Global18

N° 3 | 2023
Vers l'interopérabilité des bases de textes / Toward Interoperability of Texts Databases — Digitizing Enlightenment V

Le programme des Registres de la Comédie-Française : enjeux d'interopérabilité

Charline GRANGER

Virginie YVERNAULT

Édition électronique :
URL :
https://global18.numerev.com/articles/revue-3/953-le-programme-des-registres-de-la-comedie-francaise-enjeux-d-interoperabilite
DOI : numerev_2209
Date de publication : 12/05/2023

This submission is licensed under **CC BY-NC-ND** (Attribution - No commercial - No derivatives).

Pour **citer cette publication** : GRANGER, Ch., YVERNAULT, V. (2023) Le programme des Registres de la Comédie-Française : enjeux d'interopérabilité. *Global18*, (3). https://doi.org/10.34745/numerev_2209

Le programme «â¯Registres de la Comédie-Françaiseâ¯» (RCF) valorise les archives comptables et administratives de la Comédie-Française (1680-1899), essentielles pour étudier son fonctionnement, son répertoire et sa fréquentation. Conçu en collaboration internationale, il repose sur plusieurs bases de données (recettes, dépenses, feux) et propose des outils de visualisation sophistiqués. Il vise l'interopérabilité entre bases hétérogènes et périodes variées, malgré des défis techniques et historiques (évolution des salles, catégories de places, formats de registres). Ce projet illustre les apports des humanités numériques à la recherche théâtrale et institutionnelle.

Abstract :

The "Registres de la Comédie-Française" (RCF) program makes the most of the Comédie-Française's accounting and administrative archives (1680-1899), essential for studying its operations, repertoire and attendance. Designed in international collaboration, it is based on several databases (income, expenditure, lights) and offers sophisticated visualization tools. It aims for interoperability between heterogeneous databases and varied periods, despite technical and historical challenges (evolution of rooms, categories of seats, register formats). This project illustrates the contributions of digital humanities to theatrical and institutional research.

Mots-clés :
Humanités numériques, Théâtre, Interopérabilité, Registres, Comédie-Française

---

L'impressionnant fonds archivistique de la Bibliothèque-musée de la Comédie-Française, située sous les arcades de l'une des galeries qui bordent le jardin du Palais-Royal, constitue une source primordiale pour la connaissance de l'organisation et du fonctionnement internes de ce théâtre, mais aussi pour l'étude de ses interactions avec les autres scènes dramatiques françaises ou étrangères. Les possibilités d'accéder à de telles archives étaient cependant encore assez limitées il y a une quinzaine d'années. Forts de ce constat, plusieurs chercheurs, spécialistes de théâtre et historiens des institutions culturelles de l'Ancien Régime, se sont accordés à reconnaître l'utilité, et même la nécessité de créer un site internet global propre à la visualisation et à l'analyse de ce patrimoine archivistique d'ampleur. De ce manque d'accessibilité des données a ainsi germé l'idée d'un programme en humanités numériques consacré au

répertoire et à la programmation de la Comédie-Française, depuis sa fondation, par ordonnance royale, en 1680. La société des Comédiens-Français a conservé l'ensemble de ses archives comptables et administratives, qui ont été si minutieusement consignées qu'il est possible de connaître l'intégralité du répertoire, les recettes et les dépenses que chaque représentation a occasionnées, ainsi que l'affluence des spectateurs (dont on peut connaître précisément la répartition au sein de la salle de spectacle, les différentes catégories de places et leurs prix étant détaillés). Les comptes rendus des comités et des assemblées des comédiens renseignent également sur les activités quotidiennes de la troupe, ses rapports avec le pouvoir (notamment avec les Premiers Gentilshommes de la Chambre et l'intendant des Menus Plaisirs, puis, au XIX^{e} siècle, avec le surintendant des spectacles et, à partir de 1850, avec l'administrateur général), ainsi que sur les diverses tâches dont le personnel du théâtre était chargé – autant d'informations qui permettent souvent de nuancer certaines idées reçues ou certains mythes tenaces que continue de véhiculer l'historiographie théâtrale.

Le programme des Registres de la Comédie-Française, baptisé « RCF », a pour ambition de valoriser et de diffuser ce fonds patrimonial, qui comporte des corpus archivistiques considérables et pour partie sériels. L'inventaire de la Bibliothèque-musée, accessible en ligne sur la base Lagrange, montre en effet des documents de nature très variée. Certains ensembles forment des séries cohérentes, à l'exemple des registres comptables (registres journaliers de recettes et de dépenses), d'autres constituent des séries doubles ou incomplètes, tels les registres de feux, qui concernent la distribution des rôles par pièce, ou encore les registres d'assemblée et de comité. D'autres ensembles, enfin, sont isolés (livres d'accessoires, répertoires divers, dossiers consacrés à des auteurs et autrices ou à des comédiens et comédiennes…). Une telle hétérogénéité explique que le programme RCF se soit élaboré en premier lieu autour des ensembles sériels, en particulier ceux qui donnaient accès aux informations les plus élémentaires et les plus centrales, comme les recettes du théâtre qui ont permis d'élaborer une base calendaire : les registres de recettes ayant été remplis quotidiennement et régulièrement depuis 1680, ils renseignent, jour par jour, sur l'intégralité des œuvres du répertoire. Par la suite, le programme a pu s'étendre – et continue de s'étendre – aux autres corpus archivistiques.

Un programme commun, des bases de données hétérogènes et autonomes

Si la question de l'interopérabilité se pose d'une manière si aigue au programme des Registres de la Comédie-Française, c'est que celui-ci recouvre plusieurs volets, qui disposent chacun de leur propre base de données relationnelle. Depuis 2008, celles-ci ont été mises en place grâce au concours de plusieurs institutions partenaires : outre la Bibliothèque-musée de la Comédie-Française (BMCF), sont parties prenantes aux États-Unis le MIT et New York University (après Harvard), l'Université de Victoria au Canada, et en France, l'Université Paris Nanterre, l'Université de Rouen Normandie et Sorbonne Université, qui ont mis à disposition d'importants moyens financiers, techniques et humains en associant à leurs chantiers des laboratoires spécialisés en humanités numériques (tel l'Obvil/Obtic de Sorbonne Université) et de nombreux ingénieurs parmi

lesquels figurent des prestataires informatiques indépendants (notamment Logilab à Paris et Performance Software à Boston). Si la première phase du programme concernait les recettes obtenues entre 1680 et 1793 (RCF17-18[1]), une seconde phase a débuté en 2018, s'attachant à traiter les registres de dépenses de 1680 à 1776 qui, par leur contenu irrégulier et touffu, constituent un défi pour les chercheurs et les ingénieurs. D'une façon presque concomitante, un troisième volet du programme se mettait en place à partir de 2020, autour des archives comptables allant de 1799 à 1899 (RCF19[2]).

Afin d'exploiter au mieux les données de la base la plus « ancienne », celle des recettes des XVII^{e} et XVIII^{e} siècles, ont été perfectionnés sur un site global des outils de visualisation et d'investigation qui sont de plusieurs sortes : l'« outil découverte » s'adresse à un public de curieux et de non spécialistes et permet d'obtenir des informations générales et des statistiques autour des auteurs et autrices, des acteurs et actrices, des pièces, des genres dramatiques et des saisons théâtrales. La base est également interrogeable à partir d'instruments de recherche « par facettes » proposant un certain nombre de filtres pour affiner les requêtes, tandis que le « tableau croisé dynamique » constitue, pour sa part, un outil plus précis et plus complexe qui génère des tableaux et des graphiques comparatifs propres à fournir des données numériques et des calculs ciblés (il est, de ce fait, davantage destiné aux chercheurs aguerris) : il est possible, par exemple, d'obtenir des graphiques donnant des indications sur la répartition des pièces selon leurs genres pour une saison donnée ou bien de mesurer l'évolution du succès d'un dramaturge ou d'une pièce dans le temps à partir des recettes engrangées. La base des feux, établie à partir des registres donnant la distribution des rôles par séance théâtrale de 1765 à 1793, permet quant à elle d'obtenir divers résultats concernant la carrière dramatique des comédiens et comédiennes, ou l'évolution d'un rôle particulier (figure 1).

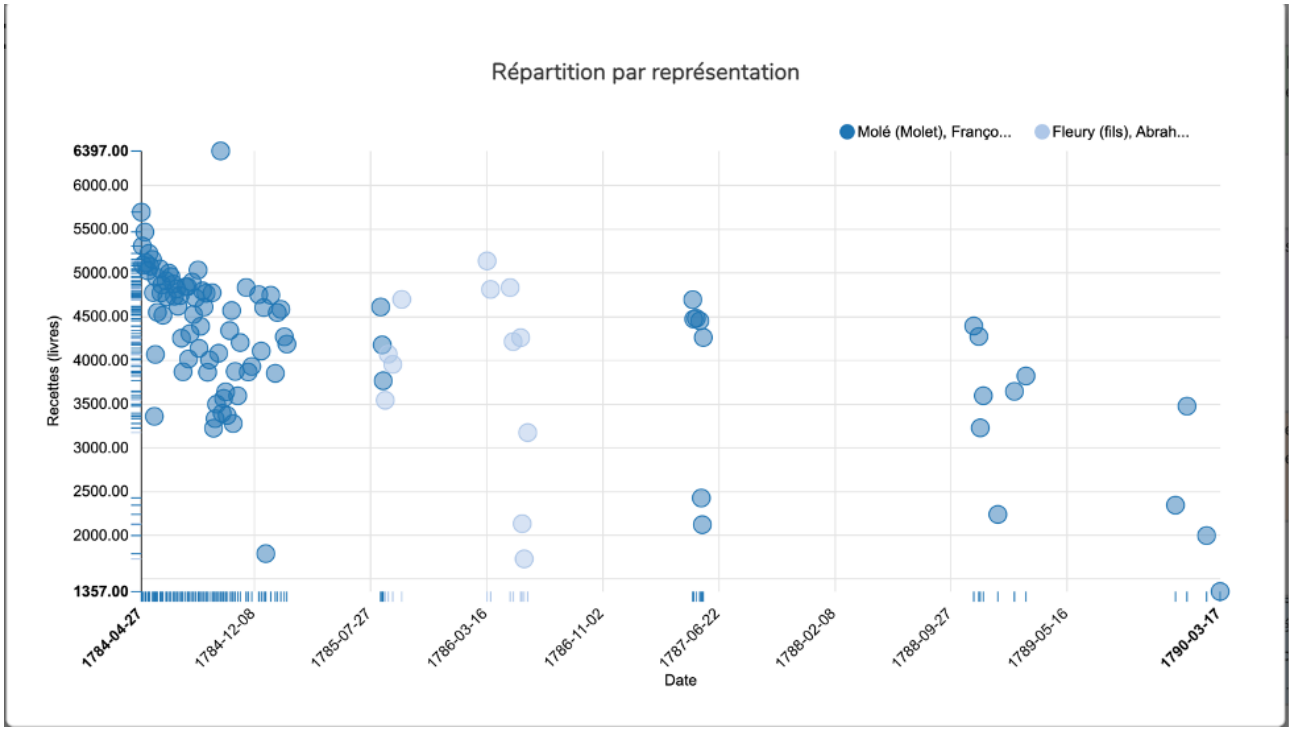

Fig. 1 – Le rôle du Comte Almaviva dans *Le Mariage de Figaro* (exemple de l'usage de l'outil « Graphe » de la base des feux).

Les défis à relever, ou comment rendre interopérables les différentes bases RCF ?

Le programme des Registres de la Comédie-Française est à présent confronté à deux défis de taille, en diachronie (comment rendre interopérables la base des recettes des XVII^{e}-XVIII^{e} siècles avec celle des recettes du XIX^{e} siècle ?), et en synchronie (comment rendre interopérables les bases des recettes, des dépenses et des feux pour la période du XVIII^{e} siècle où ces trois types d'information peuvent être croisés ?).

Interopérabilité en diachronie : recettes, catégories de place et affluence du public

Au sein même des registres de recettes des XVII^{e} et XVIII^{e} siècles, les données ne sont pas homogènes. De 1680 à 1793, la Comédie-Française occupe successivement quatre théâtres différents : l'Hôtel Guénégaud (1680-1689), la salle des Fossés-Saint-Germain (1689-1770), la salle des Machines, aux Tuileries (1770-1782), et le théâtre de l'Odéon (1782-1793). L'aménagement de la salle évolue notablement d'un bâtiment à l'autre (et parfois à l'intérieur d'un même théâtre), ce qui rend difficilement comparables les catégories de places entre elles. Par exemple, la place du « parterre » de 1681 n'a pas vraiment d'équivalent en 1793 (figure 2), de même que la catégorie des « premières loges », quoique présente dans les deux registres, ne renvoie pas à la même réalité spatiale à chacune de ces deux dates. À ce premier problème s'ajoute un second, qui concerne les places dans les loges : suivant les périodes, celles-ci sont tantôt comptabilisées une par une, tantôt comptabilisées à part entière, ce qui fausse naturellement la mesure du taux d'occupation de la salle. Relatif à la jauge, ce dernier sert de base de calcul pour tenter d'évaluer le succès ou l'échec d'une pièce. Un chercheur qui ignorerait le biais que constitue le comptage des spectateurs en loges et qui se livrerait à des comparaisons diachroniques sur l'ensemble de la période serait amené à formuler des hypothèses approximatives, voire inexactes.

Fig. 2 – Registre journalier du 14 septembre 1681 (les recettes apparaissent sur la partie haute de la page de gauche) et registre des recettes du 27 mars 1793 (sur la droite). Bibliothèque-musée de la Comédie-Française, coll. Comédie-Française.

Quant au volet consacré aux archives du XIX^{e} siècle (RCF19), il s'intéresse aux registres de recettes mais aussi aux registres de comité et d'assemblée des comédiens qui ont été transcrits par des stagiaires et des étudiants afin d'être édités sur le site global du programme. Les informations qui figurent dans les 83 registres de recettes allant de 1799 à 1899 sont, elles aussi, loin d'être homogènes ; de plus, elles gagnent en précision à mesure que l'on avance dans le siècle. Aussi les pages de registres ne se ressemblent-elles pas, comme le montrent les figures 3, 4 et 5. Un formulaire de saisie

unique a néanmoins été établi pour l'ensemble de la période à partir des informations communes aux différents registres, et qui concernent principalement la programmation journalière du théâtre et l'ordre des pièces jouées chaque soir ou chaque après-midi, en matinée (puisqu'au XIX^{e} siècle, il n'est pas rare que les séances théâtrales, d'ordinaire composées de deux pièces, en comportent trois ou même quatre[3]), les différentes catégories de places (qui évoluent de manière significative tout au long du siècle), leurs prix et, à partir de 1877, leur taux d'occupation. L'interface de consultation conçue pour exploiter les données de la base RCF19 permet d'effectuer des recherches d'une très grande précision, en croisant le prix des places, le taux d'occupation de la salle suivant les périodes, les auteurs ou autrices et les pièces. Les outils sont toutefois distincts de ceux qui ont été développés pour la base RCF17-18 (*cf. supra*). Le prochain défi auquel le programme RCF doit faire face concerne donc la mutualisation des bases de recettes et des instruments de visualisation et de recherche afin qu'il soit possible d'effectuer des requêtes couvrant la période allant de 1680 à 1899.

Fig. 3 – Page d'un registre de recettes de 1831. Bibliothèque-musée de la Comédie-Française, coll. Comédie-Française.

Fig. 4 – Page d'un registre de recettes de 1894 (recto). Bibliothèque-musée de la Comédie-Française, coll. Comédie-Française.

Fig. 5 – Page d'un registre de recettes de 1894 (verso). Bibliothèque-musée de la Comédie-Française, coll. Comédie-Française.

Interopérabilité en synchronie : recettes, dépenses et feux de 1765 à 1776

La récente création d'une base de données consacrées aux dépenses de la troupe, de 1680 à 1776, pose la question de la mise en relation de cette nouvelle base avec les autres bases existantes. Alors que la production de registres de recettes journalières est continue depuis la fondation de la compagnie, les registres de dépenses s'interrompent en 1776, tandis que les registres de feux, rédigés de 1765 à 1776, ne couvrent qu'une courte période. Par conséquent, la décennie 1765-1776 pourrait constituer un terrain d'exploration très propice pour mettre au jour de possibles corrélations entre ces différents types d'informations. Le croisement des données des registres de recettes, de dépenses et de feux est essentiel pour comprendre le fonctionnement administratif et comptable du théâtre, mais aussi pour faire apparaître les traces d'une proto-politique culturelle : quand et à quelles conditions les pièces qui coûtent le plus cher sont-elles aussi celles qui rapportent le plus ? Quelles pièces, quels genres sont concernés ? Dans quels cas la troupe investit-elle dans des mises en scène coûteuses ? Comment réagit-elle aux séances théâtrales qui génèrent des recettes peu élevées ? Comment envisage-t-elle l'équilibre budgétaire ? Comment ajuste-t-elle le répertoire et comment définit-elle

la distribution au gré des fluctuations de l'affluence du public ? Pour l'heure, les réponses à ces questions sont loin d'être évidentes, les bases de recettes et de dépenses n'étant pas encore solidaires. Mais afin que puissent être résolus les problèmes d'interopérabilités synchronique et diachronique évoqués, l'élaboration de la base de dépenses s'est accompagnée de la création d'un dépôt RDF.

Le dépôt RDF

Le programme RCF s'étoffant, il fallait que les données issues des différentes bases de recettes existantes (RCF17-18 et RCF19) soient converties dans un langage commun, le langage RDF. Un dépôt RDF des jeux de données a donc vu le jour au cours de l'année 2022, grâce à la concertation d'ingénieurs informaticiens et de chercheurs. L'alignement de l'ensemble des données a été rendu possible par la présence d'une information récurrente et commune à toutes les bases : la date de représentation, qui constitue donc l'élément pivot autour duquel a été conçue l'ontologie générale. Nécessaire à l'export des données en RDF, l'ontologie désigne le modèle de données : elle décrit avec précision les relations qu'entretiennent les données entre elles. Par exemple, à telle date sont associées une, deux, trois ou quatre pièces représentées ce jour-là, qui ont rapporté telles recettes, auxquelles correspondent tels acteurs et telles actrices associés à tel rôle.

De nombreuses difficultés continuent de se poser néanmoins. Sur le plan technique, les bases ainsi fusionnées ne peuvent être interrogées qu'en langage SPARQL, que fort peu de chercheurs en humanités maîtrisent : il nous faut donc soit penser une interface d'interrogation des données aisément consultable par tout utilisateur, soit imaginer une initiation des utilisateurs au langage SPARQL – deux solutions qui sont malaisées à mettre en place. Sur le plan scientifique, les problèmes ne sont pas moins grands, d'une part parce que certaines données font défaut (par exemple, les feux avant l'année 1765), d'autre part parce que la mutualisation des bases ouvre sans contredit de nouveaux questionnements sur la politique du répertoire, les logiques institutionnelles ou encore l'évolution des genres dramatiques (de nouvelles catégories apparaissant au cours du XIX^{e} siècle, tandis que d'autres disparaissent).

D'autres chantiers ont récemment été lancés dans le cadre du programme RCF, afin de concilier les approches qualitative et quantitative. L'Encyclopédie RCF se donne pour ambition de répertorier l'ensemble des termes utilisés dans les registres des XVII^{e}, XVIII^{e} et XIX^{e} siècles (recettes, dépenses, feux, assemblées et comités) : à ce jour, environ 180 définitions contextualisées et nourries d'exemples concrets tirés des registres la composent. Par ailleurs, la future base de données dédiée à la critique théâtrale proposera la transcription d'un choix de périodiques spécialisés, tels le *Mercure de France*, *L'Année littéraire* ou *L'Observateur des spectacles*. Dans les deux cas, il s'agit d'approcher au plus près la réalité historique des conditions matérielles de représentation des spectacles et de leur réception à la Comédie-Française en mobilisant des sources et des méthodes variées. En tout état de cause, le mouvement général qui

caractérise le programme des Registres de la Comédie-Française est celui d'une expansion, dont la poursuite n'est pas sans soulever plusieurs difficultés d'ordre technique et scientifique : quelles bornes donner à ce programme ? À partir de quels critères fixer de telles limites ? Comme parvenir à diffuser et à exploiter de manière croisée l'ensemble des données qui sont déjà disponibles ? Comment concilier la nécessaire interopérabilité des différentes branches RCF avec l'ouverture, tout aussi souhaitable, à d'autre bases calendaires existantes, dédiées à la vie des théâtres ? Les projets en humanités numériques et en études théâtrales connaissent en effet une très grande vitalité[4], en particulier ceux qui visent à nourrir une connaissance globale des spectacles et de la programmation théâtrale des siècles passés, à l'exemple de Dezède, RECITAL, Theaville ou encore Revolutionary Opera Comique. Le dialogue avec ces programmes devrait être approfondi lors de prochains colloques ou journées d'étude[5]. Les enjeux d'interopérabilité se posent donc à un double niveau interne et externe et constituent à l'évidence, avec les questions de pérennité des bases et des outils numériques, l'un des principaux défis que RCF devra relever dans les prochaines années.

---

[1] Le projet a bénéficié d'un premier financement ANR, piloté par plusieurs chercheurs des universités partenaires : Christian Biet et Tiphaine Karsenti (Paris Nanterre), Sylvaine Guyot (Harvard), Sara Harvey (Victoria), Jeff Ravel (MIT) et Agathe Sanjuan (conservatrice-archiviste de la Comédie-Française).

[2] RCF19, porté par Florence Naugrette, a bénéficié de deux années de financement Émergence, issu de l'Alliance Sorbonne Université, ainsi que du soutien matériel, financier et humain de l'Obvil/Obtic. Les premiers résultats sont pour le moment consultables à cette adresse : https://obvil.huma-num.fr/rcf/consultation/#/explore.

[3] On ne joue alors qu'une partie ou qu'un acte de la troisième et/ou de la quatrième pièce.

[4] À ce sujet, voir notamment Ioana Galleron, « Études théâtrales et humanités numériques : une introduction », *Revue d'Historiographie du Théâtre* numéro 4 [en ligne], consulté le 28 février 2023, URL : https://sht.asso.fr/etudes-theatrales-et-humanites-numeriques-une-introduction/.

[5] Le colloque, « Des archives aux données : bilan et perspectives du programme des Registres de la Comédie-Française » (Paris, juin 2023) a permis d' aborder l'interopérabilité sous tous ses aspects. Signalons aussi la journée d'étude intitulée « Explorer les manuscrits de la Comédie-Française autour des programmes RCF et ECuMe » (Rouen, avril 2023). ECuMe est un projet émergent porté par Laurence Macé qui tire profit de la numérisation des fonds manuscrits de la Comédie-Française et, en l'occurrence, de celle des manuscrits de censure. Cette journée d'étude rouennaise a permis d'interroger la spécificité des objets et des méthodes de RCF et d'ECuMe : ces deux programmes, qui ont été pensés et élaborés séparément, ont opté pour des formes et des structures différentes : bases de données relationnelles pour RCF, bibliothèque numérique pour ECuMe.