



HAL
open science

Présentation de la base de données Péniche Opéra

Augustin Braud, Thomas Bottini, Rebecca Bristow

► **To cite this version:**

Augustin Braud, Thomas Bottini, Rebecca Bristow. Présentation de la base de données Péniche Opéra. Journée d'études : “ La Péniche Opéra. 1982-2015 ”, Apr 2022, Paris, France. hal-03950377

HAL Id: hal-03950377

<https://hal-amu.archives-ouvertes.fr/hal-03950377>

Submitted on 21 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Présentation de la base de données Péniche Opéra

Augustin Braud, chercheur post-doctoral

Thomas Bottini, CNRS, responsable informatique de l'IReMus

Avec la participation de **Rebecca Bristow**, ingénieure, IReMus/CERES (SU)

Institut de Recherche en Musicologie
IReMus UMR 8223 — CNRS/SU/BnF/MC

Introduction

- Création d'une base de données ; quels objectifs musicologiques ?
- État de l'inventaire :
 - Un premier inventaire descriptif des archives de la compagnie "La Péniche-Opéra" 1975-2015. Par Cécile Auzolle (dir.), Sylvain Labrousse, Clara Roupie (2018).
 - Un inventaire numérique entamé sur un disque dur de 5To ; plusieurs problèmes posés.
- Quels sont les enjeux informatiques ? Comment les outils techniques permettent-ils la réalisation du projet musicologique ?

1. Les premières étapes du projet

- Familiarisation avec un fonds dense et hétérogène :
 - documents publics (programmes), artistiques (partitions), administratifs (comptes rendus)
 - certaines archives stockées à la Péniche, d'autres à la BNF, en ligne, sur un disque dur...
- Nécessité de délimiter certaines typologies de données à relever pouvant faire sens dans le contexte d'une base de données :
 - Tout ce qui est lié au temps : dates de création, de représentation, etc.
 - Effectifs artistiques : chanteurs / instrumentistes, mise en scène, décors, personnels techniques, etc.

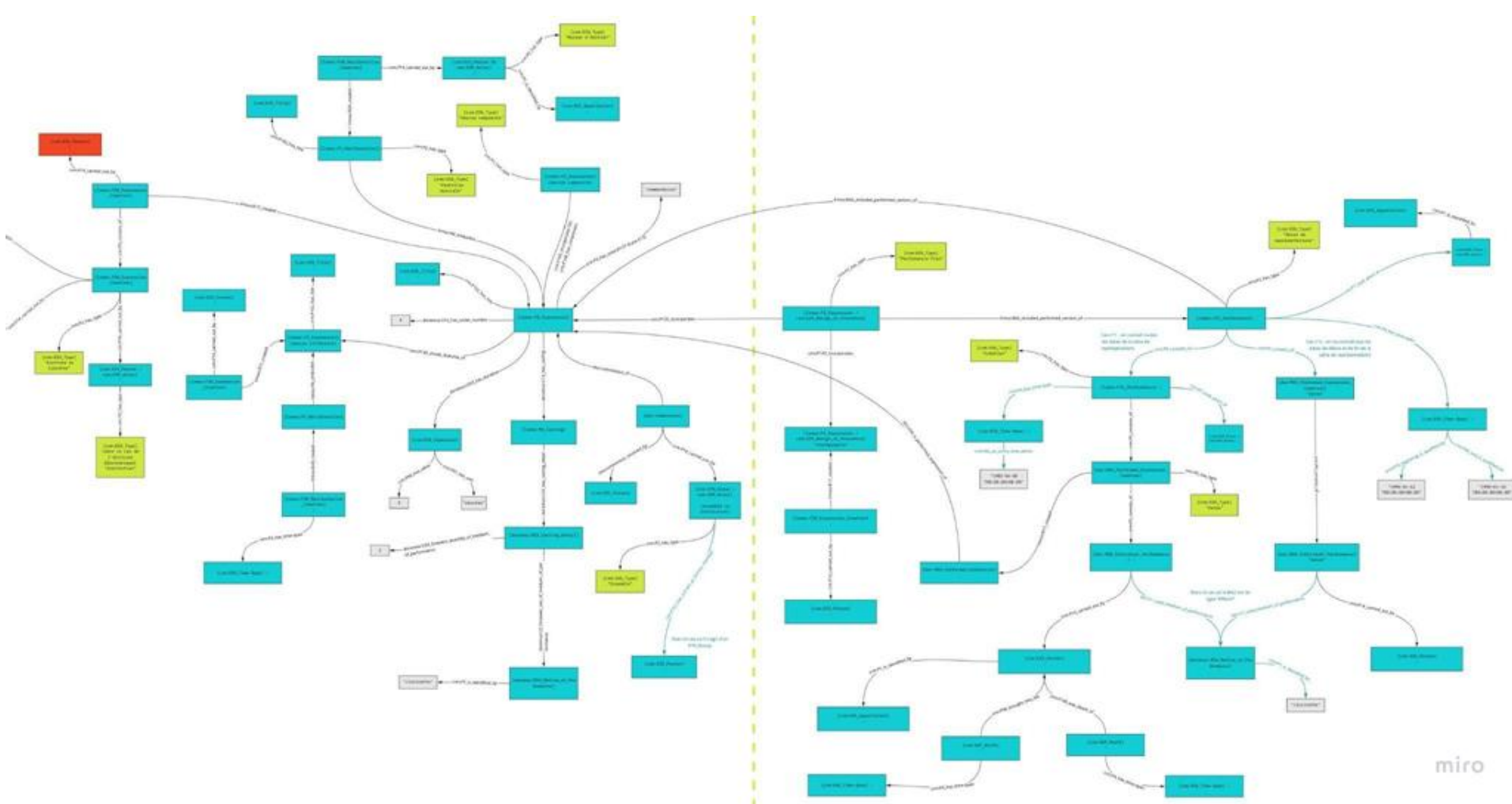
Méthodologie pour un projet d'« humanités numérique »

- **Objectifs :**

- Déployer une interface Web permettant une exploration fine du corpus (listes par types d'entités, vues synoptiques chronologiques, navigation hypertexte complète, statistiques...)
- Publier le jeu de données sous une licence libre (paradigme des sciences ouvertes)

- **Méthode & artefacts :**


- Modélisation sémantique du domaine (recenser les entités et les propriétés) :
 - Entretiens libres avec les chercheur-se-s
 - Ontologie sémantique pour le patrimoine CIDOC-CRM (sous l'égide de l'ICOM) + Extension DOREMUS (ANR, ontologie BnF/Philharmonie/Radio France)
- Déploiement d'une interface Web collaborative ergonomique pour saisir les données au fil du dépouillement (Directus <https://directus.io/>)
- Mise en place d'un processus de génération du graphe de données ouvertes & liées à partir des données de saisie
- Développement du site Web (travail amorcé en cours)
- 1 post doc + 2 ingénieur-e-s IReMus + hébergement sur la TGIR Huma-Num (CNRS)



Deux mots sur le modèle conceptuel

- Trois grandes catégories d'entités :
 - Ce qui est lié au contexte de création des œuvres et des supports liés (rôles, dates, effectifs, partition, institution...)
 - Ce qui est lié aux performances (lieux, ensembles, interprètes, premières, reprises, dates, tournées...)
 - Ce qui relève d'un regard analytique signé
- Vocabulaires contrôlés (instruments, interprètes, compositeur·ice·s) qui seront par la suite alignés sur des référentiels faisant autorité :
 - MIMO (thésaurus organologique, Musée de la musique)
 - data.bnf.fr
 - ISNI
 - ...
- Centralité de la dimension temporelle des données (CIDOC-CRM).

2. Création d'une interface de saisie

- Une interface permettant de recenser et décrire les éléments du corpus dans leur richesse et leur singularité.
 - Le remplissage se fait majoritairement à l'aide de la Bible en ligne qui convoque les informations fondamentales
 - Utilisation de l'inventaire pour identifier des besoins précis vis-à-vis de mes sources, puis une journée de travail à la Péniche pour trouver quelques éléments manquants.
 - [Démon Directus]
- L'approche exposée supra allant de l'explicitation des connaissances tirées du corpus au développement de l'interface d'exploration n'est pas linéaire. La nature des liens entre les entités nous apparaît à mesure que nous réalisons la saisie. La modélisation conceptuelle est un processus itératif, reposant sur une boucle de rétroaction féconde entre chercheur·se·s et ingénieur·e·s (HN). 

3. Exploitation critique des données

- [Démon site]
- Le corpus dépouillé depuis le début du travail (Novembre 2021) représente environ 25% des données totales :
 - Toutes les oeuvres jouées de 1980 à 1990 ont été inscrites dans la base de données, tout comme les informations s'y rapportant.
- Le site Web évolue selon les besoins scientifiques qui émergent.
- L'ensemble des données est accessible publiquement via le protocole SPARQL, ce qui favorise la réutilisation du corpus dans d'autres projets de recherche venant avec d'autres questions.

Conclusion

- Quelles sont les prochaines étapes à accomplir pour ce travail ?
 - Saisir l'intégralité des données ; retoucher certaines données déjà entrées pour une plus grande cohérence. Où s'arrêter ?
 - Finaliser le développement de l'interface Web d'exploration du corpus.
 - Valoriser le jeu de données et la méthode de modélisation dans d'autres cercles scientifiques.
- Les objectifs musicologiques à long terme :
 - Valider à l'aide d'éléments concrets des hypothèses esthétiques : témoigner de relations intra-personnelles permettant de préciser la place de la Péniche dans l'écosystème du spectacle vivant.
 - Donner accès à des sources musicologiques difficiles à trouver, analyser des oeuvres et le contexte (politique, social, historiographique) dans lequel elles s'inscrivent, etc.