

## **Doctorat en Science des Données sur le Patrimoine**

### **Le contexte**

Le doctorant sera salarié de la société Epitopos, laboratoire d'expertise et de caractérisation de matériaux du patrimoine, proposant une panoplie de prestations dans le domaine du patrimoine y compris des études climatiques, pour des projets de restauration en collaboration avec l'entreprise MES-CLA.

Le doctorant fera également partie de l'équipe du projet EquipEX+ ESPADON et interagira étroitement avec le groupe de travail " Climat, interface et patrimoine " WP6 du projet. Il/elle sera inscrit(e) à CY Cergy Paris Université et sera dirigé(e) par Nistor Grozavu (Professeur des Universités à CYU) et Ann Bourgès (C2RMF) Université Gustave Eiffel, et encadré(e) par Fabrice Surma (EPITOPOS), Martin Labouré (Mescla patrimoine) et Jean-Marc Vallet (CICRP) avec des déplacements sur les sites sélectionnés pour la recherche doctorale.

### **Information Générale :**

**Référence : thèse CIFRE - EPITOPOS**

**Lieu de travail : STRASBOURG et CERGY**

**Date de publication : 21 juillet 2022**

**Responsables scientifiques : Nistor Grozavu, Ann Bourgès**

**Responsable et co-responsable Entreprise: Fabrice Surma-Martin Labouré et Jean-Marc Vallet**

**Type du Contrat : Thèse CIFRE**

**Durée du Contrat : 36 mois**

**Début de la thèse: 1 Octobre 2022**

**Temps : Temps complet**

**Rémunération : 1957 € bruts mensuel**

### **Ecole doctorale :**

**CY CERGY PARIS UNIVERSITE**

Établissement Public à caractère scientifique Culturel et Professionnel, dont le siège social est situé 33 Bd du port, 95011 Cergy-Pontoise cedex, N° SIRET 130 025 976 00015, représenté par son Président, Monsieur François GERMINET,

Ecole Doctorale EM2PSI

L'Ecole doctorale n°407 Economie, Management, Mathématiques, Physique et Sciences Informatiques

Tél: 01 34 25 63 12

[lalia.sy@cyu.fr](mailto:lalia.sy@cyu.fr)

<https://collegedoct.cyu.fr/>

### **Affectation :**

**Nistor Grozavu, Full Profesor**

**ETIS UMR 8051, CY Cergy Paris University, ENSEA, CNRS, France**

**E-mail: [nistor.grozavu@cyu.fr](mailto:nistor.grozavu@cyu.fr)**

**Web: <https://www.grozavu.fr>**

**ET**

**Fabrice SURMA**

**Epitopos - [www.epitopos.fr](http://www.epitopos.fr) – [fabrice.surma@epitopos.fr](mailto:fabrice.surma@epitopos.fr)**

**20 rue des serruriers**

**67000 STRASBOURG**

## Description du sujet de thèse

Les **besoins** identifiés par la commission EU sur la question de l'impact du changement climatique sur le patrimoine se focalisent sur le développement de **plan d'exploitation des données**. En effet, aujourd'hui la modélisation de l'impact du changement climatique sur le patrimoine est **établie à partir d'un maillage global (échelle régionale)** impliquant l'application d'une méthodologie de conservation préventive systématique. Cependant, il s'avère indispensable **d'identifier des sites sensibles** en fonction des risques en mettant en place **des plans de monitoring systématiques sur long terme** (MH, musées, archives, réserves, sites archéologiques, jardins et parc historiques). Il faut donc centraliser des données indispensables pour aller vers **la modélisation, la prédiction** et en établir des **critères de vulnérabilité** sur des sites témoins.

L'instrumentation (thermo-hydrigue) de site se fait régulièrement lors des phases de chantiers ou d'étude. Il existe une vraie volonté de rendre ces suivis climatiques plus systématiques mais la problématique du suivi, du stockage et de l'exploitation des données se pose notamment au regard de l'environnement du site.

L'intérêt de l'approche multi-échelle que nous proposons est bien **d'affiner le maillage de récolte** de données et de coller **aux évolutions réelles de chaque site** pris dans leur intégrité et leur spécificité. Il s'agit de créer un lien entre l'évolution du climat global et ses conséquences à l'échelle locale du site et de sa réponse comportementale. Pour cela nous proposons de prendre et intégrer les évolutions passées de l'objet patrimonial et de ses matériaux en relation avec son histoire climatique pour le projeter dans un environnement actuel et futur. Il s'agit donc d'intégrer l'objet patrimonial dans une **multi-temporalité intégrant les données historiques**.

Le but de la thèse est donc de coordonner l'ensemble des données climatiques pertinentes, d'une échelle large (satellite et métrique) à l'échelle de l'objet patrimonial, en intégrant son micro-climat extérieur et intérieur, afin de tester un ou des modèles prédictifs de son état de conservation. Elle comportera deux axes :

- mettre en corrélation ces données avec différentes réponse mesurables de l'objet patrimonial face aux variations climatiques pour aller vers la modélisation de son comportement en fonction de son environnement
- mettre en corrélation de ces données avec des données issues de la bibliographie donnant les limites de transformation irréversible des matériaux constituant l'objet patrimonial, ce à différents états de vieillissement

## Methodologie

Les systèmes d'acquisition avancés permettent d'enregistrer plusieurs types de données pour une compréhension plus approfondie de scénarios complexes dans lesquels les données sont très souvent collectées en continu à partir de différentes sources et disponibles dans différentes modalités, impliquant des **données dynamiques multimodales**. Il en résulte des données multimodales hétérogènes, y compris des séries chronologiques, du texte, des images, etc. présentent des problèmes méthodologiques dans la modélisation et l'apprentissage à partir de données aussi complexes. Ce défi rend le processus de supervision presque impraticable et très incertain compte tenu de la nature dynamique inconnue des systèmes observés, pour lesquels les approches classiques d'apprentissage automatique « statique » et d'inférence statistique sont inappropriées et irréalisables [3 ; 4].

En effet, malgré cette augmentation importante de l'accessibilité aux données dynamiques complexes multimodales, il existe un manque d'approches à traiter, notamment pour l'apprentissage des représentations [1;5]. La plupart des systèmes classiques d'apprentissage automatique et d'inférence statistique dédiés aux données multimodales et/ou complexes, qu'ils soient basés sur des modèles aléatoires, des mesures empiriques ou des modèles basés sur des prototypes, reposent sur une hypothèse forte, consistant à supposer au moins la structure des données. Le processus de génération de la scène observée est fixe, bien qu'il puisse être supposé inconnu. Dans un contexte non supervisé, certains travaux existants sur les approches d'apprentissage automatique par Ensemble et Collaboratif ont été proposés mais sont limités à la même distribution de données, c'est-à-dire dans un contexte multi-vues [2;3].

Nous voulons à travers cette thèse explorer **l'apprentissage non supervisé topologique à base de réseaux de neurones et probabilistes de données multimodales complexes, à la fois au niveau spatial et temporel**. Nous nous intéressons particulièrement aux données hétérogènes dont la représentation peut avoir été renseignée de différentes manières : représentation experte qui peut être complexe (par exemple : des multi-graphes dynamiques qui peuvent éventuellement avoir des topologies différentes pour chaque observation du jeu de données et pour chaque instant) ou une représentation apprise automatique, images, signaux.

## Références :

1. Bengio Y., Courville A., Vincent P. (2013), Representation Learning: A Review and New Perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence , Volume 35 Issue 8 Publisher: IEEE Computer Society
2. Rastin, P., Matei, B., Cabanes, G., Grozavu, N., Bennani, Y.: Impact of Learners' Quality and Diversity in Collaborative Clustering. Journal of Artificial Intelligence and Soft Computing Research, 9(2): 149–165 doi:10.2478/jaiscr-2018-0030 (2018).
3. W. Guo, J. Wang and S. Wang, "Deep Multimodal Representation Learning: A Survey," in IEEE Access, vol. 7, pp. 63373-63394, 2019, doi: 10.1109/ACCESS.2019.2916887.
4. Rastin, P., Matei, B., Cabanes, G., Bennani, Y., Marty, J.M.: A new sparse representation of complex data: application to dynamic clustering of web navigation. Pattern Recognition, 91: 291–307, doi:10.1016/j.patcog.2019.02.020 (2019).
5. Bengio Y. (2013), Deep learning of representations: looking forward. SLSP'13: Proceedings of the First international conference on Statistical Language and Speech Processing.

## Informations Complémentaires

Ce doctorat s'inscrit dans le projet d'EquipEx+ ESPADON "En Sciences du Patrimoine, l'Analyse Dynamique des Objets anciens et Numériques" sélectionné par le PIA 3. Ce projet ambitieux et fédérateur est coordonné par Vincent Detalle (chercheur au C2RMF) et porté par la Fondation des Sciences du Patrimoine.

Le projet ESPADON vise à mettre à disposition de la communauté, d'une part, de nouveaux moyens instrumentaux de tomoographies 3D multi-échelles et d'imagerie multi-physique 2D devenus possibles notamment grâce à l'augmentation des puissances de calculs, et, d'autre part, des ressources et savoir-faire numériques uniques de traitement et de gestion de données massives, et de stockages en complémentarité et avec l'appui du TGIR Human-Num et les actions SoCoRe et Parcours déjà menées. Dans son ensemble, ESPADON vient conforter des dynamiques déjà installées depuis le lancement du premier EquipEx PATRIMEX. A terme ce projet va intégrer les éléments suivants :

- La mise en place d'un réseau d'appareils travaillant à différentes échelles et à différents niveaux de mobilité permettant de se déplacer vers les œuvres quand cela est nécessaire.
- Une gestion et une utilisation mutualisable et mutualisée des données.

- La structuration effective, en grande partie déjà démontrée d'un réseau de chercheurs et de laboratoires sur le territoire intégrant les éléments de formation des étudiants et futurs chercheurs.
- La mise en place d'un Archivage de données brutes via Humanum ainsi que des données « finalisées » permettant la création de l'objet patrimonial augmenté prenant en compte les interactions entre l'objet et son environnement.

Dans ce cadre le projet doctoral s'inscrit plus particulièrement dans le WP4 PASTAGA (**P**atrimoine **A**rchitectural, **S**ites, **T**erritoires **A**rchéologiques et **G**randes **A**rtéfacts) et le WP6 CLIP (**C**limat, **I**nterface et **P**atrimoine) du projet ESPADON

Pour plus d'informations: <http://www.sciences-patrimoine.org/2020/12/selection-espardon/>

**Modalités de candidature:**

Dossier de candidature complet avec le CV, lettre(s) de recommandation(s), lettre de motivation, résultats scolaires à envoyer à :

**Nistor GROZAVU**

email : [nistor.grozavu@cyu.fr](mailto:nistor.grozavu@cyu.fr)