

## 2 Curriculum Vitae

Tanguy LEFORT  
Montpellier, France      GitHub : tanglef  
tanguy.lefort@seenovate.com      Website : <https://tanglef.github.io>

**Consultant data** : expertise computer vision, deep-learning et crowdsourcing

### Expérience

---

#### Consultant data R&D

*Seenovate Montpellier, France*      *02/2025 – Présent*

- **Manager** : Frank Imbach (SeeSports, Univ. Montpellier)
- **R&D** : Création de modèles et processus liés à la vision par ordinateur appliquée au sport et la Business Intelligence (BI).
- **Chef de projet en reconnaissance d'action** pour la fédération de judo ; estimation de squelettes, segmentation de vidéo (>300K vidéos) ; cadrage et chiffrage du projet ; partenariat INRIA/France.Sport ; déploiement partiel 2025 puis complet 2026. Contient un axe de recherche sur l'identification de poses en combat de judo pour la reconnaissance de prises avec des données non balancées et des annotations bruitées
- Autres projets R&D :
  - **Lead vision** : Détection de maladies sur des photographies de cuirs traités pour une tannerie Lyonnaise avec Hermès ; déploiement effectué
  - **Lead vision** : Création d'un système de détection de plaques de BMX en compétition jusqu'à 89% de précision en temps réel (0.05-0.1s latence) ; déploiement 2026
  - **Lead tech et responsable Mise en Production** Logiciel SeeQL permettant d'obtenir des requêtes SQL et leur résultat dans un univers BO (type SAP, leader du marché sur les traitements de données en business analytics) en posant une question en langage naturel. Supervision d'une équipe de développeurs en méthode Agile.
  - Développeur : **Séries temporelles BI** en temps réel à grande échelle (de ~8minutes à moins de 30s) ; déploiement Janvier 2026

#### Post-Doctorat

*INRIA Lille, France*      *10/2024 – 02/2025*

- **Superviseur** : Odalric-Ambrym Maillard (INRIA)
- **Recherche** : Profilage d'expertise d'utilisateurs avec l'apprentissage par renforcement et collecte de donnée personnalisée via les systèmes de recommandation
- Modélisation de l'expertise d'utilisateurs pour l'annotation de données botaniques à grande échelle
- Création d'un système de recommandation pour mieux exploiter les interactions utilisateurs dans le système Pl@ntNet
- Obtention de garanties théoriques et implémentation d'algorithme pour la quantification de l'incertitude des prédictions et des profils utilisateurs en crowdsourcing

## Doctorat

*IMAG, Université de Montpellier et INRIA Montpellier, France*      *10/2021 – 09/2024*

- **Superviseurs** : Benjamin Charlier (CNRS), Alexis Joly (INRIA) et Joseph Salmon (CNRS)
- **Recherche** : Étiquettes bruyantes dans des ensembles de données de classification crowdsourcées avec retour d'experts
- Identification de l'ambiguïté des données dans les ensembles de données collaboratifs à l'aide de marges de réseaux de neurones
- Développement d'un cadre standardisé dans une bibliothèque Python : `peerannot` (<https://peerannot.github.io/>)
- Comparaison des stratégies d'agrégation en crowdsourcing dans un ensemble de données réel avec de nombreuses classes et de nombreux annotateurs avec Pl@ntNet

## Stage

*IMAG, Montpellier, France*      *03/2021 – 08/2021*

- Mémoire de Master sur l'optimisation en haute dimension pour les modèles linéaires pénalisés avec interactions
  - **Superviseurs** : Benjamin Charlier et Joseph Salmon
  - Évaluation des méthodes de descente pour les modèles linéaires avec pénalités L1 et L2 et interactions de premier ordre
  - Application de stratégies d'accélération GPU pour des ensembles de données génomiques
- IMAG, Montpellier, France*      *07/2020 – 08/2020*
- Contribution à la bibliothèque PyKeOps sous la supervision de Benjamin Charlier
  - Réécriture de la routine de gradient conjugué Fortran de Scipy pour les matrices symboliques dans PyKeOps
  - Évaluation des performances de la régularisation de Ridge-Tikhonov et travaux sur les problèmes de transport optimal

## Formation

---

*Master en Biostatistique, Montpellier, France*      *2019–2021*

- Statistiques et probabilités théoriques et appliquées
- Algorithmes de classification, analyse de survie et modélisation des populations

*Licence en Mathématiques, Dijon, France*      *2016–2019*

- Licence avec mention, spécialisée en mathématiques appliquées
- Projet final sur l'algorithme de squelettisation pour la chirurgie par rayons gamma

## Publications

---

### En cours

- *Application of the factor-analytic mixed model framework to embedded data* by A. Marchal, M. Lamure, T. Lefort and F. Imbach      *2025*

## Journaux

- *Bayesian implementation of the factor-analytic mixed model and application to embeddings* in **Wiley Stat.** by A. Marchal, M. Lamure, T. Lefort, F. Imbach (submitted) 2025
- *Cooperative learning of Pl@ntNet's Artificial Intelligence algorithm : how does it work and how can we improve it ?* in **Methods in Ecology and Evolution** by T. Lefort , A. Affouard , B. Charlier , J-C. Lombardo , M. Chouet, H. Goëau , J. Salmon , P. Bonnet and A. Joly 2024
- *Peerannot : classification for crowdsourced image datasets with Python* in **Computo** by T. Lefort, B. Charlier, A. Joly and J. Salmon 2024
- *Identify ambiguous tasks combining crowdsourced labels by weighting Areas Under the Margin* in **TMLR** by T. Lefort, B. Charlier, A. Joly and J. Salmon 2024

## Conférences

- *Weighted majority vote using Shapley values in crowdsourcing* in CAp by T. Lefort, B. Charlier, A. Joly and J. Salmon 2024
- *Cooperative learning of Pl@ntNet's Artificial Intelligence algorithm using label aggregation* in Journées des statistiques de France by T. Lefort, A. Affouard, , B. Charlier, J. Salmon, P. Bonnet and A. Joly 2024
- *peerannot : A framework for label aggregation in crowdsourced datasets* in Journées des statistiques de France by A. Dubar, T. Lefort and J. Salmon 2024
- *Weighting areas under the margin in crowdsourced datasets* in Journées des statistiques de France by T. Lefort, B. Charlier, A. Joly and J. Salmon 2023
- *Crowdsourcing label noise simulation on image classification tasks* in Journées des statistiques de France by T. Lefort, B. Charlier, A. Joly and J. Salmon 2022
- *Benchopt : Reproducible, efficient and collaborative optimization benchmarks* in **NeurIPS 2022** by T. Moreau, M. Massias et. al 2022

## Autres engagements

---

### Contributions aux bibliothèques open source

- Développeur principal de la bibliothèque Peerannot pour le traitement des ensembles de données d'images crowdsourcées
- Développeur de la bibliothèque BenchOpt pour des benchmarks reproductibles en optimisation

### Services communautaires

- Relecteur pour le journal international TMLR 2024 – En cours
- Relecteur pour le journal Computo (SFDS) 2023 – 2025
- Co-organisateur du séminaire doctoral à l'IMAG, Université de Montpellier 2022–2023

## Enseignement

---

- Formateur dans le cadre de formations inter-entreprises en IA et sciences des données à vocations d'initiations ou de présentations de domaines techniques (notamment pour la DAE – Direction des Achats de l'Etat et l'INSEP – Etablissement d'enseignement supérieur à Paris) 2025
- Vacataire dans le cadre d'un projet M2-Statistiques et Sciences des Données pour le développement d'une application de visualisation et analyse de séries spatio-temporelles. 2025
- Co-directeur de Guillaume Demoor (Mines Alès) avec F-D. Collin et G. Durif, stage de Master de 4 mois sur *la conversion d'images de formules mathématiques en encodage LaTeX avec l'apprentissage automatique* 2024
- Encadrant de 4 étudiants en Master pour la création d'un atelier de visualisation de données génomiques à la conférence Montpellier Omics Days 2023
- Chargé de TD en optimisation convexe pour la licence 3 mathématiques
- Chargé de TD en probabilités pour un cours de licence 1 SVSE (Sciences de la vie)
- Chargé de TD en bases de logique et raisonnements mathématiques pour la licence 1 mathématiques
- Chargé de TD en outils mathématiques pour la chimie (algèbre, analyse) pour la licence 2 chimie 2023

## Conférences et Séminaires

---

- *Recommending observations and aggregating annotations for plant species identification* INRIA Lille (équipe MAGNET) 2025
- *Annotation collaborative des données botaniques : à qui faire confiance sur quoi ?* INRAE Paris Saclay 02/2025
- *Keeping humans in the loop via citizen science* Univ. Lille 11/2024
- *Crowdsourcing and citizen science : how to handle our data ?* INRIA, Lille (SCOOL) 07/2024
- *Cooperative data enrichment algorithms* at Journées des Statistiques de France (JDS) Univ. Bordeaux 05/2024
- *Apprentissage collaboratif d'espèces de plantes et agrégation de labels dans Pl@ntNet*, IA-ECO seminar, UMR MARBEC. 03/2024
- *ChatGPT & co, Myths and Reality. Everything you wanted to ask about Deep Learning but did not dare to ask*, General knowledge Seminar with Francois David Collin, IMAG. 10/2023
- *Data collection from a crowd : where is the noise coming from ?* at Ph.D students seminar, IMAG. 09/2023
- *Weighting areas under the margin in crowdsourced datasets* at Journées des Statistiques de France (JDS) Univ. Bruxelles 07/2023
- *Learning from crowds : going beyond aggregation schemes* at ML-MTP seminar, Univ. Montpellier 02/2023
- *Improve learning combining crowdsourced labels by weighting Areas Under the Margin* at ML-MTP seminar, Univ. Montpellier 10/2022

- *Crowdsourcing label noise simulation on image classification tasks* at Journées des Statistiques de France (JDS) Univ. Lyon. 06/2022
- *High dimensional optimization for penalized linear models with interactions using graphics card computational power* at Probability and Statistics (EPS) team seminar 11/2021
- *Introduction to neural networks* with Joseph Salmon at ML-MTP seminar, Univ. Montpellier 10/2021
- *Paper club Ridge Regularization : an Essential Concept in Data Science* by Trevor Hastie with Florent Bascou at ML-MTP seminar 04/2021

## Compétences générales

---

- **Programmation** : Python, R, Git, L<sup>A</sup>T<sub>E</sub>X, JavaScript, HTML, CSS, Docker
- **Algorithmes de Machine learning** : régressions (linéaires, logistiques, *etc.*), arbres, SVM, KNN, réseaux de neurones (pytorch et keras/tensorflow), bandits, systèmes de recommandation, systèmes d’allocation de tâches
- **Langues** : Français, Anglais, Espagnol, Italien (ordre décroissant d’expertise)

## Références

- **Frank Imbach**, Manager Seenovate frank.imbach@seenovate.com
- **Lionel Reveret** Collaboration académie-industrie pour la FFJ lionel.reveret@inria.fr
- **Joseph Salmon**, Directeur de thèse joseph.salmon@inria.fr