

# Sophia Yazzourh | CV

Institut de Mathématiques de Toulouse, France

✉ [sophiyazzourh@gmail.com](mailto:sophiyazzourh@gmail.com) • [sophiyazzourh.github.io](https://github.com/sophiyazzourh)

Française & Marocaine

*Docteure en Mathématiques Appliquées, spécialisée en  
Machine Learning appliqué aux données médicales.*

Récemment diplômée d'un doctorat à l'*Institut de Mathématiques de Toulouse* (IMT), je travaille sur l'apprentissage par renforcement dans le cadre des séquences de traitements pour des patients atteints de maladies chroniques. Je suis également diplômée d'un titre d'ingénieure de l'*Institut National de Sciences Appliquées* (INSA) de Toulouse en mathématiques appliquées et statistique.

Domaines de recherche : Reinforcement Learning, Precision Medicine, Dynamic Treatment Regimes, Bayesian Support Vector Machine, Preference Learning

## Parcours académique

- **Thèse de doctorat en mathématiques appliquées** **IMT (France)**  
*Algorithmes de décision dans la recherche médicale* *Septembre 2021 - Octobre 2024*  
*"Apprentissage par renforcement et outcome-weighted learning bayésien pour la médecine de précision. Intégration de connaissances médicales dans les algorithmes de décision."*  
Sous la supervision de [Nicolas Savy](#) (IMT) et [Philippe Saint Pierre](#) (IMT) en collaboration avec l'*Université de Caroline du Nord* (UNC) à Chapel Hill, Etats-Unis et notamment avec [Michael Kosorok](#) et [Nikki L.B. Freeman](#).
- **Diplôme d'ingénieure, grade Master** **INSA Toulouse (France)**  
*Spécialité Mathématiques Appliquées & Statistique* *Septembre 2014 - Juin 2021*
  - Machine Learning, High Dimensional and Deep Learning, Artificial Intelligence Framework (NLP, Reinforcement Learning).
  - Projet d'Innovation et Recherche sur le thème de la modélisation d'une variable quantitative dans le cadre de la régression logistique au travers de splines d'interpolation.
- **Erasmus** **NTUA (Grèce)**  
*Réalisé dans les départements de Mathématiques et d'Informatique* *Septembre 2016 - Janvier 2017*  
Programme d'échange européen effectué pendant mon premier semestre de License 3 à la *National Technical University of Athens* (NTUA).

## Expériences professionnelles

- **Alternance en Big Data & IA pour les activités spatiales** **Thales Alenia Space**  
*Recherche et Développement en Intelligence Artificielle* *Janvier 2020 - Juin 2021*
  - Développement d'une bibliothèque Python d'algorithmes pour la détection d'anomalies, appliquée au domaine spatial.
  - Mise en œuvre et adaptation de méthodes d'Active Anomalies Discovery sur des signaux temporels : détection d'anomalies dans le cadre de l'Active Learning, notamment l'intégration du retour utilisateur dans des algorithmes tels qu'Isolation Forest ou SVM.*Compétences acquises : Python, ACP, Isolation Forest (IF), SVM, Autoencodeurs, courbes ROC, GIT, Docker, API Flask.*

- **Alternance en gestion et processus industriels des segments sols spatiaux**      **Thales Alenia Space**  
*Gestion de projets*      *Octobre 2019 - Janvier 2020*  
 Réflexion sur la transmission d'informations et la cohérence des outils jusqu'en phase de maintenance.  
*Connaissances acquises : Cycle de vie des projets segments sols spatiaux et outils associés.*
- **Stage de Master 1, GALILEO Mission Segment**      **SOGETI HIGH TECH, Groupe Capgemini**  
*Intégration et Validation*      *Juillet 2019 - Septembre 2019*
  - Intégration d'un élément prototype.
  - Analyse et modélisation des données et des flux des éléments connectés du composant Mission and Uplink Control Facility (MUCF).
  - Analyse des spécifications de L'Agence Spatiale Européenne, préparation et mise en place de scénarios de tests dans l'équipe d'assemblage.*Connaissances acquises : Fonctionnement du segment sol de GALILEO*

## Bourses

---

- **Contrat doctoral Ministère français de l'Enseignement Supérieur, de la Recherche et de l'Innovation**  
 Bourse doctorale de 3 ans obtenue après concours.

## Publications

---

- **Article soumis pour publication :**
  - Sophia Yazzourh (IMT), Nicolas Savy (IMT), Philippe Saint Pierre (IMT) et Michael Kosorok (UNC). *Medical Knowledge Integration into Reinforcement Learning Algorithms for Dynamic Treatment Regimes*, International Statistical Review, 2024. arXiv preprint [arXiv:2407.00364](https://arxiv.org/abs/2407.00364).
- **Soumissions prévues :**
  - Sophia Yazzourh (IMT) et Nikki L. B. Freeman (Duke). *Bayesian Outcome Weighted Learning*. En préparation pour Biostatistics. arXiv preprint [arXiv:2406.11573](https://arxiv.org/abs/2406.11573).
  - Sophia Yazzourh (IMT), Nicolas Savy (IMT), Philippe Saint Pierre (IMT) et Michael Kosorok (UNC). *Rewards construction based on preference learning for Dynamic Treatment Regimes*.
- **Manuscrit :**  
 Sophia Yazzourh, "Apprentissage par renforcement et outcome-weighted learning bayésien pour la médecine de précision. Intégration de connaissances médicales dans les algorithmes de décision.", Manuscrit de Doctorat, Université de Toulouse préparé à l'Université de Toulouse III - Paul Sabatier, 2024.

## Présentations orales

---

- **10ème édition des Journées Statistiques du Sud** (*Université Toulouse III - Paul Sabatier, Toulouse, 18 juin 2024*) "[Bayesian Outcome-Weighted Learning](#)".  
 Présentation sélectionnée sur abstract.
- **Séminaire Étudiant de l'Institut de Mathématiques de Toulouse** (*Université Toulouse III - Paul Sabatier, Toulouse, 6 juin 2024*) "[Introduction to Reinforcement Learning](#)".  
 Introduction à l'apprentissage par renforcement pour la prise de décision, formalisme mathématique, algorithmes clés et axes de recherche.
- **55 èmes journées de Statistique de la Société Française de Statistique à Bordeaux** (*Université de Bordeaux, Bordeaux, 27 mai 2024*) "[Construction de récompenses par apprentissage par préférences pour les modèles d'apprentissage par renforcement appliqués aux stratégies de traitements adaptatifs](#)".  
 Présentation sélectionnée sur résumé.
- **Présentation à l'équipe de recherche du PhairLab (Michael Kosorok)** (*UNC, Chapel Hill, 21 avril 2023*) "[Integration of Medical Knowledge into Reinforcement Learning for Dynamic Treatment Regimes](#)"  
 Présentation de mes recherches en cours.

- **Session posters du Séminaire de Statistique et Optimisation de l'IMT** (IMT, Toulouse, 6 décembre 2022) *"Introduction d'Expertise Médicale dans les Modèles d'Apprentissage par Renforcement appliqué aux Stratégies de Traitements Adaptatifs"*.  
Présentation de mon poster pour un séminaire interne thématique sur les Statistiques pour la Biologie et la Santé.
- **Journée Toulousaine « Statistique pour la Biologie » à l'Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE)** (INRAE, Toulouse, 15 novembre 2022) *"Apprentissage par renforcement : application aux stratégies de traitements adaptatifs"*.  
Présentation lors d'une journée d'échanges des chercheurs et chercheuses toulousain.e.s autour de « Statistique pour la Biologie ».

## Enseignements et projets pédagogiques

---

- **Doctorante chargée d'enseignements en mathématiques au sein de l'Institut Universitaire de Technologie Génie Mécanique** (2021 - 2024)
  - Chargée de travaux dirigés et travaux pratiques aux niveaux Licence 1, 2 et 3.
  - Cours/travaux dirigés en anglais pour le parcours anglophone.
  - Réalisation et traduction en anglais de supports de cours, travaux dirigés et travaux pratiques.
- **Bénévole à l'association Ouverture Rencontres Évolution (ORE)** (2014-2021)
  - Soutien scolaire de la 6e à la Terminale en mathématiques, sciences et français.
  - Mise en place d'une formation d'initiation à Python et à l'algorithmique.

## Engagement académique

---

- **Responsable du Séminaire Étudiant de l'IMT** (2022 - 2024) : Organisation bimensuelle de [séminaires](#) à destination des doctorants et post-doctorants, de moments conviviaux et réalisation [d'affiches](#).
- **Membre du comité d'organisation et scientifique des 10èmes Journées Statistiques du Sud** : Ensemble de groupes de travail ayant lieu dans le sud de la France, organisés autour de mini-cours et d'exposés.
- **Vulgarisatrice scientifique en lycée** (2022-2024) : Projet "Déclics", médiation scientifique autour d'échanges successifs et présentations de mon parcours.
- **Membre de sociétés scientifiques** : Société Française de Statistique, Groupe de Recherche Intelligence Artificielle.
- **Présidente de l'association étudiante du département de mathématiques de l'INSA de Toulouse** (2019-2020).

## Centre d'intérêts

---

- **Joueuse de Handball** :
  - Club, Niveau National 3 : en parallèle de mes études et depuis l'âge de 8 ans, je m'entraîne 3 à 4 fois par semaine et dispute un match tous les week-ends.
  - Universitaire : participation aux Championnats de France des Écoles en 2015, 2016, 2017, 2018 et 2019.
- **Responsable de la section handball et capitaine de l'équipe féminine de l'INSA** (2016-2019).
- **Trésorière de l'association sportive de l'INSA** (2016-2017).

## Compétences complémentaires

---

- **Langages de programmation** : Python,  $\text{\LaTeX}$ , R, Matlab, Java, C, C++.
- **Outils de développement** : Git, Docker, Google Cloud.
- **Langues étrangères** : Anglais, Allemand (Compétences professionnelles limitées), Chinois (Compétences professionnelles limitées) et Grec (Notions élémentaires).