

***Chronopéra* : principes de constitution d'une base conçue pour être mise au service des historiens**

Solveig SERRE¹

La base de données *Chronopéra* (<http://chronopera.free.fr>), consacrée au répertoire de l'Opéra de Paris, a été élaborée par l'une des équipes de l'Institut de recherche sur le patrimoine musical en France, unité mixte de recherche qui associe le CNRS à la Bibliothèque nationale de France et au ministère de la Culture et de la Communication. Elle est née de la volonté de créer un outil de travail destiné aux historiens de la musique, de la danse et du spectacle, qui fournirait le répertoire de l'Opéra de Paris depuis le règne de Louis XIV, qui l'a institué, jusqu'à nos jours. Le postulat initial était que la chronologie des représentations avait son propre intérêt : grâce à elle, on pourrait sûrement savoir ce que voyait le public à une époque donnée. En outre, la connaissance de toutes les représentations de chaque saison donnerait le contexte global dans lequel survenaient les créations, donc permettrait de mesurer l'effet de nouveauté avec précision. Toujours selon ce même postulat, les données matérielles feraient également apparaître, par petites touches, les intentions des directeurs de l'Opéra et des différents responsables de production : de leur confrontation ressortirait une vision d'ensemble de la manière dont l'Opéra, au fil des siècles, avait conçu sa stratégie artistique. La volonté était donc de mettre un outil statistique élaboré au service de l'interprétation historique.

Cela se justifiait d'autant mieux que l'Opéra de Paris, depuis sa création, avait constitué son répertoire comme un fonds où les œuvres restaient durablement présentes. Ainsi se dessinerait sur le court terme ou sur une période plus longue, le degré de créativité de cette institution musicale majeure.

Aspects théoriques

Une bonne base de données est une base bien pensée dès sa conception

Rappelons qu'une base de données est constituée d'un ensemble de tables, une table étant peu ou prou un tableau dont les colonnes sont les champs et les lignes les données. Les diverses tables peuvent être liées entre elles (explicitement ou implicitement selon le logiciel de base de données utilisé) en ce sens qu'elles partagent des informations communes. Il n'existe pas une façon unique d'organiser ses données et c'est sans doute l'une des difficultés majeures de la conception de base de données que de faire ici le meilleur choix : les performances et possibilités de la base en dépendront de manière cruciale.

Notre base de données *Chronopéra* pouvait ainsi être réalisée de trois manières. La première consistait à élaborer une seule table contenant les champs Date/Lieu/Liste des œuvres jouées/Recette :

Date	Lieu	Liste des œuvres jouées	Recette
13/10/1750	Palais-Royal	<i>Almasis – Ismène – Linus</i>	1143

Une seconde manière était d'élaborer une seule table contenant les champs Date/Lieu/Œuvres/Position/Recette :

Date	Lieu	Œuvres	Position	Recette
13/10/1750	Palais-Royal	<i>Almasis</i>	1	1143
13/10/1750	Palais-Royal	<i>Ismène</i>	2	1143
13/10/1750	Palais-Royal	<i>Linus</i>	3	1143

Il y avait enfin une troisième possibilité, qui consistait à construire deux tables, la première contenant les champs Date/Lieu/Recette/CléReprésentation (c'est-à-dire un identifiant unique généré automatiquement par le logiciel de base de données) et la seconde contenant les champs CléReprésentation/Œuvre/Position :

Date	Lieu	Recette	CléReprésentation
13/10/1750	Palais-Royal	1143	xxx

CléReprésentation	Œuvre	Position
xxx	<i>Almasis</i>	1
xxx	<i>Ismène</i>	2
xxx	<i>Linus</i>	3

La principale critique que l'on peut faire à la première réalisation est qu'elle mélange des données qui devraient être séparées – celles relatives aux œuvres – ce qui va rendre compliquées et lentes des manipulations telles que renommer une œuvre, rechercher toutes les représentations d'une œuvre ou bien encore afficher les représentations où une œuvre est jouée en deuxième position. Si la seconde réalisation a le mérite de régler ces problèmes, elle a le défaut de posséder de la redondance d'information (date, lieu, recette). Non seulement cela provoque un ralentissement de la base lors des requêtes, mais cela implique également un problème de mise à jour et de saisie des données : pour la mise à jour (changement de la recette), il faut corriger globalement et s'assurer lors de la saisie qu'il n'y aura pas d'erreur pour une même représentation. La troisième réalisation évite quant à elle tous les écueils précédents, aussi l'avons-nous adoptée pour *Chronopéra*. Elle illustre surtout un principe crucial qui est de séparer le plus possible les données afin d'éviter au maximum la redondance d'information.

Dans le cas de *Chronopéra*, la table principale a la structure proposée dans la troisième table et le principe de dispersion est

présent dans d'autres points : la seconde table ne contient pas les noms explicites des œuvres mais un identifiant et une troisième table contient les identifiants d'œuvres avec en regard leur titre, l'année de création, le compositeur, etc. Ainsi, si l'on souhaite corriger le titre d'une œuvre, cette correction n'intervient que dans une seule entrée de la table des œuvres alors que si nous avions choisi la réalisation de la première table, cette correction aurait dû être effectuée pour toutes les entrées où l'œuvre était jouée.

Par conséquent, on voit comment une bonne conception est bénéfique à tous : l'utilisateur obtiendra de bonnes performances en terme de temps lors de ses requêtes et le concepteur aura une maintenance simplifiée de la base.

Une bonne base de données sait à qui elle s'adresse

184

En effet, les exigences des uns et des autres ne sont pas forcément identiques ni compatibles. Ainsi, un utilisateur attendra d'une base de données qu'elle propose une interface conviviale qui lui permet d'obtenir le plus simplement possible les réponses à ses questions (Ai-je obtenu la réponse à ma question ? Cela était-il simple ?). L'administrateur de la base quant à lui privilégiera une base qui offre une interface simple de saisie et de mise à jour et, s'il est utilisateur de la base, souhaitera pouvoir faire des requêtes complexes même si celles-ci ne sont pas aisées à réaliser d'un point de vue pratique (Puis-je renseigner ma base de façon simple ? Puis-je en tirer le maximum ?). Ces critères ne sont certes pas orthogonaux, mais ils sont parfois difficiles à concilier, et ce d'autant plus que l'on a au départ tendance à privilégier l'administrateur et que l'on ne pense à l'utilisateur qu'une fois la base de données remplie.

Dans le cas de *Chronopéra*, nous avons réfléchi préalablement aux besoins des utilisateurs futurs, étant bien conscients que ce qui n'était pas prévu ne serait pas faisable (nous ne voulions pas laisser l'accès complet à la base, et cela serait de toutes les façons inutile pour un utilisateur qui ne connaîtrait pas la manière dont la base a été conçue). C'est pour cela qu'une bonne base doit s'accompagner d'une bonne connaissance des utilisateurs de cette dernière et que

l'interface doit pouvoir être régulièrement enrichie des remarques et des besoins des utilisateurs.

Une bonne base de données doit tenir compte des problèmes d'échelle

Plus le contenu d'une base est important, plus certains problèmes de performances vont apparaître : on ne peut donc pas attendre d'une base de données gigantesque les mêmes choses que d'une base de taille raisonnable. Dans le cas où une base n'est faite que pour servir à une personne, à savoir celle qui la conçoit, l'interrogation de la base peut ne pas être conviviale et peut être aussi complexe qu'elle le veut du moment que l'unique utilisateur s'y retrouve.

La diffusion de la base de données : logiciel de base de données ou interface web ?

Une première possibilité pour la diffusion d'une base de données serait de permettre de la télécharger sous un format standard (par exemple un format *Microsoft Access*). Dans le cas de *Chronopéra*, nous avons rejeté cette solution pour plusieurs raisons : propriété intellectuelle et coût de l'opération, nécessité de bien documenter la base pour que quelqu'un d'extérieur au processus de conception puisse s'y retrouver, nécessité de prévoir une interface pour qu'un utilisateur inexpérimenté puisse interroger la base, difficulté de mise à jour des données et de l'interface.

185

La méthode qui nous a semblé la meilleure a été la mise en ligne sur le web. Cette solution nous permettait en effet de garder un contrôle sur les données et de faire évoluer la base et l'interface de façon simple, c'est-à-dire sans avoir à prévenir systématiquement les utilisateurs d'une nouvelle version de la base. Cela nous offrait également l'avantage de savoir qui utilisait la base et à quelles fins, puisque les utilisateurs enregistrés pouvaient être fichés. *Chronopéra* dispose ainsi d'une interface visuellement agréable, simple et intuitive (critère certes non scientifique mais qui n'est pas dénué d'importance dès lors que l'on s'adresse à un public non

académique), offre différents niveaux d'utilisation (simple / expert), permet un échange entre les utilisateurs et les concepteurs (par exemple pour demander la mise en place de requêtes spécifiques), permet un export des données pour une utilisation statistique via un logiciel extérieur.

Présentation

Chronopéra s'échelonne pour l'instant entre 1671 et 1989, elle couvre une période de 250 ans, soit près de 90 000 jours pour plus de 1 200 œuvres. L'organisation de la base permet de rechercher des informations simples – par œuvre ou par date – mais aussi de mener une analyse poussée du répertoire de l'Opéra et d'en interpréter l'évolution artistique. Les informations peuvent être obtenues même sans être préalablement identifiées, même si l'inscription permet d'avoir accès à davantage de ressources. Tous les champs de la base sont interrogeables : date, titre, compositeur, chorégraphe, librettiste. L'utilisateur a également accès à une fiche œuvre, qui résume les caractéristiques de celle-ci.

186

Chronopéra fonctionne avec le logiciel libre *MySQL* qui permet une mise en ligne aisée et gratuite, et est interrogeable au moyen du langage de requête *SQL*. Comme cette utilisation n'est accessible qu'à l'utilisateur qui possède des connaissances avancées en *SQL* et une bonne vision de l'articulation des diverses tables de la base de données, une interface conviviale, via un site Internet, a été progressivement mise en place. Cette dernière permet d'effectuer la plupart des requêtes classiques : affichage du calendrier pour une période définie, des œuvres d'un compositeur, d'un librettiste ou d'un chorégraphe. Les résultats de ces diverses requêtes peuvent être triés selon plusieurs critères (par exemple en fonction de la recette) et exportés vers différents formats (*pdf* ou *SQL* notamment). Le mode de recherche expert via l'emploi de requêtes *SQL* permet d'exploiter toute la puissance de la base de données *Chronopéra*, à l'aide de requête multicritères sur les diverses tables constituant la base de données. Il est par exemple possible de générer une table

des compositeurs, classés selon la recette moyenne que produisent leurs œuvres pendant une période donnée. On peut enfin déterminer quels sont les jours de la semaine les plus prolifiques, si certains mois sont désertés par le public et si ces phénomènes connaissent des évolutions au cours des siècles.

Par conséquent, le chronologie des représentations nous donne accès au spectre entier de ce que pouvait voir et entendre le public d'une époque donnée : pour qui est familier du répertoire de l'Opéra, voici une manière globale de pénétrer le goût d'une époque et de suivre son évolution au fil des années. Les chercheurs intéressés par l'expérience de spectateur d'une personne particulière trouveront des informations précises sur les spectacles auxquels elle a pu assister. Enfin, la connaissance de toutes les représentations de chaque saison fera apparaître le contexte intégral dans lequel surviennent les créations, dont on pourra ainsi mesurer en connaissance l'effet de nouveauté. Toutes les données laissent apparaître, à petite et à grande échelle, les intentions des directeurs de l'Opéra et des différents responsables de production. Avec l'espoir que, de leur confrontation, ressortira une vision complète et détaillée de la manière dont l'Opéra, au fil des siècles, a conçu sa stratégie artistique.

¹ Docteur en histoire moderne de l'université de Paris I, chercheur associé à l'Institut de recherche sur le patrimoine musical en France (UMR 200) et à l'École nationale des chartes.