

# Application BASTRI

## Fiches Equipes

### MODAL (SR0518VR)

MModel for Data Analysis and Learning  
MODAL (SR0434ZR) □ MODAL

**Statut:** En cours de fermeture

**Responsable :** Cristian Preda

**Mots-clés de "A - Thèmes de recherche en Sciences du numérique - 2024" :** A3.1.4. Données incertaines , A3.1.10. Données hétérogènes , A3.2.3. Inférence , A3.3.2. Fouille de données , A3.3.3. Analyse de données massives , A3.4.1. Apprentissage supervisé , A3.4.2. Apprentissage non supervisé , A3.4.5. Méthodes bayésiennes , A3.4.7. Méthodes à noyaux , A5.2. Visualisation de données , A5.9.2. Estimation, modélisation , A6.2.3. Méthodes probabilistes , A6.2.4. Méthodes statistiques , A6.3.3. Traitement de données , A9.2. Apprentissage

**Mots-clés de "B - Autres sciences et domaines d'application - 2024" :** B2.2.3. Cancer , B9.5.6. Science des données , B9.6.3. Economie, finance , B9.6.5. Sociologie

**Domaine :** Mathématiques appliquées, calcul et simulation  
**Thème :** Optimisation, apprentissage et méthodes statistiques

**Période :** 01/01/2012 -> 30/06/2025

**Dates d'évaluation :** 19/03/2014 , 14/03/2018 , 01/12/2022

**Etablissement(s) de rattachement :** CNRS, UNIVERSITE DE LILLE  
**Laboratoire(s) partenaire(s) :** LPP (UMR8524)

**CRI :** Centre Inria de l'Université de Lille  
**Localisation :** Centre Inria de l'Université de Lille  
**Code structure Inria :** 101037-1

**Numéro RNSR :** 201020969D  
**N° de structure Inria:** SR0518VR

### Présentation

The main focus of MODAL is to design generative models dealing with complex multivariate and/or heterogeneous data. Typical instances of such data are

- nominal covariables for the multivariate setting,
- and the combination of continuous and nominal variables for the heterogeneous setting.

Obviously, other widespread complex covariables are of interest such as ordinal, ranks, and intervals data.

From these generative models, a convenient and efficient statistical analysis remains to be carried out, leading to data analysis (visualization, clustering) and data learning (supervised and semi-supervised classification, density estimation).

### Axes de recherche

MODAL is focused on generative models, that is models describing the generation process of data, unlike predictive models.

Generative models are of great interest. On the one hand, they are required in several statistical objectives such as clustering, semi-supervised classification, and density estimation, where predictive models are useless. On the other hand, these models enable data visualization. Indeed, they provide a full description of the data distribution, which gives access to several aspects of the data such as high density areas for instance.

In supervised classification, generative and predictive models directly compete with one another. However, the lack of flexibility of the generative approach, as opposed to the predictive one, is completely balanced by the use of model selection.

In addition, among generative approaches, parametric ones such as mixture models are preferred. Provided parameters are meaningful and parsimonious, mixture models allow valuable data interpretation.

### Contact

- **Responsable :** Cristian Preda
- **Tél :** 03.59.57.78.57
- **Secrétariat Tél :** 03.59.57.78.45

### En savoir plus

- Site sur [inria.fr](http://inria.fr)
- Site du [responsable](#)
- Derniers Rapports d'Activité : [2016](#) , [2017](#) , [2018](#) , [2019](#) , [2020](#) , [2021](#) , [2022](#) , [2023](#)

### Documents sur la structure

- [Intranet](#)
- [Privés](#)

### Décisions

- **8640** (07/05/2012) : création
- **10552** (26/01/2015) : prolongation
- **13244** (10/12/2018) : prolongation
- **14558** (03/12/2020) : cessation du responsable
- **14559** (03/12/2020) : nomination responsable
- **15814** (29/11/2022) : prolongation
- **16760** (16/01/2024) : prolongation
- **17109** (07/06/2024) : prolongation
- **17551** (10/12/2024) : prolongation

### Localisation

- **Adresse postale :** Centre Inria de l'Université de Lille Parc Scientifique de la Haute Borne 40, avenue Halley Bât.A, Park Plaza 59650 Villeneuve d'Ascq France
- **Coordonnées GPS :** 50.606, 3.149

# Logiciels

- MASSICCC (Massive Clustering with Cloud Computing)

## Relations industrielles et internationales

### Current collaborations

- University College London, UK
- University of Oxford, UK
- University of Bucharest, Romania
- INRS, Canada
- ALICANTE
- DECATHLON
- CEREMA
- ADEO
- SECKIOT
- WORLDLINE