

Francia – Forschungen zur westeuropäischen Geschichte

Bd. 40

2013

DOI: 10.11588/fr.2013.0.40965

Copyright

Das Digitalisat wird Ihnen von perspectivia.net, der Online-Publikationsplattform der Max Weber Stiftung - Deutsche Geisteswissenschaftliche Institute im Ausland, zur Verfügung gestellt. Bitte beachten Sie, dass das Digitalisat urheberrechtlich geschützt ist. Erlaubt ist aber das Lesen, das Ausdrucken des Textes, das Herunterladen, das Speichern der Daten auf einem eigenen Datenträger soweit die vorgenannten Handlungen ausschließlich zu privaten und nicht-kommerziellen Zwecken erfolgen. Eine darüber hinausgehende unerlaubte Verwendung, Reproduktion oder Weitergabe einzelner Inhalte oder Bilder können sowohl zivil- als auch strafrechtlich verfolgt werden.

JEAN-BAPTISTE RENAULT – BENOÎT-MICHEL TOCK

ACTES ORIGINAUX ET ACTES ÉDITÉS

Deux bases de données de textes diplomatiques français

1. Le concept de bases de données

L'expression de «base de données» a connu depuis quelques années un grand développement. Elle est récente au regard du médiéviste puisque, dans sa version anglaise *database*, elle ne date que de 1964. Pourtant, le médiéviste, et en particulier le diplomate, pourrait faire remonter le concept, et la réalisation, de bases de données bien plus haut. Car si on admet qu'une base de données (électronique) est «un lot d'informations stockées dans un dispositif informatique», organisé de manière à pouvoir facilement manipuler le contenu et stocker efficacement de très grandes quantités d'informations¹, bien d'autres systèmes, non électroniques, pourraient être qualifiés de «bases de données».

Après tout, lorsque le pape Innocent III décide de faire copier les bulles qu'il émet, ou du moins une partie des bulles qu'il émet, dans les registres dits du Vatican, il s'agit bien d'un souci de conservation pérenne des données, destiné à en faciliter la manipulation et à assurer le stockage (mais un stockage actif) de grandes quantités d'informations, pour autant qu'on accepte la relativité de cette notion de «grandes quantités»: Innocent III crée en quelque sorte une base de données manuscrite. Ce qui est vrai des *registra vaticana* l'est aussi pour chaque cartulaire, d'ailleurs: il s'agit bien là aussi d'une base de données textuelles sur support manuscrit. On peut aussi songer aux inventaires d'archives rédigés au Moyen Âge: ils peuvent également être considérés comme des bases de données, toujours sur support manuscrit, mais cette fois les données sont analytiques, et non plus textuelles.

L'édition d'actes médiévaux, à partir du XVI^e et surtout du XVII^e siècle, nous fait passer à l'ère des bases de données imprimées. L'idée est contestable pour les plus anciens volumes qui consistent souvent en la réunion d'actes opérée de manière assez aléatoire. Mais très rapidement on dispose de recueils visant, sinon à l'exhaustivité, du moins à beaucoup de diversité et de richesse².

Une autre étape, pour la France en tout cas, est franchie avec la publication d'un répertoire d'actes³: il s'agit là de la première base de données analytique, sur support imprimé, destinée à remplir (aussi) un objectif historique. Elle était liée, on le sait, au vaste projet de constitution à Paris d'un cabinet des chartes (qui nous a légué les magnifiques collections de copies d'érudits de la Bibliothèque nationale de France)⁴: la réunion à Paris de copies de toutes les chartes conservées en France en ce qui était destiné à être une gigantesque base de données textuelles sur support manuscrit.

1 Voir le site http://fr.wikipedia.org/wiki/Base_de_donnees (tous les sites cités ont été consultés le 3 janvier 2011).

2 Voir par exemple les *Monumenta Boica*, 60 vol., Munich, depuis 1763.

3 Louis DE BRÉQUIGNY, *Table chronologique des diplômes, chartes, titres et actes imprimés concernant l'histoire de France*, 8 vol., Paris 1769–1876.

4 Xavier CHARMES, *Le Comité des travaux historiques et scientifiques (histoire et documents)*, t. I, Paris 1886; Blandine KRIEDEL, *L'histoire à l'âge classique*, Paris 1996 (Quadrige).

Enfin, en ce qui concerne plus spécifiquement les bulles pontificales, comment ne pas considérer les »Regesta« de Philipp Jaffé comme une autre base de données analytiques, ou les différents volumes des »Papsturkunden in Frankreich« comme une tentative de base de données textuelles ?

2. Le but des bases de données

Ce petit jeu sur des »bases de données«, manuscrites ou imprimées, antérieures à l'avènement de l'informatique est certes un peu facile. Il n'est cependant pas gratuit ou purement intellectuel. Il permet de souligner que, en cherchant à constituer nos bases de données, nous nous plaçons dans la continuité d'Innocent III, de cartularistes médiévaux ou d'érudits modernes: comme eux, nous essayons d'utiliser les outils qui sont à notre disposition pour arriver à gérer la masse des actes, et donc des informations. Et en fin de compte nous nous retrouvons face aux mêmes interrogations que celles qui pouvaient inquiéter les cartularistes et archivistes médiévaux: comment organiser la masse d'actes dont nous disposons, de la manière la plus utile pour nous ?

Ce qui pose précisément la question du but des bases de données. À quoi ces bases doivent-elles servir ? Le premier but des bases de données est de permettre de trouver facilement les actes qu'auparavant on devait chercher avec peine dans des centaines de volumes imprimés, de cartulaires, de cartons d'archives. Le travail n'était pas impossible. Mais il était long, lourd, incertain. Il suffit de regarder, et ce n'est qu'un exemple, l'édition des actes des comtes de Flandre Baudouin VIII et Baudouin IX (1191–1206), publiée par Walter Prevenier: ce dernier a tenu à indiquer tous les fonds d'archives et tous les cartulaires dans lesquels il a cherché, souvent vainement, les actes de ces comtes⁵. De même l'éditeur d'un corpus d'actes médiévaux doit-il encore parcourir des dizaines de volumes de copies d'érudits du XVIII^e siècle à la Bibliothèque nationale de France (collections Moreau, des Provinces, Duchesne ...). Idéalement, les bases de données devraient donc permettre de trouver facilement tous les actes donnés par un même auteur précis, ou reçus par un même bénéficiaire, ou relatifs à un même sujet.

Le deuxième but est l'étude des mots. L'histoire se fait dans une large mesure à partir des mots utilisés par les rédacteurs anciens; là aussi, comme pour les actes, l'informatique permet une recherche sûre et facile. Encore qu'en la matière les risques d'erreurs et d'oublis soient nombreux, mais c'est là une autre question qu'il n'y a pas lieu de traiter ici. Cette question de la recherche des mots implique en tout cas que les bases de données ne soient pas qu'analytiques, mais également textuelles; et, de préférence, qu'elles donnent accès à un texte de qualité, alors qu'on sait que les éditions de textes médiévaux sont de qualité inégale. L'étude des formules, si importantes en diplomatique et plus largement en histoire culturelle, nous place devant une autre exigence: non seulement les bases de données doivent être textuelles, mais le mode d'interrogation du texte doit être optimisé pour faciliter la recherche, précisément, des formules.

Un autre but des bases de données est de mettre en relation des actes que, spontanément, on ne penserait pas à relier entre eux: cela signifie qu'il ne faut pas seulement mettre dans une base les actes qui ont un lien clair entre eux, mais aussi des actes qui n'ont a priori aucun lien entre eux. Ceci doit nous inciter à la quantité (c'est la richesse en actes qui permettra de multiplier les croisements et les découvertes), mais aussi à éviter de sélectionner à l'avance; en d'autres termes à éviter de limiter les bases de données à un seul type d'actes. Enfin, on ne peut plus accepter des bases de données qui risquent de disparaître ou de n'être plus accessibles ou lisibles après quelques années: il faut donc que la structure qui les héberge soit durable, et que le langage informatique dans lequel elles sont écrites le soit également.

5 De oorkonden der graven van Vlaanderen (1191–aanvang 1206), éd. Walter PREVENIER, t. III, Bruxelles 1971 (Commission royale d'histoire. Recueil des actes des princes belges), p. 4–70.

Idéalement, une base de données devrait donc, non seulement offrir des garanties de pérennité – ce qui est un préalable indispensable: le coût de réalisation des bases de données est tel qu'on ne peut accepter de base qui ne réponde pas à ce critère –, mais aussi s'appliquer à un domaine géographique très vaste – pour permettre les croisements d'actes, et pour assurer une bonne recherche des actes – et reposer sur des éditions critiques à jour. Ce sont là des exigences considérables, si on prend en compte que pour la France jusqu'à la fin du XII^e siècle on peut très grossièrement estimer le nombre d'actes conservés à, peut-être, 150 000. Il paraît difficile d'envisager l'établissement d'une telle base de données dans un laps de temps prévisible.

3. La base de données des originaux antérieurs à 1121

Dès lors, comment pouvons-nous procéder? Deux stratégies sont possibles. La première, la plus ambitieuse, consiste en l'élaboration d'une base de données faite d'éditions nouvelles, dans un cadre entièrement cohérent. La seconde stratégie est de partir de ce qui existe et qui peut, rapidement et de manière peu coûteuse, être intégré dans une base de données.

Quelles sont, en France, les bases de données reposant sur des éditions nouvelles et sur un cadre cohérent, relevant, en fait, de la première stratégie définie plus haut? On pourrait citer la base de données »SCRIPTA«, qui ambitionne de réunir en une base de données tous les actes relatifs à la Normandie jusqu'à la fin du XII^e siècle, en donnant pour chacun d'eux une édition critique⁶. Bien qu'elle progresse régulièrement et soit particulièrement prometteuse, cette base n'est cependant pas encore librement consultable. C'est donc surtout la base des actes originaux antérieurs à 1121 conservés en France, qui est librement consultable sur internet, que nous allons présenter ici.

Le projet a été lancé en 1966 à Nancy par le doyen Jean Schneider, et longtemps dirigé par Michel Parisse, puis Pierre Pegeot au sein du CRAL d'abord, de l'ARTEM (Atelier de recherche et d'histoire des textes médiévaux) ensuite⁷. Sous la direction scientifique de Cédric Giraud et Benoît-Michel Tock, les dernières relectures ont été effectuées par Jean-Baptiste Renault en 2008, et la base a été publiée en ligne le 10 juin 2010⁸. Aujourd'hui, cette base donne accès à exactement 4932 actes, dont 183 bulles pontificales. La plus ancienne de ces bulles est celle, fautive, de Jean IV pour l'abbaye de Remiremont en 642⁹. Si on ne retient que les bulles authentiques, la plus ancienne est celle du pape Benoît III à l'abbaye de Corbie en 855¹⁰.

Il s'agissait de pouvoir étudier le vocabulaire des actes en prenant comme comparaison tous les actes conservés en original (ou apparemment en original) en France, et datés ou datables d'avant 1121¹¹. L'établissement de cette base de données n'a pas été facile: l'objectif était ambi-

6 Pierre BAUDUIN, La base de données SCRIPTA, Tinchebray et l'histoire anglo-normande, dans: *Le pays bas-normand 271–272* (2008), p. 157–166. Voir aussi http://www.unicaen.fr/recherche/mrsh/document_numerique/projet/scripta.w

7 L'histoire du projet et la bibliographie concernant la base des originaux se trouvent sur la page de l'Atelier diplomatique sur le site du Centre de médiévistique Jean Schneider, <http://medievistique.univ-lorraine.fr>.

8 <http://www.cn-telma.fr/originaux/index/>.

9 JE 2046. Original: BNF, nouv. acq. lat. 2547, n° 4. Dernière édition: *Chartes de l'abbaye de Remiremont des origines à 1231*, éd. Jean BRIDOT, Turnhout 1997 (ARTEM), p. 32–36, n° 2. Adresse électronique: <http://www.cn-telma.fr/originaux/charte2478/>.

10 JL 2663. Original: BM Amiens, ms. 526, n° 1. Édition: Clovis BRUNEL, *Bulle sur papyrus de Benoît III pour l'abbaye de Corbie (855)*, Amiens, Paris 1912 (Société des Antiquaires de Picardie. Fondation Henri Debray). Adresse électronique: <http://www.cn-telma.fr/originaux/charte4803/>.

11 On trouvera une présentation du corpus des originaux, de sa richesse et de ses disparités régionales dans Michèle COURTOIS, Marie-José GASSE-GRANDJEAN, Benoît-Michel TOCK (dir.), *La diplomatique française du Haut Moyen Âge. Inventaire des chartes originales antérieures à 1121*

tieux (environ 5000 actes, et donc autant d'éditions nouvelles), les moyens ont souvent manqué, le projet a dû suivre l'évolution constante de l'informatique. Il a en effet démarré à l'époque des cartes perforées et a été publié l'année dernière sous la forme d'une publication en ligne. Grâce à une collaboration avec l'IRHT (Institut de recherche et d'histoire des textes), la publication prend place sur le serveur du centre de ressources numériques TELMA, qui a également élaboré l'interface de consultation.

La base de données des chartes originales antérieures à 1121 conservées en France offre des possibilités étendues de recherches croisées. De par la multiplicité de ses champs elle permet une approche des documents avec une grande finesse, qu'il s'agisse de leurs caractères internes ou externes. Dès les origines du projet, en effet, une attention a été réservée à la matérialité de l'acte: les transcriptions ont été attentives aux signes graphiques, à l'indication des fins de lignes, à l'utilisation de capitales ou de lettres célestes. Il y a en fait trois types de recherches possibles:

- la «liste des chartes» permet une recherche par feuilletage concernant le numéro ARTEM ou la date;
- la «recherche floue», qui donne parfois des résultats surprenants, parce qu'elle interroge tous les champs ce qui augmente les chances de trouver des choses auxquelles on ne s'attendait pas;
- dans la «recherche avancée», la plupart des champs de la base de données sont interrogeables et il est possible de les croiser. Les principaux champs sont la date (avec une possibilité d'interroger une fourchette de datation), le lieu et l'institution de conservation, la fiabilité (qui permet d'écarter les actes douteux), la date de lieu, la région actuelle et le diocèse médiéval du bénéficiaire de l'acte, le support, le scellement, l'état de tradition, l'auteur et le bénéficiaire. Enfin trois champs comprennent la transcription du texte de l'original, l'analyse et la bibliographie. Bien entendu, il est possible de trier les résultats obtenus suite à ces requêtes. Le tri peut se faire principalement par date, auteur, bénéficiaire, fiabilité ou numéro d'acte dans la base de données.

Quelles recherches cette base rend-elle possibles? Elles sont évidemment très nombreuses et très diverses. De l'étude des survivances de formulaires francs à la diffusion d'un préambule à travers plusieurs centres d'écriture, les enquêtes sur le formulaire seront facilitées par la base. Parmi les thèmes peu étudiés jusqu'à maintenant, la base pourra servir à détecter la présence des citations et des figures bibliques dans les chartes et à voir dans quelle mesure ces appels à l'histoire sainte relèvent du formulaire ou d'une intervention du rédacteur. Prochainement, lorsque les photographies seront publiées en ligne, l'enquête sur les scribes pourra jumeler l'étude des formules à l'étude paléographique.

Un des buts originels de la base était d'offrir un vocabulaire daté, une sorte de *thesaurus* du latin des chartes. Les recherches en ce sens ont été peu nombreuses¹². Pourtant la limitation aux originaux présente l'avantage d'écarter les actes potentiellement remaniés (si l'on exclut dans une requête les actes douteux) et offre un vocabulaire fiable qui présente en outre la richesse née de l'irruption de la langue vernaculaire: par des actes en vernaculaire, des actes en latin truffé ou des latinisations de mots. Enfin la base présente un corpus à exploiter pour l'onomastique,

conservées en France, t. I, Turnhout 2001 (ARTEM), p. XI–XV. Les tables de ce recueil ont été un outil précieux pendant une décennie mais sont aujourd'hui dépassées et remplacées de fait par la base de données publiée en ligne.

12 Voir cependant, par exemple, Marie-José GASSE-GRANDJEAN, Le mot *bannus* dans les actes diplomatiques originaux du Haut Moyen Âge (conservés en France), dans: Le pays de Remiremont des origines à nos jours, Remiremont 2001, p. 44–60; Benoît-Michel TOCK, Les néologismes dans les textes diplomatiques français (VII^e–début XII^e siècle), dans: Bulletin du Cange 63 (2005), p. 99–107.

compte-tenu de la richesse des toponymes ou des anthroponymes énumérés dans les chartes. L'étude de la culture matérielle et des structures économiques gagnerait à venir glaner dans la base, que ce soit pour les *realia* ou les mentions de moulins, tonlieux, péages, monnaies, etc.

La prochaine étape de la publication électronique de la base de données des chartes originales sera la mise en ligne de toutes les photos des actes. Les démarches sont en bonne voie d'achèvement pour obtenir les permissions de publier les photographies des actes conservées dans environ cent vingt institutions de conservation (Archives nationales, Bibliothèque nationale de France, archives départementales, bibliothèques municipales, musées ...). Les photos seront consultables en 2013 ou 2014.

Le gros problème posé par la base de données des actes originaux est son coût de réalisation: on peut estimer, toujours de manière très grossière, qu'elle a «coûté» 150 000 heures de travail. Chiffre qui ne doit pas étonner: la réalisation d'un *Thesaurus* de Thomas d'Aquin a requis, elle, plus d'un million d'heures de travail. Avec l'expérience, avec aussi une informatique plus stable qu'elle ne l'a été au XX^e siècle, il serait possible de réaliser cette base de données de manière bien moins coûteuse. Mais cela resterait, de toute manière, un travail important.

4. La base de données »Chartae Galliae«

C'est précisément pour cela qu'un autre projet a vu le jour: celui d'une base de données réalisée à partir d'éditions existantes. Le projet, soutenu d'abord par le réseau des Maisons des sciences de l'homme, puis par l'Agence nationale de la recherche (projet Corpus 2007), vise à la publication d'une base de données qui portera le nom de »Chartae Galliae«. Elle aussi sera hébergée par le centre de ressources numériques TELMA. Son but est de transposer sur support électronique les éditions d'actes déjà existantes. Le travail s'est fait sans passer par l'établissement de nouvelles éditions, afin de permettre la mise en place rapide de cette base de données. Certes, les éditions existantes peuvent être imparfaites – elles peuvent être partielles ou contenir des erreurs –, mais cela n'empêche pas les historiens de les utiliser tous les jours. L'espoir est d'ailleurs de susciter de nouvelles éditions, ou des corrections aux éditions existantes. Par facilité, on a pris en compte les recueils d'actes, et pas les petites publications (pièces justificatives en fin d'article, par exemple); par facilité toujours, on a privilégié les recueils anciens, le plus souvent possible libres de droits. Le projet réunit plusieurs partenaires, dans diverses universités ou centres de recherche en France; chacun a veillé à la sélection et à la saisie des actes dans la région qui l'intéressait. Il faut souligner ici le remarquable travail accompli à Dijon par le projet des »Chartae Burgundiae Medii Aevi«¹³, et à Caen par le projet »SCRIPTA«, qui préexistaient au projet »Chartae Galliae«. Réalisée, on l'a dit, sans prétention d'exhaustivité, la base »Chartae Galliae« offre aujourd'hui l'accès à quelque 37 691 actes, dont 1897 bulles pontificales, relatifs à la France actuelle jusqu'à la fin du XIII^e siècle¹⁴.

Comme pour la base de données des chartes originales, il est possible sur »Chartae Galliae« d'interroger les actes par leur date, leur auteur, leur bénéficiaire, leur genre diplomatique, le diocèse du bénéficiaire, la langue dans laquelle ils ont été écrits, leur authenticité (d'après l'édition utilisée), leur texte, l'analyse (bien que celle-ci soit parfois absente) ou la bibliographie. On peut aussi préciser que si la plupart des actes repris dans »Chartae Galliae« l'ont été d'après des éditions existantes, quelques-uns sont des actes inédits, transcrits d'après des originaux ou des copies. Le principe ici est nouveau, puisqu'il ne s'agit pas d'éditions critiques traditionnelles, mais de la simple transcription à partir du document considéré comme offrant le meilleur texte: l'original s'il est conservé, sinon une copie considérée comme la meilleure. L'idée est de ne pas

13 <http://www.artehis-cnrs.fr/page-documentaire-CBMA>.

14 La base de données est accessible à l'adresse <http://www.cn-telma.fr/chartae-galliae/index/>.

attendre que tous les actes inédits (sans compter tous ceux qui ont été mal édités) aient bénéficié d'une édition critique, mais d'en offrir déjà une première version¹⁵.

5. Perspectives aujourd'hui: les bulles pontificales

En quoi ces bases de données de chartes originales peuvent-elles être utiles à la «Gallia Pontificia»? D'une part parce que d'ores et déjà l'une d'elles offre l'accès à de nombreuses bulles pontificales originales; l'autre suivra bientôt le même chemin à partir d'éditions existantes, avec les qualités et les défauts de ces éditions. Au total, aujourd'hui plus de 40 000 actes seront à très court terme accessibles grâce à des bases de données gratuites, disponibles sur internet. Concrètement, cela comprend tous les actes originaux jusqu'au début du XII^e siècle, mais aussi, par exemple, tous les actes de l'abbaye de Cluny, ou tous les actes des différentes éditions de cartulaires du diocèse de Troyes par l'abbé Lalore, les cartulaires de l'abbaye Saint-Victor de Marseille, du chapitre cathédral d'Amiens, de l'abbaye d'Ourscamp ... D'autre part, parce qu'elle va faciliter la mise en relation entre ces bulles pontificales et les autres actes conservés dans les fonds d'archives.

La poursuite du projet «Chartae Galliae» dépendra très largement de l'obtention de financements nouveaux. Ceux-ci pourraient venir de l'Agence nationale de la recherche, mais aussi d'aides régionales – les collègues de Caen ou de Dijon ont obtenu des aides substantielles en la matière; rien n'oblige d'ailleurs à prendre comme cadre de travail la France entière. Il faudra de toute manière continuer à procéder par des projets partiels. Réaliser une base de données exhaustive de tous les actes jusqu'à la fin du XII^e siècle reste un objectif à long terme, qu'il serait impossible d'atteindre aujourd'hui.

Cependant, en ce qui concerne les bulles pontificales, le mieux serait d'envisager un projet européen. Il est vrai qu'un tel projet a déjà été déposé, malheureusement sans succès. Il serait cependant sans doute judicieux de le relancer: c'est là un projet très fort pour l'Europe, tant il n'est guère contestable que les bulles pontificales ont été, au Moyen Âge, un vecteur fort de l'unité non seulement religieuse, mais aussi juridique, intellectuelle et culturelle de l'Europe.

15 Sur cette proposition de première transcription publiée sur internet, voir Benoît-Michel TOCK, *Die Edition von Urkunden und Chartularen im Norden Frankreichs*, dans: Theo KÖLZER, Wilibald RÖSENER, Roman ZEHETMAYER (dir.), *Regionale Urkundenbücher, Sankt Pölten 2010* (NÖLA. Mitteilungen aus dem niederösterreichischen Landesarchiv, 14), p. 195–204.