

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/363405490>

Mesurer une discipline émergente : une base de données sur les thèses en calcul et en informatique (1956–1973). A Data Base of Doctorates to Measure the Emergence of Computer Scien...

Preprint · September 2018

DOI: 10.13140/RG.2.2.13967.87203

CITATIONS

0

READS

53

1 author:



Pierre Mounier-Kuhn

French National Centre for Scientific Research

46 PUBLICATIONS 163 CITATIONS

SEE PROFILE

**Mesurer une discipline émergente : une base de données
sur les thèses en calcul et en informatique (1956-1973)**

Pierre Mounier-Kuhn

CNRS, Sorbonne Université & Centre Alexandre-Koyré

mounier@msh-paris.fr

Résumé. L'article décrit un travail d'histoire quantitative en cours, visant à mieux comprendre le processus d'émergence d'une discipline : l'informatique. Les bornes chronologiques sont la mise en service des premiers ordinateurs en France et la reconnaissance de l'informatique comme discipline dans les instances d'évaluation de l'Enseignement supérieur. Après un rappel historique sur l'évolution des diplômes doctoraux dans l'enseignement supérieur français, on présente successivement les sources utilisées et les problèmes rencontrés dans ce travail, les limites des conclusions que l'on peut en tirer et les principaux résultats obtenus. Ceux-ci ne font pas que confirmer des phénomènes que l'on connaissait par d'autres approches historiques, ils font apparaître des faits invisibles autrement. Parmi ceux-ci, on découvre une chute des études doctorales au début des années 1970.

Mots-clés. Histoire de l'informatique, histoire quantitative, émergence d'une discipline, diplômes, doctorat, France.

A Data Base of Doctorates to Measure the Emergence of Computer Science (1956-1973)

Abstract. Our paper presents a quantitative analysis of doctorates in computer science, aimed at better understanding the emergence process of this then new discipline. The chronological limits are the installation of the first electronic computers in France and the official recognition of *informatique* as an autonomous discipline in the French research and higher education system. The paper recalls the evolution of doctoral diplomas in French higher education, then presents in turn the sources employed and the problems encountered in this work, the limits of its possible conclusions and the main results obtained. Those not only confirm several phenomenons that we knew or suspected through other historical approaches : They reveal facts that were otherwise invisible. They particularly lead to the discovery of a drastic fall in doctoral studies in the early 1970s.

Mots-clés. History of Computing, Computer Science, Quantitative History, Emergence of a Discipline, Diplomas, Doctorate, France.

La catégorie « informatique » est apparue progressivement au cours des années 1960, par convergence de champs scientifiques tels que les mathématiques appliquées, la logique ou la linguistique formelle avec les pratiques du calcul automatique, sous la pression croissante de la demande socio-économique. Pour mesurer ce processus de construction d'une discipline et préciser ses contenus, j'ai entrepris d'établir une liste exhaustive des thèses en calcul et en informatique soutenues en France de 1956 à 1973, soit entre la mise en service des premiers calculateurs électroniques numériques dans l'Enseignement supérieur et l'époque où l'informatique fut reconnue dans les instances supérieures de régulation académique.

Aucune base existante ne fournissant de données fiables, ce travail exigeait de procéder à un long recensement, université par université, en vérifiant autant que possible que les contenus des thèses correspondaient bien à l'objet de la recherche. On donnera plus loin la liste des types de sources utilisées.

La liste de thèses atteignait 650 titres fin 2009 quand mon livre, où elles figurent en annexe¹, fut mis sous presse – encore comprenait-elle trop de thèses en mécanique des fluides ou dans d'autres branches des mathématiques, dont on l'a « nettoyée » depuis. Elle en totalise maintenant près de 900. Cet ordre de grandeur ne devrait plus beaucoup augmenter – il peut même diminuer par élimination de thèses qui s'avéreraient trop marginales par rapport au champ *calcul et informatique*.

La base a permis d'ores et déjà d'établir des statistiques : rythmes de croissance des thèses dans ce nouveau domaine, répartition entre doctorats d'état, de 3^e cycle ou de docteurs-ingénieurs. Elle a confirmé la domination durable de quelques pôles universitaires dans cette discipline. Elle a aussi révélé des irrégularités de croissance et des disparités locales, fournissant des pistes d'investigation. Les titres des thèses montrent l'apparition de nouveaux thèmes de recherche se différenciant du calcul.

Au-delà de mon propre travail sur l'histoire de l'informatique, cette étude veut contribuer à produire une connaissance statistique sur l'enseignement supérieur français, dont le Ministère de l'Éducation nationale ne présente trop souvent que des éléments disparates. On ne dispose apparemment d'aucune statistique officielle sur les thèses en sciences dans la période considérée².

Un rappel historique sur les diplômes doctoraux dans l'enseignement supérieur français et leur évolution permet de situer l'enquête dans le cadre juridique de son objet (le lecteur pressé d'arriver aux résultats pourra sauter ce chapitre). On présentera successivement les

1 P. Mounier-Kuhn, *L'Informatique en France, de la seconde guerre mondiale au Plan Calcul. L'émergence d'une science*, Paris, PUPS, 2010, p. 597-632.

2 J.-R. Cytermann et A. Prost, « Une histoire en chiffres de l'enseignement supérieur en France » dans Jean-M. Chapoulie, P. Fridenson & A. Prost (dir.), *Le Mouvement Social*. n° sur « Mutations de la science et des universités en France depuis 1945 », 2010. Vol. 4, n° 233, p. 31-46, notamment les tableaux 10 et 11. Nos voisins britanniques ne sont pas mieux lotis, selon R. Simpson, *The Development of the PhD Degree in Britain, 1917-1959 and Since: An Evolutionary and Statistical History in Higher Education*, préface de Harold Silver. Lewiston, Edwin Mellen Press, 2009.

sources utilisées et les problèmes rencontrés dans ce travail, les limites des conclusions que l'on peut en tirer et les principaux résultats obtenus. Parmi ceux-ci, l'on découvre une chute des études doctorales au début des années 1970.

Les diplômes

Après la Révolution française qui avait supprimé les universités, Bonaparte avait rétabli les doctorats en médecine et en droit, puis, à la suite de la création de l'Université de France en 1806, un décret de 1808 en avait fixé l'organisation en ajoutant les doctorats ès sciences, ès lettres et en théologie. D'autres textes au milieu du XIX^e siècle transformaient profondément l'exercice doctoral : si, depuis le Moyen-Âge, la thèse n'était qu'un court texte de quelques pages résumant les idées que le candidat allait plus longuement argumenter lors de la soutenance suivant la tradition de la *disputatio* orale, la thèse devenait le rapport de recherches détaillé que nous connaissons maintenant³. Ces textes précisaient que, pour obtenir le *doctorat ès sciences*, il fallait remplir les deux conditions suivantes :

- être licencié ès sciences de l'ordre correspondant (mathématiques, sciences physiques, sciences naturelles), ce qui contribua d'ailleurs à séparer l'Université des grandes écoles d'ingénieurs, puisque ceux-ci devaient passer une licence s'ils voulaient s'inscrire en doctorat ;
- soutenir deux thèses approuvées devant la faculté et répondre aux questions ou objections du jury. Le sujet de la seconde était défini « sur proposition de la faculté » pour apprécier l'étendue de la culture scientifique du candidat dans sa discipline.

La préparation d'une thèse de doctorat d'état ès sciences prenait habituellement entre quatre et six ans.

Le *doctorat d'université*, créé en 1897 en s'inspirant du modèle allemand, était un titre scientifique institué par chaque université et ne conférant aucun droit d'ordre public, notamment au recrutement. Pour s'y présenter, la licence n'était pas obligatoire. Il suffisait d'avoir deux certificats d'études supérieures au choix pour les candidats français, ou des équivalences pour les candidats étrangers. Le doctorat d'université était principalement préparé par des candidats étrangers. Il sera supprimé en 1984.

Le titre d'*ingénieur-docteur*, créé en 1923, était ouvert aux titulaires de certains diplômes d'ingénieur. Le but était de favoriser les recherches concernant les applications de la science, en facilitant le rapprochement entre ingénieurs et universitaires – en fait la majorité des ingénieurs-docteurs venait des écoles d'ingénieurs intégrées aux universités, comme les

3 Sur ce sujet en général, voir Claude Jolly & Bruno Neveu (dir.), *Éléments pour une histoire de la thèse*, Paris, Aux Amateurs de livres, diff. Klincksieck, 1993.

instituts polytechniques et les ENSI. Il a été transformé en titre de docteur-ingénieur par décret du 19 mars 1948, modifié en 1959, 1961, 1963 et 1966.

Pour obtenir ce diplôme en 1966, il fallait avoir un titre d'ingénieur et trois certificats d'études supérieures ès-sciences ou deux de maîtrise, sauf dérogations pour certaines écoles. Le diplôme de docteur-ingénieur était délivré après deux ans d'études et de recherches dans un laboratoire scientifique : la première année était consacrée à la préparation d'un diplôme d'études approfondies (DEA) ; la seconde année conduisait à la soutenance de deux thèses : une étude originale « pouvant aboutir à des résultats nouveaux d'ordre technique ou à des résultats d'ordre scientifique susceptibles d'applications pratiques » ; une 2^e thèse (« sur proposition de la faculté ») pour apprécier les connaissances générales du candidat, ou constituant une autre étude originale. La durée de préparation de la thèse passe à deux ans en 1974.

Le *doctorat de 3^e cycle* ou *doctorat de spécialité* a été créé au milieu des années 1950 pour répondre au besoin de multiplier la population de chercheurs et d'enseignants, avec une formation moins longue et plus spécialisée que le doctorat d'état. En 1954 un troisième cycle d'études supérieures, dit de spécialisation, fut institué dans les facultés des sciences, puis dans les facultés de lettres (1958), de droit et de sciences économiques (1963). La durée du 3^e cycle était fixée à deux ans minimum. Pour s'inscrire en deuxième année, le candidat devait posséder un certificat d'études supérieures dans sa spécialité, obtenu dans le cadre de la licence ou en première année du 3^e cycle. Le titre de docteur de 3^e cycle précisait la spécialité, suivant une liste établie par le décret de 1964. Ainsi, en Mathématiques :

- Astronomie
- Mathématiques appliquées
- Mathématiques pures
- Mécanique céleste
- Mécanique quantique
- Mécaniques des fluides
- Probabilités
- Statistiques
- Traitement de l'information (terme apparu à la fin des années 1950).

Ce dernier terme sera remplacé par *informatique* à la fin des années 1960. Mais l'essentiel est que les spécialistes universitaires des ordinateurs, en phase avec la Direction de l'Enseignement supérieur, ont vite profité de la préparation de ce décret pour introduire dans le texte une spécialité *Traitement de l'information* distincte des *Mathématiques appliquées*. Cela en 1964, un an après l'affirmation de l'autonomie de cette science au congrès de l'Association française de calcul et de traitement de l'information, à Toulouse, et la création

d'une action concertée *Calculateurs électroniques* à la DGRST. Et trois ans après la publication d'un article fougueusement programmatique par John MacCarthy (MIT), affirmant la constitution d'une nouvelle branche du savoir autour de la *computation*.

L'automatique, comme l'électronique, était une spécialité du 3^e cycle de Physique. Ces deux disciplines, officialisées quelques années avant l'informatique, ont formé un certain nombre de spécialistes des ordinateurs : soit de leur emploi pour la conduite de processus automatiques, où le numérique remplaçait progressivement l'analogique pendant la période considérée, soit de la conception de circuits digitaux. Chacune employait des modèles mathématiques qui étaient également importants dans la formation des informaticiens. Ces disciplines émergentes ont donc fonctionné en synergie, dans une dynamique militante visant à affirmer leur légitimité académique, à obtenir des moyens et des positions de pouvoir.

Les postes universitaires de maîtres-assistants étaient accessibles aux docteurs de 3^e cycle, aux docteurs-ingénieurs et aux agrégés.

Le diplôme d'études approfondies (DEA) a été créé en 1964 dans les facultés des sciences. Le *Traitement de l'information* figurait, là aussi, parmi les 53 spécialités énumérées dans le décret fondateur. Généralement préparé à la suite de la maîtrise, le DEA remplaçait les certificats d'enseignement supérieur de 3^e cycle, sanctionnant la première année des études doctorales. La soutenance de thèse devait avoir lieu au minimum un an après l'obtention du DEA. La réforme Fouchet de 1966 plaçait, dans les facultés des sciences et des lettres, le 3^e cycle à la suite de la maîtrise avec la préparation du DEA en un an et du doctorat de 3^e cycle en un ou deux ans.

La licence ès-Sciences appliquées et le diplôme d'état de *docteur ès-Sciences mention Sciences appliquées* ont été créés par décret du 5 mai 1961 (« tendant à favoriser la formation technique supérieure dans les facultés des sciences) et modifiés en 1963 et 1966 lors de la réorganisation des deux premiers cycles d'enseignement dans les facultés des sciences. En 1966, on prévoyait de les supprimer respectivement en 1969 et 1973. La préparation du doctorat ès sciences appliquées était accessible aux titulaires d'une licence ès sciences classique, aux titulaires d'une licence ès sciences appliquées munis d'un certificat d'études supérieures complémentaire et aux titulaires de certains diplômes d'ingénieur munis d'un certificat d'études supérieures complémentaire.

Avec la multiplication des diplômes et des particularités selon les disciplines universitaires, la réglementation les concernant est devenue très complexe. D'où le souci de

les simplifier et de les unifier. Ce sera l'objet d'une série de grandes réformes, du milieu des années 1966 à la réforme LMD des années 2000.⁴

En 1966, la réforme Fouchet place les doctorats créés en 1808 et les doctorats de 3^e cycle à la suite de la maîtrise. C'est alors qu'apparaît l'expression *doctorat d'État* pour distinguer les premiers des seconds.

La loi Edgar Faure d'orientation de l'enseignement supérieur, en 1968, retouche l'ensemble des doctorats de toutes disciplines. Elle supprime la 2^e thèse dans le doctorat d'état. Une mesure assez révolutionnaire, répondant à la fois aux critiques d'extrême-gauche contre « l'individualisme bourgeois » et aux besoins de l'industrie et de la recherche technologique en chercheurs sachant travailler en équipe, prévoit que la thèse et les travaux connexes « peuvent être individuels ou, si la discipline le justifie, collectifs, déjà publiés ou inédits. Dans le cas où la thèse ou les travaux résultent d'une contribution collective, le candidat doit rédiger et soutenir un mémoire permettant d'apprécier sa part personnelle. » Notre enquête fait effectivement apparaître plusieurs groupes de doctorants ayant présenté des thèses dans le cadre de projets collectifs. C'est le cas par exemple, à la fin des années 1960, dans le centre scientifique IBM-IMAG (Grenoble) où l'on expérimente un réseau d'ordinateurs ; et à Toulouse où une dizaine d'ingénieurs de l'ENSEEIH consacre leurs thèses à développer les éléments d'un gros système d'exploitation novateur pour l'ordinateur de l'IRIA, sous la direction de l'ingénieur de l'Armement Henri Boucher.

La réforme de 1974 a pour principal effet de réunir les réglementations des trois doctorats sous des arrêtés indépendants des disciplines. Le *doctorat d'État*, correspondant au grade de docteur du décret de 1808, « sanctionne la reconnaissance, par un jury de l'aptitude du candidat à mettre en œuvre une recherche scientifique originale de haut niveau. » La maîtrise est obligatoire pour postuler à ce diplôme. Il s'obtient après la présentation d'une thèse ou d'un ensemble de travaux et se situe toujours dans l'une des grandes catégories disciplinaires : sciences, droit, etc. sans autre précision de sous-discipline. Il reste indispensable pour accéder à un poste de maître de conférences ou de professeur titulaire. Le *doctorat de 3^e cycle*, comme le diplôme de *docteur-ingénieur*, sanctionne une formation acquise dans la pratique de la recherche. La préparation dure trois ans dont une consacrée au DEA. Le doctorat de 3^e cycle fait suite à la maîtrise, tandis que le diplôme de docteur-ingénieur est destiné aux titulaires du titre d'ingénieur.

La réforme de 1984 (loi Savary) n'a laissé qu'un doctorat unique, correspondant au grade de docteur du décret de 1808. Ce doctorat unifié est devenu le diplôme de référence

4 Pour une synthèse, voir Patrick Fridenson, « La politique universitaire depuis 1968 », dans Jean-Michel Chapoulie, Patrick Fridenson & Antoine Prost (dir.), *Le Mouvement Social*. n° sur « Mutations de la science et des universités en France depuis 1945 », 2010. Vol. 4, n° 233, p. 47-67.

pour accéder au corps des maîtres de conférences, qui s'est substitué dans la même loi à celui des maîtres-assistants. L'objectif était de n'avoir plus en France qu'un seul doctorat de recherche dont la durée de préparation serait plus conforme aux « standards internationaux », c'est-à-dire aux Ph.D. La durée de préparation de la thèse de doctorat de 3^e cycle était en effet jugée trop courte (1 à 2 ans) et celle de l'ancien doctorat d'État trop longue (4 à 7 ans en sciences, souvent le double en lettres). Toutefois la loi Savary stipule que l'aptitude à diriger des recherches est sanctionnée par une habilitation (HDR), nouveau diplôme qui donne accès au corps des professeurs des universités, donc perpétue le doctorat d'État sous un nouveau nom.

Nos entretiens avec les premiers informaticiens de Grenoble (J.-P. Verjus, etc.) livrent une évaluation « vécue » de ces formations au milieu dans les années 1960. Dans les faits, le diplôme de docteur-ingénieur se faisait en trois ans, donc était plus coté que la thèse de 3^e cycle (deux ans). Les deux débouchaient sur des postes de maîtres-assistants. La thèse de docteur-ingénieur a vite été labellisée *Informatique*, tandis que la thèse de 3^e cycle était en *Mathématiques / spécialité Traitement de l'information*. L'Institut Polytechnique de Grenoble a labellisé des thèses *Informatique* plus tôt que l'Université. Louis Bolliet précise que le doctorat de Sciences appliquées se situait entre le diplôme de docteur-ingénieur et le doctorat d'état ; il y en eut quatre en informatique à Grenoble.

Sources et méthode

« D'un simple clic de souris » ? Pour effectuer une étude quantitative en histoire de l'informatique, mieux vaut ne pas compter exclusivement sur les ressources en ligne. Le système de recherche Sudoc, la base nationale des thèses soutenues en France, qui a heureusement une cellule d'assistance dévouée, ne trouve pas toutes les thèses demandées, loin de là, et comporte bien des lacunes.

Ainsi, interrogé avec le terme *informatique*, il ne donne pas celles de J. Hebenstreit (première thèse d'état répertoriée *informatique*) et d'autres fondateurs de la discipline comme J. Pitrat, M. Gross, L. Nolin, A. Lentin, J.-Cl. Boussard, L. Bolliet... En 2008, quand j'ai commencé cette enquête, la recherche par Sudoc avec le mot *informatique* faisait apparaître une masse de thèses de l'Université de Toulouse (une soixantaine pour la période considérée), presque rien de Paris ou de Grenoble (une demi-douzaine chacune seulement), encore moins des pôles secondaires. Une interrogation avec les mots-sujets : *mathématiques appliquées*, *analyse numérique*, *informatique*, sur l'Université d'Aix-Marseille 1 entre 1958 et 1973, ne fait apparaître que 10 thèses de 1972-1973, la plupart soutenues à... Toulouse ou Paris ! Même genre de résultat pour les requêtes sur Lille. Pour Nancy, Sudoc ne fait apparaître que quelques thèses d'ingénieur-docteur, en 1972-1973 seulement ; et ne répertorie pas cinq

thèses soutenues et conservées à Nancy, celles de : Roger Dunstetter, Photius Nanopoulos, Geneviève Jomier, Jean-Claude Bietry et Maryse Tommasini.

Les libellés ne sont pas totalement fiables. Il est étrange que Sudoc libelle *Informatique*, en 1960, la thèse d'université de Pierre-L. Guichet, ou encore la thèse d'état de Jean-Claude Boussard, en 1964, alors que ce libellé n'existait pas à l'époque. Généreux, Sudoc crédite la même année 1973 Joffroy Beauquier de deux thèses, l'une apparemment d'état (« Thèse : Mathématiques »), l'autre en « 3e cycle: Informatique algébrique » (Paris 7, *Caractérisation des générateurs standards du cône rationnel des langages algébriques*). Il faut une recherche complémentaire pour vérifier que la première est une erreur de Sudoc, puisque l'intéressé a passé un doctorat d'État en Mathématiques à Paris 7 en 1977 (*Contribution à l'étude de la complexité structurelle des langages algébriques*). La thèse de 3e cycle du toulousain Serge Castan, passée en 1963, est libellée « Thèse de doctorat : Mathématiques Appliquées », alors qu'il soutiendra sa thèse d'état cinq ans plus tard. Certaines thèses ne sont pas datées, comme celle de Patricia Bournai, *Un système de CAO pour la simulation des systèmes physiques contrôlés par réseaux de Petri* (je n'ai donc pu l'inscrire dans ma liste).

Les catégories disciplinaires ne sont pas totalement délimitées. Ainsi une thèse qui relève évidemment de l'Informatique de gestion, celle de Jean-Marie Aragon sous la direction de Michel Laudet (*Étude d'une méthode optimisée de gestion des fichiers sur mémoire de masse*), est classée en Automatique.

Défaut plus général, la base Sudoc ne mentionne que rarement le directeur de thèse, encore moins les membres du jury⁵. Le type de doctorat n'y est pas toujours précisé.

Quelle méthode de recherche adopter ? On a choisi de commencer par *dégrossir le sujet* dans Sudoc. Les deux critères de recherche inscrits ont été la période qui nous intéresse, 1956-1973, et la « note de thèse » qui dans Sudoc désigne la discipline ou la sous-discipline dans laquelle une thèse est libellée. Pour ce faire, il faut connaître les intitulés de ces disciplines, tels qu'ils ont été définis successivement par les textes réglementaires (voir plus haut).

Pour les thèses de 3^e cycle ou de docteurs-ingénieurs, on a utilisé les termes :

- *Traitement de l'information* (inscrit en 1964 dans la législation pour les DEA, les thèses de 3^e cycle et les thèses de docteurs-ingénieurs) : 39 thèses trouvées. [dont 22 nouvelles], dont 35 de 3^e cycle, 3 de Dr.-Ingénieurs et 1 en électronique.

- *Informatique* (terme autorisé plus tard, mais qui semble utilisé dès 1966 à Toulouse) : 186 thèses trouvées. [dont 12 nouvelles]

⁵ Il y aurait de quoi écrire un rapport sur les lacunes ou les étrangetés de Sudoc. Ainsi la recherche par Sudoc avec les mots clés *Programmation, calculateur, compilation...* ne fait pas apparaître des thèses parmi les plus importantes dans l'histoire de l'informatique universitaire : celles de Nivat, de Gross ou de Lentin, Nolin, Pair... Heureusement Sudoc emploie une équipe dévouée, qui répond aux questions mais ne peut remédier aux défauts du système.

- *Mathématiques appliquées* (terme plus général, englobant l'analyse numérique, les probabilités et statistiques) : 260 thèses trouvées, dont une fraction seulement a été retenue comme pertinente, en fonction de la place qu'y tiennent le calcul et les algorithmes, et de leur contexte institutionnel (présence ou non de l'informatique). [dont 30 nouvelles]

- *Automatique* (terme introduit en 1964 dans la même réglementation, mais l'on trouve dès 1961 quelques thèses en *Servomécanismes*) : 120 thèses trouvées, dont une fraction seulement a été retenue comme pertinente, en fonction de la place qu'y tiennent le calcul et les technologies numériques. [dont 14 nouvelles et pertinentes]

- *Électronique* (terme introduit en 1964 dans la même réglementation) : associé au terme *ordinateur*, 17 thèses trouvées [dont 5 jugées pertinentes, car portant au moins partiellement sur la conception d'un système, au-delà du niveau des composants élémentaires]

- En « mots du sujet », les noms des principaux langages de programmation de l'époque : *Fortran*, *Cobol*, *Algol*, et d'autres termes caractéristiques comme *algorithme*, *calculateur*, *compilat** ou *programm**.

On n'a pas retenu les mémoires de diplôme d'ingénieur du CNAM, qui apparaissent dans Sudoc parmi les thèses.

L'utilisation de la ligne « note de thèse » ne sert pas à grand-chose pour les doctorats d'état, où les thèses étaient libellées *Sciences*, parfois seulement avec un sous-libellé. Là, il faut donc chercher par nom de directeur et par divers mots-clés en espérant utiliser un vocabulaire assez vaste pour ne rien omettre. Notre filet était-il assez serré, ses mailles ne présentaient-elles aucune déchirure ? Nous le saurons par les retours d'information des principaux intéressés, quand ils verront cette liste sur l'internet.

Avec de l'expérience et du temps, en utilisant successivement une large gamme de mots-clés, en interrogeant alternativement par ville, par directeur de thèse (quand on le connaissait), bref en passant au pilotage manuel du système informatique, ce ratissage parvient en fin de compte à obtenir de Sudoc un nombre appréciable de titres.

On a croisé cette collecte avec celle du *Mathematics Genealogy Project* qui donne de précieuses indications, mais plus sur l'ascendance scientifique des mathématiciens que sur leur descendance, faute que les données soient complètes⁶ : on en revient toujours au même problème de sources.

Cette collecte des thèses effectuée, on les a intégrées en un fichier Excel. Puis on les a interclassées par année, en supprimant les éventuels doublons. L'ensemble obtenu n'étant pas

⁶ Le *Mathematics Genealogy Project* (<http://genealogy.math.ndsu.nodak.edu/>) est un service de la North Dakota State University, associée à l'American Mathematical Society. Il vise à dresser l'arbre généalogique scientifique des mathématiciens, les ascendants étant les patrons de thèse, les descendants les thésards. C'est un *work-in-progress* permanent, dont les principales sources sont les bibliographies en ligne et les informations remontant des internautes. Ainsi, pour Joseph Pérès, il ne trouve (à la date où j'écris) que... quatre « fils » : Pierre Huard, Henri Cabannes J. Legras et J.-M. Souriau ! (Malavard n'apparaît pas). En revanche, l'effet Saint-Mathieu joue à plein pour J.-L. Lions (1 246 descendants !).

complet, loin de là (par exemple de nombreuses thèses parisiennes n'apparaissaient pas), on a complété progressivement la liste « à la main », par des recherches complémentaires :

- En consultant la base des thèses mises en ligne par l'université de Grenoble, essentiellement les listes des thèses préparées à l'IMAG. Il existe en effet une autre ressource en ligne, le catalogue des thèses recensées par la bibliothèque de l'ex-IMAG : <http://tel.archives-ouvertes.fr/>. Les thèses de l'IMAG sont elles-mêmes en ligne en version intégrale, ce qui fournit une source de grand intérêt pour l'histoire des sciences.
- Le fonds des thèses françaises en mathématiques consultées à la bibliothèque de l'ENS Ulm. Celle-ci constitue une ressource magnifique pour le chercheur, qui peut consulter à loisir toutes les thèses de mathématiques soutenues depuis le XIX^e siècle.
- Les bibliographies annexées aux quelques monographies publiées sur l'histoire de l'informatique à Grenoble (Verjus 1988, David & Deguerri 2008, etc.), à Toulouse (Laborie 1990, Thirriot 1993) et à Nancy (article de Cl. Pair en 1988 et biographie en ligne de J. Legras). Et l'historique bien documenté rédigé par Marie-Thérèse Pourprix, Yves Leroy et Pierre Vidal, mis en ligne sur le site ASA par l'Association de solidarité des anciens de l'université des sciences et technologies de Lille⁷.
- Les bibliographies des nombreux ouvrages et thèses d'informatique consultés lors de mes recherches.
- Des listes de thèses soutenues à Nancy, à Clermont et à Toulouse, transmises par des chercheurs ou des archivistes (pour Nancy, merci à Gérard Eguether et Philippe Nabonnand).
- En obtenant la liste des thèses préparées au LIMSI, communiquée par ce laboratoire (le CCA, installé à Orsay, devenu LIMSI en 1972), puis en éliminant celles où le calcul semble marginal.
- Les rapports annuels de diverses facultés ou laboratoires, ainsi l'institut Blaise Pascal et l'institut de Programmation de Paris, qui énumèrent les thèses soutenues dans ces unités.
- Le catalogue en ligne de la bibliothèque de l'INRIA, qui offre une dizaine de thèses antérieures à 1974 (notamment quelques thèses toulousaines que l'université de Toulouse semble avoir perdues).

Ce travail fait, on a procédé à nouveau à de nombreuses vérifications sur la base Sudoc et sur d'autres sources comme les CV en ligne de professeurs et de chercheurs.

Les archives livrent parfois des bilans de formation doctorale avec une répartition entre (sous-) disciplines. Ainsi, fin 1966, la Faculté des Sciences de Grenoble et son institut d'informatique et de mathématiques appliquées (IMAG) fournissent, à l'appui d'une demande d'équipement en nouveaux ordinateurs, la liste des thèses soutenues dans les domaines concernés : en Méthodes mathématiques et calcul numérique, la plus ancienne section, 35 thèses (dont 3 d'état et 1 de Dr. ing) ; en Logique et technologie des calculateurs, 7 thèses ; en

⁷ M.-T. Pourprix, *Des mathématiciens à la faculté des sciences de Lille : 1854-1971*, Paris, L'Harmattan, 2009.

Programmation et étude des langages, 10 thèses + 7 sous la direction du CETA (Centre d'études de traduction automatique) ; en Recherche opérationnelle, 4 thèses⁸. Mais ce classement est exceptionnel et, faute d'en disposer dans la plupart des universités, l'historien doit faire ses choix et compter sur le retour d'information critique des spécialistes, notamment des auteurs de ces thèses ou de leurs anciens collègues.

Bien antérieurement à la tranche chronologique que nous étudions, on pourrait considérer que les premières thèses françaises consacrées aux techniques du calcul et aux machines à calculer sont :

- la thèse de doctorat de Louis Couffignal (1938), *L'Analyse mécanique, application aux machines à calculer et à la mécanique céleste*,
- la thèse d'ingénieur-docteur de Charles-Henri Chamard (1942), *Contribution à l'étude des machines à calculer. De l'extraction des racines. Annexe sur la division et la multiplication automatiques*. (2e thèse : *Théorie des intégrateurs mécaniques*) ,
- la « 2e thèse » du doctorat d'André Saint-Joanis, *Les Machines à calculer arithmétiques*, la 1ère thèse étant un travail de Physique, *Étude d'un nouveau fréquencesmètre, Application au réglage fréquence-puissance* (1951, Université de Grenoble, dir. L. Néel),
- la thèse de doctorat du Yougoslave Dusan Mitrovic (1953), *Étude théorique d'un principe nouveau de construction des machines électroniques servant à résoudre les systèmes d'équations algébriques linéaires*.

On ne sait ce qu'est devenu C.-H. Chamard, mais les trois autres lauréats ont poursuivi une carrière scientifique ultérieure, Couffignal portant brièvement les espoirs du CNRS après guerre en « calcul mécanique ». L'existence de ces thèses est intéressante historiquement, mais elle n'est pas significative pour la présente étude statistique.

Critères et limites de constitution de la base

La limite la plus évidente tient à l'imprécision de la catégorie « informatique » alors en construction. Si les thèses consacrées à un outil de programmation (assembleur, langage, compilateur, etc.), qui apparaissent dès 1959, en relèvent évidemment, d'autres sont à cheval sur ce nouveau domaine et sur des disciplines préexistantes, principalement en physique et en mathématiques. Pour employer le langage de la théorie de l'information, on a pris le parti de préférer l'excès de « bruit » au défaut de « signal » : on s'est efforcé de répertorier ici toutes les thèses qui paraissaient contenir un important travail sur calculateur électronique, ou qui attestaient une contribution au développement d'une machine ou d'un programme. Quitte à éliminer plus tard celles qui paraîtraient trop hors sujet.

⁸ Faculté des Sciences de Grenoble, dossier extension IBM/360-67, Ministère de l'Education Nationale, Arch. Nat. 19771326, art. 7.

Par exemple, au début de la période étudiée, on a retenu les thèses d'une physicienne (Claire-Liliane Schwartz, ép. Gaudferneau, 1955) et de trois astronomes (Jacques Arsac, Jean-Louis Rigal, Bernard Vauquois, 1958), car leurs auteurs y ont programmé des calculs sur ordinateurs, puis se sont orientés vers l'informatique.

On note l'importance des thèses en analyse numérique. Ce n'est pas de « l'informatique », mais cette branche de mathématiques appliquées apparaît comme la matrice de l'informatique, à la fois sur le plan institutionnel et, dans beaucoup de cas, sur le plan individuel : bien des docteurs en analyse numérique sont ensuite devenus informaticiens. Par exemple Michel Ritout, qui a passé en 1963 à Toulouse un doctorat d'état en mathématiques (*Contribution au problème du lissage des relevés expérimentaux*), est ensuite employé à programmer des calculs de navigation des sous-marins pendant son service militaire, puis fait carrière dans le service informatique. Cécile Courtois, qui a soutenu en 1964 à Nancy une thèse classique de mathématiques appliquées (3^e cycle, *Approximation de l'équation de la chaleur au sens des moindres carrés*), s'oriente ensuite – sous la direction du même professeur, Jean Legras – vers un sujet on ne peut plus « informatique » : l'Algol linguistique (un témoin la crédite d'une thèse d'état sur ce sujet, mais il a été impossible d'en trouver la trace).

Dans la constitution de notre base, on a donc pris le parti d'inclure les thèses en analyse numérique jusqu'à l'époque où l'informatique s'en détache ou s'en distingue dans les catégories universitaires, à la fin des années 1960. Il est évident qu'une partie importante des thèses relève des deux domaines à la fois, par exemple celle de Jean Vignes en 1969, comme celles de nombreux élèves de J.-L. Lions, d'E. Durand ou de J. Kuntzmann.

Un problème similaire de délimitation se pose avec l'automatique. Là aussi on a pris le parti d'englober des thèses d'automatique, étant donnée la proximité intellectuelle et institutionnelle avec l'informatique : ces deux sciences appliquées ont pris essor dans les mêmes pôles universitaires, ont partagé certains enseignements, se sont renforcées mutuellement – ne serait-ce que face à la réserve des « sciences pures » à leur égard. Elles ont été associées sur le plan institutionnel, d'abord dans l'IRIA créé en 1967, puis dans la commission du CNRS définie en 1975. C'est dans des laboratoires d'automatique qu'ont commencé les premières constructions d'ordinateurs en milieu universitaire, à Toulouse, à Grenoble et à Lille.

Je voudrais bien savoir, par exemple, comment classer en automatique plutôt qu'en informatique la thèse du Lillois Marc Gillant, *Langages de description de processus et d'automatismes séquentiels* (1972, 3^e cycle : Electrotechnique-automatique), ainsi résumée :

« On caractérise le fonctionnement d'un **automatisme** et on définit les caractéristiques des **processus**. Conformément aux définitions et propriétés, on envisage la construction et la réduction des **équations booléennes des systèmes** asynchrones et la définition de la structure de

bloc de commande des systèmes synchrones. On définit également les systèmes programmes. On élabore le **langage desprog** permettant la mise au point et la description des organigrammes de commande et de processus qui définissent le fonctionnement d'un processus ou d'un système cyclique. On utilise la **forme normale de Backus** [un élément de la définition d'Algol, élaboré par J. Backus, l'inventeur de Fortran] pour décrire **les règles de production du langage**. On **construit ensuite le langage Desaucyk** autorisant des études plus importantes »⁹.

Dans ce type de cas, le critère de labellisation est d'ordre institutionnel, c'est le nom et l'objet du laboratoire où la thèse est préparée (ici, la commande de processus automatiques), mais l'essentiel de cette recherche-développement porte sur du logiciel. Des travaux comme celui-ci attestent la convergence de l'automatique et de l'informatique, qui conduira en 1975 à créer au CNRS une commission rassemblant ces deux disciplines.

Au-delà des contenus scientifiques de ces thèses, qu'on pourra juger plus ou moins pertinents pour l'informatique, leur inclusion se justifie pour une raison institutionnelle : ces travaux et la présence des doctorants qui les menaient contribuaient à renforcer les équipes de mathématiques appliquées et d'informatique, à faire masse pour s'imposer dans la concurrence entre les disciplines. De plus, en formant des numériciens ou des automaticiens, par définition l'on formait des utilisateurs d'ordinateurs. Et par conséquent l'on renforçait les arguments en faveur du développement des laboratoires de « calcul électronique » qui étaient en même temps des centres de recherche, au moins jusqu'en 1970.

Des thèses soutenues dans de tout autres disciplines pourraient s'ajouter au corpus, en ce qu'elles ont pu contribuer à l'innovation – ainsi celle de Marie-France Kouloumdjian, *Le vécu de la relation homme-machine dans l'enseignement assisté par ordinateur* (3e cycle, Psychologie, Lyon 2, 1975, dir. Jean Guillaumin). Mais elles ne sont pas significatives du point de vue statistique.

⁹ Sudoc, Marc Gillant, *Langages de description de processus et d'automatismes séquentiels*, Lille 1, 1972, thèse de 3e cycle : Electrotechnique-automatique. [mots soulignés en gras par moi]

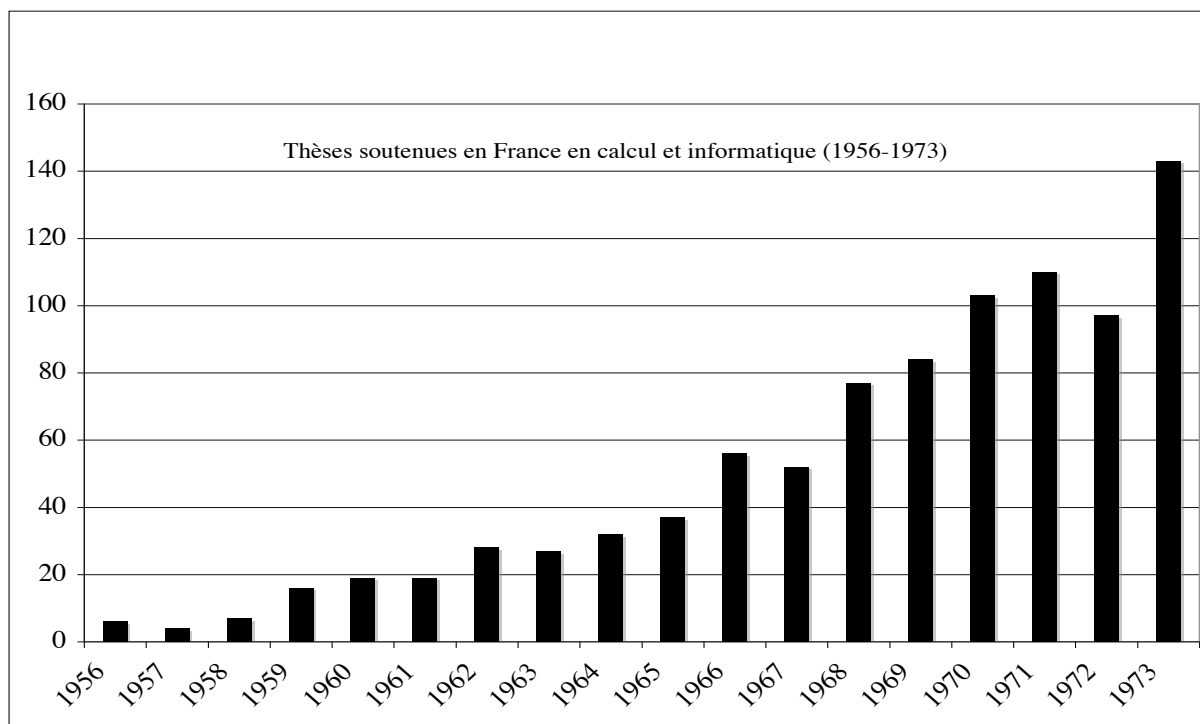


Figure 1. Thèses en calcul et en informatique soutenues entre 1956 et 1973, d'après le fichier à jour en février 2020. La silhouette générale de l'histogramme est similaire à celle de fin 2009, malgré l'ajout de plus de 250 titres.

Quelques remarques sur les limites des conclusions que l'on peut tirer de cette base

Quelle était la signification, la valeur du doctorat ?

Contrairement à l'Université que nous connaissons depuis la fin du XX^e siècle, celle des années 1960 employait une masse d'enseignants qui n'étaient pas docteurs. Comme le remarque Marie-Thérèse Pourprix dans son ouvrage sur les mathématiques à Lille, les professeurs n'encourageaient pas toujours leurs assistants à préparer des thèses. A cette époque de « démocratisation » universitaire, les assistants étaient souvent issus de milieux où personne n'avait jamais rêvé de devenir « un savant », et le simple fait d'avoir un poste universitaire pouvait être très satisfaisant. Il était souvent admis que les thèses, c'était « surtout pour les normaliens ».

Faire de la recherche n'impliquait pas obligatoirement de publier beaucoup, ni de terminer une thèse. Et l'enseignement était hautement prioritaire sur la recherche. Il est frappant de constater qu'à Caen, par exemple, un professeur aussi actif que Pham en analyse numérique ne forma en 15 ans qu'une seule doctorante (d'après Sudoc, confirmé par deux anciens assistants de Pham, qui s'inscriront eux-mêmes en doctorat à Paris).

Le doctorat n'était pas non plus un point de passage obligé pour devenir chercheur ou professeur d'informatique ou de mathématiques appliquées. Citons trois exemples.

François Genuys (ENS 1946) est entré au CNRS en 1947 pour préparer une thèse, mais le sujet d'analyse que son professeur lui a imposé l'ennuie et il abandonne (Genuys est, par goût profond, proche de Bourbaki). Il enseigne les mathématiques dans un lycée quand IBM France le recrute en 1957 pour assurer les relations entre son centre de calcul et le milieu scientifique. Genuys anime bientôt un séminaire d'analyse numérique à l'Institut d'Astrophysique de Paris, puis enseigne cette matière à la faculté des sciences dans le cadre de l'Institut Blaise-Pascal – cours qu'il transmettra ensuite à Jacques-Louis Lions. Il recentre son séminaire à partir de 1960 sur les langages de programmation et reste au long des années soixante l'un des principaux représentants de la recherche et de l'enseignement français dans ces domaines, participant notamment aux commissions du CNRS et à divers jurys.

Jean-Jacques Duby (ENS 1959), lui aussi agrégé de mathématiques particulièrement doué, commence en 1961 une thèse en logique avec Daniel Lacombe. Ce sujet l'amène à s'intéresser au langage LISP et à apprendre à se servir d'un ordinateur. Il entre chez IBM à l'occasion d'un stage au département de recherche de Yorktown Heights, y participe à la conception du langage Formac, puis dirige un groupe de recherche opérationnelle à Paris. Il a vite abandonné sa thèse pour se consacrer à la recherche en informatique. Duby est désigné en 1967 pour diriger le centre scientifique IBM-IMAG de Grenoble et ne cessera de diriger des recherches et d'enseigner jusqu'à sa nomination dans l'état-major du CNRS, puis à la direction de Supélec dans les années 1980. Il est nommé entre temps professeur des universités.

Alice Recoque (Physique-Chimie Paris, 1955), après 15 ans passés à développer des ordinateurs à la SEA, puis à la CII, commence au début des années 1970 une thèse sur les architectures massivement parallèles, sous la direction de Bolliet (Grenoble). Mais, prise en 1975 dans la bousculade de l'absorption de la CII par Honeywell-Bull, elle l'abandonne. Elle n'en continuera pas moins à enseigner dans des écoles d'ingénieurs et à animer des recherches chez Bull.

Il serait donc intéressant de compléter la présente enquête par un sondage sur les thèses entreprises, mais non abouties. Les trois exemples ci-dessus, retenus parce qu'il s'agit d'acteurs bien visibles sur la scène de la recherche informatique, ne sont sans doute pas représentatifs de la moyenne des doctorants inaboutis. Mais ils indiquent que les spécialistes ayant un niveau doctoral étaient très sollicités à la fois par les entreprises et par l'enseignement supérieur.

Premiers apports de cette enquête

L'enquête et la constitution de la base ont apporté des renseignements à la fois statistiques et non-quantitatifs.

La base fait apparaître des thèses qui représentent des « premières », des marqueurs chronologiques de l'émergence de l'informatique comme discipline se dégageant des mathématiques appliquées. Les deux premières thèses décrivant des systèmes de programmation sont celles de Louis Cariou (Rennes, 1959, docteur-ingénieur) et celle, plus importante, de Marion Créhange, qui améliore le système PASO d'IBM (Nancy, 1961, 3^e cycle). La thèse de Michel Lucas (3^e cycle, Grenoble, 1968), *Technique de programmation et d'utilisation en mode conversationnel des terminaux graphiques*, est considérée comme la première thèse consacrée à l'informatique graphique en France. La première thèse française sur le calcul formel par ordinateur est celle d'Yvon Siret (Grenoble, 1970), qui signale en conclusion le nombre croissant de publications et de réunions sur ce thème dès la fin des années 1960. Il importe de placer ces « premières » universitaires dans un contexte où des recherches non-universitaires contribuaient tout autant à l'édification de la nouvelle technique – ainsi, pour l'informatique graphique, la thèse de Michel Lucas est contemporaine des courbes de Bézier développées chez Renault.

La base montre des « grappes » de thèses (comme l'on parle de « grappes de brevets »), qui révèlent ou confirment l'existence, dans telle université, d'un programme de travail collectif. Ainsi, à Rennes puis à Nantes autour de 1960, où G. Brillouet fait développer des outils de programmation pour l'IBM 650 ; à Grenoble peu après, où une série de thèses résulte d'un ensemble cohérent de recherches sur la compilation des langages, animé par L. Bolliet ; à Toulouse à la même époque, où de nombreux sujets de doctorats sont liés au développement d'un ordinateur, la CAT ; à Marseille au début des années 1970, où trois thèses au moins portent sur la démonstration automatique, sujet porté par A. Colmerauer. L'observation de ces ensembles suscite la curiosité et demande à être croisée avec d'autres sources, notamment les témoignages des acteurs, les archives des universités en question et le contenu des thèses elles-mêmes, afin de mieux connaître ces programmes de recherches. Mais la simple lecture des titres de thèses en montre l'évolution : à côté de travaux essentiellement techniques, qui caractérisent jusqu'à nos jours cette « science appliquée », apparaissent des problèmes beaucoup plus abstraits qui révèlent à la fois le besoin de lui donner une assise théorique et le fait qu'elle commence à attirer des chercheurs fondamentalistes, généralement des algébristes.

La colonne *Diplômes* révèle l'apparition progressive de l'informatique comme discipline reconnue. La plupart des thèses soutenues en ce domaine dans les années soixante sont répertoriées en « Mathématiques appliquées », parfois en « Sciences appliquées » (mention créée par décret en 1961). Toulouse, dès 1966, indique la mention « Traitement de l'information » (mention créée par décret en 1964) ou « Informatique » sur des thèses de 3e

cycle ou plus souvent de docteur-ingénieur ; même chose à Grenoble peu après. La mention « Informatique » est admise en 1969 au niveau national, comme en témoignent les thèses d'état de Michael Griffiths (Grenoble) et de Jacques Hebenstreit (Paris).

L'identification des directeurs de thèse ou des membres du jury n'a pas toujours été possible. On a relevé leurs noms chaque fois qu'on les a trouvés, généralement sur la page de garde et dans les remerciements des thèses que nous avons pu examiner. Ces précisions confirment le rôle de quelques grands patrons de recherche, notamment dans les pôles pionniers. Elles montrent aussi des relations suivies entre plusieurs centres – ainsi la présence fréquente de Parisiens dans les jurys toulousains. Les noms des doctorants et leur nationalité (quand on la connaît par ailleurs) confirment l'ouverture des universités françaises aux chercheurs étrangers, notamment en sciences appliquées, ouverture déjà constatée au début du XX^e siècle.

La colonne « Prénoms » permet de distinguer les doctorants par sexe, en tenant compte bien sûr du fait que certains prénoms (Camille, Dominique, etc.) ne caractérisent pas un genre (là aussi, une connaissance personnelle du milieu étudié permet d'affiner les données). J'ai pu l'utiliser dans le cadre d'une autre étude, sur les femmes dans l'informatique : elles constituaient environ 10 % des doctorants à l'époque considérée – moins qu'en mathématiques, plus qu'en ingénierie classique.

Élaborée dans une perspective d'histoire des sciences et d'une discipline universitaire, cette base offre aussi des renseignements sur la recherche industrielle et sur la collaboration recherche-industrie. On trouve ainsi une thèse de 3^e cycle par un M. Behrouz Sedehi, *Assembleur et simulateur d'un calculateur en projet (Callisto)* (Paris, 1969), projet de très petit ordinateur que nous avons rencontré par ailleurs dans l'histoire du groupe Thomson. Plus importants, divers projets d'IBM, de Bull, de la CII, de l'Armement, donnent lieu à des recherches universitaires sanctionnées par des doctorats. En sens inverse, quelques développements initiés dans l'Université (projet de système de base de données Socrate, à Grenoble vers 1970, par exemple) seront ensuite industrialisés. Ce type de collaboration se pratiquait bien avant que l'informatique n'émerge, notamment en mécanique des fluides et dans d'autres branches des mathématiques appliquées.

Le tri par université de soutenance confirme l'importance des « pôles pionniers » (Grenoble, Toulouse, Paris, Nancy) et l'existence de pôles secondaires (Clermont) ou en rattrapage rapide au début des années 1970 (Lyon, Rennes), en même temps que l'inactivité de plusieurs grandes universités dans ce champ de recherche en émergence – hiérarchie révélée naguère par une enquête non-quantitative¹⁰.

10 Michel Grossetti et Pierre Mounier-Kuhn, "Les débuts de l'informatique dans les universités. Un moment de la différenciation géographique des pôles scientifiques français", *Revue Française de Sociologie*, vol. XXXVI, 1995, n° 2, p. 295-321.

Les particularismes locaux apparaissent aussi, par exemple les positions épistémiques dans le débat sur la place de l'informatique vis-à-vis des mathématiques. En 1968 à Grenoble on labellise encore « Mathématiques appliquées » une thèse sur les *Langages compilables* (Gabrielle Chambre, 3e cycle, dir. M. Nivat) qui serait certainement étiquetée « Traitement de l'information » voire « Informatique » à Toulouse. La base montre que, sur les thèses labellisées « Traitement de l'information », 11 ont été soutenues à Paris, 27 à Toulouse (de 1966 à 1973), zéro à Grenoble et à Nancy.

Il y a cependant quelques surprises. Notamment, comment interpréter le creux observé au début des années 1970 ? Reflète-t-il les restrictions budgétaires de l'après-1968 ? Elles incitent à pousser les investigations sur les particularités locales.

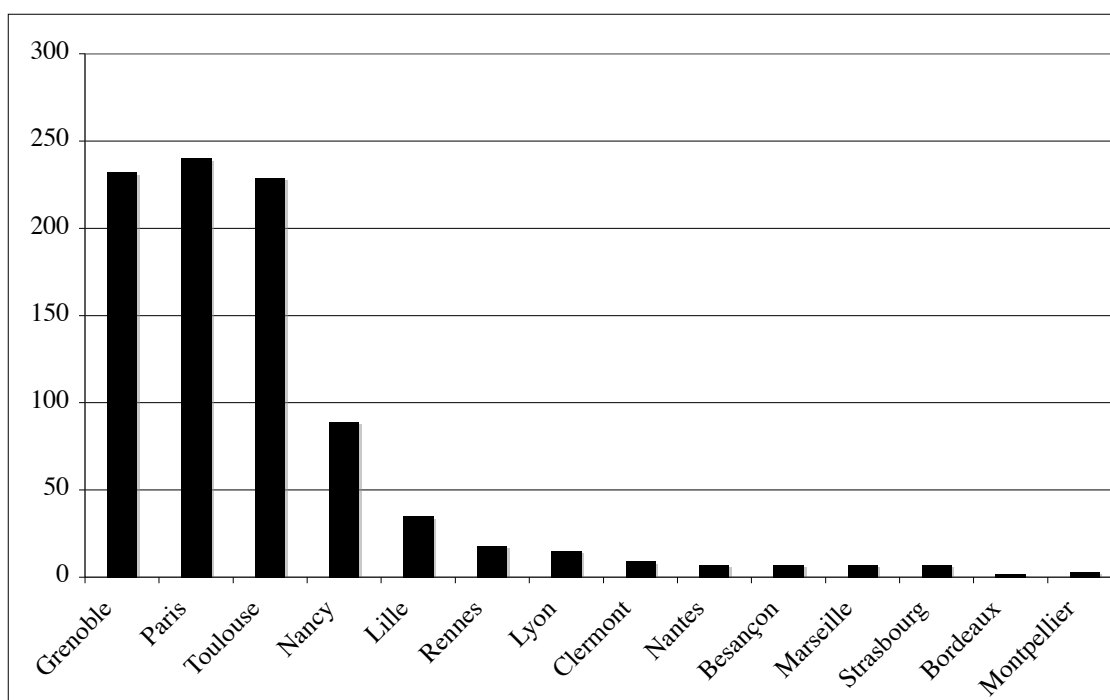


Figure 2. Nombre de thèses soutenues, par villes (1956-1973)

Particularités locales

On a réalisé des histogrammes par ville et par année, pour voir si les irrégularités au niveau national pouvaient se réduire à des accidents locaux.

À Paris (Paris-centre et Orsay), on observe une quasi-absence de thèses en 1960-1961, inexpliquée. Il se peut que des facteurs individuels, totalement contingents, suffisent à en rendre compte car il s'agit de très petits nombres. La forte augmentation à partir de 1967 résulte vraisemblablement de la mise en place du troisième cycle de l'Institut de Programmation en 1965, puis de la maîtrise d'informatique l'année suivante. Le creux spectaculaire de 1971 peut s'expliquer en partie par le départ du directeur de l'IP, Jacques Arsac, en année sabbatique au Canada.

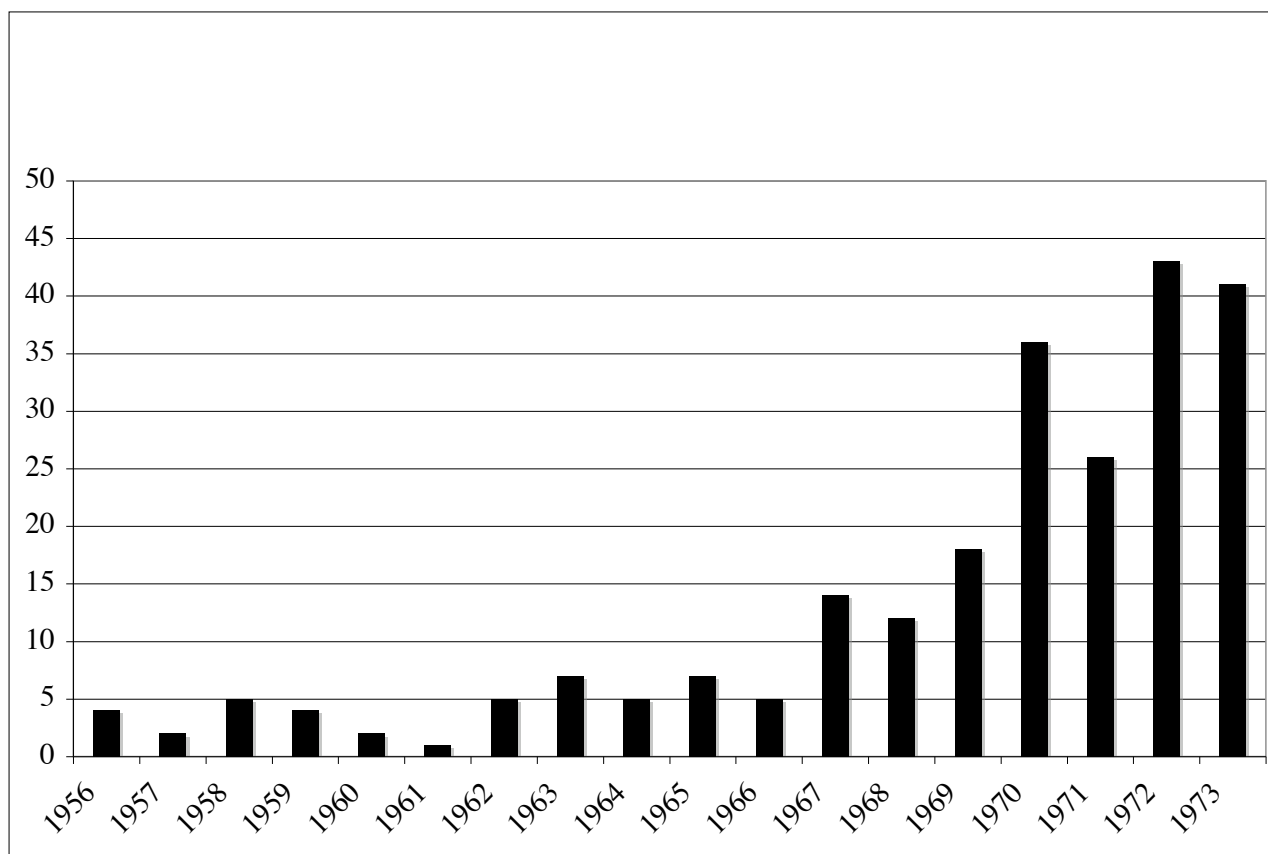


Figure 3. Thèses soutenues à Paris en calcul et en informatique

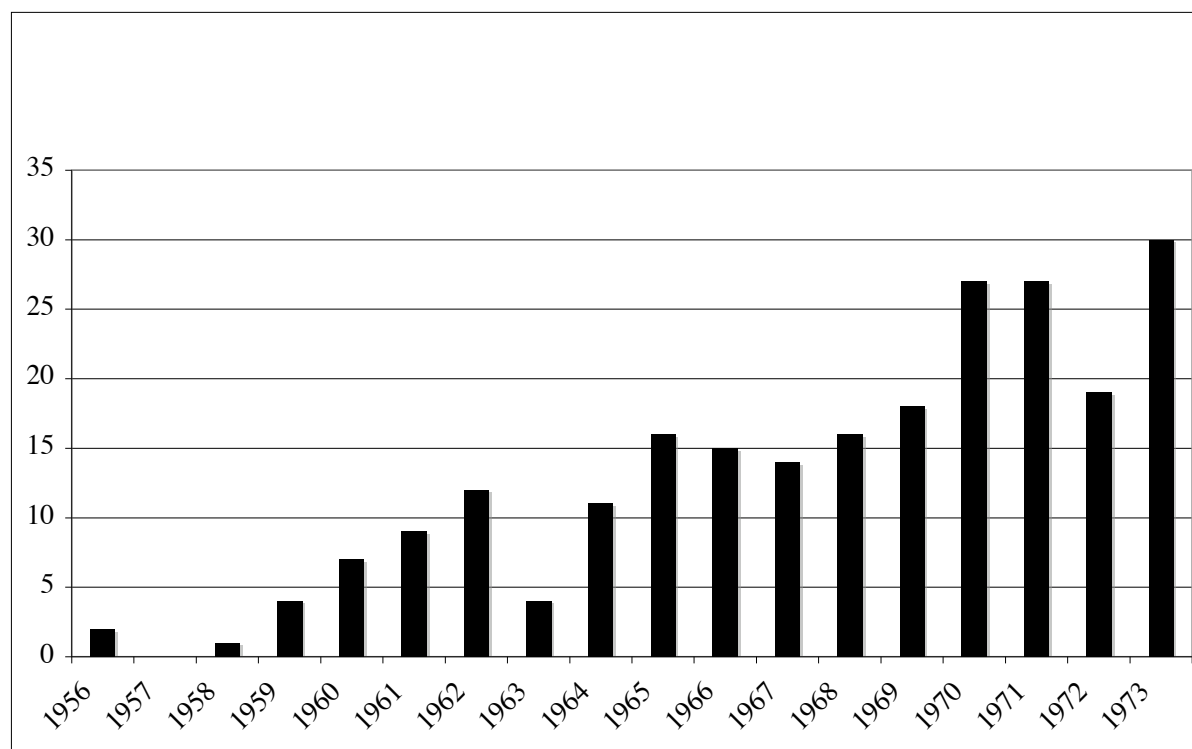


Figure 4. Thèses soutenues à Grenoble en calcul et en informatique

À Grenoble, on observe un creux en 1963 et un tassement en 1967-1969, inexplicables. La forte poussée de thèses en 1970-1971 résulte en partie de la création du centre scientifique IBM-IMAG, quatre ans plus tôt, qui a apporté de nouvelles problématiques, un encadrement supplémentaire et des moyens de calcul exceptionnels ; ce facteur local s'ajoute à la création de la maîtrise d'informatique, qui joue également à Paris, Nancy, Lille et Toulouse.

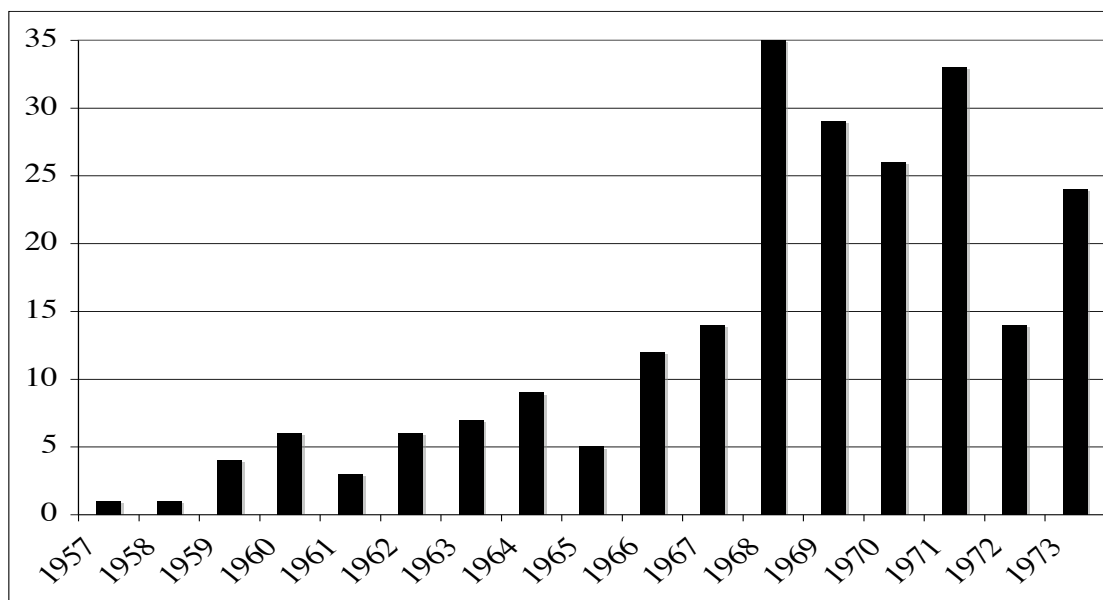


Figure 5. Thèses soutenues à Toulouse en calcul et en informatique

À Toulouse, l'histogramme révèle un plafonnement du nombre de thèses à la fin des années 1960, explicable par les tensions internes au milieu informaticien de cette faculté et le séjour de Michel Laudet à la direction de l'IRIA pendant cinq ans. De fortes variations annuelles s'observent à partir de 1968, les chiffres de 1972 retombant pratiquement au niveau de 1966-1967 (le pic de 1968 peut-il s'expliquer par le souci politique de « lâcher du lest » face à des assistants contestataires en leur décernant des doctorats ? Cette interprétation, plausible en ce qui concerne des licences ou des maîtrises, me semble douteuse concernant des thèses, pluriannuelles ; elle est de plus infirmée par le fait qu'on ne retrouve pas ce pic à Paris ou à Grenoble). On peut attribuer cette chute à trois facteurs locaux :

- Au milieu des années 1960, les équipes d'informatique toulousaines sont déchirées par une crise (conflit entre deux patrons de laboratoires).

- De 1967 à 1972, Laudet est directeur de l'IRIA, en région parisienne, et s'occupe moins de ses doctorants à Toulouse.

- En partie pour les raisons ci-dessus, le centre de calcul de l'université de Toulouse ne parvient pas à obtenir un nouvel ordinateur avant 1970 : étudiants et enseignants-chercheurs doivent se contenter de machines de 2^e génération à bout de souffle. Comme pôle de

recherche en informatique, Toulouse perd donc momentanément une partie de son attractivité, tandis que de nouvelles universités entrent sur le marché (Rennes, Lyon, etc.).

À Lille le nombre de thèses en calcul semble faible par rapport à l'activité d'enseignement et de recherche que l'on a repérée dans cette ville – tandis que les thèses d'automatique sont relativement nombreuses. Il est vrai que, jusqu'au début des années 1960, les thèses des mathématiciens Lillois ont parfois été soutenues à Paris plutôt qu'à Lille ; et que l'informatique n'y démarre vraiment qu'en 1962. D'autre part Lille ne dispose pas d'un vivier d'élèves ingénieurs susceptibles de devenir doctorants, comparable aux écoles d'électrotechnique de Toulouse, Grenoble ou Nancy. Quant à Clermont, qui démarre au même moment avec un plus petit potentiel, elle n'a aucune école d'ingénieurs dans ces disciplines.

Point commun à toutes les universités actives dans le domaine : le pic de soutenances en 1970, qui s'explique par la création de la maîtrise d'informatique en 1966, laquelle constitue dès lors une pépinière de futurs doctorants.

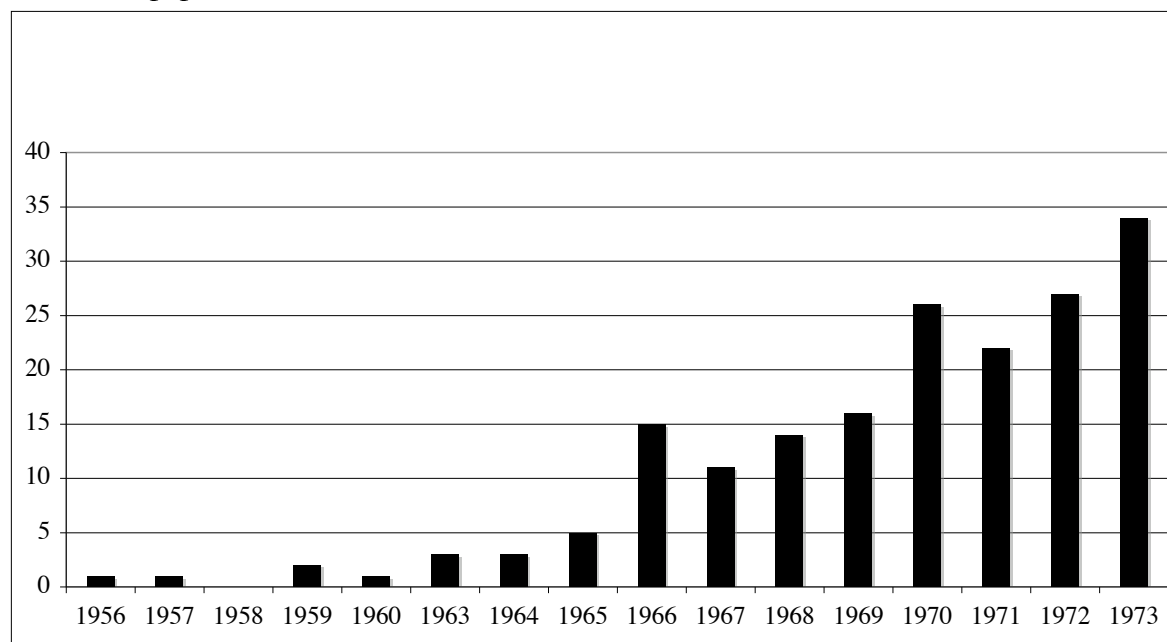


Figure 6. Thèses de docteurs-ingénieurs. Les universités comportant des écoles d'ingénieurs sont évidemment les mieux placées : 52 diplômes à Toulouse, 50 à Grenoble, pour seulement 36 à Paris.

Le creux des années 1970 : une crise des études scientifiques ?

Comment interpréter le creux observé au début des années 1970 – faible augmentation des soutenances en 1971, puis chute en 1972 ? On peut formuler trois hypothèses :

- S'agit-il d'un accident local dans l'un des grands pôles, Paris, Grenoble ou Toulouse, accident assez marqué pour faire baisser toute la moyenne ? Cette question a motivé le calcul des histogrammes pour ces trois villes. Or ils montrent une baisse à peu près simultanée dans ces différentes universités. On doit donc rejeter cette hypothèse : le phénomène est général.

- S'agit-il d'un biais introduit par le fait que, à partir de la fin des années 1960, nous ne prenons plus en compte les thèses portant essentiellement sur l'analyse numérique ou la mécanique des fluides ? C'est possible, mais cela n'explique pas le rebond de 1973 : ce facteur ne peut donc être que secondaire ou marginal.

- Pourrait-il s'agir d'une chute générale des soutenances cette année-là, au-delà de l'informatique ?

Pour le vérifier, on a interrogé Sudoc sur l'ensemble des thèses libellées « Sciences » (en « Note de thèse »), successivement sur chaque année de 1968 à 1973, en considérant simplement le nombre de réponses. Le résultat valide clairement notre troisième hypothèse :

Année	1968	1969	1970	1971	1972	1973
Thèses soutenues	1 228	1 449	1 330	1 085	1 271	1 321

Il y a donc bien une chute générale et forte du nombre de soutenances en 1971. Elle s'amorce dès l'année précédente, mais cela ne se voit pas en informatique, sans doute en conséquence de la création des maîtrises d'informatique en 1966, dont une fraction des titulaires passe un doctorat quatre ans plus tard. La reprise est progressive à partir de 1972 et c'est seulement en 1973 que l'on dépasse le niveau de 1968.

Ces statistiques sont imparfaites, elles reflètent simplement ce que ramène le grand filet Sudoc interrogé avec « Sciences » en « Note de thèse » : on y trouve non seulement les thèses en sciences, mais aussi en sciences politiques ou économiques, par exemple¹¹. Mais il n'y a aucun inconvénient à ce que le constat dépasse la catégorie des thèses libellées « Sciences ».

L'essentiel est de disposer d'une série cohérente, ce qui est le cas de la nôtre puisque la même question a été posée à la même base sur les cinq années. Elle révèle une crise des études doctorales après 1969 dans l'enseignement supérieur français. Les thèses d'informatique n'en sont qu'un cas particulier.

Cette chute résulte-t-elle des restrictions budgétaires de l'après-1968 ? Ou des profondes réorganisations de l'Université consécutives à la loi Edgar Faure, qui a pu obliger les professeurs à donner une priorité momentanée aux démarches administratives, aux négociations et aux luttes politiques pour mettre en place les nouvelles universités et les départements (« UER »), plutôt qu'aux directions de thèses ? D'autres études pourront y répondre.

¹¹ Ces chiffres fournis par Sudoc sont approximatifs, l'interrogation avec la date et le seul mot *Sciences* ou *Mathématiques* faisant aussi apparaître des thèses d'histoire et de philosophie des sciences, comme celle de Michel Serres sur *Le Système de Leibniz et ses modèles mathématiques*.

Deux générations plus tard : les années 2000

Au début des années 2000, ce sont les élèves non seulement des anciens doctorants des années 1960, mais de leurs disciples des décennies 1980 et 1990, qui se forment à leur tour. Entre temps, la recherche en informatique s'est diffusée dans la plupart des universités françaises. Cette diffusion n'aboutit pas pour autant à une égalisation : on constate toujours une hiérarchie des pôles universitaires, qui s'est modifiée mais non inversée en quarante ans.

Ici, nous avons limité l'enquête aux nombres de thèses affichés par Sudoc, les critères d'interrogation étant simplement la date (2001-2005, cinq années nous paraissant permettre un lissage suffisant des variations annuelles contingentes) et, en « Note de thèse », le mot *informatique* et le nom de la ville. On s'est limité à 17 villes universitaires jugées représentatives.

On doit garder à l'esprit les imperfections de Sudoc et les approximations induites par ce questionnement volontairement simple. Notamment, l'interrogation fait chaque fois apparaître quelques thèses qui ne relèvent pas de la recherche en informatique proprement dite, tandis que Sudoc omet des titres¹². On espère qu'un défaut compense l'autre... L'essentiel est de présenter une série cohérente sur 5 ans.

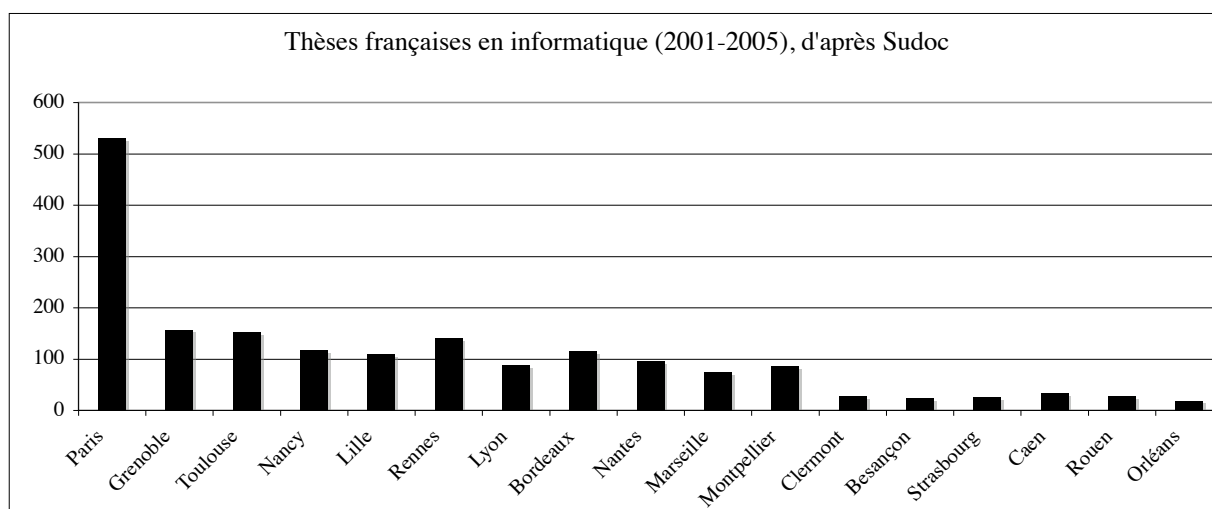


Figure 7. Thèses françaises en informatique (2001-2005).

12 Ainsi, à Grenoble à la fin du 20e siècle (1998-2000), Sudoc ne fait apparaître qu'un nombre ridicule de thèses en informatique, de l'ordre de 2 ou 3 / an, contre une vingtaine en temps normal. Tandis que les mêmes termes (informatique Grenoble) demandés pour « Tous les mots » révèlent 52 thèses pour l'an 2000. http://cuiivre.sudoc.abes.fr/xslt/DB=2.1/SET=1/TTL=1/CMD?ACT=SRCHM&ACT0=SRCH&MATCFILTER=Y&MATCSET=Y&PARSE MNEMONICS=N&PARSE OPWORDS=N&PARSE OLDSETS=N&IMPLAND=Y&screen_mode=Recherche&IKT0=1004&TRM0=&ACT1=*&IKT1=8063&TRM1=&ACT2=*&IKT2=63&TRM2=&ACT3=*&IKT3=63&TRM3=informatique+grenoble+&SRT=RLV&ADI_TAA=fre&ADI_LND=fr&ADI_JVU=2000&ADI_MAT=Y&ILN_DEP_BIB=DEP&NOT_USED_ADI_BIB

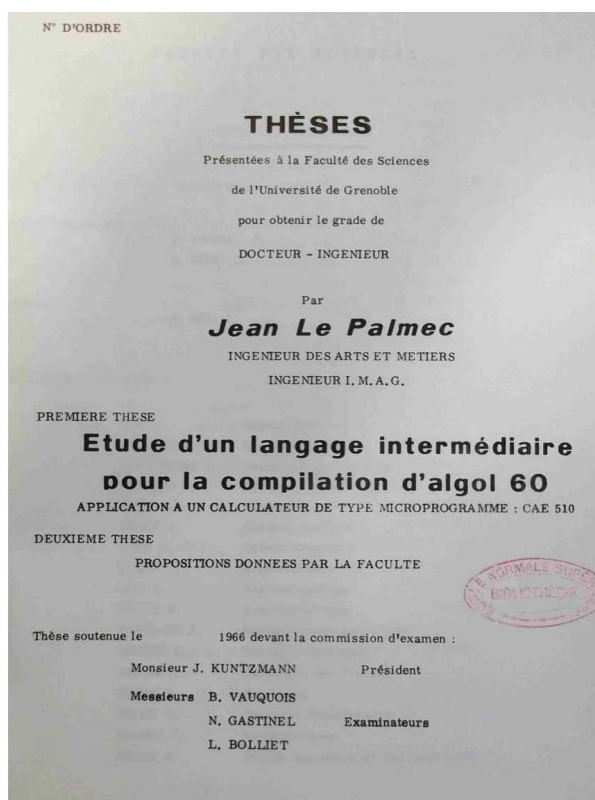
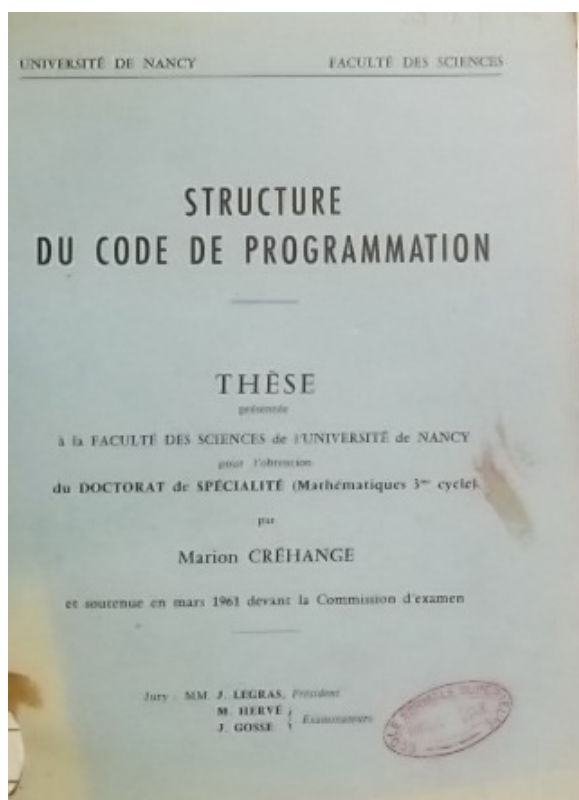
On observe que la hiérarchie des pôles universitaires pionniers s'est maintenue, Paris, Grenoble et Toulouse restant en tête (de l'ordre de 150 thèses pour ces deux dernières), Nancy et Lille se maintenant à un bon niveau (de l'ordre de 110 thèses chacune). Les principaux changements sont :

- L'explosion du nombre de thèses parisiennes, découlant notamment de la multiplication des pôles de recherche académique dans cette discipline depuis les années 1970 : outre la pléiade d'universités scientifiques constituées autour de la capitale, les écoles normales supérieures et les grandes écoles d'ingénieurs se sont investies dans la recherche informatique. Il y a donc une forte re-centralisation par rapport aux années 1950 et 1960, où les initiatives provinciales relativisaient le rôle de Paris.
- Le rattrapage de quelques universités depuis les années 1970, soit mécaniquement par « effet capitale » (Lyon et Marseille), soit par la grâce d'une politique gouvernementale décentralisatrice (Rennes), soit par une combinaison d'initiatives locales et de transferts d'équipes parisiennes (Bordeaux, Montpellier). En fait, toutes les universités qui « rattrapent » bénéficient de transfert d'équipes formées dans les pôles pionniers, principalement Grenoble et Paris.

On peut distinguer deux groupes, en mettant à part l'énorme agglomération d'universités et d'écoles parisiennes :

- Les universités qui produisent une vingtaine ou une trentaine de thèses d'informatique / an. Cette classe comprend les pionnières (Grenoble, Toulouse, Nancy, Lille ainsi que Nantes qui, malgré un accident au début des années 60, a toujours maintenu une activité en informatique et en automatique) ; et celles qui ont rattrapé le plus vigoureusement (successivement Rennes, Montpellier, Bordeaux).
- Les universités qui produisent moins d'une dizaine de thèses d'informatique / an. Ce sont notamment les universités qui avaient raté leur entrée dans la recherche informatique au cours des années 1960 (Strasbourg, Besançon) et ne s'en sont toujours pas remises ; et les universités qui ne faisaient pas d'informatique jusqu'au milieu des années 1970 et s'y sont lancées plus tard, soit qu'elles fussent restées longtemps cantonnées dans les mathématiques (Caen, Orléans), soit qu'elles fussent de création récente (Rouen).

À ma connaissance, cette approche quantitative par la statistique des thèses n'a jamais été appliquée à l'étude d'une discipline « émergente » à l'époque contemporaine¹³. En tout cas elle n'a pas été employée dans les nombreux travaux menés depuis trente ans en histoire de la *Computer Science*. Les résultats présentés ci-dessus me semblent justifier les minutieuses recherches qu'elle exige.



13 Le beau travail de Juliette Leloup, *L'Entre-deux-guerres mathématique à travers les thèses soutenues en France* (dir. Catherine Goldstein et Hélène Gispert, UPMC, 2009), vise un objectif différent : analyser l'état de la discipline mathématique à une époque où elle était solidement installée dans l'Université. Il est utilisé notamment dans Hélène Gispert & Juliette Leloup, « Des patrons des mathématiques en France dans l'entre-deux-guerres », *Revue d'histoire des sciences*, septembre 2009, vol. 62, n° 1, p. 39-118.

Mon texte était rédigé depuis longtemps quand a été organisé le colloque *La thèse et le doctorat. Socio-histoire d'un grade universitaire (XIXe-XXIe siècle)*, (8-9 septembre 2022 en Sorbonne), par les porteurs du projet ès lettres, Pierre Verschueren et Cécile Obligi, avec le soutien de la Bibliothèque interuniversitaire de la Sorbonne et le Centre Lucien Febvre (université de Franche-Comté). On y trouve une approche comparable à la mienne notamment chez Jean-François Condette (IRHIS, université de Lille), « La thèse de doctorat comme outil d'affirmation d'un champ disciplinaire : la difficile promotion de la géographie (1808-1939) ».