

PORTADA



Un projet alliant sciences humaines numériques et histoire maritime

Section: [La cale](#)

Numéro de magazine: [#15](#)

AUTOMNE 2025

TEXT: Inma González Sánchez. Conservatrice du Musée maritime de Barcelone. Olga López Miguel. Conservatrice du Musée maritime de Barcelone.

FOTO: Fonds MMB.

Le Musée maritime de Barcelone participe à PortADa – Ports Arrivals Data , un projet visant à créer une base de données sur les arrivées de navires dans les ports de Barcelone, Marseille, Buenos Aires et La Havane entre 1850 et 1914.

En novembre 2023, le projet [PortADa](#) (Port Arrivals Data. Collecte automatique de données pour une histoire comparative à grande échelle du trafic maritime au 19e siècle : une approche basée sur les sciences humaines de l'héritage maritime) a été lancé, qui a reçu un financement de l'Union européenne par le biais du programme « Actions Marie Skłodowska-Curie » de l'Agence exécutive européenne pour la recherche,

qui se déroulera jusqu'en 2027.

Quatre équipes, composées de spécialistes en histoire maritime et en informatique appliquée aux sciences humaines, travaillent ensemble, pour la première fois, afin de faire progresser la connaissance du trafic maritime dans les ports de Barcelone, Marseille, Buenos Aires et La Havane, entre 1850 et 1914, période qui coïncide avec la transition de la voile à la généralisation du transport par navires à propulsion mécanique.

Un projet, deux objectifs

Lors du passage de la navigation à voile à la navigation à vapeur au XIXe siècle, les villes portuaires furent le théâtre d'une activité commerciale maritime florissante et de la création de réseaux s'étendant à travers le monde. Dans ces ports, l'arrivée des navires constituait un événement économique, culturel et politique, dynamisant non seulement la ville elle-même, mais aussi un réseau qui rayonnait jusqu'à l'arrière-pays grâce au trafic local et régional.

Ainsi, les nouvelles concernant l'arrivée de navires en provenance de ports lointains et tous les mouvements portuaires que cela impliquait bénéficiaient d'une attention privilégiée dans la presse locale de chaque port.

Le projet PortADa se concentre sur quatre des ports les plus importants de l'époque (Barcelone, Marseille, Buenos Aires et La Havane), avec un double objectif : le premier, constituer une base de données rassemblant les informations fournies par la presse locale concernant les arrivées de navires dans ces ports et publiées dans les journaux de ces villes. Toutes ces informations, triées et organisées, faciliteront de nouvelles études sur des questions économiques, sociales et culturelles.

Pour ce faire, les membres du projet exploitent plusieurs publications de l'époque (*Diari de Barcelona* ; *Le Sémaphore de Marseille* ; *La Gaceta Mercantil* , l'hebdomadaire *British Packet et Argentine News* , ainsi que les quotidiens *El Nacional* et *La Prensa* , tous quatre publiés à Buenos Aires, et enfin le *Diario de la Marina* , de La Havane) afin d'en extraire les informations relatives à l'arrivée des navires dans les ports susmentionnés. Cette tâche, auparavant effectuée manuellement, est désormais automatisée dans le cadre du projet, ce qui permet le traitement de volumes importants de données.

Le second objectif du projet est la création d'équipes de travail mixtes, composées de spécialistes en histoire maritime et en sciences humaines numériques, qui s'attaqueront ensemble aux défis et rechercheront des solutions aux problèmes rencontrés lors du traitement de ce type de sources d'information. Cette « alliance » permet de rompre avec les vieux clichés (« les historiens et les informaticiens ne se comprendront jamais ») et surtout, ouvre la voie à des collaborations durables et encourage d'autres institutions à suivre l'exemple de PortADa.





Techniques innovantes

PortADa offre un outil fondamental pour mieux comprendre l'évolution du trafic maritime dans chaque port et sa place dans la division internationale du travail. Ainsi, le commerce international entre les ports peut être reconstitué avec un niveau de détail inédit. Parallèlement, l'étude du cabotage, activité jusqu'ici peu explorée par la recherche, permet d'analyser plus facilement les connexions des réseaux commerciaux locaux et leur rôle dans la circulation des marchandises.

Mais les informations relatives à l'arrivée des navires ne se limitent pas aux seuls enseignements de l'histoire économique et commerciale. Les détails qu'elles contiennent nous permettent d'élargir notre

champ d'étude à d'autres disciplines : la typologie des navires nous permet de retracer les évolutions technologiques et la transition de la voile à la vapeur, tandis que l'analyse des noms des capitaines et des officiers de marine nous permet de reconstituer leurs parcours de vie et professionnels et, à plus grande échelle, d'étudier la mobilité sociale de ce groupe et l'impact des mutations technologiques sur leurs vies.

PortADa utilise des outils informatiques existants et développe des technologies innovantes pour faciliter la transcription automatique de la grande quantité de données obtenues à partir des sources documentaires traitées et pour intégrer ces informations dans la base de données en cours de création.

Cependant, la grande nouveauté de PortADa réside peut-être moins dans les aspects technologiques que dans un élément plus « prosaïque », sa méthode de travail. Le projet a en effet adopté une dynamique fondée sur la création de quatre équipes mixtes qui travaillent sur les ports étudiés et se réunissent une fois par an pendant deux mois lors des *écoles d'été*. Ce dispositif permet de partager les avancées, de résoudre les difficultés et de transmettre les connaissances sur les aspects historiques et informatiques liés à la recherche.

Grâce à ce mécanisme, chacun des objectifs du projet est atteint : la formation d'une équipe de spécialistes des sciences humaines numériques ayant vocation à la stabilité et à devenir une référence pour des initiatives similaires.

Outre l'adoption des pratiques habituelles de la science ouverte, la méthodologie du projet repose sur la conception d'un système modulaire d'éléments logiciels, dans le but de faciliter son application ultérieure à d'autres sources documentaires présentant des caractéristiques et des problèmes similaires.

Le stockage des données tirera parti des ressources disponibles à l'Université de Barcelone, qui dispose d'un dépôt institutionnel accessible à tous, avec plusieurs collections, dont la base de données, ainsi que la procédure méthodologique pour la générer et les premières études qui analysent ces données.

La base de données ainsi créée sera disponible sous une licence libre afin que d'autres chercheurs, ainsi qu'un public non universitaire intéressé, puissent en explorer le contenu. Il est prévu d'organiser son exploitation libre en tenant compte des différents profils d'utilisateurs potentiels (des experts possédant de solides connaissances techniques aux utilisateurs non experts) et en garantissant sa compatibilité avec les normes [d'Europeana](#).

Sur le plan technologique, le projet assure la transcription automatique des ressources numériques existantes et nouvellement numérisées, grâce à des procédures et des logiciels reproductibles pour d'autres ports et extensibles chronologiquement. L'objectif est de contribuer à la reconstitution du trafic maritime mondial sur la plus longue période possible. Dans certains cas, la numérisation servira non seulement à l'exploitation prévue par ce projet, mais garantira également la préservation et l'accessibilité des documents numérisés.

PortADa propose l'utilisation d'outils informatiques existants et le développement de nouveaux outils facilitant le traitement du volume important de données collectées. La transcription est réalisée par extraction automatisée d'informations, à la fois par application de processus basés sur des expressions régulières (lorsque les données apparaissent toujours dans le même ordre dans les articles) et par utilisation d'outils de traitement automatique du langage naturel et de techniques d'apprentissage automatique pour les textes moins structurés.



Vue du port de Barcelone. Phototype (photolithographie). Fin du XIXe siècle. Auteur : inconnu. Parera y Cía. Éditeurs. Fonds MMB.



Vue panoramique du port de Barcelone. Impression photomécanique 1912. Auteur : Àngel Toldrà Viazo. Collecte MMB.

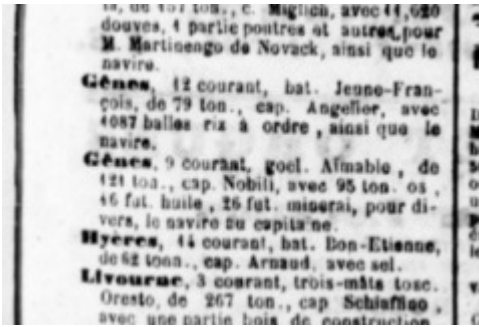
Une base de données ouverte

La [base de données PortADa](#) rassemble 1 600 000 articles de presse sur les arrivées de navires dans les ports de Barcelone (330 000 fiches), Buenos Aires (480 000), Marseille (540 000) et La Havane (242 000), publiés dans la presse de ces villes entre 1850 et 1914.

Chaque fiche fournit des données sur le port de départ, les ports d'escale, la durée du voyage, le type, le nom et le tonnage du navire, le nom du capitaine et la nationalité du navire. Elle comprend également des informations sur la cargaison transportée : type de marchandises, poids et destinataires.

À la fin du projet, la base de données sera disponible sous une licence ouverte afin que la communauté scientifique et le public intéressé puissent explorer son contenu.

Les images montrent des pages du journal *Le Sémaphore de Marseille* où sont publiées les entrées des navires au port, ainsi qu'un détail de ce journal où figurent les articles dont les informations sont extraites vers la base de données. Cela permet de constater la complexité de l'extraction des données de ces sources.



Détail du journal où apparaissent les articles dont les informations sont saisies dans la base de données. Photo : Fonds MMB.



Image extraite du journal *Le Sémaphore de Marseille*, où sont publiés les enregistrements des navires entrant dans le port. Photo : Fonds MMB.

Spécialistes des sciences humaines numériques

Le projet PortADa s'articule autour de quatre *écoles d'été* de deux mois, organisées dans chacun des ports étudiés. Lors de ces rencontres, la plupart des équipes se réunissent pour partager leurs avancées, résoudre ensemble les difficultés rencontrées et échanger des connaissances sur les aspects historiques et technologiques liés à la recherche. L'objectif est de constituer des équipes mixtes composées de spécialistes en humanités numériques. Le calendrier des *écoles d'été* prévoit des sessions à Mar del Plata (2024), La Havane (2024), Barcelone (2025), Réthymnon (2026) et Barcelone (2027).

Sur l'image, nous voyons une séance de travail de l'équipe lors de la première *école d'été*, au siège du Conseil national de la recherche scientifique et technique à Mar del Plata (Argentine).



Séance de travail en équipe lors de la première école d'été, au siège du Conseil national de la recherche scientifique et technique à Mar del Plata (Argentine). Photo : MMB - PortADa