Supervisors: Morgan MAGNIN (Prof., LS2N / Centrale Nantes) and Jérémie POSCHMANN (Dr., CR2TI, INSERM UMR 1064, Nantes Université)
With the collaboration of: Tony RIBEIRO (Dr., LS2N) and Antoine ROQUILLY (Prof., MD-PhD, CR2TI – LUNA PARK coordinator)
Host laboratories: Laboratoire des Sciences du Numérique de Nantes (LS2N, CNRS UMR 6004) and Centre de Recherche en Transplantation et Immunologie Translationnelle (CR2TI, INSERM UMR 1064, team 6)
Institutions: École Centrale de Nantes and Nantes Université
Place: Nantes, France
Research project: LUNA PARK – LUNg Atlas from symbiosis to Pneumonia dysbiosis: a multi-compartment, multiomics state and transition framework
Funding: Labex IMMUNE – 2026 Call for Projects (3-year doctoral contract shared between the two partner teams)
Research teams: VELO (Vérification pour l'Environnement et le LOgiciel, LS2N) and Host-Pathogen Interactions, Inflammation and Mucosal Immunity (CR2TI, team 6)
Doctoral school: MASTIC (https://ed-mastic.doctorat-paysdelaloire.fr/)

PhD position – Symbolic AI and multiomics for modelling lung dysbiosis in pneumonia

A 50/50 computer science — bioinformatics / biological data analysis thesis

Context

The general goal of this PhD thesis is to bring new contributions to symbolic artificial intelligence and inductive logic programming, and to deploy them on a major health issue: understanding how the lung ecosystem moves from health to pneumonia. The work is deliberately balanced, with one half dedicated to computer science (the design and implementation of learning algorithms) and one half to bioinformatics and the biological interpretation of multiomics data. The successful candidate will be jointly trained and co-supervised by the LS2N digital-sciences team and the CR2TI immunology team.

More specifically, this thesis aims at better understanding the dynamic links between the respiratory microbiome, host immunometabolism and the progression toward hospital-acquired pneumonia in critically ill patients. The lung is best described not as a stable community of micro-organisms but as a low-biomass flux ecosystem held in check by mucociliary clearance and immune surveillance. In intensive care, this constrained symbiosis is rapidly reshaped by physiological stress and by treatments such as antibiotics and immunomodulatory therapies, creating a perturbed intermediate state that may either stabilise or progress to pneumonia.

The thesis seeks to build longitudinal, interpretable models of this evolution over time, with the aim of (1) formalising the transitions between three clinically relevant configurations — healthy lung symbiosis, ICU-associated perturbation, and infection-associated dysbiosis; (2) identifying the molecular and microbial markers that drive progression toward pneumonia; and (3) contributing to state-informed risk stratification and to the rational design of prevention strategies.

This work is part of the **LUNA PARK** research project (*LUN*g *A*tlas from symbiosis to *P*neumonia dysbiosis), funded by the Labex IMMUNE 2026 call and coordinated by Professor Antoine

Roquilly (Nantes Université – CR2TI). The project brings together two complementary partners: the CR2TI immunology team, which provides the clinical cohorts, the multiomics data and the biological expertise, and Centrale Nantes / LS2N, which provides the formal modelling and symbolic-AI methodology.

Mixing inductive learning approaches, statistics-based methods and multiomics data analysis, the work will consist in designing and producing new logic programs able to capture the dynamic reorganisation of the lung ecosystem and to expose, in an explainable way, the checkpoints that precede pneumonia.

Background

The LS2N team has long been developing methods for symbolic learning and logical analysis of the dynamics of living systems, with a focus on explainable dynamic models. For a decade, their research has centered on Learning From Interpretation Transition [Ino2014, Rib2015] (LFIT), an inductive logic programming paradigm that automatically builds a model of the dynamics of a system from the observation of its state-transitions. LFIT outputs explainable dynamical models from time series without any prerequisite about the semantics of the model: from time series data it builds a comprehensible set of logical rules (see Figure 1), which can equivalently be read as a Boolean network.

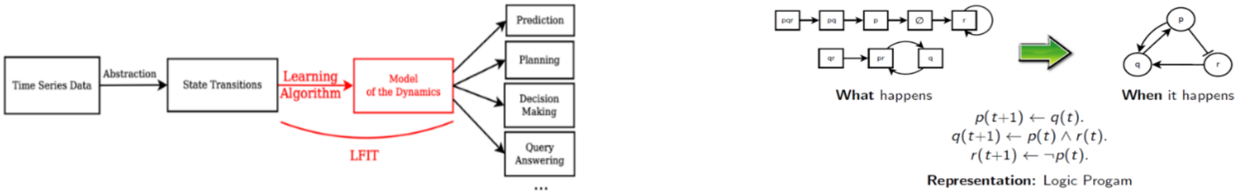

Figure 1 : Left: general workflow for using LFIT; Right: illustration of the learning process highlighting the logic program (bottom) learned by LFIT from the state-transition sequences given as input, which can be represented as a Boolean network (top right)

Unlike statistical or deep-learning predictors, LFIT produces explicit transition logic rather than a black box, which is decisive for clinical interpretation. Recent theoretical advances [Rib2025a] extended LFIT to learn from transitions between partial states, where some variable values are unknown, which is a critical requirement for ICU cohorts in which data are frequently missing, either because some variables were not measurable or were not considered of interest when the cohort was assembled. By modelling the unknown, this extension yields a controlled overestimation of the real system, both for its dynamics and for the interactions between variables.

On the biological side, the CR2TI team has assembled a unique set of resources: a healthy-volunteer multiomics cohort with upper and lower airway sampling, and several longitudinal ICU cohorts (IBIS, PrevHAP, HAPDex) sampled at standardised day-1, day-3 and day-7 timepoints. These datasets combine bacterial and virome metagenomics and metatranscriptomics with lipidomic, proteomic and metabolomic layers. The team and others have shown that pneumonia patients can be grouped into reproducible airway community patterns [Montassier2023], that viral community changes can precede hospital-acquired pneumonia [Anani2025], and that longitudinal metabolic response patterns are associated with HAP risk and can condition the response to immunomodulatory (interferon-gamma) intervention [Petrier2026]. What is still missing is a unified, interpretable representation linking host metabolism, immune regulation and microbial function, and making the transitions between system states explicit. Bridging symbolic AI with multiomics analysis is precisely the gap this thesis addresses.

Two major scientific locks must be overcome. The first is the design of a learning approach robust enough to cope with the strong heterogeneity, noise and partial observation of multiomics ICU data; this calls for the integration of statistics-based methods [Ker2003] into the inductive logic programming framework LFIT. The second is the need to adapt the current LFIT so that it can identify the specific logic "master" rules that govern the transitions between the three clinically relevant configurations that are healthy lung symbiosis, ICU-associated perturbation, and infection-associated dysbiosis. This could be envisioned by extending recent findings about counterfactual analysis [Rib2025b] which paves the way to simulate in silico how the removal of specific perturbations would alter the probability of evolving towards pneumonia.

Research subject and work plan

The thesis plan is structured to reflect the interplay between formal developments in computer science (guided by the LS2N team) and data analysis and biological interpretations (drawing on the expertise of the CR2TI team); this complementarity is essential to the project's success.

The work follows the three work packages of the LUNA PARK project.

WP1 – Reference state of healthy lung symbiosis. Contribute to building a multiomics, transkingdom interactome of the healthy lung from the 75-volunteer cohort, then use LFIT to encode multiomics modules as discrete states and infer interpretable rules capturing the stable couplings and checkpoints of healthy symbiosis. (Bioinformatics: feature harmonisation, module inference; computer science: discretisation and rule learning.)

WP2 – Mechanisms of lung dysbiosis. Build stage-specific transkingdom networks for HAP-free and HAP ICU patients across the IBIS, PrevHAP and HAPDex cohorts, and apply LFIT to learn rules linking ICU baseline configurations, perturbation exposures (antibiotics, interferon-gamma, dexamethasone) and transitions toward infection-associated dysbiosis, testing reproducibility across cohorts.

WP3 – Formalising states and transitions. Turn LFIT prototypes into a production-grade pipeline for noisy, partially observed longitudinal multiomics, including a feature-influence module with confidence metrics based on rule support and stability; learn an explicit three-state transition model (symbiosis → ICU perturbation → pneumonia); and deliver clinician-facing visualisation and counterfactual (in-silico perturbation) analyses of the transition logic.

Methodologically, the work will be decomposed into three main tasks: (1) identify and analyse the existing contributions that link logic programming with statistics and machine learning; (2) propose learning algorithms that take longitudinal multiomics time-series as input and output models competitive with the best predictors while remaining explainable — that is, the underlying rules about the dynamical behaviour of the system can be produced and made understandable to biological and clinical experts; (3) assess the efficiency of the algorithms on the real cohorts of the LUNA PARK project. The resulting algorithms will be implemented in a tool released under a free-software licence with a graphical user interface and clinician-facing visualisation.

Candidate profile

Because the thesis is shared equally between two teams, the ideal candidate sits at the interface of computer science and life sciences. We welcome applications from candidates with a primarily computational background who are eager to work on biological data, as well as from candidates trained in bioinformatics or computational biology who have solid programming and formal-reasoning skills.

- Solid skills in computer science, both theoretical (logic programming, formal methods, machine learning) and practical (implementation, software engineering).
- Competence in data analysis and bioinformatics (handling of omics or other high-dimensional biological data, statistics, Python/R), or a strong willingness to acquire it.
- A genuine interest in health sciences and immunology; experience in systems biology or microbiome analysis would be a clear plus.
- Good scientific writing skills and an aptitude for teamwork in an interdisciplinary, multi-site environment.

Contract

- Status: paid thesis under contract.
- Fields: computer science / artificial intelligence, bioinformatics, systems biology, modelling.
- Supervision: jointly supervised across LS2N (Centrale Nantes) and CR2TI (Nantes Université).
- Contract duration: 36 months.
- Desired start date: ideally September–October 2026 (negotiable).

How to apply?

Applications will be reviewed on a rolling basis until the position is filled. Send (1) a cover letter, (2) a CV and (3) Master (first- and second-year) transcripts to the two thesis supervisors, in copy to their collaborators:

- Morgan MAGNIN (supervisor) – morgan.magnin@ls2n.fr
- Jérémie POSCHMANN (supervisor) – jeremie.poschmann@univ-nantes.fr

References

[Cro2020] A. Cropper and R. Morel. “Learning programs by learning from failures.” arXiv preprint arXiv:2005.02259 (2020).

[Ino2014] K. Inoue, T. Ribeiro, and C. Sakama. Learning from interpretation transition. Machine Learning, 94(1):51–79, 2014.

[Ker2003] L. De Raedt and K. Kersting. “Probabilistic logic learning.” ACM SIGKDD Explorations Newsletter 5.1 (2003): 31–48.

[Rib2015] T. Ribeiro, M. Magnin, K. Inoue, and C. Sakama. Learning Delayed Influences of Biological Systems. Frontiers in Bioengineering and Biotechnology, 2, 81, 2015.

[Rib2025a] T. Ribeiro, M. Folschette, M. Magnin, K. Okazaki, K.-Y. Lo, A. Roquilly, J. Poschmann, K. Inoue. Learning From Interpretation Transition with Unknowns. In Learning and Reasoning (IJCLR 2024 / ILP 2024), Nanjing, China, Springer, pp. 66–85, 2025.

[Rib2025b] T. Ribeiro, M. Folschette, M. Magnin, K. Inoue, T. Nguyen, K. Okazaki, K.-Y. Lo, A. Roquilly, Counterfactual Explanations for Learning From Interpretation Transitions. In the 5th International Joint Conference on Learning & Reasoning (IJCLR 2025). University of Surrey, Guildford, UK, September 11-14, 2025. Full proceedings to appear.

[Anani2025] H. Anani et al. Lung virome convergence precedes hospital-acquired pneumonia in intubated critically ill patients. Cell Reports Medicine, 6:102289, 2025.

[Montassier2023] E. Montassier et al. Robust airway microbiome signatures in acute respiratory failure and hospital-acquired pneumonia. Nature Medicine, 29:2793–2804, 2023.

[Petrier2026] M. Petrier, D. Sinha, F.P. Martin, ..., J. Poschmann, A. Roquilly. Identification of a robust metabolic signature associated with hospital-acquired pneumonia and response to interferon-gamma treatment in critically ill patients. Critical Care, 2026. DOI: 10.1186/s13054-026-05946-6.