combine des phénomènes morpho-pédologiques, agricoles et d'occupation humaine au sens large, favorise l'approche interdisciplinaire nécessaire à la production de cartes utiles pour le développement rural. La télédétection permet un suivi des variations saisonnières et des évolutions pluri-annuelles de ces phénomènes de surface.

tion trop forte entre sciences de la nature et sciences sociales, entre géographie physique et géographie humaine qui, au cours des dernières décennies, tendaient à s'éloigner l'une de l'autre. Il faut dans l'interdisciplinarité reconstruire une vision globale de la surface terrestre pour interpréter correctement ces images. On est obligé de concevoir de nouveaux types de cartographie qui combinent souvent des phénomènes d'origines différentes. C'est ainsi qu'on est amené à réintroduire, en la repensant, la notion de paysage et, dans le cas qui est cité ici, la notion de paysage agro-écologique, c'est-à-dire qui associe des phénomènes du milieu biophysique, de son utilisation agricole et de son occupation par les hommes. Les états de surface se combinent par analyse de texture en unités de paysages agro-écologiques qui expriment à la fois des données morpho-pédologiques, agricoles et d'occupation humaine au sens large. Ces paysages agro-écologiques se regroupent eux-mêmes (par association de couleurs par exemple) en « ensembles régionaux » qui sont la traduction dans l'espace des systèmes agraires. Je rejoins ici les phénomènes que j'avais abordés plus haut par l'analyse multivariée de variables socio-économiques à l'échelle du village.

On a pu ainsi établir une série de cinq cartes en couleurs (voir p. 3 et 4 de couverture) permettant de suivre sur une région de 18 700 km², au cours d'une période de dix ans (1972-1982), les ressources en eau, que ce soit sous la forme d'eau libre (lacs, zones inondées) ou sous la forme d'humidité superficielle des sols (variations saisonnières), ainsi que la couverture arborée, tant dans ses aspects forestiers qu'agaires. On a pu également aboutir à une régionalisation de l'agriculture de cet espace (Nord-Est de la Thaïlande).

Plus de rigueur et d'interdisciplinarité

De ces deux expériences différentes impliquant le recours à l'ordinateur, seule la seconde a débouché sur un véritable programme de coopération, parce qu'elle correspondait à une demande de la Thaïlande (traitement d'images spatiales en vue d'applications utiles au développement rural) et à la politique spatiale de la France (satellite SPOT). Elles ont cependant toutes les deux contribué à modifier le rapport du géographe que je suis à son terrain et à l'interprétation des phénomènes étudiés sur trois points précis :

- l'outil informatique, que ce soit dans l'analyse de données socio-économiques ou dans le traitement d'images satellitaires, m'a obligé à une plus grande rigueur et précision d'observation pour le recueil des données : définition des variables de l'enquête villageoise, fiche d'observations et éventuellement de mesure des paramètres définissant chaque parcelle d'échantillonnement ou zone-échantillon sur l'image spatiale ;
- les analyses multivariées peuvent prendre en compte un très grand nombre de variables ou d'informations (pixels d'une image), ce qui m'a permis de ne pas trop rapidement simplifier la réalité et de lui conserver plus longtemps sa complexité. Il a ainsi été possible, dans une enquête villageoise, de balayer un large spectre de variables, allant de la démographie aux techniques agricoles et aux relations sociales, pour dégager, à la suite du traitement, lesquelles avaient les plus fortes valeurs interprétatives et rendaient le mieux compte du dynamisme d'une société rurale sans nécessairement occulter les autres variables. L'image classée, ou carte des états de surface, carte par points (pixels), était d'une complexité telle qu'elle était

difficilement lisible par un œil non exercé. Elle devait donc être simplifiée en une carte par plages (carte des paysages agro-écologiques). Chacune de ces plages ou unités spatiales, définie par une combinaison d'états de surface, traduisait un ensemble de phénomènes naturels et sociaux qu'exprimait le concept de système agraire. Ainsi, le réel, après traitements et simplifications successives, conservait tout de même une grande partie de sa complexité et pouvait ainsi faire l'objet d'une interprétation plus riche et nécessairement interdisciplinaire ;

● enfin, l'usage de l'ordinateur exigeait une plus grande rigueur et une plus grande homogénéité dans les interprétations. L'analyse des correspondances, appliquée successivement à chacun des dix villages, puis à leur ensemble, avait permis de déterminer objectivement lesquelles des 89 variables choisies rendaient le mieux compte de la physionomie de chacun des villages et de la société rurale représentée par leur ensemble. De même, la classification de l'image Landsat en états de surface, en fonction d'échantillons observés sur le terrain, puis leur regroupement par analyse des textures en unités de paysage, donnait une interprétation rigoureusement homogène sur l'ensemble de l'image (c'est-à-dire répondait aux mêmes critères objectivement définis). Il était évidemment impossible d'obtenir par photo-interprétation à l'œil humain une telle rigueur dans la définition des critères et dans l'homogénéité de l'interprétation sur la vaste superficie que recouvre une image Landsat.

Mais ce qui reste capital et fonde la valeur de la démarche scientifique du géographe à l'amont comme à l'aval des traitements informatiques, c'est la connaissance qualitative des phénomènes qu'il étudie, connaissance qu'il n'acquiert que par une imprégnation minimum dans et par le milieu. Il n'y a pas de règle générale pour l'acquisition de cette connaissance qui dépend de beaucoup de facteurs propres au chercheur (qualité de son intuition, expérience antérieurement acquise, rapports humains...) et au milieu étudié (certains pouvant être plus rapidement appréhendés que d'autres). C'est cette connaissance préalable qui, dans le cas analysé ici, a conditionné le choix des villages et des variables de l'enquête, ainsi que celui des zones-échantillons de l'image et des paramètres qui y ont été observés. C'est également cette connaissance qualitative qui a guidé le traitement d'image dans sa phase interactive et les interprétations finales qui ont été données aussi bien aux analyses multivariées des systèmes agraires qu'à leur traduction cartographique à l'issue du traitement d'image. Rien ne remplace pour cette approche qualitative la vie dans le pays, en particulier dans les villages. La connaissance de la langue et un certain recul historique.

BIBLIOGRAPHIE

- Bruneau (M.), *Recherches sur l'organisation de l'espace dans le Nord de la Thaïlande*, thèse de doctorat d'État, Paris, H. Champion, 1980, 1044 p., 4 cartes couleur h.t.
- Bruneau (M.), « Paysages, rapports sociaux de production et éco-géographie » in *Terrains vagues et terres promises*, Paris, PUF, 1981, pp. 283-299.
- Bruneau (M.), Kilian (J.), « Inventaires agroécologiques, paysages et télédétection en milieu tropical, essai méthodologique », *L'Espace géographique* (3), 1984, pp. 215-224.
- Bruneau (M.), Kilian (J.), Le Men (H.), Mongkolkeha (C.), « Identification et dynamique des milieux agricoles dans le Nord-Est de la Thaïlande », *Travaux et documents de géographie tropicale*, CEGET, 1986, 72 p., 5 cartes en couleurs h.t.

L'outil informatique permet de conserver aux phénomènes étudiés une plus grande complexité grâce à l'abondance des données qui peuvent être traitées et d'introduire plus de rigueur et d'homogénéité dans les interprétations. Mais la connaissance qualitative nécessaire en amont est indispensable en aval des traitements informatiques.

Géographie : un renouveau assisté par ordinateur

Le renouveau que connaît aujourd'hui la géographie n'est dû qu'en partie à l'ordinateur mais son utilisation l'a facilité en permettant l'expérimentation de modèles et de théories.

Denise PUMAIN

On surévalue généralement le rôle de l'informatique dans la modernisation de la géographie.

S i quelqu'un vous dit que, grâce à l'ordinateur, tout a changé et que l'on ne saurait désormais plus faire de géographie sans lui, ne le croyez pas tout à fait : il s'agit sans doute d'un technocrate mystificateur. Mais ne croyez pas, non plus le traditionaliste entêté qui, tout en l'accusant de toutes les perversions, réduirait à cette seule innovation technique le foisonnement nouveau de la pensée géographique, qu'il récuse. Attribuer un rôle décisif à l'ordinateur dans la modernisation et le renouvellement scientifique de la géographie au cours des vingt ou trente dernières années, n'est peut-être qu'une interprétation trop facile de bouleversements plus profonds.

Diffusion des ordinateurs...

En géographie, les ordinateurs ont été utilisés plus tôt et plus systématiquement que dans d'autres sciences humaines.

Dans l'apparence des choses, le technocrate a raison. Parmi les sciences humaines, la géographie est certainement celle qui a entamé le plus tôt et le plus complètement sa « révolution informatique ». Les géographes sont devenus, dès le début des années soixante-dix, des clients assidus et actifs des centres de calcul des universités et du CNRS. Aujourd'hui, les matériels informatiques sont présents dans la quasi-totalité des laboratoires de géographie, et ils tournent ! L'aventure a commencé avec les programmes d'analyse des données, elle s'est poursuivie par la réalisation de logiciels de cartographie automatique, le traitement des données de télédétection, la constitution de bases de données, la modélisation et la simulation, le contrôle d'appareils de mesure sur le terrain, etc.

Mais le renouveau épistémologique de la géographie est bien plus large et plus ancien.

Désormais, dans la plupart des laboratoires, l'informatique assure l'enchaînement des opérations de recherche les plus courantes : stockage des données, traitement statistique et mathématique, expression cartographique. L'amélioration des procédures et l'innovation technique sont assurées par des équipes spécialisées — dans la représentation cartographique, ou la télédétection, ou la modélisation. Ces équipes sont, depuis deux ans, réunies en un réseau (le Groupement d'intérêt public RECLUS*), rare exemple de mise en commun des savoir-faire, probablement favorisé et soutenu par la diffusion préalable de l'outil informatique dans cet ensemble de chercheurs.

...et réaffirmation de la méthode scientifique

Ce ne sont pas seulement les outils et les méthodes de travail qui ont changé. Dans le même temps, une véritable « révolution scientifique »

La construction théorique et l'élaboration de modèles ont relancé les monographies régionales et la géographie descriptive traditionnelle.

■ Denise Pumain, professeur à Paris XIII, conseiller technique à l'Institut national d'études démographiques (INED), membre de l'équipe « Pour l'avancement des recherches sur l'interaction spatiale » (PARIS). INED, 27, rue du Commandeur, 75675 Paris Cedex 14.

a eu lieu. En schématisant à outrance, on pourrait dire que la géographie française d'entre les deux guerres s'était un peu trop repliée sur le succès de ses descriptions régionales, appuyées sur une problématique des relations entre l'homme et le milieu ainsi que sur un type d'explication d'ordre « génétique », faisant une large place à l'enchaînement historique des causes. La contestation d'une approche jugée trop « naturaliste », morphologique et paysagère, et la volonté d'élargir le champ de la discipline aux questions des théories économiques et sociales, la prise en compte réfléchie de l'intervention des idéologies et de la subjectivité dans l'organisation des phénomènes territoriaux et dans les travaux des géographes, ont fait éclater le champ thématique de la discipline. Les contacts renouvelés avec d'autres disciplines, ainsi qu'avec les écoles de géographie étrangères, ont aussi contribué à son enrichissement.

De l'idiographique au nomothétique

Mais le moteur de la révolution est ailleurs. Il réside dans un changement d'optique, dans l'adoption d'une autre attitude scientifique. Au lieu de trouver l'explication de la physionomie et de l'organisation d'une certaine région dans son sol et dans son histoire, on s'est mis à la recherche de lois, de modèles généraux qui soient susceptibles d'en rendre compte. Avec l'approche nomothétique, le champ d'étude s'est déplacé de ce qui fait la beauté de ce qu'on ne verra jamais deux fois vers des éléments plus banaux des paysages et des relations « verticales » entre milieu physique et événements accumulés en un même lieu, vers des relations « horizontales », régularités de l'espace et de la disposition relative des objets géographiques. Des modèles, parfois anciens, ont été éprouvés et améliorés, qu'il s'agisse de l'organisation des villes en une trame de « lieux centraux », de l'utilisation agricole du sol en anneaux d'intensité décroissante, des flux de toute nature selon des schémas de gravitation, ou encore des règles strictes de la diffusion des innovations sur un territoire... Par rapport à ces modèles, les particularités locales étaient interprétées comme des « résidus ».

L'ordinateur : un frein...

Il est clair que la plupart de ces renouveaux épistémologiques ne doivent guère, dans leur origine, et peu dans leur déroulement, à l'ordinateur. On pourrait même prétendre que les premières applications de l'informatique au traitement des données géographiques ont au tant freiné que favorisé la diffusion des idées nouvelles : le choix des données était d'abord restreint aux indicateurs les plus accessibles, les relations mathématiques cherchées étaient mal

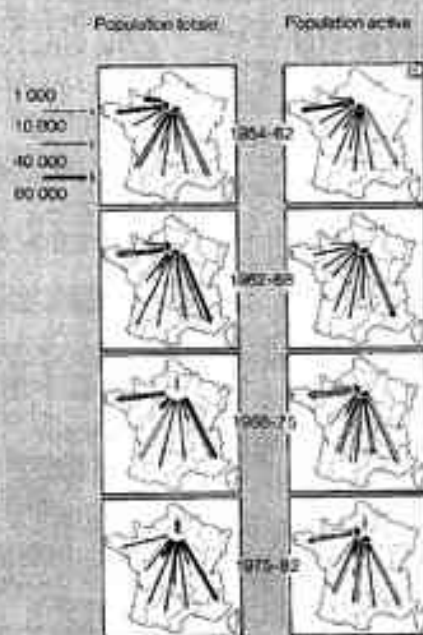
* Réseau d'étude des changements dans les localisations et les unités spatiales (Montpellier).

Les migrations interrégionales

Les flux de migrants échangés par les régions sont très fortement liés à l'importance des populations en présence et aux distances qui les séparent. Dès 1885, Ravenstein avait fait le rapprochement entre la gravitation universelle et les « lois de la migration ». On a donc fait l'hypothèse qu'un modèle gravitaire simple pouvait rendre compte du nombre de migrants entre deux régions à partir de leur population et de la distance entre elles. Les données pour plusieurs périodes intermédiaires ont été ajoutées à ce modèle, selon un programme mis au point par un chercheur belge, M. Poulain. Plusieurs expressions de la fonction de la distance ont été tentées et les « résidus » (ici cartographiés) représentent les écarts possibles entre flux observés et flux calculés pour les échanges entre l'Ile-de-France et la province. Ils montrent des nombres de migrants réels en excès par rapport aux hypothèses d'homogénéité et de symétrie faites sur la mobilité par le modèle, donc des courants préférentiels. On voit ainsi progressivement s'établir un renversement de l'attraction gravitaire vers les régions du sud et de l'ouest et dans certaines régions du Bassin parisien (Centre puis Picardie). Cette inversion a été plus précoce et massive pour la population dans son ensemble

que pour les seuls courants migratoires de population active. Cependant, l'accroissement relatif des migrations à longue distance et les fondances à la déconcentration réduisent la pertinence du modèle — par ailleurs statique et descriptif. Pour affiner l'analyse et la prévision des échanges interrégionaux, on peut utiliser d'autres hypothèses, qui ne forment pas intervenir à priori la distance ou la population de la région d'accueil. En utilisant les acquis de la synergie, on peut construire un modèle dynamique qui relie les probabilités individuelles de migrer à des équations de mouvement décrivant les flux migratoires agrégés. Les tests de ce nouveau modèle sont en cours, en collaboration avec des chercheurs du Laboratoire de physique théorique de l'université de Stuttgart.

Poulain (M.), « Contribution à l'analyse spatiale d'une matrice de migration interne », « L'Annuaire de la Nouvelle Calédonie, Recherches démographiques », 1981, Cahier n° 9, 225 p.
Pumain (D.), « Les migrations interrégionales de 1954 à 1982 : directions préférentielles et effets de barrière », *Population*, 1986, n° 2, pp. 276-288.



Courants préférentiels entre l'Ile-de-France et les autres régions. (Document extrait de *Population*, 1986, n° 2. INED.)

adaptées aux problèmes posés par la contiguïté et la dépendance géographiques, les programmes de traitement trop frustes et les interprétations parfois incertaines ou trop ambitieuses. Les résultats pouvaient apparaître soit bien décevants, soit obtenus au prix d'un énorme détour d'explicitation et de formalisation qui n'apportait guère de progrès immédiat par rapport à l'état antérieur des connaissances.

...puis un formidable outil d'expérience

Pourtant, déjà, la valeur ajoutée par l'ordinateur était irremplaçable — non seulement sur le plan technique de la facilité et de la rapidité des calculs, et des possibilités de traitement simultané de grandes masses de données — mais aussi sur un plan épistémologique : en autorisant enfin la mise à l'épreuve de modèles parfois élaborés plusieurs décennies auparavant, le test systématique d'hypothèses dans des conditions contrôlées, l'ordinateur a contribué à développer en géographie la méthode expérimentale. Dès cette première phase, une conclusion énoncée s'appuyait désormais sur une procédure répétée par un autre expérimentateur. Surtout, la voie était ouverte à la modélisation qui, seule, permet de raisonner en géographie toutes choses égales par ailleurs, de choisir entre plusieurs conjectures et, par la simulation, de faire des expériences irréalisables sur le terrain et, donc, d'étayer des prévisions.

Actuellement, le dialogue ouvert avec les statisticiens, les mathématiciens, les informaticiens, permet aux géographes d'aborder scientifiquement des domaines où l'approche littéraire semblait il y a peu, encore, la seule possible : description des paysages, traitement de données qualitatives, reconnaissance du contour flou des régions... Ils ont aussi renoué avec la « géographie mathématique » du siècle dernier, qui traitait des représentations planes de la surface de la terre, pour dessiner, cette fois, à l'aide de l'ordinateur, non seulement des projections classiques, mais aussi des images plus ou moins

« déformées ». La forme même de la carte est transformée selon les perceptions individuelles, ou en fonction de l'accessibilité des lieux dans les réseaux de transport, de la façon dont les habitants perçoivent leur territoire ou selon des critères aussi variés que le nombre d'enfants par famille ou l'intensité de la pratique du rugby... L'image obtenue par télédétection, après différents lissages et filtrages, se substitue parfois avantageusement à la carte pour l'analyse et la communication.

S'il fallait résumer cette histoire, déjà trop brièvement évoquée, nous dirions que l'ordinateur, par ses exigences techniques, a heureusement obligé les géographes à reconsidérer avec beaucoup d'attention la nature, le contenu et les limites de l'information qu'ils utilisent : par son efficacité, il leur a permis d'élargir leur champ d'investigation et d'accélérer les procédures banales ; par sa souplesse actuelle, il devrait à la fois soutenir la diffusion, dans le grand public, des acquis de la recherche (on pense à la multiplication des « systèmes d'information géographiques », bases de données localisées, avec transcription cartographique) et autoriser des développements nouveaux d'une géographie française réenracinée dans un passé mais appuyée sur des bases scientifiques plus solides.

L'enthousiasme des pionniers doit être relayé par la formation de personnels compétents en informatique. Enfin, le savoir-faire technique, et notamment la production des cartes et des images, ne saurait se substituer à l'effort de construction théorique. Telles sont les conditions de succès.

BIBLIOGRAPHIE

- Racine (J.-B.), Raymond (H.), *L'analyse quantitative en géographie humaine*, Paris, PUF, 1973, 316 p.
L'Espace géographique, 1984, n° 2 « Modélisation et simulation » et n° 3 « Télédétection spatiale ».
Les Annales de géographie, 1983, n° 511, par thèmes.
Mathématiques et sciences humaines, 1981, n° 50, « Géographie urbaine » et n° 51 « Géographie rurale ».

D'abord rudimentaire et réducteur, l'emploi de l'ordinateur ouvre aujourd'hui, par la souplesse et la diversité de ses applications, de vastes possibilités au progrès et à la diffusion de la discipline.

La lexicométrie socio-politique

L'ordinateur ne permet pas de traiter directement le contenu des discours politiques, mais, en analysant les mots qui les composent, il fait apparaître bien des choix et des habitudes, des partis pris.

Maurice TOURNIER

Les mots occupent l'espace médiateur entre la violence et l'obéissance.

Comment éviter le piège des mots pris aux pièges du sens ?

Examiner les mots avant le sens.

Quand la violence des pouvoirs ou des contestations remise ses armes pour recourir à des moyens symboliques, elle entre en politique. Devenant politique, elle se fait mots. Les mots sont des armes reculées d'un pas.

Ce retrait est aussi, bien évidemment, celui des duplicités rendues possibles. Rien n'est davantage propre aux dissimulations et aux diffracteurs que les mots qui font la paix alors qu'ils sont nés d'un conflit.

Cette ruse fondamentale du discours socio-politique semble, à première vue, à mille lieues de tout traitement informatique, tant ses objets constitués d'ambiguïtés et de flou sont réfractaires par nature aux oui/non des rangements systématiques. Cette première vue est juste, s'il s'agit de sens. Les mots-sens ne peuvent se classer; il faudrait raconter l'histoire pour cerner au plus près le pourquoi de chacun et rendre compte de ses alliances de contexte et de situation. Et en sait bien qu'à la limite, vu l'instabilité des référents et l'infinité variée des sites d'emploi, aucun mot ne se superpose à lui-même de l'une à l'autre de ses occurrences, et

jusqu'à l'intérieur d'un même texte, puisque le jeu est justement là. En outre, avant d'en arriver aux sens, l'initiation est longue. Remontant d'un texte à l'autre - aucun mot, disait Bakhtine, qui ne soit *déjà habité* -, le chercheur s'implique à chaque pas davantage dans cette histoire qu'il choisit, se déneutralisant. Comment décrire ensuite, lorsqu'on a déjà interprété ?

C'est ce refus de décrire une interprétation qui conduit à l'idée de prendre les choses à l'envers, à savoir examiner d'abord l'agencement du tissu textuel apparent avant d'accéder à ses sens profonds.

Trois principes président à ce renversement de la perspective scientifique: surface, clôture, chiffage.

- analyser un texte comme une suite de signes graphiques sur un papier permet d'éviter momentanément le problème du sens: il faudra donc rechercher des critères d'échantillonnage qui tiennent moins compte du contenu que des conditions de l'énonciation;
- utiliser un corpus clos permet de réduire les effets d'évolution en construisant un ensemble daté et situé, dont chaque terme sera jugé selon les mêmes critères de probabilité, sans faire appel à des normes qui lui soient extérieures;
- la comparaison chiffrée autorise la confrontation des emplois d'après un nombre réduit de variables socio-historiques.

L'ordinateur est ainsi devenu, depuis une dizaine d'années (1, 2, 3) l'outil majeur de ce type de recherche, car il sait explorer, signe à signe,

■ Maurice Tournier, directeur de recherche au CNRS, responsable de l'unité de recherche linguistique « Lexicométrie et textes politiques », URL: LP6862 03, de l'Institut national de la langue française (INALF), LP 6862, Ecole normale supérieure de Saint-Clément, 2, avenue de la Grille d'Honneur, Parc de Saint-Clément, 92211 Saint-Clément.

Banque d'information politique et d'actualité (BIPA)

Les déclarations présidentielles sont enregistrées dans la banque de données LOGOS de La Documentation française depuis mai 1974, quelle que soit leur forme: interview, conférence de presse, discours, lettre.

Tous les mots du texte sont interrogables, ainsi que les champs signalétiques, constants depuis 1974, tels que: auteur, titre, typologie, date, lieu, circonstances, qui sont respectés pour chacun des unités documentaires référencées.



Le président de la République, François Mitterrand, à la Sorbonne, lors de la 1^{re} Conférence internationale sur l'arbre et la forêt (5-6-7 février 1986).

Une banque de textes et de mots

Il ne faut pas manquer de signaler la création à l'INALF* de Nancy, sous la direction de B. Guimada, de la plus grande banque informatique de données textuelles et lexicales du monde (en cours de mise en ligne et sur réseau): plus de cent soixante millions de mots y ont déjà été enregistrés, d'abord dans l'optique de la réalisation du TLF (*Treasure of the French language*, dictionnaire en voie d'achèvement; onze volumes parus, CNRS-Gallimard) puis en vue de bases documentaires permanentes. J. Dondin a mis au point un logiciel d'interrogation, STELLA*, qui permet à tout utilisateur d'accéder par console à une multiplicité de renseignements linguistiques sur le français. Une succursale de cette banque est depuis quatre ans implantée à Chicago; d'autres se préparent (Stockholm). Des terminaux sont envisagés dans les URL* et à des postes de consultation universitaires.

Autour de ce système serveur, un réseau d'observatoires* fait le guet de la langue à divers points de la francophonie.

Les recherches d'E. Brunet (URL 9) portent sur cette masse documentaire: Le vocabulaire français de 1789 à nos jours, Genève, Slatkine, 1981, 3 vol. et les *Cahiers CUMFID*.*

Le recours à la machine.

Le recours aux contextes.

Quand la machine calme le jeu.

Systématiser pour neutraliser.

Le sens des mots c'est d'abord le système de leurs relations d'emploi et de leurs relations statistiques.

une surface, recenser et comptabiliser ses unités formelles jusqu'au bout, ainsi que leurs liens statistiques, calculer les probabilités intrinsèques de ces formes et de ces liens, classer et hiérarchiser, opposer ou corréler, selon les partitions qu'on lui donne à traiter. Le regard qu'il permet de poser sur les mots dans le cumul de leurs contextes s'avère riche de distance, de mémoire et d'imagination.

L'ordinateur, outil de neutralisation

Une distance entre le matériau et le chercheur est imposée par le codage (des unités textuelles comme des variables de situation) et la mise en place d'une opération de saisie. L'analyste ne peut plus effectuer cette plongée immédiate dans le document, ni cette fouille personnelle qui rendaient la quête des mots si gratifiante. Il se trouve placé devant l'obstacle d'une systématisation préalable à mettre au point. Et comme cette opération s'applique, par définition, de manière identique à tous les éléments du corpus (et pour tout corpus parallèle ou ultérieur), elle a pour conséquence de remplacer par des arbitraires (de découpage, par exemple) des choix intuitifs ou raisonnés, trop susceptibles de modulations. Neutralisation de l'enquête? En partie certainement, car l'effort de non-intrusion du chercheur au sein du matériau à inventorier se prolonge dans l'application de programmes eux-mêmes systématiques.

L'ordinateur, outil de totalisation

La banque de textes et de mots est au bout du processus de sa mise en mémoire avec tous les avantages qu'elle présente, dont le moindre n'est pas son utilité collective. Chacun garde le droit d'enrichir l'enregistrement minimal standard, mais au moins dispose-t-il au départ, comme son voisin, de ce type d'enregistrement et des logiciels qui le gèrent (voir encadré 1).

* Les sigles marqués d'une astérisque sont développés dans le glossaire.

C'est ainsi que l'accumulation homogénéisée plaide pour une première phase d'études, quantitatives. Dresser le bilan formel — non pas de la langue, intolérable autant que non-fréquentielle, mais de certains discours situés — est le premier objectif de la lexicométrie. Avant d'être des idées, les mots sont des occurrences, c'est-à-dire des gestes ou des objets mis en place.

Leur présence — et plus encore leur co-présence — dans tel discours et à tel moment du texte ou de l'histoire n'a aucune innocence, ni aucune gratuité. Repérer leurs lieux d'emphase, ou d'absence relative, discerner le caractère spécifique ou banal de leurs distributions, saisir la cadence de leurs occurrences, inventorier et cumuler leurs voisinages proches comme lointains, apprécier en probabilité tous ces chiffres et, au-delà, similitudes ou contrastes d'emploi, cela fait aujourd'hui partie des bilans nécessaires, parce que devenus systématiquement possibles. Il ne suffit plus de glaner son bien dans le ça-et-là des fiches recueillies, d'aller, comme le disait G. Gougenheim, à la pêche à la ligne de citations intéressantes, pour faire avancer l'étude des vocabulaires politiques. L'ordinateur double la monographie partielle par l'analyse exhaustive d'échantillons-témoins serrés dans leur corset d'invariants.

L'ordinateur, outil de représentation

Outil heuristique, la machine l'est aussi par l'image qu'elle peut donner des fonctionnements des mots en texte et en situation. Toute science progresse lorsqu'elle débouche sur un

Programme des « spécificités »

Une probabilisation des fréquences locales classe automatiquement, par l'application d'un seuil arbitraire (5%, en général), les formes d'un texte constitué d'un corpus en trois catégories :

- les « spécificités positives », dont la fréquence et toute fréquence supérieure, cumulées, n'atteignent pas ce seuil de probabilité (on ne les attendait pas si nombreuses);

- les « spécificités négatives », dont la fréquence et toute fréquence inférieure, cumulées, n'atteignent pas non plus ce seuil (on ne les attendait pas si rares);

- le « vocabulaire banal », dont rien de caractéristique ne peut être dit, sinon qu'il est relativement mieux distribué dans ce texte que les éléments des deux ensembles précédents. Lorsqu'une forme ne s'avère ni sur-employée, ni sous-employée, dans aucune des parties du corpus, elle est banale partout. On dit alors qu'elle appartient au « vocabulaire de base » du corpus. Et cela est valable à tous les niveaux de fréquence, si l'on fait appel pour le calcul des probabilités à la distribution hypergéométrique (Mots, n° 1 et 2).

Plusieurs thèses de lexicométrie politique (3, 4) et des recherches collectives (5, 6) ont pratiqué ces analyses de « spécificités », tant sur les discours socialistes, communistes ou syndicaux que quarantehuitard ou soixantehuitard. Pour un ensemble d'éditions de l'Humanité des années trente, par exemple, la méthode a permis de corriger ou pas, selon les temps forts du renouvellement lexical (de « Front unique » à « Front populaire »), leur confrontation à l'évolution politique montrée qu'il n'y a pas forcément coïncidence entre les mots et les faits. Le discours politique est toujours une stratégie. (D. Fieschanski, Discours communiste et grand tournant. Étude du vocabulaire de l'Humanité (1934-1936), thèse de 3^e cycle, université de Paris I, 1981).

Programme des «cooccurrences»

Ce programme (4 et Mots, n° 3, pp. 95-148) permet de sélectionner les couples de formes les plus pertinentes, c'est-à-dire mathématiquement les plus inattendus (d'après la probabilité hypergéométrique calculée pour chaque co-fréquence dans le cadre du nombre des phrases analysées). Des couples représentent soit les éléments de stéréotypes textuels les plus soudés (sont détaillés dans les ISF*), soit les éléments de véritables « téléstéréotypes », lorsque à la fois récurrentes et conjointes, lesquelles peuvent passer inaperçues à la lecture. À partir des constats effectués sur la co-fréquence, la probabilité et la distance moyenne qui s'interpose entre les deux pôles des couples pertinents, on peut construire automatiquement des lexico-graphes d'attractions directes ou des graphes d'attractions successives. Se dessine ainsi autour d'un pôle un résumé de ses voisins préférentiels ou des échos lexicaux qui s'en suivent. Retour au sens selon les voies d'une « grammaire statistique ».

Les stéréotypes textuels

Prenons le cas du corpus des résolutions confédérales votées à la CGT et à la CFDT pendant la période 1971-1976(5). En sélectionnant les couples récurrents dans les phrases de l'une ou l'autre centrale sur le critère d'une probabilité infime et en excluant de la représentation — mais non du calcul — les formes-outils, on obtient deux ensembles de couples. L'ordinateur repère ainsi organisations → syndicales (couple 89 fois présent à la CGT avec une distance moyenne D de 3 mots interposés et 55 fois à la CFDT avec D = 0,5), unité → action (42 et 41 fois avec D = 2), pouvoir → achar (57 et 35 fois avec D = 4 et à 23), équipements → collectif (16 et 17 fois avec D = 4 et 0,5), fait → compte (13 et 11 fois avec D nul), maître → œuvre (19 et 15 fois avec D = 1), mise → œuvre (26 et 25 fois avec D = 1), doit → être (59 et 57 fois avec D proche de 1), doivent → être (75 et 57 fois avec D proche de zéro), etc. Couples non discriminants par eux-mêmes, qui se retrouvent d'ailleurs dans les résolutions d'autres confédérations. L'intérêt grandit avec les couples à variations et les couples moins conjugués. Alors que la CGT, nous dit la machine, privilégie masses → populaires, c'est l'inverse → populaires qui, à la CFDT, prend la même ampleur. Le premier utilise fortement mouvement → syndical, le second la préférence mouvement → ouvrier. Défense cooccure en priorité avec intérêts à la CGT et avec libertés à la CFDT. Objectifs → revendicatifs culmine à CGT et objectifs → transformation à la CFDT (mais, tandis que pour le premier couple l'indice D a une valeur nulle, pour le second il s'approche de 5). Avec ce dernier cas, nous touchons aux lois statistiques à distance, univers étonnant où l'on trouve, par exemple, à la CGT des couples comme emploi → ressources ou financement → état, 13 et 10 fois présents avec des D supérieures à 20. Le pôle CFDT, lui, 134 fois suivi par travailleurs, l'est en général de loin, avec D supérieures à 17 et le couple patronat → chômage, 12 fois présent, possède une valeur D supérieure à 18. Nous voilà bien éloignés des séquences de mots immuablement soudées. Que de médiateurs et de variants viennent se glisser entre les pôles et participer à leur association réciproque! Comment rendre compte de tous ces phénomènes? En organisant les cooccurrences en lexico-graphes, puis en suivant leurs chemins d'attractions.

Un « lexico-gramme » est une micro-structure hiérarchisée des cooccurrences statistiquement.

Lexico-graphes directs

Un « lexico-gramme » est une micro-structure hiérarchisée des cooccurrences statistiquement.

travailleurs (CGT) I = 576	travailleurs (CFDT) I = 624
1 intérêts	1 immigrés
2 défenses	2 masses
3 droits	3 peuples
4 répondre	4 population
5 appeler	5 entreprise
6 participation	6 moyens
7 masses	7 sociaux
8 solidarité	8 entreprises
9 entreprise	9 professionnelle
10 revendications	10 français
11 millions	11 syndicales
12 besoins	12 revendications
13 large	13 monde
14 lutte	14 retraités
15 catégories	
16 faire	
17 discussion	
18 organisations	
19 CGT	
20 prises	
21 vie	
22 changements	
23 permettre	
24 scolaires	
25 action	
26 pouvoir	
27 lutte	
28 assurer	
	1 CFDT
	2 appels
	3 aspirations
	4 revendications
	5 situation
	6 problèmes
	7 expression
	8 misère
	9 action
	10 entend
	11 droits
	12 fois
	13 ensemble
	14 organisations
	15 participation
	16 jeune
	17 but
	18 lutte
	19 solidarité
	20 prendre
	21 conseil
	22 refus
	23 donner
	24 satisfaction
	25 implique
	26 l'été
	27 permettre
	28 doit
	29 volonté
	1 immigrés
	2 revendications
	3 retraités
	4 entreprises
	5 entreprises
	6 travailleurs
	7 syndicales
	8 élaboration
	9 collective
	10 ans
	11 organisations
	12 temps
	13 conscience
	14 enjoints
	15 formes
	16 personnel

les plus improbables (i.e. attirés) qui font couples avec un « pôle » à la gauche ou à la droite duquel ils se distribuent. Ce graphe donne une certaine représentation de ses principales habitudes de co-fonctionnement.

Exemple: voici les lexico-graphes du pôle travailleurs dans les résolutions votées à la CGT et à la CFDT pendant la période 1971-1976, régies par les règles suivantes: probabilité < 1%, fréquences et co-fréquence > 10, élimination des « formes-outils » (à partir d'une liste-type de 333), classement dans l'ordre inverse des valeurs de probabilité.

Des règles moins contraignantes (avec p < 5% et l'exclusion des formes-outils) construisent des lexico-graphes 5 fois plus importants. Dans les uns comme dans les autres, on constate que 40% au moins des cooccurrences sont du même type: ce qui dénote dans travailleurs un terme peu discriminant, quant à ses emplois majeurs dans les deux Confédérations (les intersections cooccurrences pour les substantifs n'excèdent guère 30%). Le déséquilibre constaté entre la gauche et la droite des lexico-graphes ne reflète pas une préférence du pôle pour les fins de phrase (l'examen des distances le situe en moyenne au centre des phrases) mais un figement plus fort sur sa gauche que sur sa droite: travailleurs fonctionne en effet souvent comme un terme de complémentation (le couple des → travailleurs arrive en tête de tous les couples avec «02 réalisations à la CFDT et 368 à la CGT). Cette observation montre l'intérêt de poursuivre les graphes au-delà des cooccurrences directs en

suivant les relations d'attraction de proche en proche: chaque cooccurrence peut devenir pôle à son tour. Nous construisons ainsi des « lexico-graphes étapés ».

Lexico-graphes étapés

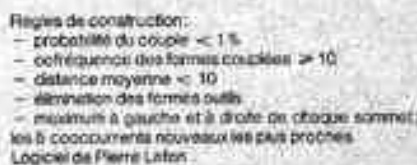
Ainsi que les listes précédentes, le graphe de travailleurs ci-après est le résultat d'une sélection et d'une itération automatique gérées par des règles. Il représente une sorte de résumé statique des rencontres qu'effectuent non seulement le pôle, mais ses cooccurrences les plus proches et les cooccurrences de ses cooccurrences. L'intérêt de cette représentation réside d'abord dans l'examen des chemins ou des circuits propres à ce graphe et dans sa comparaison aux graphes tirés d'autres organisations ou d'autres périodes. C'est par cette comparaison qu'un système des emplois commence à faire sens.

Les programmes du « Lexicloud ».

système de représentation. Or, s'il est un domaine où l'ordinateur peut « inventer » une figuration représentative d'un phénomène sans que celle-ci tienne à autre chose qu'à des règles de construction préalablement posées, c'est bien celui de l'étude des composants textuels et de leurs relations. Nous n'évoquerons pas les analyses automatiques du discours, qu'elles aient pour but la mise en évidence de schémas argumentatifs profonds ou qu'elles s'efforcent

de repérer classes morpho-syntaxiques, catégories fonctionnelles ou lexies de langues (AAD*, DEREDEC*, TA*, etc.).

Nous nous en tiendrons aux programmes d'analyse lexicométrique actuellement actifs au laboratoire de Saint-Cloud. Ils partent de la notion de « relation statistique » mais pour la dépasser par des représentations (tris, sélections, hiérarchies, graphes) qui s'efforcent de clarifier certains fonctionnements textuels majeurs.



sons principales, tout en conservant sa stéréotypie de base (B : pp. 193-219; 7 : pp. 841-853; 8 : pp. 156-187).

D'autres régies (sans limitation de la distance moyenne) par exemple) donnent d'autres représentations (à sommets moins contigus, munis d'une génération plus touffue d'arcs d'attrance). L'appréciation générale demeure, voire s'empêche. Se multiplient même les chemins statistiques qui n'ont aucune réalité dans les textes : surconstructions évidentes, qui sont des résultantes, inexistantes dans l'énoncé, des phénomènes d'attrance employés. L'analyse quantitative écrase en effet les séquences d'emplois les unes dans les autres. Artefact, dira-t-on...

Mais n'est-ce pas ainsi que les phrases se présentent dans leur rétention, le discours terminé ? Le texte politique ou syndical n'est-il pas façonné dans le soubçon, inconscient, de cet écrasement

final? L'impression de lecture ou d'audition tient aux grands martèlements de fond, aux fausses locales les plus ricurientes, phrase dans phrase, aux logiques de ricochets. En outre, un terme n'arrive pas seul à l'énoncé, dans la nuit de sa signification dictionnaire: illusion de linguiste. Il s'y impose avec toute la grappe de ses habitudes de voisinage qui font sens avec lui. Ainsi les mots sont les ou entendus les uns à travers les autres. Mieux alors que la relation syntaxique locale asservie à la linéarité, c'est peut-être la relation statistique globale qui rend le plus justement compte de ces messages sociaux à imprégnation progressive. Filigrane des confusions d'emploi, plus vrai derrière les textes que l'énoncé même de chaque emploi, filigrane de ces durcissements significatifs du discours qui nous arrosent: «désinformation».

• Rates •

Ainsi de l'analyse des « specificités » due à P. Lafon, il s'agit de déterminer quel est le vocabulaire dont l'emploi est quantitativement le plus caractéristique d'une partie du corpus rela-

Le rythme d'arrivée de chaque forme dans un texte fait, quant à lui, l'objet d'une analyse de ses « rafales ». On dénomme par ce terme la méthode qui décrit chaque cadence, parce qu'on a constaté que la plupart des formes s'organisent beaucoup plus par attroupements d'occurrences

« Segments répétés ».

ces que selon une dissémination régulière tout au long du texte. P. Lafon a imaginé un indicateur statistique qui hiérarchise toutes les formes à répétition, de la plus « rafaleuse » à la mieux répartie (*Mots*, n° 2, pp. 157-188).

Un texte est aussi le lieu où les formes viennent à la rencontre les unes des autres. Deux sortes de programmes s'efforcent de cerner ces rapprochements.

Les uns, développés par A. Salem, font l'inventaire de tous les « segments répétés » (ISR*) d'un corpus et les classent selon leur longueur, en ordre hiérarchique ou alphabétique, avec l'indication de leurs localisations, de leur fréquence et des fréquences de leurs constituants (*Mots*, n° 6, pp. 161-177).

Les ISR* vont être soumis aux analyseurs statistiques existants (de spécificités, de rafales, de cooccurrences, de correspondances), ainsi qu'aux analyseurs grammaticaux automatiques expérimentés à Saint-Cloud. Tels quels, ils permettent d'étudier en détail les habitudes stéréotypiques d'un usage: lexies soudées, slogans, locutions et tout autre type de cooccurrence formelle réitérée. (7: P. Fula, P. Lafon, A. Salem, pp. 381-390, 531-540, 759-769 et A. Salem, *Pratique des segments répétés*, (Nancy, INALF - Paris, Klincksieck, à paraître).

Une autre série de programmes, élaborée par P. Lafon, saisit le phénomène de cooccurrence des formes dans leur co-présence orientée à l'intérieur de contextes communs, actuellement identifiés à la phrase. L'ordinateur relève tous les couples de formes qui s'y répètent et probabilise leurs rencontres. Un listage hiérarchise les seuls couples dont le nombre de rencontres (appelé « co-fréquence ») n'atteint pas les 5 % de probabilité, en indiquant quelle distance moyenne sépare les éléments de chaque couple (voir encadré 3).

Retour au sens: la sloganisation, message secondaire ou primaire ?

Que signifie « sens » dans le cas d'un discours socio-politique ? D'une part, l'information et l'argumentation véhiculées au long de sa séquence produisent un sens linéaire; d'autre part, en bout de ligne, ce qui demeure des énoncés lus ou entendus (ou plus exactement ce que l'énonciateur croit pouvoir espérer de la rétention) forment un sens second, où ne commandent plus l'articulation et l'analyse mais l'accumulation et le syncrétisme. Faut-il dire qu'il s'agit d'un sens second ? N'est-ce pas pour lui en définitive que le discours est délivré ? Toute réflexion, au contraire, instaure un retour critique sur le texte. C'est l'analyse qui est secondaire.

Poser la question: que reste-t-il du message reçu ? Les gens vous répondront des mots et meubleront eux-mêmes les interstices. Ce qu'ils auront retenu du discours, c'est l'assise des rythmes qui le martèlent et ce qu'ils finissent derrière les marques cadenciées et cooccurrenceuses. Le sens se reconstruit d'abord sur les traces des échos perçus. C'est pourquoi le message le mieux entendu est celui où l'on se retrouve le mieux, tant il est vrai qu'on ne restitue spontanément que sa propre cohérence. Dès lors, qui a raison, quant à l'efficacité espérée d'une action publique de langage, de l'analyse explicative ou de la sur-construction statistique ?

Le jeu politique, dans le geste de parole, est de marquer et de tracer, marquer des inflexions verbales, tracer des chemins d'échos. La grammaire de ce jeu n'est pas, on le voit, du ressort de la syntaxe mais de la statistique. Celle-ci a pour tâche de décrypter la sloganisation des textes, cette assumption de récurrences et co-récurrences

Glossaire des sigles

AAD: Analyse automatique du discours (M. Pichoux, 1982, Cf. *Mots*, n° 4, 9, 13).

CUMFID: Cahiers des utilisateurs de machines à des fins d'information et de documentation (revue de l'URL, Nice, Dr. E. Brunet).

DEREDEC: description et reconnaissance sur le DEC-SYSTEM-10 de l'université du Québec (P. Plante, Montréal).

INALF: Institut national de la langue française, laboratoire propre du CNRS (Dr. B. Quémar).

ISR: inventaire des segments répétés, logiciel d'A. Salem.

LEXICLOUD: Lexicométrie, URL, Saint-Cloud (logiciel intégré par A. Salem).

LISH: Laboratoire d'Informatique en sciences humaines, Paris, CNRS (Dr. M. Hansworth).

MOTS: Mots, Ordinateurs, Textes, Sociétés, Paris, Presses de la FNSP (revue de l'URL, Saint-Cloud).

STELLA: Système de textes en ligne (un libre accès logiciel de J. Dendin, URL, Nancy).

TA: traduction automatique/traitement automatique (cf. *T.A.-Information*, revue de l'ATALA, Association pour le traitement automatique du langage, secrétaire: P. Lafon).

URL: unité de recherche linguistique, CNRS, équipe constitutive de l'INALF. Il y a 10 URL dans l'INALF (Nancy, Paris, Saint-Cloud, Besançon, Lyon, Nice, Toulouse).

URL3: laboratoire « Lexicométrie et textes politiques », Ecole normale supérieure de Saint-Cloud, 3^e URL de l'INALF (Dr. M. Tournier).

ces qui fait message au-delà des énoncés, ou en deçà.

Modalisations, atténuations, voire dénégations, fondées dans la variété rhétorique, laissent la place aux rapprochements brutaux de la répétition associative. Et peut-être l'effacement du discours sloganisé réside-t-il dans le comblement de rêves non-dits faits avec nos mots. N'est-ce pas en oubliant les négations et les repentirs que le psychanalyste écoute dans le flot des obsessions les cooccurrences d'une vérité ?

Cooccurrence contre séquence ou quand la statistique trahit la grammaire.

Où est le sens ?

BIBLIOGRAPHIE

- (1) Cottier (J.-M.), Emeri (C.), Gerslè (J.), Morau (R.), *Giscard-Mitterrand: 54 774 mots pour convaincre*, Paris, PUF, 1976.
- (2) Prost (A.), *Vocabulaire des proclamations électorales de 1881, 1885 et 1889*, Paris, PUF, 1974.
- (3) Tournier (M.), *Un vocabulaire ouvrier en 1848. Essai de lexicométrie*, ENS de Saint-Cloud, 4 vol., 1975 (thèse, Sorbonne).
- (4) Lafon (P.), *Dépouillements et statistiques en lexicométrie*, Genève, Slatkine-Paris, Champion, 1984 (thèse de 3^e cycle, université de Paris III).
- (5) Bergounioux (A.), Lannay (M.-F.), Mouriaux (R.), Sueur (J.-P.), Tournier (M.), *La parole syndicale. Etude du vocabulaire confédéral des centrales ouvrières françaises (1971-1976)*, Paris, PUF, 1982.
- (6) Demonet (M.), Geffroy (A.), Gouazé (J.), Lafon (P.), Mouilland (M.), Tournier (M.), *Des tracts en Mai 68. Mesures de vocabulaire et de contenu*, Paris, Presses de la Fondation nationale des sciences politiques, 1975.
- (7) *Méthodes quantitatives et informatiques dans l'étude des textes*, Actes du colloque ALLC de Nice (Association for Literary and Linguistic Computing), Genève, Slatkine, 1986, 2 vol.
- (8) Tournier (M.), « Texte propagandiste et cooccurrences. Hypothèses et méthodes pour l'étude de la sloganisation », *Mots*, n° 11, octobre 1985, pp. 155-187.

La linguistique informatique

L'ordinateur, qui nous aide à comprendre notre propre langage, s'adapte, également, au dialogue que nous tentons d'engager avec les machines.

Jean-Pierre DESCLÉS

La modélisation se fonde sur une théorisation préalable et aboutit à des représentations formelles.

La simulation d'un fonctionnement consiste à implanter dans l'ordinateur des représentations formelles en engendrant des représentations internes à la machine.

La simulation de la compréhension du langage doit être, dans des résultats observables, analogue à ce que l'on observe dans la compréhension par l'homme.

La linguistique informatique est un domaine inter-disciplinaire. Elle se développe sous la pression de la recherche fondamentale et de la recherche appliquée. Dans le premier cas, sa démarche théorique permet de mieux comprendre le fonctionnement du langage naturel. Dans le second cas, une démarche pratique aboutit à une véritable « industrie de la langue » trouvant ses applications dans des domaines comme la traduction automatique, le traitement automatique de la parole, les dialogues homme-machine ou encore les langages quasi naturels de consultation de bases de données.

Modélisation et simulation

En linguistique, l'ordinateur est un outil qui peut être destiné à tester des hypothèses que le chercheur formule à propos du langage. Cette tâche consiste à mettre en œuvre des représentations du langage et à établir des relations entre elles. La démarche implique, au préalable, de modéliser l'objet traité (pour établir des représentations formelles), puis d'engendrer des représentations internes capables d'être, ensuite, traitées par l'ordinateur, par exemple, sous forme de listes.

Donnons deux exemples. La modélisation du mot « chanterons » sera obtenue en le décomposant ainsi : chant + er + ons. Les deux derniers éléments indiquant « le futur » + « la première personne du pluriel ». Quant à la simulation, dans le cas de la phrase « Jean déplace le rocher », la liste sera : (déplace (le rocher) Jean).

Bien sûr, diverses procédures algorithmiques précises permettent, ensuite, d'accéder aux éléments dans la liste afin d'effectuer toutes sortes de manipulations sur la liste et sur ses constituants. Le langage de programmation LISP est très utilisé en linguistique informatique et en intelligence artificielle puisque ses expressions représentent directement des listes et des instructions pour manipuler les listes.

La simulation du fonctionnement du langage naturel n'est effectuée, en général, que pour un ensemble de tâches précises. C'est ainsi, par exemple, que l'on peut chercher à simuler le processus de compréhension du langage.

Mais pour résoudre l'ensemble des tâches qui entrent dans la « compréhension », on élabore, depuis quelques années, des systèmes de représentations sémantiques. Ces systèmes permettent de fournir des schémas abstraits décrivant la signification qui serait commune, par exemple, à une classe de verbes ou de prépositions. De tels schémas sont représentés dans un langage de programmation (comme LISP, cité plus haut).

L'identification des valeurs sémantiques des différentes catégories de mots, leurs relations grammaticales et leurs classements relèvent principalement de la linguistique. En revanche, la constitution de base de connaissances générales ou spécialisées répond à des objectifs différents et très ambitieux. Encore faut-il, pour les atteindre, disposer d'un formalisme capable de représenter les connaissances. L'élaboration de tels formalismes relève beaucoup plus des techniques informatiques.

Le traitement du langage naturel

Comment fonctionne un système informatique de traitement du langage naturel (figure 1) ? Un texte (la séquence d'entrée) est soumis à l'analyseur qui en construit une représentation phonique (s'il s'agit d'un discours oral), morphologique, syntaxique, sémantico-grammaticale et sémantique d'entrée après avoir consulté un dictionnaire. En interrogeant la base de connaissance, le module d'inférence procède alors aux déductions et construit d'autres représentations qui vont engendrer, en sortie, des énoncés. Pour établir des liens grammaticaux entre les phrases (Jean aime Marie ou Marie, Jean l'aime, etc.), les représentations construites à l'entrée sont soumises au module de paraphrase (sans consultation de la base de connaissance) qui construit de nouvelles représentations équivalentes se réalisant par des variations paraphrastiques.

Ce traitement des paraphrases et des ambiguïtés indépendant de toute connaissance externe, reste, pour cette raison, un des objectifs centraux de toute analyse linguistique.

Divers modèles linguistiques de la paraphrase (Z. Harris et S.K. Shaumyan) sont déjà bien élaborés, empruntant leurs formalismes à la logique. Il reste encore beaucoup à faire pour

■ Jean-Pierre Desclés, professeur d'informatique et de linguistique à l'université de Strasbourg II, membre du Centre d'analyse et de mathématiques sociales (CAMS), UM 17, 54, boulevard Raspail, 75270 Paris Cedex 06.

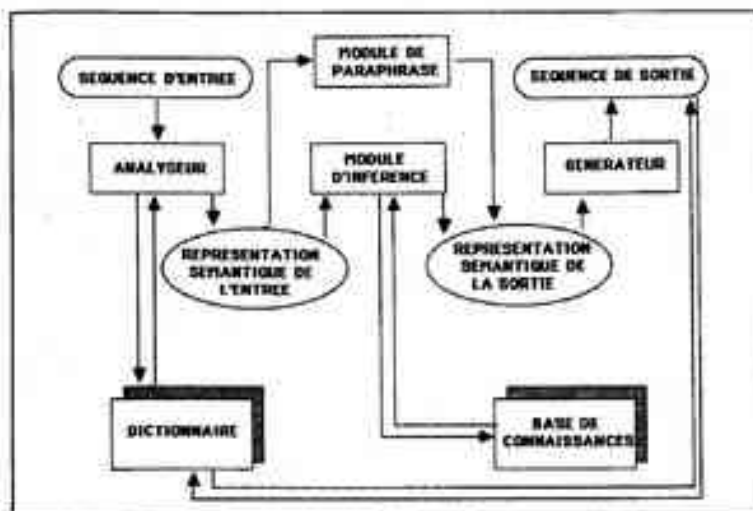


Fig. 1 : Architecture d'un système informatique de traitement du langage naturel.

réaliser une simulation automatique de la paraphrase.

La traduction automatique

Un autre domaine auquel s'intéresse la linguistique informatique est celui de la traduction automatique. Ainsi que le montre la figure 2, la traduction automatique relève d'une architecture adaptée pour le traitement du langage naturel mais comportant, en plus, des grammaires de transfert aux divers niveaux de traitement. Un système « de haute qualité » doit, à la fois, résoudre les problèmes de compréhension et de paraphrase du texte à traduire (texte source). De plus, il doit transférer une représentation en une autre. Ce qui constitue un problème délicat, en particulier pour les transferts grammaticaux d'une langue à l'autre. Par exemple, si les verbes français disposent de trois temps (simples) au passé, les verbes anglais n'en ont que deux. De même, le russe, contrairement à notre langue, ne connaît pas les articles mais il utilise des cas morphologiques.

Pour résoudre ces problèmes de transfert, il faut arriver à représenter, en fonction du contexte, la valeur sémantique exacte des catégories grammaticales de la langue source puis indiquer les procédés pour les exprimer dans la langue cible.

A l'heure actuelle, la traduction automatique se heurte à plusieurs obstacles, en particulier dans l'élaboration de dictionnaires de grande taille (analogues au dictionnaire informatisé du français conçu par le Laboratoire d'automatique documentaire et linguistique), la description des valeurs sémantiques des catégories grammaticales et l'écriture des grammaires de transfert.

Construire des représentations

On peut dire qu'il existe trois stratégies générales pour construire les représentations internes des données linguistiques. La première conduit à utiliser des représentations formelles qui appartiennent à un formalisme, lui-même constitué à partir d'une théorie linguistique déjà bien déterminée. Dans ce cas, le problème revient à définir des représentations informatiques associées aux représentations formelles

et aux données linguistiques. Par exemple, il a fallu adapter aux exigences informatiques les grammaires formelles de N. Chomsky (1956), ce qui fut fait dans les années soixante-dix par divers laboratoires.

La deuxième stratégie de construction des représentations informatiques consiste à adopter directement des implantations informatiques, sans recourir à une théorisation préalable très posée. Cette démarche revient à fournir une interprétation linguistique aux représentations informatiques et à en justifier l'éventuelle pertinence théorique. En 1973, R. Schank s'est engagé dans cette voie et a proposé directement des représentations sémantiques d'énoncés. On peut cependant s'interroger sur sa capacité de généralisation à un grand nombre d'exemples.

A ces deux stratégies s'en ajoute une troisième avec la programmation logique (avec le langage PROLOG). Le groupe de programmation PROLOG, dû au professeur A. Colmerauer (chercheur en informatique à Marseille), est de plus en plus utilisé en intelligence artificielle. Cette dernière stratégie consiste à partir d'un formalisme emprunté à la logique (le calcul des prédicats) puis, d'une part, à chercher à l'adapter à la structure informatique et, d'autre part, à tenter de représenter les énoncés du langage naturel dans le système de la logique empruntée au départ. Des mécanismes de déduction sont intégrés au langage de programmation lui-même, d'où son utilisation pour l'interrogation de bases de données. PROLOG est l'un des langages de programmation qui est retenu dans la cinquième génération des ordinateurs.

BIBLIOGRAPHIE

- Desclés (J.-P.), « Représentations des connaissances : archétypes cognitifs, schèmes conceptuels, schèmes grammaticaux », in *Actes sémiotiques*, Paris, INaLF/URL 7, 1985, t. VII, n° 69-70, 51 p.
- Harris (M.D.), *Introduction to Natural Language Processing*, Reston, Virginia, Reston Publishing Co., 1985.
- Rustin (R.), ed., *Natural Language Processing*, New York, Algorithmics Press, 1973.
- Tennant (H.), *Natural Language Processing*, Princeton, Petrocelli Books, 1981.
- Winograd (T.), *Language as a Cognitive Process*, Reading, Massachusetts, Addison-Wesley, vol. 1 : *Syntax*, 1983.

Le problème des grammaires de transfert.

Trois stratégies.

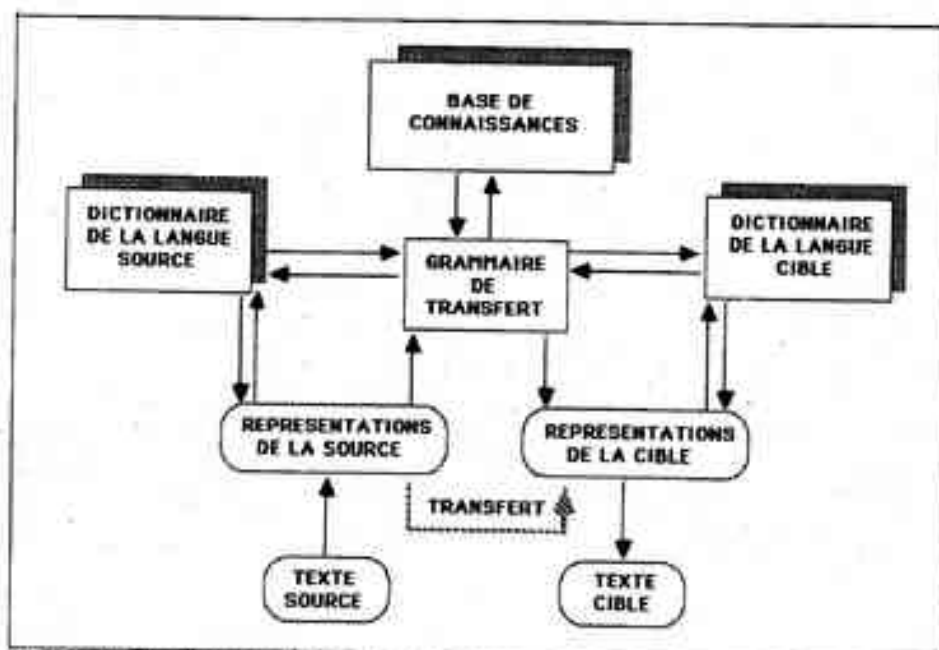
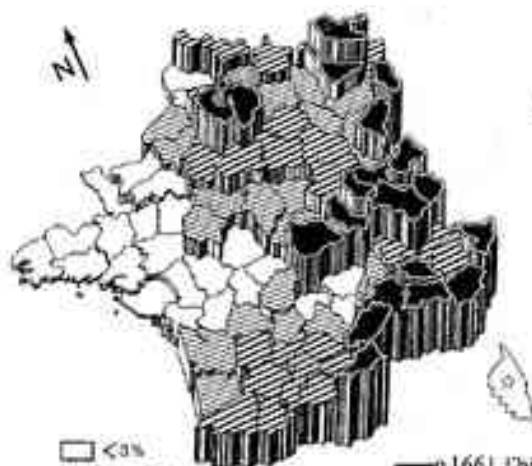


Fig. 2 : Architecture d'un système de traduction automatique.



Proportion d'étrangers dans la population active (recensement de 1982). La hauteur des départements est fonction de cette proportion. Les teintes correspondent aux quatre quartiles de la distribution.

Dès son texte fondateur, la démographie est confrontée à de redoutables problèmes numériques.

Les progrès de l'actuariat au XVIII^e siècle entraînent un grand raffinement dans les calculs.

Les progrès de la collecte au XIX^e siècle permettent de disposer de résultats agrégés exacts.

Deux siècles d'expérience et de progrès du calcul créent une tradition hostile à l'ordinateur.

L'informatique se développe alors aux frontières de la démographie.

Démographie : l'âge de raison

Une tradition hostile à l'ordinateur explique la conversion tardive des démographes mais, depuis 1975, cet outil acquiert petit à petit ses lettres de noblesse.

Hervé LE BRAS

En 1661, l'histoire de la démographie commence à Londres, avec la publication des *Observations naturelles et politiques* de John Graunt. Ce livre original contient déjà une furieuse mêlée de chiffres et de calculs. Erreurs de transcription et erreurs d'addition rivalisent avec les erreurs d'observation : l'assemblage des données forme, en effet, la première difficulté de la statistique démographique. L'ouvrage de Graunt bute sur une seconde difficulté, les calculs, lorsqu'il tente de construire la première table de mortalité. Il contourne l'obstacle au prix de simplifications drastiques. Il faudra attendre trente ans pour qu'Halley s'empare du problème et en triomphe : astronome, il craint moins les calculs que le marchand drapier Graunt.

Les deux sources des calculs démographiques

Dès la fondation, les deux problèmes majeurs de la technique démographique sont donc présents : les procédures d'agrégation des données d'une part, les calculs actuariels et mathématiques d'autre part. Pendant deux siècles, l'histoire de la démographie sera ainsi associée à la résolution de ces deux questions. Au XVIII^e siècle, les calculateurs l'emportent : Simpson, de Moivre, Deparcieux, d'Alembert, Euler et d'autres simplifient les calculs par le choix de lois judicieuses, affinent les concepts et améliorent la disposition des opérations. Au XIX^e siècle, la toute nouvelle statistique officielle réduit progressivement les imperfections de la collecte et de l'agrégation. On renonce encore au recensement de 1805, tant les erreurs d'addition (et les retards des envois départementaux) faussaient les résultats, mais, en 1851, un premier recensement moderne est admirablement réalisé et publié. La qualité des résultats tient alors essentiellement à l'organisation des opérations de collecte et de fabrication des tableaux. Toute une tradition se met en place que les recensements actuels et les sources d'état civil respectent encore malgré l'introduction d'automatismes (classificateurs de L. March inspirés des procédés Hollerith, au début du XX^e siècle, et ordinateurs après 1960). Après ce XIX^e siècle de l'enthousiasme statistique comme l'a appelé Westergaard, la mathématisation de la démographie reprend, soulevant de délicats problèmes de calcul souvent résolus avec une grande imagination (on oublie l'importance qu'a eue la recherche de simplifications numériques dans la découverte

en mathématiques : Lotka, Volterra, Gini, Leslie, puis les premiers polytechniciens de l'Institut national d'études démographiques (INED), jonglent avec des tables de logarithmes, puis des « rouleaux à calcul », sortes de règles à calcul cylindriques.

La tradition face à l'ordinateur

Cette longue habitude du calcul et ces traditions, loin de favoriser l'utilisation des ordinateurs vont, d'abord, constituer un obstacle à son introduction. Pensons à Kepler qui ne voulait jamais utiliser de lunette astronomique, confiant en sa seule acuité visuelle : les démographes des années soixante furent ses émules. La très grande standardisation des données acquise durant le XIX^e siècle, la régularité des calculs et des ajustements, pour lesquels toute une série de procédés ingénieux avait été mise au point, empêchaient l'ordinateur d'être aussitôt compétitif. En d'autres domaines, l'introduction d'images synthétiques et du dessin automatique s'est heurtée à des résistances analogues. Dès lors, les premières utilisations de l'ordinateur en démographie ont été développées aux marges du domaine, presque à son extérieur sous l'impulsion de personnalités originales, elles-mêmes parfois marginales.

Au centre de la discipline, c'est-à-dire en matière d'analyse démographique et de prévisions de population, l'ordinateur n'a été régulièrement utilisé qu'à la fin des années soixante-dix. Ainsi, en 1969, l'INED fit une étude pour remplacer les rouleaux à calcul par des machines électromagnétiques qui étaient capables d'effectuer les quatre opérations (les divisions demandaient près d'une minute durant laquelle on craignait que la machine ne se démantibule tant ses rouages souffraient). Le choix se porta sur une machine de 3 000 F de l'époque (environ 15 000 F actuels) dont dix exemplaires furent acquis. On eût l'impression de s'équiper pour les dix prochaines années, sans imaginer aucunement le phénomène des calculettes et des micros. Bien qu'un ordinateur existât déjà à l'INED, il était impensable qu'il serve à résoudre des problèmes courants, simples, répétitifs. Il semblait destiné au traitement de problèmes mathématiques inaccessibles, tels que l'inversion d'immenses matrices, les ajustements statistiques multivariés ou les simulations.

Et ce fut la résolution de ces problèmes qui entraîna une amélioration substantielle des connaissances démographiques. Dès 1956, S. Ledermann, à l'aide des ordinateurs du CEA, effectuait des analyses en facteurs communs et spécifiques de la mortalité (comme Jöreskog n'avait pas encore découvert sa méthode, de longues approximations étaient nécessaires). A partir de 1966, A. Jacquard et P. Vincent se lancent dans des simulations de Monte-Carlo

■ Hervé Le Bras, directeur de recherche à l'Institut national d'études démographiques (INED), chef du département « Méthodes et prévision », 27, rue du Commandeur, 75675 Paris Cedex 14.

Des progrès sont réalisés dans la compréhension des mécanismes de fécondité grâce aux simulations de Monte-Carlo.

L'analyse multi-variée renouvelle l'idée ancienne de loi démographique en lui substituant la notion de table-type.

À partir de 1975, le domaine traditionnel de la démographie (analyse et enquêtes) se convertit à l'informatique.

Actuellement, plusieurs directions de recherche mises sur l'ordinateur : bases de données historiques, cartographie interactive, reconstitutions géantes du développement des populations.

des comportements féconds qui allaient permettre de dégager les « composantes » de la fécondité et de chiffrer l'effet des politiques de limitation des naissances. Par ces travaux précurseurs, les démographes échappèrent à l'euphorie informatique des années soixante-dix (analyse factorielle - version composantes principales - et analyse des systèmes - version club de Rome -).

En revanche, des recherches élaborées, menées en général en collaboration internationale, poursuivirent les deux directions fondatrices : l'analyse multivariée pour construire des « tables-type » destinées à l'estimation et la prévision de la mortalité ; les modèles de simulation des structures familiales étendus soit aux structures de la parenté (H. Le Bras), soit aux structures et cycles de vie des ménages (P. Laslett et K. Wachter).

Ces recherches, en plein essor, rejoignent souvent les modèles markoviens inhomogènes (tables d'extinction multiple, par exemple) pour lesquels l'ordinateur permet d'effectuer des estimations et d'anticiper des résultats mathématiques. De tels modèles servent, dès maintenant, à analyser les flux interrégionaux de migrations, à prédire l'évolution des familles sous l'effet des changements de mœurs matrimoniales et fécondes (U.-L. Rallu), ou encore à analyser les flux d'activité.

La conversion générale à l'informatique

À partir de 1975, les domaines traditionnels de la démographie qui s'étaient protégés en refusant toute informatisation, s'y hasardèrent timidement, puis avec plus d'ardeur. On put, alors, craindre le pire : travaux répétitifs et sans imagination, enquêtes énormes et insipides dont les traitements les plus longs ne tiraient qu'un maigre jus. Après les physiciens, les démographes découvraient dans l'ordinateur un champ de bataille idéal pour régler leurs conflits de pouvoir. L'arrivée des micros, la multiplication des liaisons, l'accroissement des capacités de stockage et d'accès, notamment sur le centre de l'INED, sont en train de guérir cette maladie infantile.

Tant de progrès et tant de conflits rendent incertaine la prévision de l'évolution future. Quelques tendances se dégagent cependant, et sont largement souhaitées :

- l'accès rapide aux données anciennes et modernes : constitution de banques de données de démographie historique (quarante villages de l'INED, généalogies TRA du Laboratoire de démographie historique de l'EHESS, qui coordonne l'enquête TRA sur les 3 000 familles - voir encadré -), données de la Statistique générale de la France* à Grenoble dans la Banque de données socio-politiques** (accès aux sondages des recensements (millième des ménages) et aux grandes enquêtes (enquête sur la formation et la qualification professionnelle***));

* Statistique générale de la France - SGF - : organisme créé par Thiers, chargé jusqu'à sa disparition durant la Deuxième Guerre mondiale, des traitements statistiques et plus particulièrement des recensements. Ces fonctions ont été reprises et élargies depuis par l'Institut national de la statistique et des études économiques (INSEE).

** Banque de données socio-politiques - BDSP - : (GRIECO-4) du CNRS, ce laboratoire entre autres actions, met à la disposition d'utilisateurs externes des données historiques et contemporaines, sur support magnétique, dans les domaines de la sociologie et des sciences politiques.

*** L'enquête Formation Qualification Professionnelle (EQP) est réalisée par l'INSEE, environ tous les sept ans ; son taux de sondage est variable et s'élève jusqu'à 1/200^e pour les jeunes adultes. Elle permet de connaître la formation, l'origine sociale, l'insertion professionnelle et les associations de l'emploi exercé non seulement à la date de l'enquête mais aussi avec un recul de cinq ans.

L'enquête TRA

L'objectif est d'étudier la mobilité géographique et sociale en France, de 1800 à nos jours, en reconstituant la généalogie descendante, en ligne masculine, de 3 000 couples repérés sous le Premier Empire et répartis sur tout le territoire national, proportionnellement à la population présente, dans les limites actuelles des départements, au recensement de 1806.

L'enquête est menée avec le concours bénévole des généalogistes, sous la responsabilité de correspondants départementaux.

Dans un premier temps, on a relevé dans les tables décennales de la période 1803-1832, mention de tous actes concernant des individus dont le patronyme commence par le groupe de lettres TRA (groupe choisi après consultation d'experts, dans le but de n'exclure aucun groupe ethnique, y compris les étrangers fixés en France).

Ensuite, on reconstitue les généalogies patronymiques descendantes des familles de l'échantillon, avec couverture d'une fiche de famille pour chaque couple.

Par la suite, on échangea par département à département, les informations concernant les migrants.

Des fiches de famille, mises anonymement en ordinateur, on tirera des statistiques entièrement nouvelles intéressant l'histoire de la population française, sa mobilité, les processus de l'exode rural, les structures sociales, le renouvellement des élites, etc.

- la visualisation interactive : les tableaux de données et de résultats sont arides à lire et défient la synthèse. Pour les comprendre et les faire comprendre, les représentations graphiques et cartographiques (abstraites ou concrètes) forment un remarquable outil. Les recherches portent autant sur les techniques que sur leur signification (sémantologie) ;

- les reconstitutions géantes : il s'agit de reconstituer, année par année et âge par âge, la population d'une ville, d'une région voire d'un pays entier, dont on ne connaît que quelques traces statistiques, séries de décès, recensements très espacés et incomplets, dénombrements de certaines catégories, etc. Pour y parvenir, les modèles mathématiques de la dynamique des populations sont estimés statistiquement.

Au total, l'âge de raison semble atteint. L'ordinateur n'exerce plus un attrait ou une répulsion immodérés. Sans qu'ils en aient toujours conscience, il a libéré les démographes de leur plus lourde contrainte.

BIBLIOGRAPHIE

Biraben J.-N., Bonneuil N., « Population et économie au pays de Caix aux XVI^e et XVII^e siècles », *Population*, 1986, n° 6 (à paraître).

Jacquard A., Hodmer W.F., « La variance de la dimension des familles selon divers facteurs de la fécondité », *Population*, 1968, pp. 869-879.

Le Bras H., « Evolution des liens de famille au cours de l'existence », *In Les âges de la vie*, Paris, PUF, 1983, pp. 27-43.

Ledermann S., *Nouvelles tables-type de mortalité*, Paris, PUF, 1970.

Rallu U.-L., Festy P., « Constitution et reconstitution des familles françaises », *Population*, 1981, pp. 63-92.

Wachter K., Laslett P., Le Bras H., *Statistical studies of historical social structure*, New York, Academic Press, 1979.

Le numéro 51 (1984) de *Histoire moderne et contemporaine. Informatique* contient une majorité d'articles consacrés à l'application de l'informatique en démographie.

Sciences économiques : quand l'outil valide la parole

Les sciences économiques peuvent attendre beaucoup de l'ordinateur. Encore faut-il que les données fournies soient fiables, que les calculs de simulation puissent valider les théories et, surtout, qu'ils permettent, dans la pratique, une aide à la décision.

Jean BOURDON

Pour les sciences économiques, le recours à l'outil ordinateur permet de prendre en compte plus d'informations donc de mieux s'approcher des interactions complexes du réel.

L'annonce d'une baisse du prix de l'essence conduit chaque automobiliste, connaissant peu ou prou la capacité du réservoir de son véhicule, à simuler l'économie qu'il réalisera lors du prochain passage à la pompe. L'expérience aidant, il prend en compte l'effet des variations du dollar sur son budget carburant.

Cet exemple peut apparaître futile, en fait, il montre la grande utilité du calcul automatique pour la simulation économique. Car si nous désirons poursuivre notre exemple, il faudrait prévoir le cours du dollar et la dimension du problème serait tout autre, tant en informations à prendre en compte qu'en terme d'écriture de systèmes d'équations reliant plusieurs centaines de variables. D'où le nécessaire appel au calcul automatique.

Dans ces conditions, l'utilisation de l'ordinateur en sciences économiques n'aurait aucun caractère original empruntant au calcul numérique et aux diverses méthodes statistiques (voir l'article de J.-M. Legay, « Qu'est-ce que la biométrie ? », *Le Courrier du CNRS*, n° 64, janvier-mars 1986, pp. 56-61).

Plus original est, en revanche, le renouvellement de méthodes pour les sciences de l'économie qui accompagne cette utilisation et les font progressivement échapper au qualificatif de « sciences des mondes de l'imaginaire ».

En permettant une multiplication des possibilités de calcul, l'ordinateur peut-il dans les multiples interprétations des faits, d'une science pleine de contradictions, trier celles qui rendent le mieux compte des événements ? L'utilisation du calcul automatique peut-elle assurer un certain mouvement dans les pratiques de la discipline, du deductif vers l'inductif ? Les hypothèses de base de l'économie sont-elles transposables sans risque par les outils de calcul disponibles ? Les données traitées sont-elles fiables et l'information qu'elles contiennent est-elle pertinente ? Etc.

Ces questions restent actuellement posées et ceci pour bien longtemps. Aussi, le principal apport du calcul automatique pour les sciences économiques peut-être une remise en cause méthodologique, minimisant le système de pensée théorique préexistant pour privilégier une démarche heuristique. A une pratique deductive peut se substituer une démarche scientifique où le rôle des hypothèses s'efface par rapport à la capacité de cette démarche à expliquer, et surtout à prévoir, les faits économiques.

Cette capacité nécessite que la mesure des phénomènes économiques par les données soit

fiable pour que le calcul automatique ne crée pas une illusion de précision. Mais elle nécessite aussi deux autres conditions : les calculs de simulation doivent permettre de valider une théorie par rapport à une autre et, surtout, ils doivent conduire à des applications hors recherche, pour l'aide à la décision ; ces trois points seront discutés tour à tour. Remarquons d'entrée que ces points appartiennent aux questions de l'information : sélectionner l'information, la traiter correctement et savoir la diffuser.

Le terrain et les données : le mythe de la précision

Alors qu'elle restait en partie paralysée par ses conflits internes entre corps de pensées antagonistes, l'économie a pu disposer d'un matériel statistique plus adéquat à son sujet d'étude. Matériel venant du rôle croissant de la puissance publique dans les rouages de l'économie, donc du désir de mieux connaître les faits pour assurer la régulation des grands équilibres (emplois, commerce international, prix).

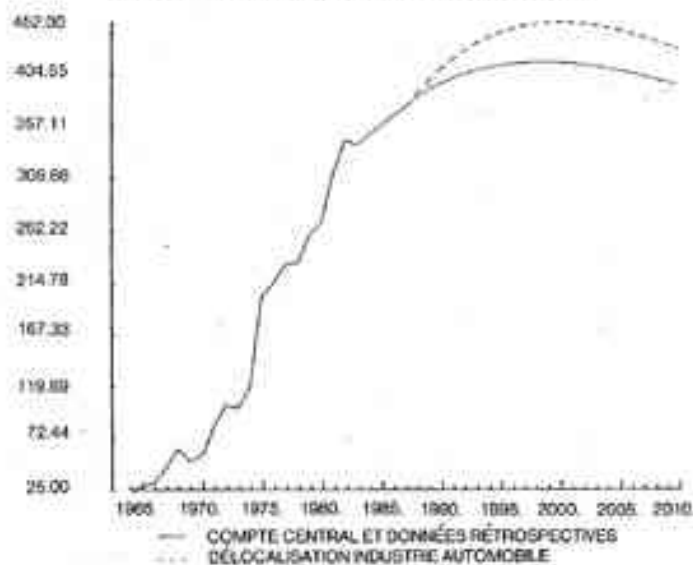
Dans l'ensemble des sciences sociales, l'accès au calcul automatique a surtout permis le traitement de lots de données de plus en plus conséquents. Face à cette constatation, il faut se demander si la démarche de l'économie se rattache à un terrain. En fait, ce terrain existe, mais pour l'essentiel il n'est pas directement observable par le chercheur : il lui parvient à travers un système statistique où compteable qu'il ne maîtrise pas. L'initiative de ce système revient à des organisations spécialisées qui élaborent des statistiques pour des besoins administratifs ou de gestion courante, donc éloignés du besoin d'information pour la recherche. Cette question statistique interfère largement avec l'utilisation de l'outil informatique, ceci d'un double point de vue :

- l'utilisation d'une donnée économique pour la recherche est essentiellement une utilisation secondaire. Les grands fichiers de données ont pour origine des obligations légales de recensement (recensement de la population), des obligations administratives ou fiscales (statistiques douanières, comptabilité des entreprises). S'il est certain qu'une grande partie de cette information ne peut être dissociée des outils de suivi de la politique économique, l'utilisation au niveau de la recherche n'a plus de rapport avec l'information qui était recherchée lors de la décision administrative de création de l'outil statistique. Ainsi, les données les plus détaillées sur l'évolution, en France, des types d'emploi et des qualifications viennent d'enquêtes liées au suivi des obligations d'emploi, par les entreprises, de travailleurs handicapés ;

- le couple informatique-statistique induit souvent une illusion de précision. Les contraintes

■ Jean Bourdon, chargé de recherche au CNRS, Centre de documentation sciences humaines (CDSH), 54, boulevard Raspail, 75270 Paris Cedex 06.

Impact d'une politique industrielle sur le chômage en Ile-de-France, application d'un modèle (graphique obtenu par le logiciel Ariane du Centre interrégional de calcul électronique du CNRS-CIRCE).



Ce graphique donne le résultat d'un changement de politique économique réalisée avec un modèle de la région Ile-de-France. Le graphique présente en trait continu les valeurs observées jusqu'en 1984 et la simulation centrale jusqu'en 2010, cette simulation centrale est par hypothèse celle dont on peut estimer la réalisation possible si les conditions restent identiques à leur état présent. La variante (représentée en traits discontinus) donne la mesure d'un impact, ce dernier correspond ici à une politique de délocalisation, hors de la région Ile-de-France, des principaux centres de production d'automobiles entre 1985 et 1995. Le chômage est, ici, traduit par les demandes d'emploi non satisfaites (DENS) avec pour unité le millier. Pour plus de détails, cf. F. Bourdon et J. Bourdon, « Un modèle macroéconomique de l'Ile-de-France », Cahiers d'économétrie appliquée, 1985, n° 1 « Espace, régions, santé ».

L'illusion de précision du chiffre encouragée par le calcul automatique conduit à dangereusement minimiser les questions d'obtention et de contenu exact des données.

La modélisation est vite apparue comme le point de passage obligé pour sélectionner l'information adéquate par rapport à des références théoriques exigeantes.

Ceci conduit à un débat sur la finalité de la recherche économique appliquée. Doit-elle être appréciée par rapport à la justesse de ses prévisions indépendamment de tout contexte théorique, ou à l'inverse doit-elle se limiter à la validation d'un cadre théorique préalable ? La réponse ne peut être qu'une perpétuelle interaction entre ces deux options.

du droit imposent aux comptabilités d'être tenues au centime près. Ainsi, pour une firme comme Electricité de France dont le chiffre d'affaire s'exprime en dizaines de milliards de francs, la précision illusoire serait de l'ordre du rapport 10^{-12} , soit à un niveau comparable aux meilleures mesures de l'étalon temps dans les sciences physiques. Pour les comptes de la Nation* l'ambition comptable est moindre, l'unité est en général du million de francs, soit un ordre de précision de 10^{-6} puisque le « Produit national » en France est de l'ordre de 10^{12} francs.

Hélas, même ces normes de mesure sont trop ambitieuses et si les variations d'une période à l'autre sont assez bien mesurées, on estime que l'imprécision sur les niveaux de certaines variables est de l'ordre de 1 à 5 %. De plus, la « donnée économique » n'est pas obtenue par une mesure unique, elle s'élabore lors d'un processus complexe de traitement de l'information. Il se crée ainsi un processus « vertueux » : le calcul automatique permet d'obtenir des données précises qui elles-mêmes seront plus facilement utilisables dans des travaux d'analyse qui conduisent eux-mêmes à une meilleure couverture statistique.

La donnée économique répond à plusieurs fonctions qui tendent à contraindre l'activité de recherche, mais surtout la donnée est elle-même une représentation subjective de la réalité qui ne peut être directement observée pour l'ensemble du domaine couvert.

Aussi l'utilisation du calcul économique tend à estomper la frontière dans la recherche entre l'observé et le simulé, puisque la donnée est elle-même un modèle, ce qui est source d'ambiguïté (M. Barbut, 1985).

* Les termes marqués d'une astérisque sont définis dans le glossaire.

• Les modèles économiques

Associer modélisation et pratique du calcul automatique apparaît trivial, tant cette forme d'étude sur maquette est désormais dépendante du calcul automatique. Toutefois, le rôle du modèle dans l'utilisation de l'ordinateur est particulièrement éclairant pour les sciences de l'économie.

On aurait pu croire que le modèle pour les sciences économiques servirait d'élément réducteur pour les différents corps de pensée. Pourtant, même s'ils utilisent un lot de données commun, les modèles économiques se différencient largement. On rejoint ici le structuralisme, le modèle économique est l'image d'une réalité « transformée » (irréalité formelle au sens de Cl. Lévi-Strauss) pour être conforme au corps de pensée auquel il se rattache.

En profitant des progrès du calcul automatique, les modèles économiques appliqués apportent l'immense avantage, dans un domaine scientifique où l'expérience réelle est très limitée de pouvoir offrir des scénarios alternatifs, en permettant un cadre de cohérence rigoureux. Le modèle apparaît ainsi l'outil privilégié de la science économique en aidant à réaliser ses deux objectifs essentiels : la prévision et l'évaluation des politiques.

Mais cette utilisation n'est pas sans objet de contestation. Celle-ci ne vient pas essentiellement d'une insuffisance des outils d'analyse. Si complexes soient ces derniers, la « boîte à outil » du calcul numérique offre des réponses et n'est pas un facteur de blocage.

Ainsi, une des avancées marquantes de ces dernières années en économie est la théorie des anticipations rationnelles* : ses applications et vérifications se trouvent largement contraintes par la nécessité d'intégrer dans les procédures d'estimations statistiques des systèmes dynamiques non linéaires complexes (D.K. Begg, 1982). De même, pour les formulations en terme de déséquilibre*, les difficultés pour définir l'algorithme d'optimisation font que les applications sont encore excessivement rares, mais très prometteuses.

La question essentielle est d'ordre épistémologique. L'introduction du calcul automatique pose la cohérence de l'objet d'analyse et des buts poursuivis (modèle économique) par rapport à l'ensemble des descripteurs possibles (données disponibles). On touche ici un point d'importance rattaché à la philosophie des sciences : le modèle est-il une vérification gratuite d'un corps de pensée ou permet-il d'infirmer (au sens de K. Popper) une théorie ? Le débat est très actuel dans les sciences économiques, une synthèse en est donnée par M. Blaug (1982). Il présente deux visions assez antagonistes :

- a - les outils d'analyse doivent être sans cesse améliorés, ce que permettent les progrès du calcul automatique, la validation des données et des outils doit conduire aux meilleures prévisions. Il s'agit d'une démarche « déontologique » où l'induction domine ;
- b - les données utilisées par la science économique resteront toujours imparfaites de même que les structures de modèle. De ce fait, le calcul ne peut valider une théorie. Seule compte, *ex post*, l'infirmité d'une théorie par rapport à la prévision qui en est issue. Cette affirmation avancée par M. Friedman (1953) a pris une importance considérable avec la pratique du calcul informatique dans la recherche économique.

• L'informatique comme outil d'aide à la décision

Une activité d'étude économique aussi centrale que le choix des investissements use de la

La facilité d'accès de plus en plus grande au calcul automatique, conduit à discerner de plus en plus difficilement ce qui est recherche de ce qui est action d'aide à la décision.

Le spectaculaire développement des outils de bureautique est une chance pour la valorisation de la recherche économique et de gestion, cela en intégrant les règles de connaissances acquises par la recherche dans ces outils d'aide à la décision.

recherche opérationnelle, technique associée depuis longtemps à l'ordinateur puisque s'étant développée en parallèle avec celui-ci.

Même si une vision rigoriste conduit à placer les développements en recherche opérationnelle dans le cadre de l'analyse numérique, certaines extensions sont directement liées aux applications économiques. Aussi, la pratique des analyses en terme de dualité est liée aux applications économiques de la recherche opérationnelle; de façon identique, l'optimisation en nombre entiers et le contrôle optimal sont des techniques en partie développées autour d'une formulation économique de l'optimisation.

Malgré les résultats obtenus et même si le recul du temps est trop faible pour qu'on soit assuré de la conclusion, on remarque une certaine divergence entre recherche économique et calculs d'optimisation. Peut-être que les questions difficiles posées par ces traitements, à la théorie économique (détermination de la valeur en particulier), sont-elles la cause de cet état.

Une place particulière doit être réservée aux systèmes interactifs d'aide à la décision (SIAD). Employés pour des simulations de gestion d'entreprises, ces systèmes sont de véritables « boîtes noires » qui réalisent diverses fonctions, telles que traitement de données, interrogations assistées de fichier, calculs de simulation.

Leur expression la plus simple est constituée par le « tableur », dont la pratique est maintenant commune aux utilisateurs de micro-informatique. Parfois, ils peuvent intégrer des modules qui les rendent très proches, d'un point de vue technique, des outils de modélisation économique. Mais ici, la question centrale pour le chercheur est la mise au point d'un outil incorporant des fonctions de dialogue et de formation de l'utilisateur, car ici le chercheur n'est plus l'utilisateur final. Aussi les questions du dialogue homme-machine, du vocabulaire qui sert pour ce dialogue sont ici centrales.

Le produit de la recherche, la « boîte noire », doit s'adapter aux règles et associations de connaissances suivies par le gestionnaire d'en-

treprise/utilisateur. On est alors très près des constructions de base de connaissances de l'intelligence artificielle (IA). La multiplicité des cas d'espèce et la complexité des mécanismes décrits font que les réalisations concrètes en ce domaine sont encore à venir; toutefois, il est évident que dans l'ensemble, recherche en économie et gestion, l'IA obtient les résultats les plus probants sur ces systèmes d'aide à la décision U. Trahand et D. Hopen, 1985).

Cette situation ne va pas sans poser des problèmes pour les sciences de l'économie et de la gestion. Alors que dans un passé encore proche, une grande partie du temps de la recherche était liée à des activités de développement logiciel, le chercheur obtient actuellement les produits directement utilisables du marché de la bureautique (gestionnaires de banque de données relationnelles, aide à la création de systèmes experts). Il existe donc des choix à faire dans la division des tâches, si l'on ne désire pas réinventer la roue.

Même si ce texte pose plus de questions qu'il ne donne de réponses, trois niveaux de conclusion sont nécessaires:

- les sciences de l'économie utilisent largement l'outil « ordinateur » et cela dans des voies de recherche diversifiées, ce qui est un témoignage de vitalité;

- l'introduction du calcul automatique a surtout comme conséquence de second ordre d'entraîner une interrogation sur les statistiques et le mode de validation des théories. A ce propos, E. Malinvaud précise: « L'utilisation de techniques qui étaient autrefois difficiles à mettre en œuvre ne semble plus poser maintenant aucun problème majeur de calcul... »

Le retard de nos connaissances n'est pas lié à un retard des méthodes économétriques mais tantôt à un retard de l'information factuelle de base, tantôt à un retard dans le développement des théories.

- une interrogation reste entière quant à l'impact sur l'organisation de la recherche. Pour la recherche économique, on peut penser que les tâches liées au calcul automatique de diffusion, de validation et de conservation des données sont aussi fondamentales que les activités liées à la « mesure » dans les sciences physiques, activités qui justifient souvent un « grand équipement ». Alors, peut-être, la recherche économique, ainsi équipée, ne sera plus considérée comme une activité en partie ludique? Mais ceci est une autre histoire.

Glossaire

Comptes de la Nation (ou comptabilité nationale): retracent les transactions économiques entre agents. Par exemple, la transaction impôt qui va de l'agent « ménage » à l'agent « administrations ».

Dualité: conduit à interpréter un problème en permettant les critères et les contraintes, par exemple, passer d'un coût minimum pour une recette donnée à une recette maximale pour un coût donné.

Tableur: c'est un logiciel d'aide à la décision qui permet de manipuler des tableaux de données en définissant des opérateurs sur une case ou tant plusieurs cases entre elles.

Théorie des anticipations rationnelles: suppose laquelle la décision des acteurs économiques ne se réfère pas à un état présent, mais sur l'anticipation du futur en tenant compte du délai de réalisation de l'action à la décision.

Théorie du déséquilibre: montre qu'un équilibre général du système économique peut se réaliser avec des situations de profonds déséquilibres sur des marchés particuliers. L'inflation tendue sur le marché des loyers peut donc coexister avec un chômage durable (dépression sur le marché du travail).

BIBLIOGRAPHIE

Barbut (M.), « Ambiguïté et statistiques: quelques remarques », *Brèves*, octobre 1985, n° 7, pp. 41-44.

Begg (D.-K.), *The rational expectations revolution in macroeconomics*, Baltimore, John Hopkins U.P., 1982.

Blaug (M.), « La méthodologie économique », *Economica*, Paris, 1982.

Deleau (M.), Leras (C.), Malgrange (P.), « Stabilisation efficace des systèmes économiques en présence d'incertitude: expérimentation avec une maquette du modèle DMS », *Revue économique*, mai 1984, vol. 35, n° 3.

Friedman (M.), *Essay on positive economics*, Chicago, Chicago U.P., 1953.

Lenoir (R.), « Rapport au président de la République sur l'information économique », Paris, La Documentation française, 1980.

Malinvaud (E.), « Conclusions du colloque « Emploi-chômage », Dijon, mars 1984, in *Actes du colloque*, Vitry (D.) et Maréchal (B.) eds., Dijon, Cahiers de l'IME, 1984, n° 28.

Trahand (J.), Hopen (D.), « Les générateurs de systèmes d'aide à la décision pour la planification budgétaire et financière », *Economie et sociétés*, coll. « Sciences de gestion », 1985, n° 6, pp. 201-227.

Les nouveaux outils de l'archéologie

De l'ère de la gestion à celle de l'exploitation des données

Henri DUCASSE

Les réseaux peuvent être un élément de diffusion.

Les laboratoires du CNRS font aussi de la conception de logiciels.

La documentation et la bibliographie informatisées comme aides à la connaissance.

C'est un fait, depuis quelques années, l'emploi de l'ordinateur en archéologie se confirme et s'intensifie. Après l'informatique de gestion de données, apparaît, aujourd'hui, celle des traitements. C'est clair, l'archéologue ne se contente plus d'une machine à dénombrer (combien y a-t-il d'objets de type « x » dans mon fichier ?), mais désire un outil de test d'hypothèses, de réflexion et de décision, qui utilise aussi bien les moyens de l'analyse de données que ceux de la représentation spatiale.

L'explosion de la micro-informatique permet ce développement. Que ce soit pour les données de fouille ou de laboratoire, la constitution d'un fichier informatisé ne pose plus de problème. Ce qui demeure, c'est la difficulté à maîtriser les analyses fonctionnelle et organique nécessaires à un bon choix des données et des moyens informatiques. De même, la réflexion sur les objectifs de l'informatisation, en partie contenus dans ces analyses, reste actuelle parce qu'elle conditionne l'avenir de toute application.

L'informatique appliquée

L'ordinateur, par le biais de l'informatisation des méthodes, apporte beaucoup à l'archéologie. Grâce à l'analyse des données, une réflexion à partir de modèles raisonnés est possible : cela se voit en céramologie, en ostéologie et, en général, dans toute étude du matériel archéologique. La représentation spatiale, avec l'atout sélectif, permet de visualiser des situations globales ou particulières. Des méthodes plus sophistiquées, empruntées à l'intelligence artificielle, comme l'emploi des systèmes experts, montrent l'étendue du champ d'application.

Les applications sont souvent réalisées avec des logiciels conçus par des laboratoires du CNRS (propres ou associés) : je citerai Microbase (Philippe François, Groupe de représentation et traitement des connaissances - GRTEC - Marseille), Datavision (Maurice Roux, CNRS, Montpellier), Pending (Pierre-Jean Texier, Unité de recherches archéologiques - URA 28 - Valbonne), Texto (Cheridata-CNRS, Lyon), Cenirar (G. Lambert, O. Buschenichutz, CNRS-URA 33), Sox (Centre de recherches archéologiques - CRA -), etc. La plupart de ces logiciels sont implantés sur des micro-ordinateurs.

Mais l'archéologie sans bases de connaissances n'avancera pas. La constitution de banques de données bibliographiques et documentaires permet l'accès à un fond général ou spécialisé. Outre une partie de Francis (CNRS-CDSH), il existe en archéologie la banque de données Frantig, née de la fusion des fichiers de plu-

sieurs bibliothèques (voir encadré 1). Dans le domaine de la documentation, le *Répertoire de l'archéologie française (RAF)*, informatisé depuis cette année, apportera toute information documentaire sur les institutions de la communauté archéologique (voir encadré 2). D'autres applications, comme l'informatisation de la cartothèque du CRA, témoignent de l'enjeu de la diffusion de l'information.

L'archéologie peut espérer plus que l'informatisation particulière ou locale. Elle peut devenir nationale ou couvrir le champ opératoire d'une micro-discipline sous la forme de réseaux : nous l'avons vu avec Frantig, pour les données bibliographiques, nous le verrons bientôt en archéo-zoologie (voir encadré 3). C'est dire l'ouverture apportée par l'informatique dans cette discipline, qui permettra un jour la confrontation ou la fusion des données. En micro-informatique, signalons l'utilisation d'un

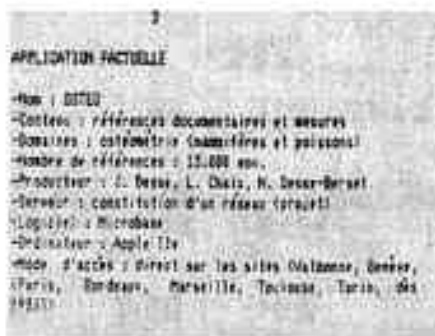
1. RÉPERTOIRE INFORMATISÉ

Nom : FRANTIG (Fichier de Recherche automatisé sur l'ANTIquité)
- Contenu : références bibliographiques, localisation, thèmes disciplinaires, indexation.
- Domaines : pré- et protohistoire principalement européenne et méditerranéenne, Egyptologie, Archéologie et histoire du Proche-Orient ancien, Archéologie et civilisation pré-classique et classique, Archéologie et histoire de la péninsule ibérique, Archéologie et civilisation brésilienne, Patristique, Sciences appliquées à l'archéologie.
- Nombre de références : 40.000 env.
- Producteurs :
26 Mars de l'Orient Méditerranéen (Lyon)
26 Mars des Pays Ibériques (Bordeaux)
Centre Pierre Paris (Grenoble)
26 Techniques Humaines et Sciences de l'Homme (Besançon)
Centre d'histoire Ancienne (Besançon)
Centre de Recherches Archéologiques (Valbonne)
- Serveur : SUDIST (Serveur Universitaire National pour l'Information Scientifique et Technique)
- Logiciel : Texte
- Ordinateur : Bull DPS Omicron
- Mode d'accès : Micro

2. DOCUMENTATION INFORMATISÉE

Nom : RAF (Répertoire de l'Archéologie Française)
- Contenu : références documentaires sur les institutions archéologiques
- Domaines : Écoles, laboratoires, musées, associations
- Nombre de références : 2000 env.
- Producteur : DM
- Serveur : en instance
- Logiciel : Texte
- Ordinateur : Bull DPS Omicron
- Mode d'accès : en direct, par le DM

■ Henri Ducasse, chargé de recherche au CNRS, Centre de recherches archéologiques (CRA), I.P. 7520, Sophia Antipolis, 06565 Valbonne Cedex.



Les applications informatiques, en dépit de quelques restrictions techniques, permettent une exploitation rapide et rigoureuse des données, mais l'homme doit encore interpréter les résultats.

micro-ordinateur-serveur; par exemple, sous le système Vidos, il est possible de consulter un fichier de datations, en se connectant simplement avec un Minitel.

Les limites

Cependant, nous ne pouvons pas passer sous silence les difficultés actuelles, les unes qui sont dues aux incompatibilités des matériels et des logiciels, les autres qui limitent la portabilité to-

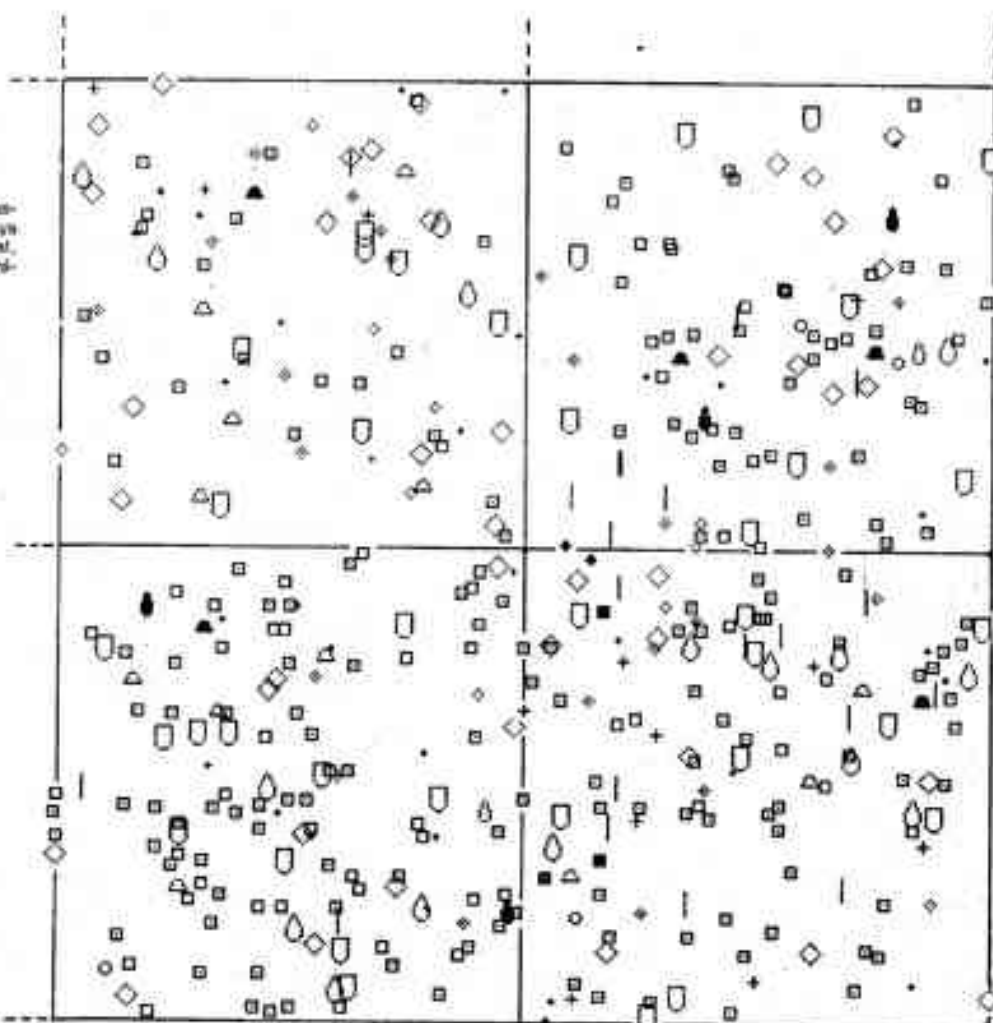
tales des données, malgré les normes de standardisation qui se développent. Nous sommes dans une phase de transition: l'archéologue de demain sera certainement mieux loti que les pionniers d'hier et d'aujourd'hui.

D'autre part, l'ordinateur a des limites; s'il a notablement modifié les mentalités (par l'imposition de sa rigueur, et cela se voit dans la formalisation même des données), il n'en reste néanmoins qu'un outil d'aide à la connaissance. Par exemple, un logiciel d'analyse de données réalisera toujours un graphe factoriel, même sur des données erronées, mais ne donnera pas une interprétation archéologique d'une situation; c'est à l'archéologue - et c'est heureux - de savoir quelles sont les données pertinentes qui lui permettront d'interpréter, après l'étape automatisée, une représentation factorielle, qui n'est, rappelons-le, qu'une façon de « voir » autrement les objets étudiés.

Malgré ces derniers constats, réalistes, nous avons toutes les raisons d'être optimistes. Les expériences actuelles seront les leçons pour demain. L'archéologie, avec l'aide de l'informatique, doit continuer sur la voie de la recherche de la connaissance. Elle en a aujourd'hui le pouvoir, il faut lui en donner les moyens. ■

Police de symboles créée pour l'illustration des plans et coupes d'Itenya (Kenya), d'après H. Roche et al., 1986, « Rapport de la mission préhistorique au Kenya ».

OBJET(S)	SYMBOLE
petit éclat	□
grand éclat	◻
éclat particulier	•
petit nucléus	◇
grand nucléus	◊
percuteur ou balle	○
petit outil retouché	△
biface	◊
hachemau	◻
biface hachemau	◻
fragment pièce biface	◊
médaillon	◊
pièce taillée	◊
bloc	+
os du crâne	•
dent ou mandibule	•
vermillon, coque	•
ceinture ou membre	—
phalange ou métapode	—
esquille	—



Répartition au sol, sur 4 m², des vestiges archéologiques du niveau VI de la fouille d'Itenya (Kajindo-Kenya). Document réalisé par P.-J. Texier (URA 28 du CRA du CNRS), auteur du logiciel PENDIAG. Mission préhistorique au Kenya.

Archéologie et calculateurs : vers une archéologie théorique

Où l'analyse du savoir archéologique rejoint, l'informatique aidant, l'archéologie du savoir réclamée par Michel Foucault dans les sciences de l'homme et de la société.

Jean-Claude GARDIN

Les applications de l'informatique à l'archéologie ont conduit à des recherches fécondes sur les raisonnements que nous pratiquons pour interpréter les vestiges du passé : l'on s'efforce d'élucider les systèmes de représentation et les règles d'inférence qui sous-tendent ces raisonnements, au-delà de leur formulation en langue naturelle.

Il y a exactement trente ans, l'Institut français d'archéologie de Beyrouth, qui dirigeait Henri Seyrig, faisait paraître une brochure intitulée *Le Fichier mécanographique de l'outillage* : on y décrivait une banque de données archéologiques avant la lettre, concernant les « outils de l'âge du bronze des Balkans à l'Indus », où le rôle de l'ordinateur était tenu par une trieuse mécanique. L'innovation principale, en vérité, n'était pas la machine, quelle qu'elle fût, mais bien plutôt la recherche d'une façon de décrire les matériaux archéologiques qui palliât certaines impropriétés des langues naturelles pour les besoins de la documentation automatique. L'on mettait ainsi l'accent sur ce que l'on nomme aujourd'hui les problèmes de représentation (fig. 1), à propos des catégories d'objets les plus variées : architecture, céramique, monuments figurés, outillage, etc. (2, pp. 75-94). La tentation était forte, dès lors, d'appliquer aux bases de données ainsi constituées les méthodes mathématiques de l'analyse des correspondances, de manière à faire apparaître des configurations formelles justifiables, espérait-on, d'interprétations variées : groupes typologiques, filiations temporelles, parentés stylistiques, schémas diffusionnistes, etc. Représentation et calcul, telles furent les deux catégories de problèmes soulevés par les premières applications de l'informatique en archéologie et en histoire de l'art : on les trouve l'une et l'autre constamment évoquées dans les Actes du colloque international organisé par le Centre d'analyse documentaire pour l'archéologie (CNRS) en 1969, sur le thème « Archéologie et calculateurs », avec ce sous-titre éloquent : « problèmes sémiologiques et mathématiques » (1).

De l'analyse du discours archéologique aux systèmes experts

Quinze ans après, cette dualité est toujours sensible : on la reconnaît dans des expressions comme « représentation et traitement des connaissances », dont l'une des équipes françaises historiquement issues du Centre d'analyse documentaire pour l'archéologie a fait son intitulé (Groupe représentation et traitement des connaissances, CNRS, Marseille). L'originalité des recherches conduites depuis quelques années sur ce sujet, en archéologie, tient essentiellement à deux orientations nouvelles :

■ Jean-Claude Gardin, directeur de recherche au CNRS, directeur d'études à l'École des hautes études en sciences sociales (EHESS), 23, rue du Maroc, 75940 Paris Cedex 19.

A) en premier lieu, une attention accrue portée aux écrits de cette discipline tels qu'ils se présentent *in vivo*, sous forme d'articles ou de livres imprimés, dans le but d'en dégager les deux composantes ci-dessus : systèmes sémiologiques de représentation, d'une part, opérations logiques ou mathématiques, d'autre part, constituant ensemble une « construction » ;
B) en second lieu, le recours aux systèmes experts pour simuler ou reproduire des constructions comparables, sur ordinateur, à propos d'objets ou de monuments quelconques. Les applications dans l'une et l'autre voie sont encore peu nombreuses ; elle suffisent néanmoins à soulever des questions d'épistémologie qui constituent à nos yeux l'intérêt principal de ces recherches, comme nous essaierons de le faire comprendre par l'évocation de quelques cas.

Commençons par les travaux d'informatique proprement dit, touchant l'emploi de systèmes

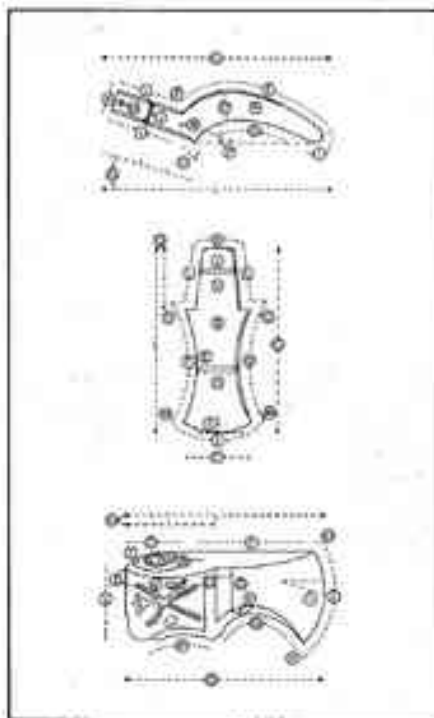


Fig. 1 : Un des premiers exemples (1956) des systèmes de représentation d'objets archéologiques (ici armes et outils), dans une perspective informatique. Les lettres désignent les différentes variables pour l'analyse des formes. Dessin : Institut français d'archéologie de Beyrouth.

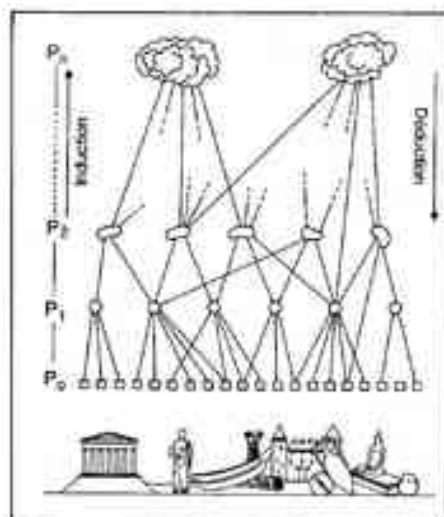


Fig. 2 : La schématisation des raisonnements archéologiques : suite d'opérations $P_0 \rightarrow P_1 \rightarrow P_2 \dots P_i \rightarrow P_j$ reliant des propositions descriptives (P_0) à des thèses interprétatives (P_n), dans le sens ascendant (induction) ou descendant (déduction).

experts en archéologie (voie B). Un livre, sous presse, présente plusieurs applications de cet ordre, dues à des auteurs différents (4). Les matériaux archéologiques considérés sont des plus divers, comme aussi les régions et les époques : figurines chypriotes de l'âge du bronze, monnaies et poteries des royaumes grecs fondés par Alexandre en Asie centrale, amphores romaines, monuments médiévaux de l'Europe chrétienne ou de l'Orient musulman. L'objectif et la démarche, cependant, sont comparables d'une étude à l'autre : il s'agit d'exprimer dans le genre de formalisme propre aux systèmes experts les mécanismes d'une interprétation « intelligente » des matériaux en question, de telle sorte que l'ordinateur puisse non seulement reproduire celle-ci mais aussi appliquer les mêmes mécanismes pour l'interprétation de matériaux semblables, ultérieurement. La référence au domaine de l'intelligence artificielle, dont les systèmes experts font partie, trouve ainsi une certaine justification, même s'il est vrai que les mécanismes en question ont d'abord été formulés et communiqués à la machine par des « experts » humains.

Exemple : les relations entre la Grèce et l'Asie centrale à l'époque hellénistique

Ainsi, l'une des études évoquées plus haut porte sur un vaste ensemble de poteries utilitaires recueillies lors de la fouille d'une ville d'époque hellénistique (3^e-2^e siècles av. J.-C.), en Afghanistan; elle aboutit à une conclusion relative à l'existence de relations directes et fréquentes entre la Méditerranée orientale et l'Asie centrale à cette époque, en dépit des arguments habituellement invoqués contre cette hypothèse (trajet long et difficile, voisins hostiles, etc.). Le raisonnement qui fonde cette construction peut se traduire sous forme d'une suite d'opérations conditionnelles : « Si p ALORS q », où p et q représentent des groupes de propositions exprimées dans une langue plus ou moins naturelle. Par exemple :

• « si les analogies sont fortes et nombreuses entre les poteries d'époque hellénistique en Asie centrale et les poteries contemporaines de Grèce ou d'Asie mineure, c'est que celles-ci ont servi de modèle à celles-là ».

• « si l'existence entre deux aires de production céramique un rapport de modèle à copie, et que ces aires sont très éloignées l'une de l'autre, l'hypothèse de relations de proche en proche doit être écartée, au profit de l'hypothèse inverse affirmant des contacts directs entre les deux régions ».

• « si ce rapport de modèle à copie se manifeste plusieurs fois dans le temps, à des intervalles relativement courts (ici, guère plus de vingt-cinq ans) et que les poteries concernées sont d'un usage courant, c'est que les relations entre les deux aires en question étaient fréquentes, régulières et non pas liées à des circonstances exceptionnelles », etc.

L'architecture symbolique des raisonnements : problèmes de fondement

La place manque pour décrire la façon dont ces règles d'inférences sont transcrites dans le formalisme d'un système expert, comme aussi les observations elles-mêmes dans la base de données. L'essentiel au demeurant n'est pas là, du point de vue de l'archéologue ou de l'historien : quel que soit le « moteur d'inférence » choisi pour l'application informatique, il faut d'abord que l'on s'entende sur le choix, la définition et la désignation des données, puis sur la nature des opérations d'inférence recevables, dans tel ou tel contexte, sous réserve de certaines conditions qu'il convient également d'explicitier. Ce sont là des tâches qui commandent le progrès épistémologique de nos disciplines, avec ou sans informatique : où l'on en revient aux recherches de la voie A, qui ont précisément pour but d'élucider l'architecture symbolique de nos écrits, et par là même leurs fondements. La démarche consiste à déterminer les données de base (P_0) – composante sémiologique : systèmes de représentation, – puis les chaînes d'opérations discursives ($P_i \rightarrow P_j$) reliant de proche en proche la base au sommet de la construction – hypothèses, conclusions : composante logique, justiciable le cas échéant d'un calcul ou d'un traitement informatique, dans le sens ci-dessus – (fig. 2).

L'archéologie, dans cette entreprise, n'est qu'un univers de discours particulier : rien n'interdit de poser les mêmes problèmes fondamentaux dans d'autres domaines des sciences humaines, par exemple les sciences de la littérature, où l'abondance des constructions interprétatives ne laisse pas de soulever des questions d'« épistémologie pratique » particulièrement insistantes (3). L'intérêt que suscitent aujourd'hui les systèmes experts est un aiguillon puissant dans cette direction; l'on ne peut que s'en féliciter, quitte à rappeler de temps à autre que notre objectif propre n'est en l'espèce ni l'informatique ni l'intelligence artificielle mais bien une connaissance plus alerte des fondements de nos constructions, dans les sciences de l'homme et de la société, dont nous ne faisons encore qu'entrevoir les conséquences probables, à plus ou moins brève échéance. ■

BIBLIOGRAPHIE

- (1) Gardin (J.-C.), éd. *Archéologie et calculateurs : problèmes sémiologiques et mathématiques*, Paris, Éditions du CNRS, 1970.
- (2) Gardin (J.-C.), *Une archéologie théorique*, Paris, Hachette, 1979.
- (3) Gardin (J.-C.) et al., *La logique du plausible : essais d'épistémologie pratique*, Paris, Éditions de la Maison des sciences de l'homme, 1981.
- (4) Gardin (J.-C.) et al., *Systèmes experts et sciences humaines : le cas de l'archéologie*, Paris, Eyrolles, (à paraître en 1986).

L'archéologie n'est, en l'espèce, qu'un cas particulier : la démarche suivie pour exprimer la structure des raisonnements et pour les reproduire sur ordinateur est applicable à d'autres disciplines dans les sciences de l'homme et de la société.

LA RECHERCHE SCIENTIFIQUE ET TECHNIQUE AU JAPON

Une superficie d'environ quatre départements français pour développer son industrie, son agriculture et loger plus de 100 millions de personnes, une déficience totale en ressources naturelles : pour compenser ces handicaps, assurer sa croissance économique et faire face à la concurrence mondiale, le Japon a choisi de tout miser sur son développement scientifique et technique.

Georges RIEU

Dès les années cinquante, alors qu'il est en pleine période de reconstruction, le Japon met en place, en particulier au niveau du Premier ministre, des organismes qui vont tracer les grandes lignes d'une politique de développement et d'applications technologiques, fortement soutenue par un plan de formation de chercheurs et d'ingénieurs.

Ainsi, durant les années soixante, le nombre d'étudiants en facultés de sciences et de technologie va passer de dix mille environ à soixante mille ; le gouvernement va augmenter sa part de financement, se charger de promouvoir la recherche dans les grands secteurs scientifiques, financer une grande part de la recherche fondamentale, définir des priorités en particulier grâce à l'appui d'organismes de collecte et d'analyse de l'information et coordonner les efforts de recherche appliquée. A fin d'inciter les entreprises à investir en recherche et développement (R et D), le gouvernement va les autoriser, dès 1967, à déduire jusqu'à cinquante pour cent de leurs frais de R et D (selon les secteurs) de leurs impôts. L'industrie privée, de son côté, financera presque les deux tiers de l'effort global de recherche et développement, se concentrant avant tout sur la recherche du maximum d'applications des technologies nouvelles importées.

Dans les années soixante-dix, le Japon ressentira fortement le premier choc pétrolier, aussi le gouvernement accentuera-t-il ses efforts de recherche sur le nucléaire (moitié des dépenses totales) et les économies d'énergie. L'investissement global de

R et D sera seulement maintenu, les industries concentrent leurs efforts sur les technologies de pointe : le tiers de ces efforts porte sur l'électronique, un autre tiers est partagé à égalité entre la chimie et le matériel de transport.

Depuis la fin des années soixante-dix, on assiste à une accélération de l'investissement global de R et D dans les secteurs de pointe : informatique, électronique, biologie, matériaux, robotique, nucléaire, espace, transport, avec des efforts soutenus de création de nouvelles technologies et une volonté affirmée d'accroître la part de recherche fondamentale — à peine plus de quinze pour cent actuellement — de façon à ne pas se faire surprendre par une nouvelle génération technologique.

Anneau d'accumulation du Laboratoire national de physique des hautes énergies (KEK) (© Monbusho)

Afin de mieux comprendre les principales caractéristiques de l'effort de recherche japonais, qui atteint actuellement, avec plus de quatre cent mille chercheurs et deux cent quarante milliards de francs d'investissement, le niveau de l'effort européen, nous examinerons tout d'abord l'organisation de la recherche japonaise.

Organisation administrative de la recherche japonaise

La structure administrative de la recherche japonaise présente trois caractéristiques principales (fig. 1) :

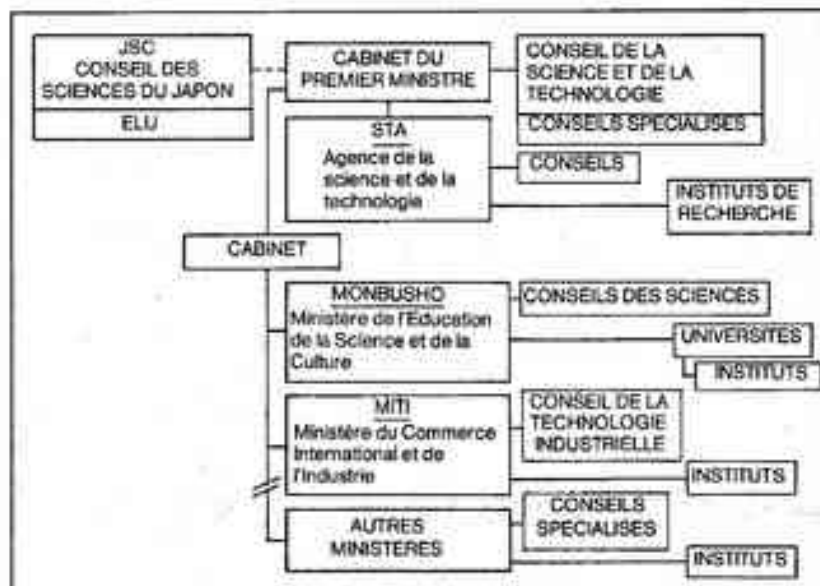


Fig. 1 : Schéma de la structure de la recherche publique au Japon.

■ Georges Rieu, directeur de recherche au CNRS, Centre de recherche de mathématique de la décision (UA 749), chargé de mission à la direction des relations et de la coopération internationales du CNRS, 15, quai Anatole France, 75700 Paris.

• un « Parlement de la science » de deux cent dix membres élus par l'ensemble de la communauté scientifique. — Le Conseil des sciences du Japon (1) — représente, auprès du Premier ministre, la communauté scientifique. Dans la pratique, il rassemble et analyse les opinions des scientifiques sur la politique de développement de la science et de la technologie et fait des recommandations à partir des résolutions prises lors d'assemblées générales. Il reçoit une part du budget national de recherche.

• les choix de politique scientifique et technique globale et les orientations à long terme sont faits au niveau du Premier ministre dans le cadre du Conseil de la science et de la technologie (2) — conseil interministériel élargi — assisté de quatre commissions scientifiques spécialisées (3) et de l'Agence de la science et de la technologie (4). Cette agence va assurer dans la pratique la coordination de l'activité de recherche des ministères et agences gouvernementales (à l'exception des recherches faites par les universités). Elle assure également le secrétariat du Conseil de la science et de la technologie (conjointement avec le ministère de l'Éducation, de la Science et de la Culture quand la recherche universitaire est en jeu) et le secrétariat des quatre commissions spécialisées.

• les ministères ont pour rôle de promouvoir le développement de la recherche et de la technique dans leur domaine propre. Pour cela, ils disposent d'instituts de recherche propres dont ils gèrent l'activité et auxquels ils procurent le financement, et de « conseils scientifiques » regroupant de nombreuses personnalités (jusqu'à un millier), sur lesquels leur administration va s'appuyer pour définir leur politique scientifique propre (fig. 2). Ainsi, le rôle de chaque ministère technique — Agriculture, Santé, Transports, etc. — dans la R et D est-il clairement défini.

Trois ministères chargés de fonctions plus générales jouent cependant un rôle très important : le Monbusho, ministère de l'Éducation, de la Science et de la Culture, qui pilote toute la recherche universitaire. Il représente le tiers du nombre total des chercheurs japonais en sciences exactes et la plupart des chercheurs en sciences sociales et humaines; le Miti, ministère du Commerce international et de l'Industrie,

qui a la responsabilité de la recherche industrielle; l'Agence de la science et de la technique chargée en particulier de la R et D dans de grands domaines scientifiques. Dans la pratique, les ministères jouent un rôle primordial; le Conseil de la science et de la technologie entérine le plus souvent lors de ses réunions annuelles les orientations présentées par les ministères et agences; le Conseil des sciences du Japon n'exerce aucun pouvoir réel de décision.

• le Monbusho (ministère de l'Éducation, de la Science et de la Culture) est responsable en particulier de la promotion de la recherche scientifique dans tous les domaines de la science, comprenant les sciences sociales et humaines, les sciences exactes, les sciences de l'ingénieur, l'agriculture et la médecine. Dans ce ministère, la Direction des affaires scientifiques et internationales est spécifiquement chargée de la recherche scientifique, alors que la Direction des enseignements supérieurs a la charge de l'enseignement universitaire. Ces directions sont assistées par le Conseil des sciences qui comprend vingt-sept personnalités scientifiques nommées par le ministre et des comités et sous-comités regroupant plus de mille personnes. De plus il existe un corps de quinze conseillers scientifiques qui participent à l'élaboration de la politique scientifique et aux délibérations du Conseil des sciences. Pour promouvoir la recherche scientifique, le Monbusho, finance les universités nationales et les instituts de recherche qui y sont rattachés (5); subventionne les universités privées et les universités publiques; finance les instituts nationaux inter-universités (6) qui possèdent en particulier les gros équipements; soutient l'activité de l'Académie du Japon (qui comprend cent cinquante membres et délivre chaque année le « Prix de l'Académie du Japon ») procure des subsides à la Société japonaise pour la promotion des sciences (USPS) qui anime différents programmes de promotion de la recherche et des programmes de coopération internationale.

Deux réussites particulièrement remarquables : celle du Laboratoire de physique des hautes énergies (KEK) à Tsukuba et celle de l'Institut des sciences spatiales et astronautiques qui dès 1970, alors qu'il était encore institut de l'université de Tokyo, lança ses premiers satellites scientifiques et fut le pionnier des activités spatiales au Japon.

• le Miti (ministère du Commerce international et de l'Industrie) est chargé de promouvoir la recherche industrielle, en particulier dans les domaines de l'énergie, de l'informatique et de l'électronique, de la robotique, de la chimie et des biotechnologies, en liaison avec l'université et l'industrie. Pour cela, il est assisté d'un Conseil de la technologie industrielle et dispose de deux agences : l'Agence des ressources naturelles et de l'énergie et l'Agence de la science et de la technologie industrielles

(AIST) qui regroupe seize instituts nationaux de recherche. L'AIST joue un rôle particulièrement important à travers ses laboratoires de recherche, en aidant l'industrie dans des domaines à hauts risques financiers et technologiques ou demandant d'importants investissements. Le Miti a parfaitement réussi également dans son rôle consistant à rassembler l'information et à dégager des priorités et des créneaux de développement industriel.

• l'Agence de la science et de la technologie est chargée, outre de la coordination interministérielle des politiques de développement scientifique et technologique, du développement de la R et D dans des domaines scientifiques ou un large effort national doit être fait : énergie atomique, espace, océan, sciences de la vie, matériaux, prévention des catastrophes, aviation, contrôle et lasers et doit promouvoir l'information scientifique. Pour cela, elle a sous sa tutelle six instituts de recherche et huit organismes à statut juridique spécial dont l'Agence nationale pour le développement spatial du Japon (NASDA), l'Institut japonais de recherche sur l'énergie atomique (JAERI), le Centre japonais de la science et de la technologie marine (JAMSTEC) et le Centre japonais d'information sur la science et la technologie (JICST).

L'Agence de la science et de la technologie a su s'affirmer en particulier dans le domaine nucléaire — en complémentarité avec le Miti et le Monbusho, dans les domaines de l'espace et l'océanologie — en complémentarité avec le Monbusho.

Chaque ministère établit ses programmes et son budget qu'il négocie directement avec le ministère des Finances. Les programmes sont examinés par l'Agence de la science et de la technologie — à l'exception des projets concernant les recherches effectuées dans les universités et les instituts inter-universités — qui va principalement essayer de les harmoniser pour éviter les duplications et mettre l'accent sur certaines orientations retenues par le Conseil de la science et de la technologie.

Les moyens d'action : hommes et crédits

Le Japon a la volonté de fonder son développement économique sur le développement de sa recherche. Un effort maximum est donc fourni dans ce domaine. Afin de mieux comprendre l'évolution de cet effort au cours des deux dernières décennies, nous avons reporté sur la figure 3, l'évolution de la contrainte principale — contrainte énergétique — et l'adaptation de l'économie japonaise à cette contrainte — courbe d'évolution du produit intérieur brut —, adaptation qui va directement se répercuter sur les moyens mis en œuvre en R et D (7).

Ainsi, la croissance très forte, durant les années soixante, des investissements de R et D, sera brisée par la première crise pétrolière, le choc se faisant très fortement

(1) Le Conseil des Sciences du Japon, fondé en 1949, comprend un « conseil » de deux cent dix membres élus tous les trois ans et réparti en sept divisions scientifiques : 1- Sciences sociales et humaines, 2- Lois et science politique, 3- Économie, 4- Sciences exactes, 5- Sciences et techniques de l'ingénieur, 6- Agriculture, 7- Médecine et pharmacie. Ce « conseil » est assisté de cent quatre-vingt comités de coordination et de sept cent trente commissions.

(2) Créé en 1959, le Conseil de la science et de la technologie est présidé par le Premier ministre lui-même et comprend quatre ministres : Finances, Éducation, Science et Culture, Planification économique.

(3) Science et Technologie, ainsi que le Président du Conseil des sciences du Japon et cinq personnalités de la communauté scientifique et de l'industrie privée nommées par le Premier ministre.

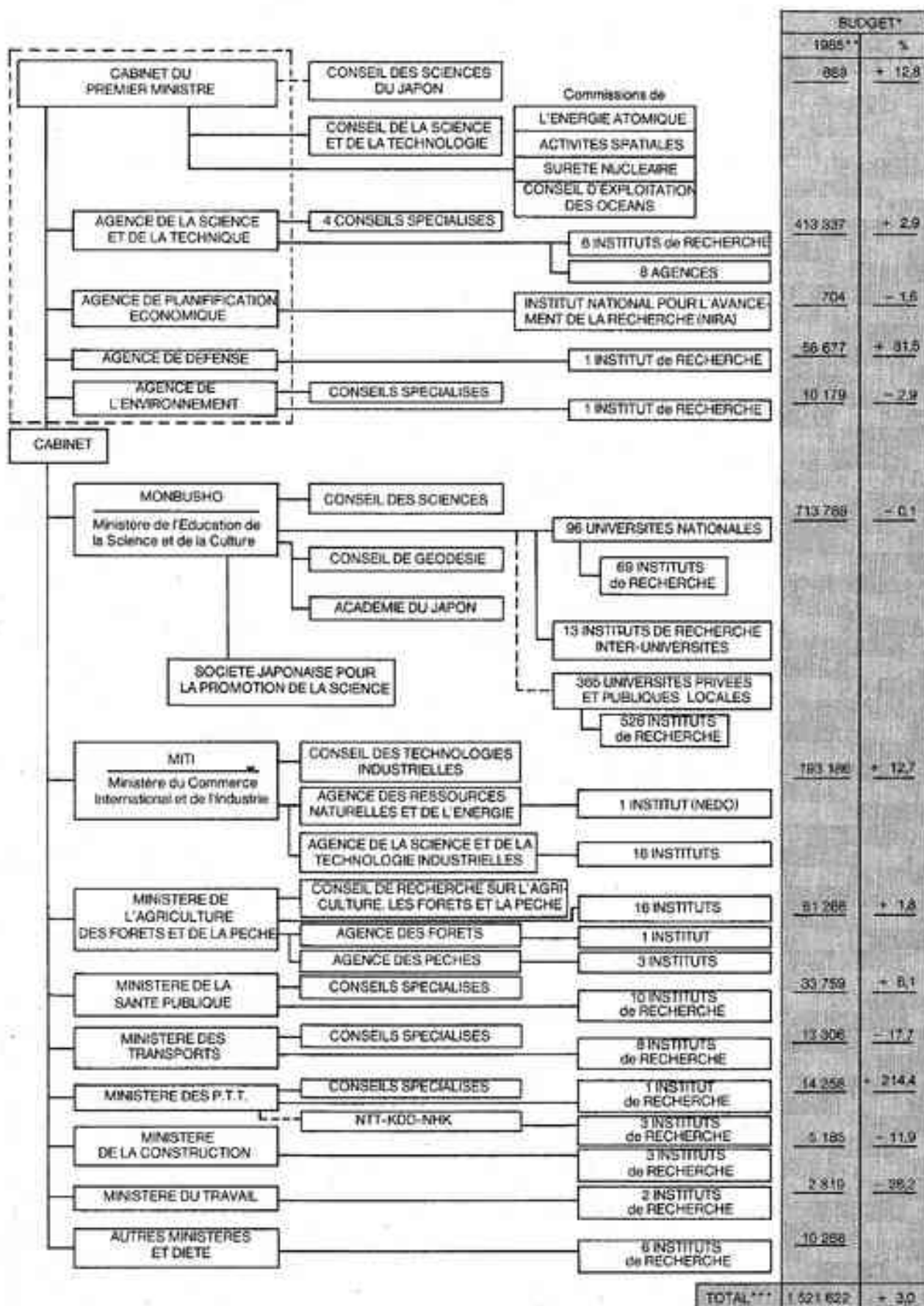
(4) Connaissance des activités spatiales, de l'énergie atomique, de la sûreté nucléaire, du développement océanique. Ces commissions, rattachées au cabinet du Premier ministre, assurent la coordination dans leur domaine respectif.

(5) Fondée en 1956, elle est rattachée au cabinet du Premier ministre et correspond à un « ministère de la Science et de la Technologie » : son directeur est d'ailleurs ministre d'État.

(6) Soit actuellement quatre-vingt-seize universités nationales et soixante-neuf instituts de recherche.

(7) Créés récemment à partir d'instituts de recherche dépendant d'une université, ces instituts regroupent surtout les grands équipements qui peuvent être utilisés ainsi par l'ensemble des universités. Il existe actuellement treize instituts inter-universités.

(7) De nombreux autres facteurs de contrainte entrent en jeu. Nous avons choisi une des contraintes principales à titre d'exemple.



* en millions de yens. ** l'année fiscale débute le 1^{er} avril. *** déduction faite des « comptes-doubles ».

Fig. 2: Structure administrative de la science au Japon (simplifiée).

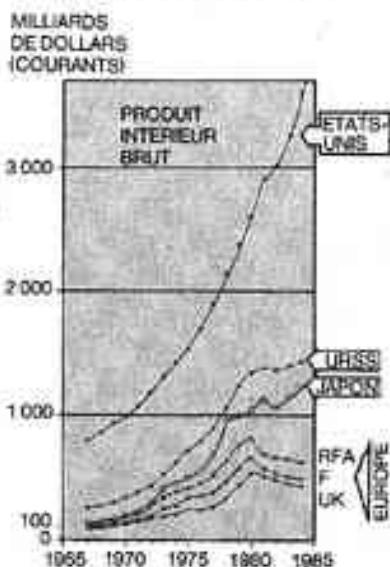
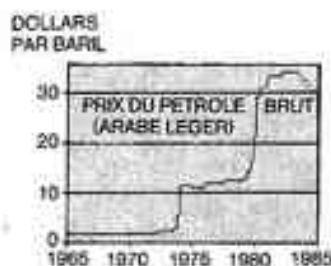


Fig. 3: Evolution de la contrainte énergétique et adaptation de l'économie japonaise.

sentir dans l'industrie où l'investissement va nettement fléchir, alors que le secteur public tentera de résister. En revanche, et contrairement aux pays européens dont les produits intérieurs bruts vont tous s'écrouler sous le deuxième choc pétrolier (fig. 3), le Japon va, dès 1975, systématiquement

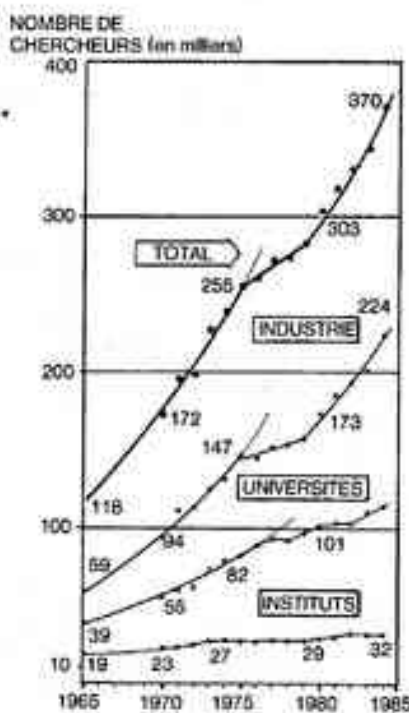
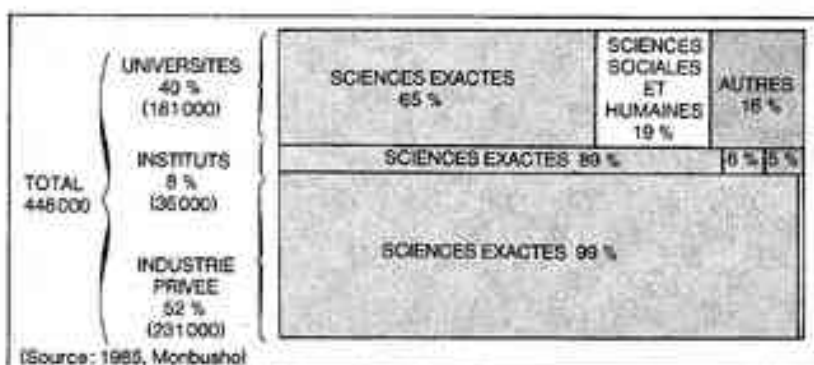


Fig. 4: Augmentation annuelle du nombre de chercheurs au Japon.



(Source: 1985, Morbushal)

Fig. 5: Répartition des chercheurs par secteurs d'exécution et par domaines scientifiques.

raffermir ses efforts de recherche dans les secteurs industriels, sans relâche jusqu'à aujourd'hui, parvenant ainsi à résister au deuxième choc pétrolier et à conserver la croissance de son produit intérieur brut.

Ainsi, l'augmentation annuelle du nombre total des chercheurs (fig. 4), qui était de 14 000 par an en moyenne entre 1965 et 1975, va être divisée par deux à partir de 1976 - 6 750 chercheurs de plus par an jusqu'en 1979 -, pour ensuite tripler - 17 600 chercheurs de plus par an - dès 1980, et être maintenue à ce niveau jusqu'à présent.

Alors que le gouvernement va essayer de maintenir son effort (tout en fléchissant quelque peu) - 4 300 chercheurs de plus par an dans les universités entre 1965 et 1975, 3 750 entre 1975 et 1979, 3 400 de 1980 à 1984 - l'effort industriel fléchira brutalement sous le choc en 1975 - tombant de 8 800 chercheurs de plus par an (deux fois plus que pour l'université) à 2 500 par an entre 1975 et 1979 - mais se ressaisit et multiplie par cinq cet effort dès 1980 : 13 400 chercheurs de plus par an (soit quatre fois plus que dans l'université). Ainsi, entre 1965 et 1984, le nombre des chercheurs japonais va-t-il tripler - 118 000 en 1965, 370 000 en 1984 -, l'accroissement étant depuis 1980 particulièrement important dans l'industrie : la répartition industrie - université - instituts

non universitaires passant respectivement de 50 % - 33 % - 16 % en 1965 à 60,5 % - 30,8 % - 8,6 % en 1984. Ces chiffres n'incluant pas les chercheurs en sciences sociales et humaines (considérées au Japon comme externes aux sciences et techniques) nous avons représenté sur la figure 5 la répartition des chercheurs en incluant ces domaines.

Si l'on considère maintenant la dépense de R et D (fig. 6) on constate qu'elle va être encore plus touchée : alors qu'elle atteignait une croissance de 16,1 % par an en valeur réelle sur la période 1965-70, elle va difficilement maintenir une croissance nulle entre 1976 et 1978 pour reprendre ensuite à un taux de 8,6 % entre 1980 et 1983.

Ce choc sera encore plus marqué pour les industries de fabrication (fig. 7) où les dépenses de R et D chuteront brutalement, ne reprenant leur niveau de 1971-72 que vers 1978 pour les recherches directement liées à la production, vers 1980 pour les recherches appliquées et vers 1982 pour les recherches fondamentales.

En dépit de cette situation, la dépense de R et D va être multipliée par quatre (en valeur réelle) entre 1965 et 1983, passant de 430 milliards de yens en 1965 à 6 500 milliards de yens en 1983 (208 milliards de francs). L'évolution de la répartition des efforts montre l'importance de l'effort industriel de recherche fait ces dernières années : l'industrie finance 58 % de la recherche en 1965, l'université 25,6 % et les instituts non universitaires 16,3 %. La

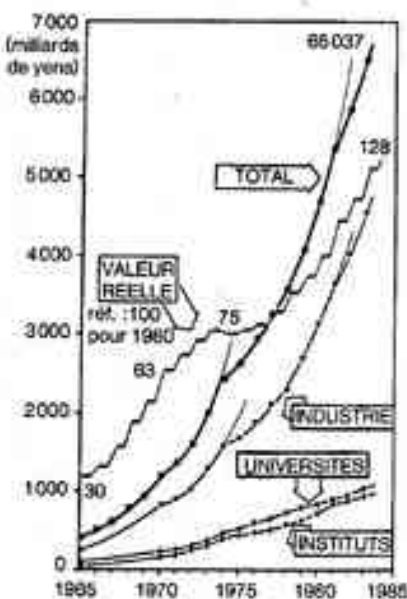


Fig. 6: Répartition de la dépense de recherche et développement (valeurs nominales).

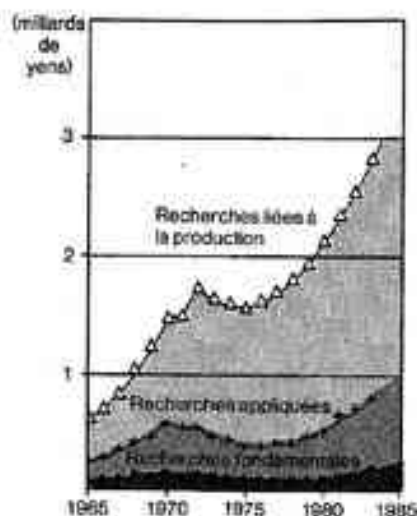


Fig. 7: Dépenses de recherche et développement dans l'industrie manufacturière.

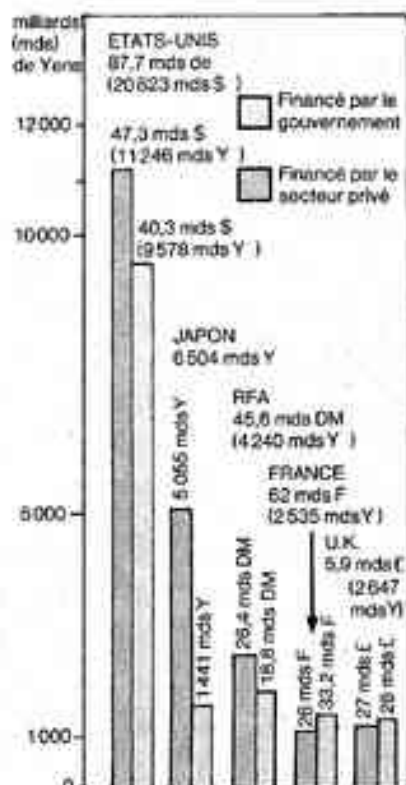


Fig. 8: Répartition du financement de la recherche en 1983 (en 1981 pour la France et la Grande-Bretagne).

part de l'industrie augmentera jusqu'à 66 % environ en 1970 et sera maintenue jusqu'en 1980 (celle de l'université restant de l'ordre de 19 %) pour ensuite augmenter à nouveau et atteindre 70,15 % dès 1983, contre 15,8 % pour l'université et 14,1 % pour les instituts non universitaires.

Cette évolution doit être mise en parallèle avec les objectifs de recherche: une recherche surtout consacrée aux applications et au développement de nouvelles technologies importées dans les années 1960-70, devenant plus originale dans les années 1970. Cependant, ayant rattrapé dans de nombreux domaines les pays occidentaux, le Japon tente maintenant de frayer lui-même son chemin vers des technologies nouvelles, ce qui nécessite d'importants investissements.

Ence qui concerne l'origine du financement de la recherche, une comparaison internationale (fig. 8) souligne l'importance du financement privé dans le cas du Japon par rapport aux autres pays: respectivement 77,7 % pour le Japon (71,7 % financé par l'industrie), 57,8 % pour la République fédérale allemande, 54 % pour les Etats-Unis, 45 % pour la Grande-Bretagne, 41,7 % pour la France.

La distribution des sources de financement (fig. 9) montre également la très faible part que consacre le gouvernement japonais au financement direct de la recherche industrielle: 1,7 % en (1983) contre 24,5 % (1981) pour la France, 17,7 % pour la RFA (1983), 29,99 % pour la Grande-Bretagne (1981) et 32,35 % pour les Etats-Unis (1973).

Le détail des répartitions du budget gouvernemental de recherche japonais (1 521 milliards de yens en 1985) est

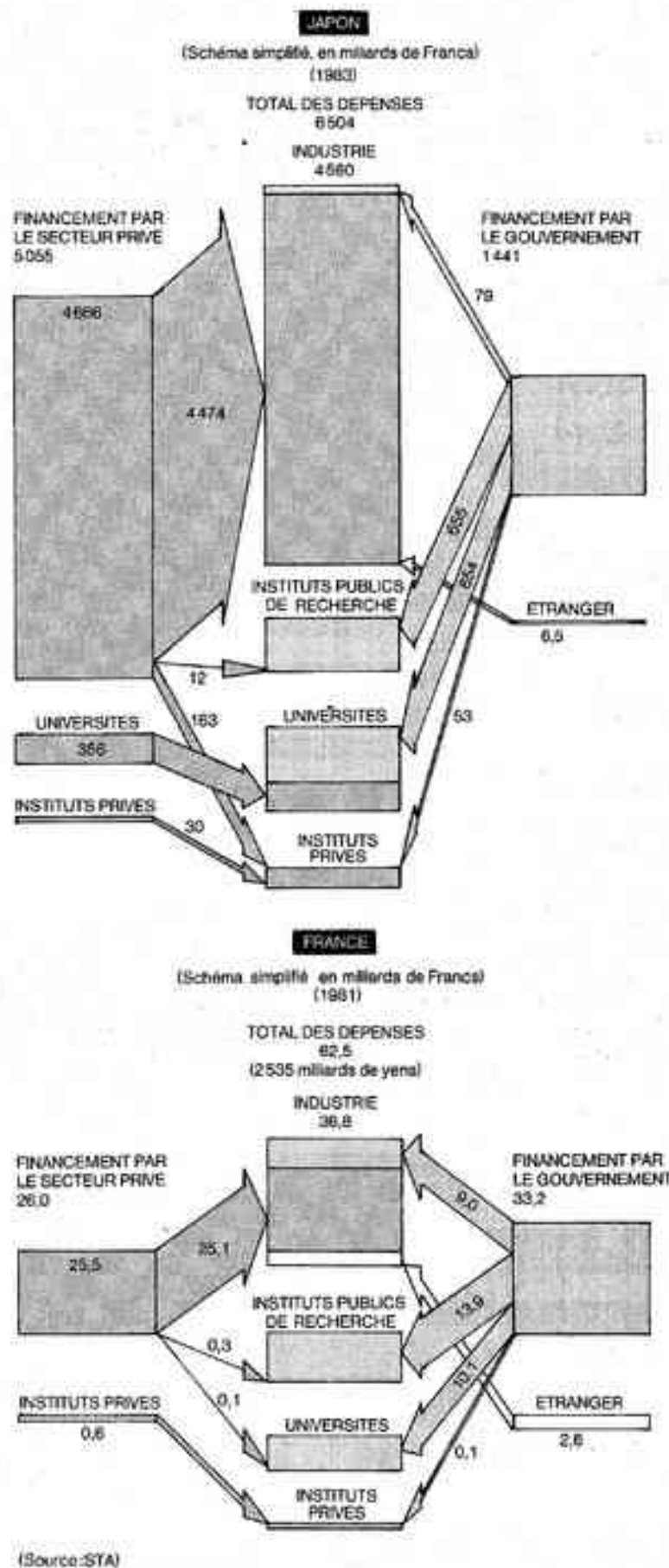


Fig. 9: Comparaison des flux de dépense de R et D pour la France et le Japon (même échelle).

donné sur la figure 2 (colonnes de droite) et est schématisé sur la figure 10. La moitié environ est affectée au Monbusho, le quart à l'Agence de la science et de la technologie, le huitième au Miti. Malgré la faiblesse du financement gouvernemental par rapport au financement privé (moins du quart) les ministères japonais apportent une aide appréciable à la recherche industrielle tant au travers de leurs instituts propres que par la collecte et la diffusion des informations et la définition de priorités et de créneaux. La figure 11 souligne également l'importance de la recherche appliquée et du développement même dans le cadre de la recherche menée dans les universités.

Le Japon a-t-il gagné son pari ? Jusqu'à présent, son produit intérieur brut a su résister aux chocs pétroliers, et les deux tiers de sa croissance économique ont été réalisés, selon l'Agence de planification économique, par le progrès technologique.

Gagnera-t-il également son pari sur l'avenir en devenant à son tour pionnier et créateur de nouvelles générations technologiques ? ■

BIBLIOGRAPHIE

- Dupuis, (M.), « La politique scientifique et technique du Japon », *La Documentation française* n° 493-494, Sept. 1984.
- *An outline of the University Based Research System in Japan*, Monbusho, Tokyo, 1986.
- *Science and Technology Agency: Its roles and activities*, Tokyo, 1985.
- *Transforming Scientific ideas into innovations: Science Policies in the United States and Japan*, Japan Society for the Promotion of Science, Tokyo, 1985.
- *Science policy perspectives: USA-Japan*, Academic Press, 1982.
- *Agence de la science et de la technologie (STA) Livre blanc 1985* (en japonais).

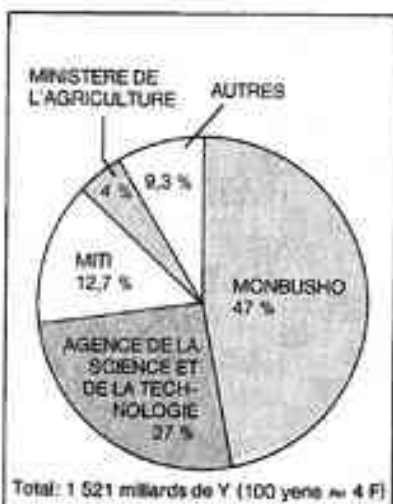


Fig. 10 : Budget gouvernemental 1985 pour la science et la technologie.



Lancement d'un satellite scientifique par un Centre de recherche inter-universitaire : l'Institut des sciences spatiales et astronautiques (© Monbusho).

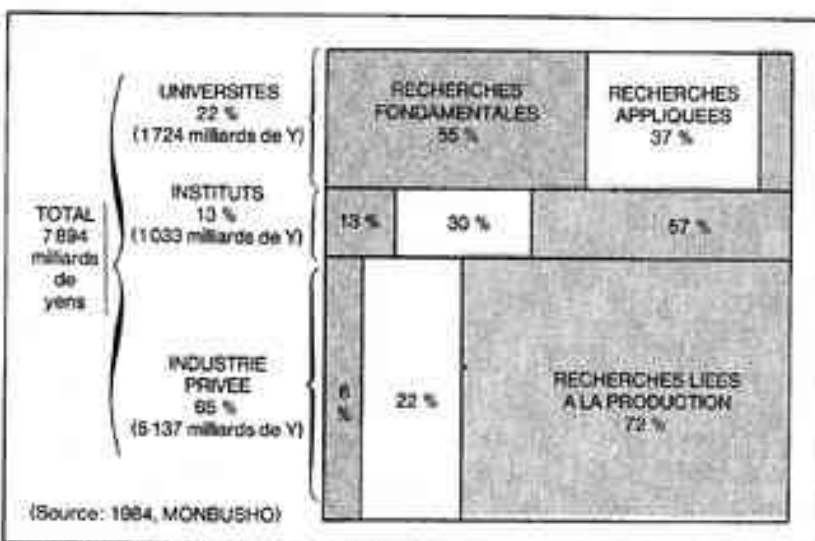


Fig. 11 : Répartition des dépenses de recherche par type.

INITIATION A L'ASTRONOMIE

Cinquante élèves de première scientifique au lycée Joliot-Curie d'Aubagne parrainés par un chercheur du CNRS dans leur découverte de la recherche en astronomie.

Monique MOUNIER-KUHN

Lo CNRS a lancé, en liaison avec le ministère de l'Éducation Nationale, une vaste opération de rencontre entre les scientifiques et les élèves du secondaire: 1 000 chercheurs pour 1 000 classes.

En Provence-Côte d'Azur, choisie comme région pilote, le Laboratoire d'astronomie spatiale (LAS) collabore depuis deux ans avec le lycée Joliot-Curie d'Aubagne, près de Marseille.

Pour les cinquante élèves de première, passionnés par la compréhension de l'Univers, Jean-Pierre Sivan, 40 ans, chargé de recherche au CNRS, s'est transformé en animateur, en pédagogue, en guide pour organiser des nuits d'observation à la Sainte-Baume, au pic de Bertagne, pour faire connaître l'Observatoire de Haute-Provence et pour expliquer les travaux du LAS, laboratoire qui conçoit et réalise des expériences embarquées à bord des navettes américaines et des sondes soviétiques.

Ainsi a-t-on fait un jeu de piste scientifique autour de la Comète de Halley, des Pléiades, de la galaxie d'Andromède. Les nuits magiques vouées à l'observation du ciel sont suivies de séances de travail et de réflexion, les élèves s'initiant ainsi au processus de recherche.

Le projet proposé aux élèves a consisté à « rentrer » un « catalogue » d'étoiles sur les micro-ordinateurs du lycée et à étudier leur répartition en fonction de leur distance à la Terre, de leur brillance et de leur couleur.

Le proviseur du lycée, M. Guy Chazeau, le responsable du club informatique, M. Gilles Abecassis, et le professeur de mathématiques, M. Pierre Robert ont soutenu activement cette initiative: un technicien du LAS, M. Louis Castinel, a apporté son concours. Cette ouverture de l'école sur la vie, avec la dimension scientifique de la culture moderne, est jugée par tous bénéfique.

Jean-Pierre Sivan consacre ses recherches à la matière interstellaire dans notre galaxie et dans les galaxies voisines, en particulier sur sa composante ionisée et les sources d'ionisation de cette matière

interstellaire. Ceci est basé sur des observations au sol, à l'Observatoire de Haute-Provence, à l'Observatoire européen austral au Chili, et au plus grand télescope du monde, le télescope de 6 mètres de Zelenchuck, au Caucase. Il participe à l'interprétation des photographies du ciel dans l'ultraviolet, obtenues avec la caméra grand-champ du LAS, à bord de Spacelab.

Cette expérience avec les jeunes le réjouit. « Par rapport à mes activités habituelles, c'est une distraction utile qui répond à l'enthousiasme de dizaines de lycéens. C'est aussi une action de vulgarisation dont je mesure immédiatement l'impact: dans les regards attentifs, dans les questions suscitées et dans la pression de ces jeunes, qui veulent assouvir leur soif de connaissance. A travers cette étude, ils peuvent percevoir tout le travail de dépouillement informatique propre à l'astronomie moderne. »

Les astres, de toute éternité, ont exercé une fascination sur les hommes. Les jeunes de 1986 n'y échappent pas, affirme Jean-Pierre Sivan. « Ils éprouvent un regain d'intérêt pour certains aspects spectaculaires des explorations spatiales. La télévision a joué un rôle en faisant entrer dans la vie quotidienne les données de la recherche spatiale. »

Le bénéfice de cette expérience d'animation scientifique pour les jeunes est considérable, estime-t-il. « Il faut partir des connaissances scientifiques qu'ils peuvent ainsi acquérir, approfondir. Nous avons abordé, non le système solaire,

mais la dimension stellaire: l'évolution stellaire n'est pas bien connue des élèves. On a travaillé sur leurs ordinateurs, ils ont tracé le diagramme de Hertzsprung-Russell, qui est l'un des diagrammes fondamentaux de l'astrophysique moderne puisqu'il retrace toute l'évolution des étoiles. En parallèle, à partir des observations visuelles qu'ils effectuent au cours des nuits d'observation, j'essaie de leur montrer la démarche de l'astronome d'aujourd'hui qui décrit l'Univers, explique son évolution sur des bases d'observation, en faisant appel à tout l'arsenal des différentes branches de la physique. » Pour la recherche, il s'agit d'une opération à long terme: « On montre à des jeunes qui choisiront très vite leur orientation universitaire, la réalité de la recherche au quotidien, ce qu'est le métier de chercheur, tel que chacun le vit, loin des mythes du savant sur son piédestal », souligne Jean-Pierre Sivan.

Les jeunes lycéens aubagnais qui participent à cette expérience avec zèle ont donc fait le premier pas. Certains d'entre eux, deux ou trois sûrement, se considèrent déjà comme des astronomes en puissance. Leur vocation, ils l'ont cultivée dans un club d'astronomes-amateurs et dans les claires nuits provençales de la Sainte-Baume et de Saint-Michel, avec leurs professeurs et avec « leur » chercheur.

Lorsqu'on les voit ensemble, on sent une compréhension mutuelle, une complicité qui repose sur un trait commun: la curiosité.



■ Jean-Pierre Sivan, chargé de recherche au CNRS, Laboratoire d'astronomie spatiale, CNRS, Traverse du Siphon, Les trois Lacs, 13012 Marseille.

■ Monique Mounier-Kuhn, chargée de mission pour les affaires culturelles auprès du directeur de l'information scientifique et technique du CNRS, 15, quai Anatole-France, 75700 Paris.

NICOLAS FRANCESCHINI OU L'EXPLORATION DES PETITS CERVEAUX

Nicolas Franceschini a reçu en 1985 une médaille d'argent du CNRS pour ses recherches sur le système visuel chez la mouche. Ses travaux, effectués in vivo, et ses découvertes en font aujourd'hui un des plus grands spécialistes en neurocybernétique.

Régine FERRÉ

Avoir vingt ans, un diplôme d'ingénieur électronicien, un autre d'ingénieur automaticien et sentir l'éveil d'une grande passion pour tout ce qui touche au traitement de l'information dans le système nerveux, tel est Nicolas Franceschini, au seuil d'une vie d'études et de recherches qui va l'entraîner aux quatre coins du monde.

De la physique à la biologie

Il s'agit tout d'abord, pour lui, d'approfondir ses connaissances en automatique dont les éléments théoriques lui permettront – il en est sûr – de mieux appréhender les systèmes nerveux. Cependant, en France, la cybernétique biologique n'a pas encore la place qu'elle mérite et après avoir cherché vainement pendant un an à poursuivre ses études dans différentes formations de recherche françaises, Nicolas Franceschini obtient une bourse pour l'Institut Max-Planck de Tübingen (RFA).

Pendant près de quatorze ans, Tübingen va être le centre de la vie et des recherches du jeune Français. Au sein d'un milieu international de scientifiques pluridisciplinaires, il se plonge dans les études les plus diverses qui toutes se complètent : automatique, neurophysiologie, cybernétique biologique. Au fur et à mesure, d'autres domaines s'ouvrent à lui, l'optique, par exemple.

La rencontre avec cette dernière discipline est décisive. Elle lui permet de préparer sa thèse d'Etat soutenue avec succès, en 1972, à Grenoble : « Sur le traitement optique de l'information visuelle dans l'œil composé de la drosophile ». Le jury est présidé par le Prix Nobel, Louis Néel. La physique et la biologie étaient réunies. Assistant de recherche, dès 1970, à l'Institut Max-Planck de cybernétique biologique, il fait également des stages de recherche au Japon, en Australie et au Canada. Pendant trois ans, il fait partie

du Conseil scientifique de la Max-Planck-Gesellschaft. Puis, il décide de rentrer en France, avec l'idée d'y créer « quelque chose » lui permettant de poursuivre et de développer ses travaux. Nommé maître de recherche au CNRS, en 1978, il dirige, aujourd'hui, l'équipe de neurocybernétique dans le cadre du Laboratoire de neurobiologie du CNRS à Marseille.

Le temps des découvertes

Des études et recherches ardues, diversifiées, mais toujours complémentaires sont de réels témoignages d'un chercheur d'une très grande exigence morale et intellectuelle, entièrement consacré à la connaissance de ce qu'il appelle lui-même *circuit intégré biologique* et que représente le système visuel de la mouche, pris comme modèle d'un système nerveux évolué.

L'observation de l'œil de la mouche révèle une structure nerveuse parfaite, hautement répétitive, construite à partir d'un million de neurones seulement et qui ne pèse qu'un milligramme.

Chaque œil est composé de 3 000 facettes et de 24 000 cellules photoréceptrices bien distinctes. Là résident précisément l'originalité et l'intérêt de l'étude de cet insecte : chaque cellule peut être analysée et stimulée individuellement *in vivo*. On peut espérer ainsi découvrir le plan de câblage précis sous-jacent et montrer comment la nature a réalisé son traitement de l'information dans un but spécifique. Cette caractéristique, grâce à de remarquables techniques micro-optiques conçues et mises au point par lui-même, a permis à Nicolas Franceschini de faire d'importantes découvertes à retombées multiples tant sur le plan de la biologie cellulaire que sur celui du traitement neuronal de l'information.

On peut citer :

- découverte de la fluorescence des cellules photoréceptrices, qui a permis de déterminer, à l'aide de microélectrodes, la sensibilité spectrale des six types de photo-capteurs présents dans la rétine de la mouche ;

- découverte d'un mécanisme de capture de substances par les cellules photoréceptrices sous l'effet de l'éclairement.



Nicolas Franceschini. (D.R.)

Ce phénomène pourrait ouvrir de nouvelles voies en neurobiologie, en permettant l'introduction de substances, curatives ou toxiques, dans une population de cellules, sélectionnées simplement par focalisation de lumière :

- découverte des propriétés élémentaires du microcircuit neuronal responsable de la vision du mouvement ;

- découverte d'un câblage neuronal propre au mâle, servant, apparemment, à piloter l'animal vers la femelle lors du vol nuptial.

Cette recherche sur les réseaux de neurones au service du comportement débouche sur un vaste champ d'applications, concernant, par exemple, la vision artificielle et la robotique, ou, plus proche de nous encore, le guidage élémentaire des aveugles.

Devant ces perspectives, on saisit toute la passion de Nicolas Franceschini et son ardeur – ainsi que celle de son équipe – à poursuivre ses expériences dans le but d'analyser le savoir-faire de la nature en matière de traitement spatio-temporel de l'information et de la transposer en utilisant les ressources de la technologie électronique.

Quelques publications récentes

Kirschfeld (K.), Franceschini (N.), Minke (B.), « Evidence for a sensitizing pigment in fly photoreceptors », *Nature* (Lond.), 1977, n°269, pp. 388-390.

Franceschini (N.), Kirschfeld (K.), Minke (B.), « Fluorescence of photoreceptor cells observed in vivo », *Science*, 1981, n°213, pp. 1264-1267.

Franceschini (N.), Hardie (R.), Ribi (W.), Kirschfeld (K.), « Sexual dimorphism in a photoreceptor », *Nature* (Lond.), 1981, n°291, pp. 241-244.

Rieble (A.), Franceschini (N.), « Movement detection in flies: parametric control over ON/OFF pathways », *Exp. Brain Res.*, 1984, n°54, pp. 360-394.

Wilcox (M.), Franceschini (N.), « Illumination induces dye incorporation in photoreceptor cells », *Science*, 1984, n°225, pp. 851-854.

Franceschini (N.), « Early processing of colour and motion in a mosaic visual system », *Neurosci. Res. Suppl.* 2, 1985, pp. 17-49.

■ Régine Ferré, ingénieur d'études au CNRS, rédacteur en chef-adjoint du *Courrier du CNRS*, 15, quai Anatole-France, 75700 Paris.

LES RESEAUX INFORMATIQUES

L'ordinateur devenu aujourd'hui moyen de communication est un outil indispensable à la recherche. Il est donc essentiel que les chercheurs s'équipent en moyens informatiques qui soient connectés en réseaux pour être intégrés à l'espace européen de la recherche.

Max FONTET, Guy PUJOLLE

Le sous-équipement chronique des laboratoires a laissé le monde de la recherche quelque peu en retrait de la réalité informatique actuelle. L'ordinateur est passé en vingt ans de l'état de simple calculateur à celui d'instrument de communication. Cette mutation qui suit le rythme du développement des réseaux numériques de transport d'informations et des ordinateurs postes de travail individuels reste bien évidemment étrangère au chercheur qui ne dispose même pas d'une console vidéo individuelle.

Quelle informatique pour la recherche ?

La politique d'équipement informatique des centres de recherche doit évoluer vers une informatique répartie où les réseaux jouent un rôle essentiel. La solution à la pénurie en moyens informatiques des laboratoires de recherche français ne peut plus consister exclusivement à développer des centres de calculs équipés de gros ordinateurs. Certes, il ne faut pas négliger les moyens de calculs lourds que sont les superordinateurs, mais il semble encore plus urgent de prendre en compte l'évolution de l'outil informatique.

L'image du chercheur qui utilise uniquement un terminal pour soumettre une tâche à un ordinateur par l'intermédiaire d'un système de traitement par lots est complètement révolue. Le fait de travailler à plusieurs utilisateurs en temps partagé sur une même machine par l'intermédiaire d'un terminal vidéo assurant une interface

raisonnable, conduit tout naturellement le chercheur à utiliser l'ordinateur pour communiquer. Très vite, il s'est développé des logiciels de messagerie pour la communication entre les utilisateurs présents au même moment sur l'ordinateur (deux utilisateurs dialoguent via l'ordinateur), ou de courrier (chaque utilisateur répertorié sur une machine à une boîte aux lettres destinée à recevoir à tout moment les messages envoyés par tout autre utilisateur). Ces outils de communication ont été très vite complétés par les forums : chacun dépose des messages (chacun s'exprime) dans une boîte aux lettres collective, et chacun peut lire (écouter) tous les messages qui s'y trouvent et y répondre s'il le souhaite. Cet outil s'est tout de suite révélé indispensable pour le travail en coopération : il assure une mise en commun des connaissances et une meilleure interactivité entre les équipes d'un domaine travaillant sur la même machine. Le Groupement de recherches coordonnées « programmation » du CNRS a montré comment utiliser un tel outil pour accélérer la validation des logiciels développés par ses équipes.

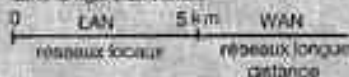
Cette situation pouvait être considérée comme satisfaisante il y a deux ans encore. Mais l'évolution des techniques de communication avec l'ordinateur a bouleversé les habitudes. Un gros ordinateur offrant des puissances de calcul de plusieurs MIPS (millions d'instructions par seconde) est indispensable pour effectuer de gros calculs — l'ordinateur dévoreur de nombres — ou pour gérer de très grosses bases de données ; en revanche, il offre bien peu de services lorsqu'on lui demande de gérer à distance les écrans vidéo « bitmap » (1) mettant à la disposition de l'utilisateur une véritable interface graphique interactive ; de tels terminaux améliorent considérablement les possibilités de dialogue avec les logiciels ; la souris ou le manche à balai qui asservissent le curseur à la main de l'utilisateur ne sont pas des gadgets ; ils permettent de manipuler les informations affichées sur l'écran par simple pointage. Malheureusement, il est nécessaire de disposer en permanence d'une puissance d'environ 0,5 MIPS pour gérer un tel écran « bitmap ». Conjointement, la notion d'ordinateur poste de travail individuel est devenue une réalité. On ne compte plus ces microordinateurs construits autour d'un microprocesseur 16 ou 32 bits, munis d'un écran « bitmap » et qui supportent des systèmes d'exploitation évolués dont le système UNIX (2) sous toutes ses versions est le flambeau. De telles machines ont une puissance de calcul suffisante pour gérer de façon autonome des éditeurs sophistiqués manipulant textes, dessins et graphiques dans un univers de multifenêtrage. Cet outil individuel ne doit ni détourner le chercheur de l'accès à une grande puissance de calcul et à de grands volumes de stockage de données, ni l'isoler en supprimant les outils de communication dont il avait l'ha-

(1) Écran se comportant pour l'utilisateur comme une matrice de points-image modifiables individuellement.

(2) Le système UNIX, développé aux Bell Laboratories d'AT&T, est un système universel d'exploitation portable sur toute classe d'ordinateurs allant des plus petits aux plus gros.

Réseaux locaux

Les réseaux locaux (ou LAN Local Area Network) sont différenciés des réseaux longue distance (ou WAN Wide Area Network) par la portée maximum entre les deux points les plus éloignés. L'ordre de grandeur de ces distances est indiqué dans la figure suivante :



Les réseaux locaux peuvent être de trois types :

1. Les réseaux locaux informatiques Basband Local Network qui ont été développés par les industriels de l'informatique. Le meilleur exemple en est Ethernet, développé par Xerox, qui spécifie en réalité la couche 1 et une partie de la couche 2 de la norme internationale ISO. Sur cette base Ethernet de nombreux réseaux locaux se sont constitués. On peut citer les réseaux Xerox, Net-One, FAC-TOR... qui possèdent leurs propres normes de niveau 3, 4, 5, 6, 7.

2. Les autocommutateurs privés ou PABX (Private Automatic Branch exchange) qui offrent des téléphones et qui utilisent le câblage téléphonique à l'intérieur des entreprises ou des campus. Les autocommutateurs étant aujourd'hui des ordinateurs, rien n'empêche de transporter sur des câbles téléphoniques des données qui seront gérées par les processeurs. Des logiciels d'émulation et de transcodage sont également disponibles dans l'autocommutateur.

3. Les réseaux large bande (ou Wideband Local Network) proviennent des diffuseurs. Ils ont été mis au point par les industriels du câble télévision. Ils permettent de superposer dans un même câble, sur des longueurs d'ondes différentes, plusieurs canaux d'informations. Un exemple est donné par le réseau Wangnet.

Il n'y a pas aujourd'hui de réseaux locaux obéissant complètement aux normes internationales.

■ Max Fontet, professeur à l'université Pierre et Marie Curie (Paris VI), Laboratoire d'informatique théorique et programmation (U.A. 248), directeur scientifique adjoint du département des sciences physiques pour l'ingénieur au CNRS, 15, quai Anatole France, 75700 Paris.

■ Guy Pujolle, professeur à l'université Pierre et Marie Curie (Paris VI), responsable de l'unité associée « Méthologie et architecture de systèmes informatiques » (U.A. 818), université de Paris VI, Institut de programmation, Tour 55, 4, place Jussieu, 75230 Paris cedex 05.

■ Ce texte a été soumis le 13-11-85.

bitide. Il faut donc connecter ces postes de travail à des réseaux pour rétablir l'accès à des ressources complémentaires et préserver les possibilités de communication entre les utilisateurs. Pour être efficaces, de tels réseaux doivent avoir une structure hiérarchisée; au premier niveau, on trouve les réseaux locaux assurant une interconnexion entre les matériels d'un même laboratoire; on peut penser à un réseau local de haut débit (10 mégabits par seconde) de type ETHERNET reliant tous les postes de travail à une mémoire de masse commune (serveur de fichiers) et à un ou plusieurs ordinateurs locaux de moyenne puissance du type miniordinateur universel 32 bits. Il est tout à fait imaginable d'inclure dans un tel réseau une machine dédiée à une application particulière: machine base de données, machine à inférences, ou bien, machine vectorielle (3). L'un des miniordinateurs du réseau doit servir de passerelle vers le niveau suivant du réseau. Ce maillon doit se situer à l'échelle du campus accueillant les laboratoires; les moyens de calcul lourds d'un campus se retrouvent à ce niveau du réseau. Le niveau supérieur du réseau interconnecte tous les sites entre eux et assure la connexion avec les réseaux des autres pays. Une caractéristique essentielle d'une telle hiérarchie de réseaux est d'être conçue pour interconnecter des matériels hétérogènes et d'être ouverte vers l'ensemble de la communauté. En effet, un réseau doit assurer avant tout une fonction de communication et d'échange tant au niveau local que national ou international entre des chercheurs travaillant dans le même domaine.

Vers une mutation du support de communication

Tous les grands pays industrialisés misent sur le développement d'industries de haut niveau technologique telles que l'aéronautique, les biotechnologies, les télécommunications, l'électronique et bien entendu l'informatique pour assurer la croissance économique durant les deux prochaines décennies. Il est illusoire d'espérer participer honorablement à cette compétition sans accroître immédiatement le potentiel de recherche dans les disciplines scientifiques concernées afin de renforcer à la fois l'effort de recherche fondamentale et d'accélérer les transferts recherche-développement-industrie. Tous ces pays, États-Unis en tête, ont choisi d'investir massivement dans l'équipement informatique de leurs laboratoires de recherche. Leur objectif est de faire croître la capacité de recherche des équipes existantes en leur fournissant un environnement informatique performant conçu pour les seconder efficacement dans leur travail quotidien. On peut aussi penser que le dialogue entre équipes de développement en milieu industriel et équipes de recherche dans des laboratoires sera d'autant plus aisé que l'environnement technique sera comparable; cette remarque s'applique tout particuliè-

rement aux domaines où la conception assistée par ordinateur est devenue un outil indispensable.

Un autre point fondamental est le rôle de la communication dans le développement des activités de recherche. La vitesse de diffusion des connaissances est un facteur déterminant de l'accroissement du niveau des recherches dans un domaine. Tout chercheur sait très bien qu'il doit établir un réseau de communication avec les collègues travaillant dans le même domaine. De telles collaborations peuvent se développer de manière plus efficace en utilisant les réseaux informatiques. Un premier niveau est fourni par la messagerie électronique qui permet de contacter quasi instantanément un collègue sans risque d'échec: le téléphone n'est efficace que si l'on est sûr d'atteindre son correspondant. Ce seul service accède considérablement l'écriture d'un article par deux coauteurs ne pouvant pas se rencontrer régulièrement; ceci arrive très souvent lors de collaborations établies au cours d'un séjour dans un laboratoire étranger; ce n'est en général qu'au moment de retourner dans son laboratoire d'origine que l'on est prêt pour écrire un article avec le chercheur qui vous a invité. La mise en commun d'informations sur un réseau grâce à un service de forum réparti est aussi une autre forme de diffusion de l'information qui évite l'isolement et permet de connaître plus vite l'état de l'art sur une question précise. Toutes ces nouvelles sources d'information pour le chercheur complètent le service offert par les bases de données bibliographiques ou techniques qui doivent être, bien entendu, disponibles à travers le réseau. On peut raisonnablement espérer voir apparaître des revues scientifiques et techniques dont le

support sera informatique; elles seront consultables par chacun à partir de son poste de travail et surtout les auteurs écriront directement leurs articles sur un support informatique; le support papier va disparaître à terme pour l'ensemble de la communication scientifique. Il faut préparer les chercheurs à cette mutation qui peut être comparée à ce qui s'est passé lors de l'apparition de l'imprimerie. Il est dès à présent imaginable d'organiser un congrès pour lequel la soumission d'articles se ferait par courrier électronique; le travail du comité de programme n'en serait que plus facile et plus efficace. Pour ne pas nous trouver marginalisés, il faut qu'en France aussi nous acceptions de doter les laboratoires de recherche des mêmes moyens informatiques que nos partenaires et concurrents. La mise en place d'un réseau de la recherche français interconnecté aux différents réseaux internationaux est certainement un objectif prioritaire, mais il doit être accompagné d'un plan cohérent d'équipement informatique des laboratoires de recherche; ce réseau ne peut jouer son rôle que si toute unité de recherche dispose d'un ensemble d'ordinateurs, postes de travail individuels, reliés entre eux par un réseau local à haut débit ainsi qu'à des serveurs locaux du type miniordinateur 32 bits.

Les projets de réseaux chez nos principaux partenaires européens

Un tel plan d'équipement est certainement très ambitieux et ne peut pas être mené à son terme sans une coordination des efforts des différents partenaires au niveau national. La difficulté majeure ne se

La normalisation internationale des réseaux informatiques

Les réseaux informatiques sont des systèmes particulièrement complexes qui nécessitent une structuration en éléments directement réalisables. Les deux grands organismes de normalisation, l'ISO et le CCITT, ont fait adopter en tant que standard une décomposition en sept couches:

7 couche application	partie concernant la gestion des informations à transporter
6 couche présentation	
5 couche session	
4 couche transport	partie concernant le transport de l'information
3 couche réseau	
2 couche liaison	
1 couche physique	support physique

1. La couche physique assure le transport des signaux sur le support de transmission.

2. La couche liaison est responsable de l'acheminement sans erreurs de blocs d'information sur les liaisons entre nœuds du réseau. Les supports de transmission introduisent des erreurs dans les informations transportées et le but de cette couche 2 est d'assurer un taux d'erreurs tout à fait négligeable.

3. La couche réseau est responsable de l'acheminement des paquets de données

de bout en bout. Ses fonctions principales sont le routage, le contrôle de flux et l'adressage.

4. La couche transport est responsable du contrôle du transport des informations de bout en bout, au travers du réseau. Cette couche doit assurer que les messages des utilisateurs connectés à un réseau informatique sont correctement parvenus à leurs destinataires. Une fonction importante sera le découpage et le rassemblement des messages en paquets.

5. La couche session est responsable de la mise en place et du contrôle du dialogue entre tâches distantes. Avant d'émettre des messages une session devra être ouverte entre correspondants.

6. La couche présentation est responsable de la présentation syntaxique des données échangées par les applications. Ceci pour que tous les matériels connectés puissent s'échanger des informations.

7. La couche application est responsable de la présentation sémantique des données échangées. Ceci pour que les applications puissent s'échanger des informations.

(3) Machine à inférences: machine spécialisée pour les systèmes d'intelligence artificielle dont l'opération de base est l'inférence logique.

Machine vectorielle: machine parallèle dédiée au calcul scientifique capable d'effectuer simultanément des calculs sur les composantes d'un vecteur.

situé pas au niveau de l'infrastructure d'un réseau de transport national; le réseau TRANSPAC offre un service qui, dans un premier temps, permet de couvrir les besoins du réseau de la recherche. En revanche, il n'existe pas à l'heure actuelle sur le marché l'ensemble des logiciels de communication nécessaires pour construire un réseau hétérogène ouvert. On peut évaluer à cinq ou six ans le délai minimum pour mener une telle opération dans l'hypothèse favorable où le financement de l'opération serait assuré sans problème. Il faut savoir que nos principaux partenaires européens ont pris conscience bien avant nous de l'enjeu scientifique que représente un réseau national pour la recherche. Nous devons transformer cet handicap en un avantage en sachant profiter de leur expérience.

Les anglais en lançant en 1979 le projet JANET (Joint Academic Network) ont été les premiers à mener une politique concertée au niveau national. Sous l'égide du Department of Education and Science, ils ont entrepris d'interconnecter les réseaux universitaires existant et les réseaux reliant les laboratoires du Science and Engineering Research Council et du Natural Environment Research Council. Il a été choisi de mettre en place un réseau de transport privé aux normes X.25, interconnecté au réseau public PSS (Packet Switch Stream) de British Telecom. Compte-tenu de l'absence de normes de communication pour les réseaux au moment du lancement du projet, une part essentielle de l'effort initié a été de définir un ensemble de normes et de les implémenter sur les différents matériels. Ces protocoles de transmission ne sont pas aux normes ISO. Cette avance britannique risque d'être pénalisante lorsque les normes ISO seront devenues le standard de communication. Près de 150 unités de recherche doivent être ainsi reliées à travers JANET depuis juin 1985. La mise en place du réseau national est accompagnée d'une politique incitative pour encourager le développement de réseaux locaux sur les campus universitaires reliés à JANET à travers des passerelles.

Le projet de réseau de la recherche allemand (Deutsches Forschungsnetz) est plus récent. Les premières études n'ont commencé qu'en janvier 1982. Le réseau DFN se propose de relier l'ensemble des équipes de recherche, qu'elles appartiennent aux universités, aux instituts de recherche ou à l'industrie par un réseau ouvert sous les protocoles ISO-OSI utilisant, comme réseau de transport, le réseau public allemand DATEX-P offrant un service conforme aux normes X.25. La réalisation du DFN s'appuie sur une « Association DFN » qui est un groupement d'utilité publique rassemblant les universités, instituts de recherche, constructeurs de matériels informatiques, industriels du logiciel chargés chacun de réaliser une partie du travail de développement. Ce projet qui est soutenu par le BMFT (Bundesministerium für Forschung und Technik) est évalué à 100 millions de Deutschmarks. En plus des services usuels de messagerie, transfert de fichiers, connexion à distance (soumission de travaux par lots, dialogue interactif en mode ligne ou via la définition d'un terminal virtuel), il est prévu de développer des protocoles d'échange de données graphiques en mode interactif sur la

base de GKS (Graphical Kernel System). Le projet DFN prévoit d'inclure des services de communication compatibles avec les réseaux locaux soit par installation de passerelles, soit par réalisation des protocoles implémentés dans le DFN à l'intérieur des réseaux locaux. Il est enfin prévu, à plus long terme d'inclure des services de transmission à haut débit utilisant des satellites de communication.

La situation en France

Cet handicap français n'est pas une fatalité. En effet, la France a été un des tous premiers pays à développer un réseau pour la recherche: le réseau CYCLADES. Celui-ci résulte d'une suite d'études entreprises depuis 1970 dans le cadre du Comité de recherche en informatique. Une entente générale des divers organismes qui ont participé à son élaboration, s'est concrétisée autour d'un projet de réseau d'ordinateurs principalement axé sur le partage de bases de données. Le projet a démarré début 1972 et mi-78, le réseau CYCLADES comportait 7 nœuds, 13 concentrateurs de terminaux et 20 ordinateurs de toutes marques. La France était alors au même niveau d'avancement que les Etats-Unis. Pour diverses raisons, le réseau a été abandonné et, en 1984, on se retrouve dans la situation de 1970: un groupe d'étude est

chargé de définir les grandes orientations d'un réseau français de la recherche. Bien sûr, les techniques ont évolué et le futur réseau pourra satisfaire directement les normes les plus récentes, mais un trou de six ans est toujours difficile à combler.

Quelles que soient les conditions de réalisation d'un tel réseau de la recherche en 1985, il est clair que les laboratoires ne peuvent certainement pas attendre que ce grand projet arrive à son terme pour disposer de moyens de communication. Il faut accepter toute solution provisoire de réseaux donnant à certains laboratoires dès à présent des accès à des réseaux existant. Le pragmatisme doit l'emporter afin d'éviter que les laboratoires français n'accumulent les handicaps vis-à-vis de leurs collègues américains ou européens. Il faut développer et promouvoir les réseaux que certaines collectivités de recherche ont mis en place pour pallier l'absence d'un véritable réseau national de la recherche encore pendant quelques années. Les services télématiques souhaités dès à présent par les laboratoires se définissent ainsi: messagerie et courrier électronique, transfert de fichiers, connexion à distance, forum réparti.

Un service de messagerie et de courrier dans un contexte hétérogène constitue la première étape à mettre en place très rapidement; des échanges de petits fichiers sont dès lors possibles; cette première phase pourrait être atteinte très vite en développant des logiciels de transfert de messages à partir de maquettes déjà opérationnelles tels COSAC, développé par le CNET.

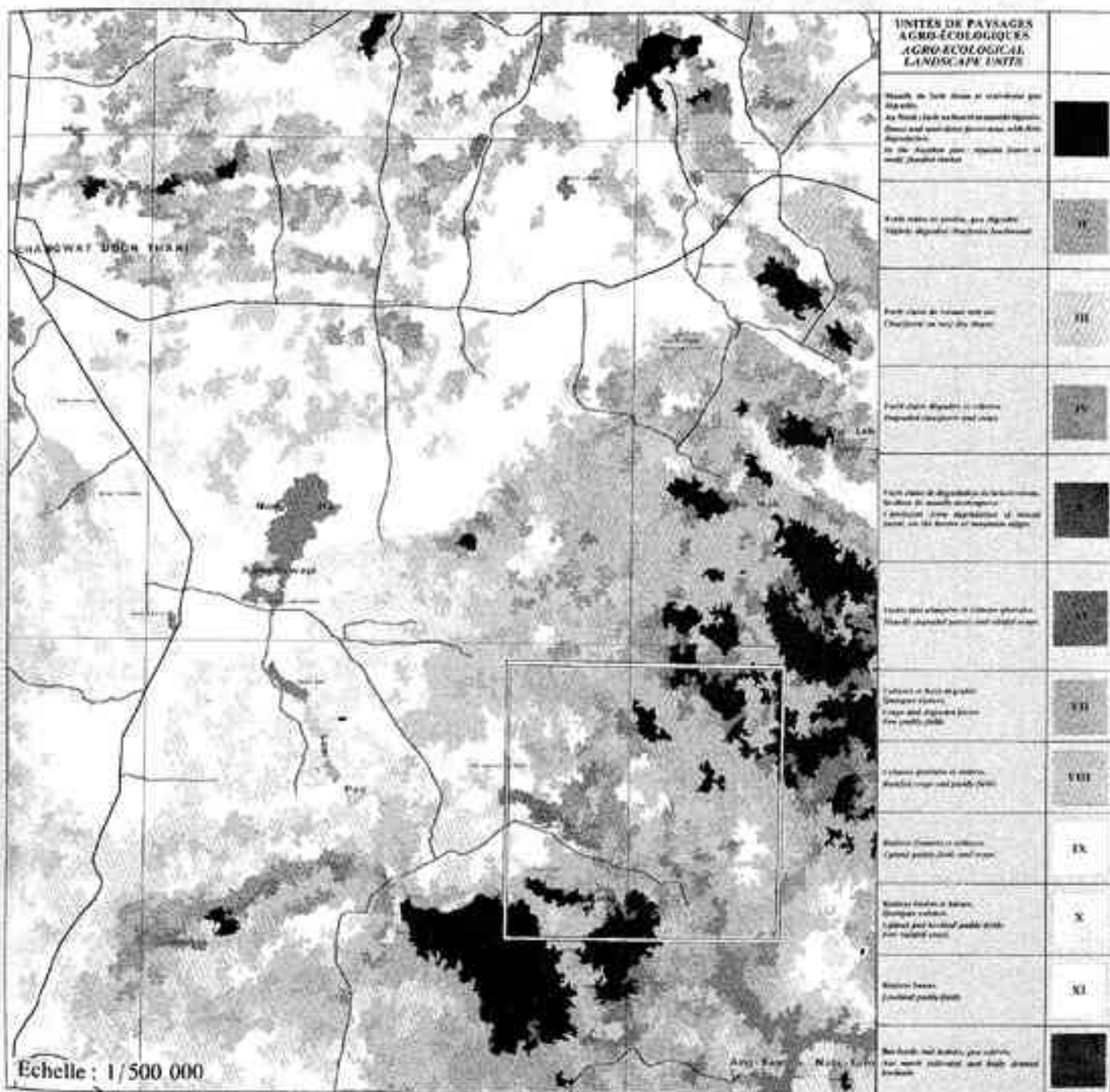
Il faut tenir compte des relations existant déjà au niveau des activités de recherche avec nos collègues américains ou européens pour mettre sur pied des solutions provisoires donnant la possibilité aux chercheurs de se retrouver sur les réseaux qu'ils fréquentent. Une deuxième priorité est donc l'ouverture de ce réseau vers le monde extérieur par des passerelles vers tous les grands réseaux publics ou privés ou sont connectés des chercheurs. Tout le monde convient du rôle essentiel qu'a pu jouer le réseau ARPANET dans le développement de recherches coordonnées aux Etats-Unis.

Le contexte de la recherche est propice à la mise en œuvre des moyens de communication les plus novateurs. C'est certainement là que l'on sera le plus motivé et le plus compétent pour les utiliser et expérimenter leur usage. A un moment où l'on se propose de créer un espace européen de la recherche, il est primordial de donner aux équipes de recherche les moyens de collaborer et de s'intégrer dans le contexte européen évitant d'être en état d'infériorité. L'équipement des équipes en moyens informatiques et leur interconnexion via des réseaux est donc urgent. Surtout si l'on veut être prêt pour le défi suivant: celui de l'informatique répartie où non seulement les ressources sont mises en commun mais où elles sont aussi exploitées via un système unique qui les gère au niveau du réseau lui-même. Il ne faudrait pas que le retard que nous accumulons dans ces domaines, faute d'avoir su investir à temps, ne devienne insurmontable et que nous ne soyons pas prêts pour ce défi des années quatre-vingt-dix.

Les problèmes techniques du réseau de la recherche

Le réseau français de la recherche ne peut que suivre la norme internationale ISO et utiliser le réseau national TRANSPAC pour les couches 1, 2 et 3. Les véritables difficultés se placent aux niveaux des couches 4, 5, 6 et 7 que nous pouvons traiter en deux ensembles. Les couches 4 et 5 étant normalisées, le problème est de savoir s'il faut adopter l'ensemble des possibilités offertes par ces normes, ou s'il conviendrait d'adopter des logiciels à prendre en compte sur les machines connectées ou sur de nouveaux fronts, ou simplement une porte, ARCHITEL par exemple. La prise OSI/ISO pour connecter le matériel Bull étant dans ce dernier cas non adaptée aujourd'hui. Les autres constructeurs acceptent-ils de développer leur propre prise de raccordement sur un sous-ensemble particulier de la norme? Les réseaux locaux existants ont également bien du mal à s'interconnecter car ils ne présentent pas d'interface standard: les réseaux Ethernet s'appuient la plupart du temps sur la norme TCP/IP. Les couches supérieures 6 et 7 présentent le défaut de n'être pas encore normalisées, exception faite de la messagerie électronique. Il faut donc, à l'heure actuelle, attendre les normes, des propositions particulières, surtout pour les transferts de fichiers et les travaux transactionnels.

3 / Dédire la carte des paysages agroécologiques après traitement combinant les états de surface



Cette réalisation par traitement numérique à partir des données Landsat 15/7/88 des 25/10/77 - 17/01/78 et 21/01/82 © IGN 1985

Ces unités dégagent les relations privilégiées existant entre les activités humaines, notamment agricoles, et leur milieu physique. Cette connaissance cartographique est nécessaire pour mettre en œuvre une politique régionale de développement rural et pour en suivre le déroulement.

Programme de coopération en télédétection : France (CNRG-CEGET, CRAD-IRAT, IGN) et Thaïlande (Université de Khonkaen-Centre de télédétection, Bangkok)

Traitement d'image effectué à l'IGN (Service des applications nouvelles)

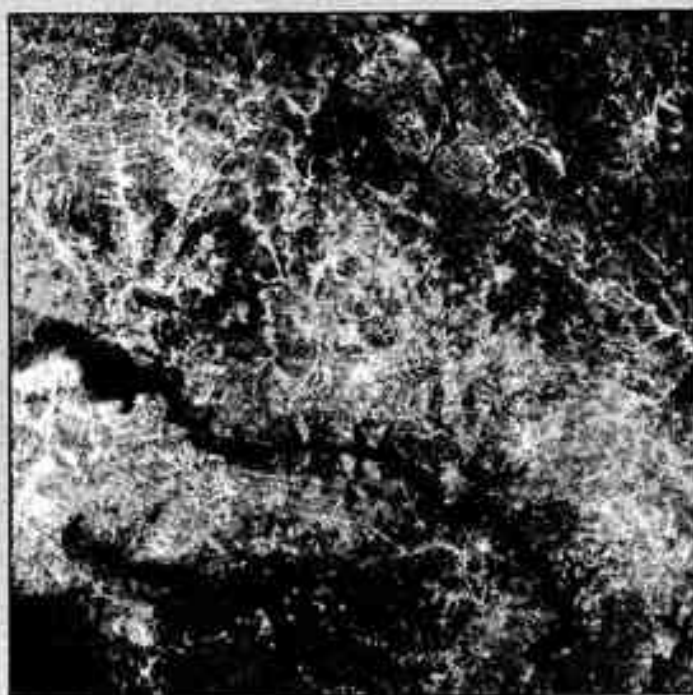
Extrait de la pléiade Télédétection du CRAD

CONNAISSANCE DES MILIEUX AGRICOLES

Région de Sakon-Nakhon, Udon-Thani, Nord-Est de la Thaïlande

Les données satellitaires LANDSAT ont permis de :

1 / Produire une visualisation qui fournit l'image d'une réalité qu'il faut comprendre



2 / Etablir la carte des états de surface après missions de terrain et classification radiométrique



- Eau libre claire.
Clear deep water.
- Eau turbide.
Turbid water.
- Bas-fonds marécageux et végétation aquatique.
Lower swampy areas and aquatic vegetation.
- Forêts denses ripicoles inondables.
Dense riparian easily flooded thickets.
- Forêt dense à tendance sempervirente.
Dense mostly evergreen forest.
- Forêt semi-dense à tendance caducifolée.
Semi-dense semi-deciduous forest.
- Forêt claire rabougrie de versant.
Low woodland on slopes.
- Forêt claire en perchis.
Clearforest brushwood.
- Forêt claire en perchis, dégradée (cultures, jachères, parcs).
Degraded clearforest (crops, fallows, pastures).
- Cultures pluviales à couvert arboré.
Wooded rainfed crops.
- Cultures pluviales à faible couvert arboré.
Slightly wooded rainfed crops.
- Rizières à couvert arboré moyen et traces d'humidité en saison sèche.
Fairly wooded paddy-fields and wet areas in dry season.
- Rizières à faible couvert arboré et sols nus.
Paddy-fields with few trees and bare soils.
- Sols nus, constructions, carrières.
Bare soils, constructions, quarries.