

2025 *Activity Report*

RESEARCH CENTRE: Inria Lyon Centre

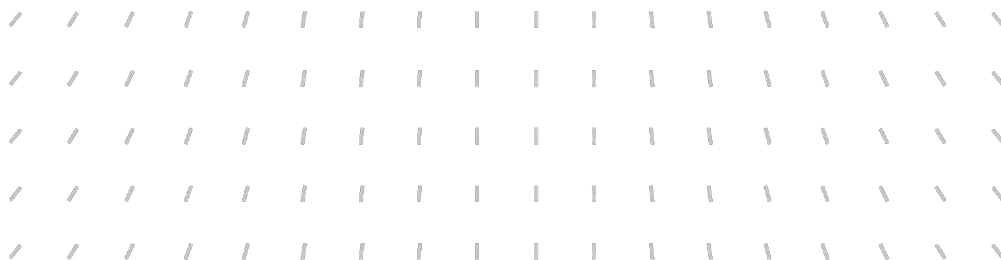
IN PARTNERSHIP WITH: Université Claude Bernard (Lyon 1), Hospices Civils de Lyon - Centre Hospitalier de Lyon, Theranexus


Project-Team

AISTROSIGHT

Viewing neuron-astrocyte pharmacology through
digital sciences





Project-Team AISTROSIGHT

Creation of the Project-Team: 2023 January 01

Each year, Inria research teams publish an Activity Report presenting their work and results over the reporting period. These reports follow a common structure, with some optional sections depending on the specific team. They typically begin by outlining the overall objectives and research programme, including the main research themes, goals, and methodological approaches. They also describe the application domains targeted by the team, highlighting the scientific or societal contexts in which their work is situated. The reports then present the highlights of the year, covering major scientific achievements, software developments, or teaching contributions. When relevant, they include sections on software, platforms, and open data, detailing the tools developed and how they are shared. A substantial part is dedicated to new results, where scientific contributions are described in detail, often with subsections specifying participants and associated keywords. Finally, the Activity Report addresses funding, contracts, partnerships, and collaborations at various levels, from industrial agreements to international cooperations. It also covers dissemination and teaching activities, such as participation in scientific events, outreach, and supervision. The document concludes with a presentation of scientific production, including major publications and those produced during the year.

Keywords

Computer sciences and digital sciences

- A3.1.1. – Modeling, representation
- A3.2.2. – Knowledge extraction, cleaning
- A3.2.4. – Semantic Web
- A3.3.2. – Data mining
- A6.1.1. – Continuous Modeling (PDE, ODE)
- A6.1.2. – Stochastic Modeling
- A6.1.3. – Discrete Modeling (multi-agent, people centered)
- A6.1.4. – Multiscale modeling
- A9.2. – Machine learning
 - A9.2.1. – Supervised learning
 - A9.2.6. – Neural networks
 - A9.2.8. – Deep learning
- A9.4. – Natural language processing
- A9.8. – Reasoning
- A9.10. – Hybrid approaches for AI
- A9.14. – Evaluation of AI models

Other research topics and application domains

- B1.1.2. – Molecular and cellular biology
- B1.1.7. – Bioinformatics
- B1.1.8. – Mathematical biology
- B1.1.10. – Systems and synthetic biology
- B1.2.1. – Understanding and simulation of the brain and the nervous system
- B1.2.3. – Computational neurosciences
- B2.2.2. – Nervous system and endocrinology
- B2.2.7. – Virtual human twin
- B2.6.1. – Brain imaging
- B2.6.3. – Biological Imaging

Contents

Project-Team AISTROSIGHT	1
1 Team members, visitors, external collaborators	5
2 Overall objectives	6
2.1 Failures in drug development for neurological diseases	6
2.2 Main deliverables	6
2.3 Overview of the AIstroSight roadmap	6
2.4 Principles of the AIstroSight partnership	7
3 Research program	8
3.1 Characterizing the mechanisms of action of candidate drugs	8
3.2 Multi-patient query for care pathway characterization and clinical trials	10
4 Application domains	12
4.1 Targeted Pathologies	12
4.2 In vitro and in vivo experimental models of neurology diseases	12
4.3 Identification of multi-source multi-scale biomarkers	13
5 Social and environmental responsibility	15
5.1 Footprint of research activities	15
5.2 Impact of research results	16
6 Highlights of the year	16
7 Latest software developments, platforms, open data	16
7.1 Latest software developments	16
7.1.1 TanaT	16
7.1.2 Soft-ECM	17
7.1.3 Synacomp	17
7.1.4 MotifLocatorAIstroSight	17
7.1.5 Pubmed Target Finder	18
8 New results	18
8.1 Investigating the ultrastructural properties of the endoplasmic reticulum in perisynaptic astrocytic processes and its impact on signaling	18
8.2 Evaluating PDE discovery methods for multiscale modeling of biological signals	18
8.3 Enhancing Fluorescence Correlation Spectroscopy with machine learning to infer anomalous molecular motion	19
8.4 Single-cell multi omics data integration for gene regulatory network inference	19
8.5 Vascular Segmentation of Functional Ultrasound Images using Deep Learning	19
8.6 Methods and frameworks for the analysis of care pathways	20
8.7 Automatic analysis of negation cues and scopes for medical texts in French using language models	21
9 Bilateral contracts and grants with industry	21
9.1 Bilateral contracts with industry	21
10 Partnerships and cooperations	21
10.1 International initiatives	21
10.1.1 Participation in other International Programs	21
10.2 International research visitors	22
10.2.1 Visits to international teams	22
10.3 National initiatives	22

10.4 Regional initiatives	25
11 Dissemination	25
11.1 Promoting scientific activities	25
11.1.1 Scientific events: organisation	25
11.1.2 Scientific events: selection	26
11.1.3 Journal	26
11.1.4 Invited talks	26
11.1.5 Leadership within the scientific community	27
11.1.6 Scientific expertise	27
11.1.7 Research administration	27
11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach	28
11.2.1 Teaching	28
11.2.2 Supervision	28
11.2.3 Juries	28
11.2.4 Educational and pedagogical outreach	29
11.3 Popularization	29
11.3.1 Specific official responsibilities in science outreach structures	29
11.3.2 Productions (articles, videos, podcasts, serious games, ...)	29
11.3.3 Participation in Live events	30
11.3.4 Others science outreach relevant activities	30
12 Scientific production	30
12.1 Major publications	30
12.2 Publications of the year	31
12.3 Cited publications	32

1 Team members, visitors, external collaborators

Research Scientists

- Hugues Berry [Team leader, INRIA, Senior Researcher, HDR]
- Audrey Denizot [INRIA, Researcher]
- Thomas Guyet [INRIA, Researcher, from Sep 2025, HDR]
- Thomas Guyet [INRIA, Associate Professor Detachement, until Aug 2025, HDR]
- Leonardo Trujillo Lugo [INRIA, Starting Research Position, until Apr 2025]

Post-Doctoral Fellow

- Maelle Moranges [INRIA, Post-Doctoral Fellow]

PhD Students

- Ismail Bachchar [ORANGE]
- Schayma Ben Marzougui-El Garrai [INRIA]
- Andrea Ducos [INRIA]
- Florian Dupeuble [INRIA]
- Arnaud Hubert [INRIA]
- Hana Sebia [INRIA]

Technical Staff

- Arnaud Duvermy [INRIA, Engineer, from Jul 2025]
- Lucas Perret [INRIA, Engineer, until Apr 2025]
- Jan-Michael Rye [INRIA, Engineer]

Interns and Apprentices

- Zoe Koenig [INRIA, Intern, from Feb 2025 until Jun 2025]
- William Peoc'H [INRIA, Intern, from Jun 2025 until Sep 2025]

Administrative Assistant

- Noemie Rodrigues [INRIA]

External Collaborators

- Arnaud Duvermy [AP/HP, until Jun 2025]
- Arthur Skowronek [UNIV LYON I, from Mar 2025 until Aug 2025]
- Leonardo Trujillo Lugo [INSA LYON, from Nov 2025]
- Leonardo Trujillo Lugo [POLE EMPLOI, from May 2025 until Nov 2025]
- Luc Zimmer [UNIV LYON I]

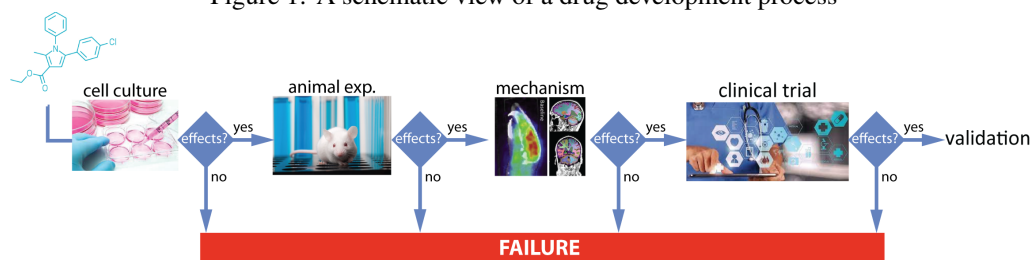
2 Overall objectives

2.1 Failures in drug development for neurological diseases

The set of available drugs for neurological diseases is both aging and lacking in effectiveness. There remains a very high unmet medical need for treatments in neurology, despite heavy historical investment in the field [29, 60]. A typical drug development process includes a set of successive stages (fig. 1), sequentially evaluating the effects of the candidate drug in vitro on cell culture models, then in vivo on animal models (pre-clinical), after which the mechanistic origins of the candidate drug effect can be assessed in animals (sometimes in humans). The next stage consists of studies in humans based on clinical trials that are themselves structured in successive phases: phase I to test for safety, phase II to determine the effect of the candidate on a small set of patients and phase III that includes large cohorts of patients and control people in a randomized setting. Each and every stage of this process has significant probability to fail and break off the development of the drug under process. With the recent public health issues (HIV, Covid-19) the general public has mostly been made aware of failures between the successive phases of the clinical trial stage. However, in the field of neurology, the difficulties in developing drug candidates are mainly due to a high failure rate in the clinic: the activity of the drug candidate in in vitro cell cultures or in animal models is very often not confirmed in humans [44, 48].

In recent years, numerical approaches have been proposed in the field, either with mechanistic modeling to predict the response of the cell to the candidate molecule (quantitative systems biology/pharmacology) [55, 61] or with machine learning to identify the impacted (sub)cellular systems or the effects of the candidate drug [69, 50]. However, these approaches are still inefficient to meet the above challenges because they often address a unique scale or modality of interest (e.g., molecular, cellular, preclinical) and lose their predictive power at other scales (e.g. clinical, i.e. the patient). The main methodological objective of AIstroSight is to develop quantitative systems biology and Artificial Intelligence (AI) approaches able to embrace several of these scales.

Figure 1: A schematic view of a drug development process



2.2 Main deliverables

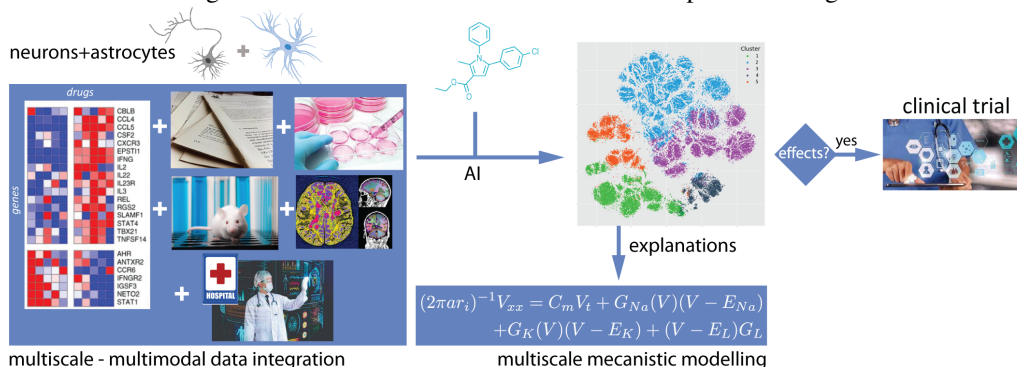
Our overall goal is to develop innovative numerical methods for neuropharmacology that will provide us with levers to accelerate and derisk the early stages of drug design. As a main deliverable and proof of concept of the efficiency of these methods, our ambition for the first four years of the project is to identify a handful (2 to 4) of new candidate drugs against neurological diseases.

2.3 Overview of the AIstroSight roadmap

To improve the probability of success of drug candidates in neurology, we integrate complementary information offered by data harvested at different spatio-temporal scales (fig. 2): from the inside of the cell (molecular and cellular biology) to the whole brain (imaging) and even to a population of patients (hospital data), using numerical tools coupling mechanistic models with dedicated AI approaches. In a way, our strategy is to break down the classical stratification silo of Fig. 1, in which literature search, in vitro cell culture, in vivo preclinic studies and in vivo clinic studies are viewed as a sequential multi-stage process. Instead, we propose an integrated machine learning framework into which all those data are combined to predict the effect of a candidate drug molecule.

AISTroSight develops innovative numerical approaches to integrate these information sources into a coherent stream of data and expert knowledge, combining the analysis of experimental observations with reasoning (of different kinds). Currently, these tasks are carried out in isolation and their reconciliation is devolved to biologists/physicians. The originality of the AISTroSight contributions are approaches that automatically carry out this reconciliation to assist biologists/physicians.

Figure 2: A schematic for the scientific roadmap of AISTroSight



Since AI algorithms are often black-box tools, we also develop mechanistic modeling approaches (multiscale quantitative systems biology/pharmacology) to produce explanations for the predictions of the AI algorithms, that can be rooted in neurobiology. Another important aspect of AISTroSight is to widen the focus of neuropharmacology beyond neurons, that constitute only one half of the nerve cells in the brain, and also take into account the other half, that is made up by glial cells and their interactions with neurons. In particular, we consider the pharmacology of astrocytes [71], one major subtype of glial cells, in interaction with the pharmacology of neurons.

2.4 Principles of the AISTroSight partnership

To accelerate cross-fertilization between digital science and medical research, AISTroSight will be located on the East Hospital Campus of the Lyon University Hospital, the “Hospices Civils de Lyon” (HCL), from 2024. We will also benefit from our strong association with **CERMEP**, the preclinical and clinical in vivo imaging platform of the HCL. In 2024, the whole team is indeed expected to move to Lyon’s neurology hospital, located just across tA joint team with the HCLhe street of CERMEP.

CERMEP is also affiliated with University Claude Bernard Lyon 1, Inserm and CNRS. This provides us with an exceptional environment for the engineering of brain biochemical imaging methods that allow the study of the effect of molecules on the whole brain (fMRI, PET, fUS) and the analysis methodologies for these images. The CERMEP also hosts team **BIORAN** of the “Centre de Recherche en Neurosciences de Lyon” (CRNL) laboratory, that has expertise ranging from the chemistry of candidate-molecules to their biochemical assays, from radiolabelling to animal PET/MRI imaging and from preclinical models to first-in-man studies in patients. The modeling expertise on the binding between candidate molecules and receptors (structural biology, docking) is also present at the CERMEP.

As a joint team with the HCL, part of AISTroSight technology developments is intended to be integrated into the **hospital information system** developed during the last decade by the HCL for patient management. This is in particular the case of the development of “Multi-patient query for care pathway characterization and clinical trials”. Beyond participating in the HCL’s mission as an innovation leader in digital health, AISTroSight also represents an opportunity for the HCL to reinforce its infrastructure for the organization of clinical trials, for instance in cooperation with pharma/biotech companies like Theranexus. Like other teams of Inria Lyon, AISTroSight is intensively implicated in the “AI innovation department” (“Pôle de Développement IA”) that Inria Lyon and HCL are supporting.

Finally, to ensure the impact of our works on pharmacology and provide it with potential industrial exit routes, the AISTroSight partnership also includes an industrial partner, **Theranexus**, a French biopharmaceutical company (an SME) that develops drugs for the treatment of nervous system diseases with an original

focus on both neurons and astrocytes. Theranexus is listed on the Euronext Growth market in Paris and its social headquarters are located in Lyon. A biotech company, the expertise of its members is entirely in the experimental aspects, not in the digital. Theranexus brings to AIstroSight experimental data (experimental cell biology and brain imaging for pharmacology), and provides their expertise for the development of the digital tools needed to analyze these data. In return, the objective is that the output of these digital tools reveal novel drug targets or novel candidate molecules that Theranexus may decide to use to develop new treatments, starting with the necessary clinical trials. Importantly, the fact that these candidate drugs have been selected from an innovative numerical approach strongly consolidates the credibility of their development on the pharmaceuticals market. In addition, Theranexus brings to AIstroSight its know-how and industrial expertise on the development of drug candidates up to the market and its strategic knowledge of the neuropharmacology industry. In this operating scheme, Theranexus is therefore the preferred partner for the early-phase transfer of the molecules that AIstroSight could identify.

Independently from AIstroSight, Theranexus and BIORAN have a longstanding collaboration together, in particular in the framework of an ANR- and AURA Region-funded joint laboratory (LabCom) called « NeuroImaging for Drug Discovery (NI2D) » that aims at the development of gliopharmacology using preclinical imaging tools (PET/MRI/brain ultrasound). This LabCom is also hosted within the CERMEP premises. NI2D aim at developing preclinical neuroimaging techniques (in animals, mainly fMRI, PET and fUS = Functional Ultra-Sound) for the evaluation of drug-candidates. AIstroSight develops numerical methods capable of integrating data at multiple scales for pharmacology, data that include imaging but also molecular data (intracellular signaling, omics data) or clinical data (biology, treatments, medico-administrative). We thus benefits from the imaging data and methodologies of NI2D. The two are therefore complementary, especially since we both share a strong interest in neuron/astrocyte interactions.

3 Research program

3.1 Characterizing the mechanisms of action of candidate drugs

Drug screening, either in vitro or in silico, generally does not provide an explanation of the mechanism by which the identified drug acts at the cellular level. However, this information is crucial (e.g., with respect to patients or health agencies), and the algorithms used for the screening must be made explicable. Our goal here is to use mathematical models and their hybridization with machine learning to provide explanations on the mechanisms of action of a candidate molecule.

We develop mechanistic models of regulatory networks or intracellular signaling pathways specific to the action of the candidate drugs identified by the screening. Those models predict the spatio-temporal evolution of the concentrations of the molecular species involved in the modelled pathways using classical reaction terms from biochemical kinetics and mass-action laws (first order reactions, bi-molecular reactions, Michaelis-Menten, Hill kinetics. . .). Depending on the importance of intracellular spatial gradients and biochemical noise, space and stochasticity is accounted for, thus resulting in models based on reaction-diffusion equations, stochastic or ordinary differential equations, or other related formalisms (Gillespie algorithm, flow analysis). These models allow us to simulate in time and/or space of the cell the mechanisms that govern the dynamics of the implicated molecules and how this dynamic is altered by the selected drug. The aim is to use such mechanistic modeling to produce explanations for the predictions made by the statistical learning techniques that are used in the other sections. It is unlikely that these mechanistic models in themselves allow us to decipher the totality of the molecular mechanisms involved, but they provide critical information to properly adjust the laboratory and clinical experiments.

To be efficient, this approach demands that we maintain an effective knowledge basis and expertise on the fundamental molecular mechanisms at play at these spatial scales and their modelling. To build and maintain this expertise, we rely on existing long-term collaborations between AIstroSight members and experimental neuroscientists -electrophysiologists- or neuropharmacologists on the intracellular signaling networks at play in neuron function or neuro-astrocyte interactions.

- *Agonist bias in GPCR*: G protein-coupled receptors (GPCR) are currently the largest family of molecular targets for potential new drugs [26]. GPCR are cell-membrane receptors ubiquitously found in all mammalian cells, but in particular in brain cells (neurons and astrocytes), where they control a large repertoire of neuronal and astrocytic responses to a variety of external stimuli and molecules. Bias antagonism refers

to the observation that two ligands of the same GPCR can activate very different cell responses [51]. This phenomenon is still not understood, but it is one possible path towards the development of new drug discovery and has been already proposed and stated to be explored, in particular by members of AIstroSight (B. Vidal, L. Zimmer) [72, 59]. Our objective here is to build realistic mechanistic models of GPCR-based cell signaling in the neuronal intracellular space. We plan then to use this/these models to propose molecular mechanisms to explain the experimentally observed biases. A first idea to explore is the hypothesis of a local subcellular compartmentalization of the signaling molecules over and close to the cell membrane (so-called nanodomains). Experimental validation of the main model predictions is then to be performed using brain imaging modalities available at CERMEP (TEP, MRI, fUS).

- *Synaptic plasticity*: Synaptic plasticity, the long-term adaptation of the efficacy of a synapse according to the activity of the neurons and astrocytes composing this synapse, is thought to underly learning and memory at the cellular scale [68]. We have been enjoying a very fruitful collaboration with Laurent Venance's lab (INSERM U1050, CIRB, Collège de France, Paris) on the subcellular mechanisms at play in learning and memory formation by synaptic plasticity [36, 38, 39, 37, 45, 74, 49]. Current work focuses on the control of synaptic plasticity mechanisms by endocannabinoids and its implication in fast learning and on metabolic regulation of synaptic plasticity by astrocytes. This collaboration is funded by ongoing ANR project EngFlea (see below).

- *Calcium signaling in astrocytes*: Calcium signaling in the terminal branchlets of astrocytes is thought to be crucial for astrocytic functions and neuron-astrocyte interactions [28]. We are studying the local dynamics of calcium signaling in terminal branchlets of astrocytes and their interaction with synaptic activity in collaboration with U. Valentin Nägerl's lab (CNRS UMR 5297, Bordeaux) for experimental (subcellular) data with supra-resolution microscopy [40]. Recently, a collaboration with Erik de Schutter's lab (Okinawa Institute of Science and Technology, Japan) has also been set up to develop new efficient modelling tools (stochastic reaction-diffusion systems) in realistic 3D geometric meshes based on the simulation framework they develop, STEPS [27, 41].

- *Multiscale modelling of the effects of a candidate drug on neurons and astrocytes*: To model the cellular effect of a candidate drug, the main molecular systems impacted by the drug are isolated from the cellular signature data (from e.g., transcriptomics) and literature exploration. Imaging data, by specifying the brain areas and structures mainly targeted by the candidate drug, helps refine these models using specific parameters. Whereas in the first models, the observation (and modelling scale) corresponds to a subcellular domain (one synapse, +/- a dendrite or the main astrocytic process in the neighborhood), we search to progressively scale up those mechanistic models from the intracellular scale of a single cell to the scale of a population of interacting brain cells, neurons and astrocytes. To do so, we explore model simplification /reduction methods, including those combining machine learning and dynamical systems modelling (see below). In the long run, this large-scale mathematical model will produce a digital twin of the pathology that will allow us to explain why the candidate drug has a positive effect on the disease. Calibration is based on fUS and fMRI imaging data in rodents obtained in the framework of the NI2D LabCom. This data provides us with quantitative measurements of the effects of microscopic perturbations by pharmacological agents or by external stimuli (e.g., visual) on the variation, correlation and spreading of cortical activity over the whole brain.

- *Astrocyte roles on brain imaging signals*: Although it is now widely accepted that astrocytes play a role in brain processes and pathologies, the exact perimeters of their roles remain to be delimited. For instance, variations of the signals measured by brain imaging methods (fMRI, PET, fUS) are still largely interpreted as variations of neuronal activity. Available experimental data however indicate that astrocytes also impact those signals but it not clear yet how they do it. A precise and quantitative answer to this question would allow us to use brain imaging to monitor not only the local activity of the neurons but also of the astrocytes. Such a feature would be precious in our framework of astrocyte pharmacology but it demands the development of new mathematical models. Existing models of fMRI signals, for instance, are either too crude to incorporate a separate astrocyte action (balloon models [47]) or are limited to the role of astrocytes as energy suppliers of the neurons (astrocyte-neuron lactate shuttle [53]). Our objective here is to start from a microscopic and mechanistic model of neuron-astrocyte-blood vessel interactions and use multi-scale modelling methodologies to obtain a large-scale model of astrocyte-neuron impact on a subset of brain imaging technics (fMRI, fUS), with explicit parametrization of local neuronal and astrocyte activities. Here again these models are calibrated using fUS and fMRI imaging data in rodents, in particular using pharmacological agents that are known to specifically silence the astrocytic population or a neuronal population in a given brain area. A crucial step is the development of a detailed, microscopic model of the

astrocyte endfeet, the specialized astrocyte processes that respond to and control local vascular diameters. This model will provides us with causal mechanisms able to interlink neuronal electrical activity, astrocytic calcium activity and local blood flow. It is to be seen as a first stage towards understanding the implication of astrocytes in variations of neuroimaging signals.

Methodological challenges: The biological systems to be considered to explain drug effects on pathologies are not only very complex but also only partly understood by neurobiologists themselves. Therefore, the available biological knowledge on these systems is constantly evolving. Since we cannot know in advance what systems are affected by the candidate drug, a major difficulty for modelers is preparedness, i.e. maintain a level of expertise on the biology and modelling state-of-the-art of a wide range of those systems. This is the reason why the first three projects above are crucial to the success of our proposal.

The most important challenge we face is that of multiscale: causal data are mainly molecular but many observations are macroscopic (e.g. brain imaging). Traditionally, linking these two scales requires the development of new theories (e.g., homogenization, population boundaries, etc.), a slow and rather hazardous process. The availability of more and more important computing resources allows to consider "brute force" approaches in which all scales of time and space are numerically simulated (cf [Blue Brain Project](#)). But the results are often as difficult to interpret as the animal experiments that these simulations emulate. Instead we consider recent advances in hybrid digital-AI systems (physics-informed neural networks [64]), in particular equation discovery methodologies [63]. These methods usually use sparse regression techniques to select in a library of nonlinear terms and operators, those that, when composed, provide the best description of the data [32]. Our idea is to generate a large number of numerical simulations at the microscopic scale of the kinetics of the biochemical reactions concerned, for example by the spatial Gillespie algorithm, and then to aggregate them at a higher spatio-temporal scale using, for example, averages at a space and time grain much higher than the spatio-temporal resolution of the initial microscopic simulations. The idea is then to use equation-discovery algorithms to infer a set of differential equations (and associated parameters) capable of describing these higher scale space-time kinetics. The resulting reduced model is then replicated in each cell of the cell population model. If successful, this model reduction process can even be reiterated at the upper scale to simulate the effect of the molecule on large brain areas. Of course, the risky and difficult nature of this objective makes it a long-term goal. If need be, alternative meta-modelling technics will also be considered when applicable (RKHS, model-order reduction).

3.2 Multi-patient query for care pathway characterization and clinical trials

Real-world data, especially the data that is routinely collected by hospitals (medical reports, hospital records. . .), provides rich information about possible links between patient information (demographic, pathological, life style), drug exposures and health events. In the context of drug development, this data can be useful at three stages. During the search for a new drug, they can be used to enrich cell culture data or imaging data. In this case, one can query patients that have been treated for the pathology in question and integrate their clinical data in the in-silico screening. This approach is presented below in the framework of data integration.

Electronic Health Records query algorithms:

Efficient patient query can also be used at the very initial stage of drug discovery: the assessment of the feasibility of drug development projects. Indeed, part of the pathologies we target are rare diseases. In this context, one has to make sure at the very early stages that the pathology in question is not so rare that the number of patients is too low to allow clinical trials, or that its description in terms of physiopathology is mature enough for the clinician to be able to diagnose it with good probability. We thus develop patient query algorithms on clinical data from hospitals (electronic health records, EHR), in particular of the HCL, that allow us to characterize the care pathways of the patients before and after diagnosis. They provides us with answers to many questions related to the clinical picture of the pathology, its genetic underpinnings, its prevalence rate, the typical care pathway of a patient with this pathology, the delay of diagnostic, the frequency of diagnostic errors etc. Answers to these questions are crucial to determine early on whether the drug discovery project is feasible. We aim at developping query algorithms and software pipelines for EHR that can provide us with tools able to answer these questions efficiently.

Efficient EHR query algorithms are also very useful at the final stage of the clinical trial itself (Fig. 2), where they can be used to finely select what patients should be integrated in the trial. Indeed, a major change of paradigm in medicine in recent years is the acceptance that the response of a group of patients to

a drug treatment exhibits strong variability. The source of this variability is diverse [73]. The definition of the pathology itself as a unique coherent class can be misleading and actually incorporate a range of different sub-classes of pathologies/disorders. The response to a drug also depends on how the patient's body affects the drug before it reaches its target organ/cells (pharmacokinetics). At the cellular level, the response can also vary because of inter-individual differences in gene sequence and receptor/protein structure (pharmacodynamics). Therefore, individual drug responses depend on the patient genes (pharmacogenetics) but also on more social factors (age, sex, anterior medical record, lifestyle, habits, exposure to pollution. . .). In any case, the strength of this variability is believed to be a major cause of failure for clinical trials, in particular in neurology and psychiatry [60, 54]. The goal of "stratified medicine" in this perspective is to subdivide the available group of patients into a number of subgroups so that the response of each subgroup is less variable than the whole [42]. Our objective is to develop computational tools and software packages able to stratify hospital data to assist in the selection of patients to be included in an evaluation protocol for a clinical trial or the building of a research cohort.

Computational phenotyping:

The task of querying patients according to a predefined criterion from a large population of EHRs is sometimes referred to as "computational phenotyping" [65]. It remains a time-consuming and challenging task with complex criteria because the query is to be addressed within multiple document types and across multiple data points, in EHRs that usually comprise both structured and unstructured data. The computational challenges raised by patient query with complex criteria are therefore considerable (integration, query, analysis, privacy). Software tools (i2b2, ACE [33]) have been proposed to query patients for cohorts or clinical trials based on EHRs but they can hardly be used by most of the physicians because they require advanced knowledge of the data in computer science terms (format, encoding). Moreover, our objective is to provide clinicians with tools able to manipulate these complex data together with medical concepts (e.g., exposure to a drug, treatment, or occurrence of a pathology). Data abstraction capabilities must therefore be integrated to automatically enrich the data using phenotype libraries that can be intuitively mobilized by the clinician. In analogy with bioinformatics workflows, we create workflows for computational phenotyping.

In cases where we already know how to stratify, the issue is not a learning problem but rather a query problem. On the other hand, when this is not the case, we have to develop methods to build these homogeneous subgroups, and in this case it is a question of (unsupervised) learning: the training criterion becomes a measure of cluster homogeneity. Two competing approaches can be thought of in order to create the building blocks of the workflow: 1) machine learning approaches that allow the construction of abstract patient phenotypes from massive data; 2) approaches inspired by both timed systems modeling and knowledge reasoning that rely on formal descriptions of computational phenotypes to enrich the data. The interest of formal descriptions is to be able to represent the whole data transformation in a formal way. This abstract representation of the construction of a cohort facilitates its understanding by users and its reproducibility (FAIR principle). On the other hand, they allow again to exploit intimately the formalized knowledge of the domain, but they also become objects that can be manipulated by reasoning tools. The use of semantic web technologies can therefore be an interesting tool for representing data, knowledge and their processing in order to propose query tools that guide the clinician through the knowledge.

Methodological challenge:

The challenge is to make these formal descriptions highly expressive and to ensure efficient processing of massive data. On the long run, we plan to take inspiration from the approach called "Ontology-Mediated Query Answering" which consists in using ontologies to mediate the query of a database by ontologies [31]. In this context, a computational phenotype is seen as a query. The difficulties encountered with observational data is the semantic gap between the available data and the medical concepts that are interesting to manipulate. This gap may be bridged by automatic reasoning that exploits expert knowledge to relate different abstraction levels.

Since computational phenotypes are difficult to formalize, the challenge is to support clinicians in defining them. In other words, the challenge becomes to abstract phenotypes from clinical data. We plan to combine automatic reasoning methods and data analysis. The first research direction we propose is the exploration of a symbolic approach parallel to the work by Tijn de Bie [30] or by Silberschatz [67] on the notion of "subjective measure of interestingness". This approach was developed to identify user-relevant statistical analysis results by means of a statistical model to evaluate the novelty of the extracted patterns (a priori knowledge model). Symbolic approaches can be combined in a similar way by using symbolic data analysis methods such as pattern mining, and by relying on formal models of the system as a priori knowledge. Patterns that are not

"explainable" by the formal model are potentially new or of particular interest to the user and will thus be extracted. This approach offers an original entry point to deeply integrate knowledge-based reasoning into pattern extraction methods. The research challenge here lies in combining formalized knowledge with experimental data. It may be implemented using the declarative pattern mining paradigm, that uses solvers to address the pattern mining task. The proofs of concept on the notion of novelty will open the way to more complex reasoning such as planning that can be used to integrate complex behaviors in biological systems, such as interaction networks. The second research direction we propose is based on recent machine learning techniques. Unsupervised ML has been applied to patient phenotyping, i.e. the discovery of phenotypes from EHR data, including temporal phenotyping [75]. Our objective is to combine such kinds of algorithm with semantic knowledge to guide the discovery toward meaningful computational phenotypes. Indeed data embedding techniques can integrate ontologies to enhance data semantics [46]. On the long run, the methods developed above may be reunified to address the problem of drug discovery at both the biological and the body scale. This justifies the coherence of the methodological approaches (Semantic Web and machine learning) that are developed in the two objectives.

4 Application domains

4.1 Targeted Pathologies

The list of pathologies that are of interest for Theranexus in the framework of AIstroSight is given in the stand-alone "convention d'équipe-projet commune" of AIstroSight. It comprises roughly 30 rare diseases of the central nervous system, including lysosomal pathologies, neurological genetic diseases, rare diseases due to α -synuclein accumulation or rare demyelinating pathologies. This list may be updated, subject to the prior joint written consent of AIstroSight partners. For the pathologies in this list, a specific regimen is defined in terms of IP and legal affairs. AIstroSight members are allowed to work on pathologies outside this list without any restriction but with a different legal regimen vis-à-vis Theranexus. In agreement with the company, we have selected two pathologies from this list as priority objectives, on which we will start our work: Rett syndrome and Niemann-Pick Type C disease. Both pathologies are neurodevelopmental diseases caused by mutations in a single gene: mutations in methyl-CpG-binding protein 2 (MeCP2) for Rett [34] and in NPC for Niemann-Pick C [62]. MeCP2 mutation in Rett syndrome causes slowed brain growth, a progressive loss of movement, motor control abilities and language in the children and can also cause heavy breathing problems, epileptic-like seizures or intellectual disabilities, among others [52]. NPC mutations in Niemann-Pick type C causes accumulation of cholesterol and other fatty acids inside lysosomes, including in brain cells. The symptoms are highly variable, ranging from defects of developmental and motor progression, difficulties in learning, speech or swallowing, to cognitive impairment or psychiatric symptoms [70]. Restoring NPC expression in astrocytes significantly increases survival of mice models of the disease, suggesting that astrocyte dysfunction is involved in disease progression [76]. The two diseases also have in common that they are rare neurological diseases of children (frequency $1/10^4$ and $1/10^5$, respectively) and that there is no known effective treatment. This rarity has strong consequences on the numerical tools that can be used to find potential pharmacological targets.

4.2 In vitro and in vivo experimental models of neurology diseases

Many of the diseases that are of interest for AIstroSight are rare diseases. This means that the volume of experimental data and the basic understanding of the pathology at the (sub-)cellular level may be too limited for the machine learning or mechanistic modeling tools that we plan to use. For example, it is known that the NPC mutation in Niemann-Pick type C induces morbid cholesterol accumulation in cells but the molecular function of NPC in cholesterol metabolism is not clearly understood [62]. Similarly, MeCP2, the gene mutated in Rett syndrome, is an epigenetic regulatory factor (DNA methylation) whose mutation theoretically impacts the expression of a large number of genes but it is not clear which ones are most involved in the symptoms of the disease [52]. Although molecular (omic) studies have been published for both diseases [35, 66], their molecular contexts are still unclear.

Our goal here is to generate additional preclinical molecular and imaging data to better delineate the perturbations that these diseases cause at a cellular and tissue level. We introduce into cultured cells the

same deficits as those observed in patients. Transcriptomic analysis of the effect of this manipulation gives us information on the implicated molecular networks and its major molecular consequences. In parallel, we induce these same perturbations *in vivo* in rodents. Observing these animals using brain imaging techniques (fMRI and fUS, possibly PET) gives us a more macroscopic view of the effect of the mutation (affected brain areas, nature and amplitude of the modifications, change in response to treatments or stimuli etc, see below). **Methodological challenges:** Developing experimental models of pathologies can be a very difficult task for pathologies that are due to the conjunction of multiple factors, when the molecular alterations at the origin of the pathologies have effects over a very large range of cellular processes or when comparison of the phenotype of the experimental model with its human counterpart is ill-defined (psychiatric diseases, for instance). To mitigate this risk, we develop experimental models only for pathologies that are well-defined in molecular terms, like for Rett or Niemann-Pick type C for a start. We use viral vector strategies (mostly shRNA-mediated gene silencing or possibly CRISPR-based gene editing via adeno-associated viruses, AAV) to manipulate the sequence or expression of the target gene. We start with cell lines that are easy to grow and analyze using omics approaches, and then use neurons and astrocytes differentiated from human pluripotent stem cells. This approach is also used *in vivo* by locally injecting the viral vector into a given brain region of an animal model, to genetically modify a particular cell type by using a specific promoter. We should therefore be able to control the area of the brain in which the genetic manipulation will be induced (e.g. visual cortex or cerebellum) as well as the type of cells targeted (neurons vs. astrocytes, for example). Of course, like all experimental models, each model taken separately has its limitations: the genes expressed by cells in culture are not necessarily those expressed by these same cells *in vivo*, the effects of gene silencing in a rodent are not necessarily transposable to humans, etc. However our hypothesis is that by combining these different modalities and scales of data (see above), it should be possible to better predict the effect of a potential treatment. The molecular and cellular biology technologies to be mobilized here (*in vitro* and *in vivo* mutagenesis, cell culture, proteomics) are tools routinely used by Theranexus. The expertise on the use of medical imaging to observe the effects at the brain level is provided by CERMEP and benefits from the advances of the NI2D LabCom.

4.3 Identification of multi-source multi-scale biomarkers

Recently, Theranexus changed its pharmacological strategy, from a strategy mainly based on the repositioning of pre-existing drugs to a technology based on antisense oligonucleotide drugs. These technologies rely on the ability to design on demand short RNA sequences that specifically bind the mRNA of a gene target, and knock it down after recognition by the RNase H1 enzymes present in all cells, or modulate its translation or splicing via steric hindrance. Pharmacological intervention thus consists in searching for a gene target able to correct the molecular perturbation caused by the disease and to synthesize an antisense oligonucleotide able to specifically bind this target gene. Note that the technology currently in clinical use does not (yet) provide ways to specifically target a cell type or a brain region.

Our first objective is to develop digital tools to model the molecular networks perturbed by the pathology of interest, and use this model to identify a gene or protein in the network the modulation of which would correct the perturbation caused by the pathology. These models are based on molecular data, in particular transcriptomics and metabolomics data. The set of data includes data derived from cell cultures as described above, that we augment with molecular data from the literature related to the pathology or more generic public, open access databases of transcriptomic responses to perturbing molecules, like **CMap** or the **LINCS L1000 data repository**. The latter, for instance currently includes the effect of close to 40,000 small perturbing molecules on 12,000+ genes of **more than 200 cell types**. We aggregate these data and use them to infer the gene interaction network, the metabolic network and/or the signaling network impacted by the pathology. Metabolic networks are important for instance for Niemann-Pick type C, to conciliate perturbations of the lipid metabolism with those of the gene expression network. This provides us with a view of the pathology at the molecular scale.

Integration of neuroimaging data:

A major objective of AIstroSight is to augment these molecular data with medical data, in particular brain imaging data and hospital data. We complement molecular data with data coming from the analysis of brain imaging (fMRI, PET, functional ultrasound brain imaging) i.e. with functional networks between brain areas targeted by the molecule or quantitative measures of radioligand binding. Most of this imaging is done in rodents (preclinic, see above) but a subset of human imaging data is also used. These imaging data are

obtained by our collaborators from the CERMEP platform.

These different neuroimaging methods provide meaningful and complementary information for understanding the functional or molecular effects of drugs in the brain:

- Positron emission tomography (PET) enables to visualize and measure the concentration of a specific radiotracer, with nanomolar to picomolar sensitivity. Numerous brain PET radiotracers have been developed over the years, enabling to study various molecular processes (such as the synthesis and release of endogenous neurotransmitters, the density of receptors, transporters or proteins aggregates, neuroinflammation and cerebral metabolism) both in animals and humans in a non-invasive way. We especially focus on the measurement of cerebral glucose consumption using [¹⁸F]FDG, a radiolabeled glucose analog, while also taking advantage of the rich information that can be obtained using other brain radiotracers when needed. Indeed, the asset of [¹⁸F]FDG-PET imaging is to be relevant for virtually all drugs that are expected to be active in the brain, since the cerebral glucose uptake is related to neuron and astrocyte activities. In addition, this neuroimaging technique can be used in awake freely-moving animals, getting rid of the anesthesia or stress confound in preclinical studies.
- Functional magnetic resonance imaging (fMRI) enables to follow the dynamic changes in the BOLD signal (for “blood-oxygen level dependent”), which is also related to the brain activity, in a non-invasive way. It has a high temporal resolution (2 to 3 seconds) as compared to PET imaging (several minutes) and therefore can be used either to measure the time series of BOLD signal changes after injection of a drug (called “pharmaco-MRI”) or the changes in functional connectivity occurring after neuropharmacological stimulation. Functional connectivity is defined as the correlation in activity over time between different brain areas; this concept has largely become prominent in neurosciences over the years, and it is known to be modified in many physiological or pathological states. fMRI can be used in animals and humans, similarly to PET, but usually requires anesthesia for preclinical applications.
- Functional ultrasound imaging (fUS) is a much more recent imaging technology. It provides access to the dynamic measurement of cerebral blood volumes changes, which are more straightforwardly related to the brain activity as compared to the BOLD signal. Moreover, its spatial, temporal resolution and sensitivity are unmatched by the previous techniques, and it can be applied to freely-moving awake animals in real-time, in correlation with a particular behavior. However, it is currently limited to imaging in 2-dimensions (one brain plane at a time) and is mostly suitable for animals. Therefore, fUS imaging is a complementary way to study brain activity in small animals in the context of preclinical neuropharmacology.

Integration of clinical data:

We also plan to integrate hospital data from the Hospices Civils de Lyon according to availability and pathologies. Hospital data provide access to rich information on possible links between patient information (demographic, pathological), drug exposures, health events or biological sample analysis (e.g., blood markers). Our goal is to integrate brain imaging and hospital data with cellular signatures to enrich them with information at the individual scale in a form that can be analyzed with machine learning (clustering, classification) or data mining (pattern matching) methods.

Methodological challenges:

A first challenge resides in the nature of action of antisense oligonucleotides, that often work by knock down/loss of function. It is not straightforward to design such a strategy in the case of a pathology that is due to a mutation that already suppressed the effect of a gene. That is precisely where numerical models of the involved gene expression and metabolic networks are important because they can be systematically assessed for the effect of gene suppression, thus providing a quick *in silico* screening of the potential targets. However, part of this program implies typical bioinformatics processing steps: analysis of transcriptomic networks, network reconstruction, conciliation between transcriptomic and metabolic networks. . . We currently do not have this expertise in the team. Therefore we leverage collaborations with local experts of the field to get the necessary operational knowledge, including experts of brain transcriptomics analysis (MeLiS lab in Lyon). Another difficulty lies in the heterogeneity of these multiscale data, their highly categorical character, the

large dimension of the corresponding variable space and often, the small number of observations. Moreover, cellular signature data are intrinsically very noisy and can have low reproducibility [56], a caveat that feature selection may improve, at least in part [43]. Class imbalance can also be strong. Finally, each type of available observation (molecular networks, imaging, hospital) gives a partial, fragmented and incomplete view of an abstract complex biological system. This is a partial view because each type of observation provides data at a given spatio-temporal scale, for a certain locus. This is a fragmented view because the data will be collected from different patients, and even from very different living systems (cell cultures, animals, patients). Each patient contributes to the description of the abstract system on only few types of observations. This is incomplete because there will be many gaps to bridge the different kinds of information related to functioning of the studied biological system.

To reach our objective, we explore the use of Semantic Web (SW) formalism which attracts a lot of interest in bioinformatics, to formalize knowledge and data. Data are observations of biological systems acquired within controlled experiments or in real life. Formalized knowledge is a representation of facts and rules acquired in a scientific domain, here medicine or life sciences. Applying machine learning techniques on data supports knowledge discovery, but it is only one particular source of knowledge. The methodological challenge is first to formalize the different types of available data within an abstract model of the biological system, and to integrate formalized knowledge in the model coming from medical literature and our medical expertise, including imaging or hospital data. By gathering a wide range of formalized data and knowledge within the same tool, we aim at creating a kind of abstract numerical twin that may be queried to infer new knowledge to assist drug design or drug repositioning.

On the longer run, the second challenge is to develop query answering at the abstract level but based on fragmented data. The objective is to answer queries about the numerical twin by exploiting the data coming from multiple patients. One of the difficulties is to detect groups of patients whose numerical twins are “similar to each other” (in a sense that remains to be defined). Semantic Web offers a natural framework for querying formalized data with multiple facets but may be limited by the time-efficiency of the query engines on a large number of patients. In such a context, numerical approaches (embedding) is more time-efficient but may lack accuracy. The challenge is to construct numerical representations in order to embed the data in a space in which the distances are both efficient to compute and semantically consistent with the applied notion of “similarity”. Numerical machine learning techniques turn out to be an interesting perspective to address this challenge [58]. Recent research on advanced machine learning, such as representation learning, offers new perspectives to address our challenge. Our objective is to initiate collaborations with teams having strong backgrounds in machine learning (e.g. Ockham Inria Team) to propose innovative solutions. Another important point is the need for logic programming methodologies able to express complex queries, especially on heterogeneous or multimodal data. For neuroimaging, the availability of **NeuroLang** for logic programming with heterogeneous data or **NeuroQuery** for query result consolidation based on automatic literature meta-analysis, for instance, should be very useful.

In most of the cases, the methodologies that we use to reach the above objectives are related to knowledge management/mining, formal reasoning, data mining or learning. Machine learning or deep learning approaches are probably less useful here. The main reason is related to the volume of available data. For rare disease like Niemann-Pick type C, for instance, the low prevalence means that 5 to 10 new patients are diagnosed in France each year, a number too low for deep neural networks. However, advances in transfer learning might be helpful here. For instance, a large number of brain pathologies come with dysfunction of intracellular cholesterol metabolism and storage. This is for instance the case of multiple sclerosis [57], for which large cohorts and databases are available worldwide. As a long term project, an interesting idea will be to leverage the large volume of data on multiple sclerosis to identify biomarkers of cholesterol dysfunction, e.g., in neuroimaging, and use transfer learning to adapt the network to Niemann-Pick type C patients.

5 Social and environmental responsibility

5.1 Footprint of research activities

The team aims to have as low a footprint as possible. For instance, we try to maximize the lifespan of our computing equipments. Travel is done by train as much as possible. We regret that the tools currently available are not designed to facilitate this type of choices when traveling abroad, which results in a significant

additional workload falling on researchers and research team assistants. We believe there is significant room for improvement in strengthening the social and ecological responsibility of Inria’s research teams and its evaluation, and we hope that work will be done to address this gap.

5.2 Impact of research results

In a long term effort to move the impact of our work closer to clinical applications, the team decided to relocate out of the premises of Inria’s Lyon research center. The whole team is located since March 2025 in new offices in the basement of Lyon’s hospital for neurology, in the Lyon Est - Bron medical campus.

This move has already started to considerably amplify our interface with the local ecosystem of academic and clinical research in neuroscience. Actually, the move allows us to interact more not only with the clinical teams of the neurology hospital, it also considerably improved our visibility towards all the clinical research teams of the Hospices Civils de Lyon, Lyon’s University hospital. Further, it also very significantly improved our engagement within the whole Lyon’s research ecosystem, since all the local neuroscience research units (CRNL, SBRI, INMG,ISC-JM) are located within the Lyon Est / Bron campus, a few steps away from our new premise.

The impact of this move will of course become fully visible in our production after some delay, since research takes time. But it has already been obvious in 2025 via e.g., the obtention of new grants with local neuroscience research units (SBRI, ISC, MeLIS) and clinical research teams (CICLY).

6 Highlights of the year

The major highlight of 2025 has obviously been the relocation of AIstroSight within the building of Lyon’s Neurology Hospital. This is a major transformative change for us, of which we are only starting to see the effects (see section 5.2).

Another important highlight has been the appointment of Hugues Berry as director of Inserm’s Office of AI and Digital Sciences in March 2025, since this new responsibility within Inserm’s headquarters now accounts for 60% of his time (under a “Mise à disposition temporaire”).

7 Latest software developments, platforms, open data

7.1 Latest software developments

7.1.1 TanaT

Name: Temporal analysis of Trajectories

Keywords: Temporal data, Data management, Temporal clustering

Scientific Description: TanaT is an open-source framework for temporal sequences analysis. Temporal sequences are made of complex events, described by qualitative and quantitative features, with a contiguous temporal footprint. Such kind of data are encountered in a wide range of applications (medicine, social science, traces analysis, education, etc.) and their analysis requires taking into account the longitudinality of the data. The proposed framework aims to empower data analysts with a coherent toolbox for handling such temporal sequences at all stages of the analysis process: data loading, data pre-processing and transformation, data analysis and data visualization. In this article, we introduce our framework, focusing on the distance-based clustering of temporal sequences. The complex nature of temporal entities requires versatility in defining distances between sequences and we highlight how TanaT addresses this challenge.

Functional Description: TanaT is an extensible Python library for temporal sequence analysis. It originates with a focus on patient care pathways but can be used in a wider range of applications (learning analytics, demographic studies, etc.).

This library aims to provide data wrangling and data analysis facilities, an equivalent to pandas, for temporal sequences data. It gathers a collection of tools related to the analysis of timed sequences (or trajectories).

The originality of TanaT is to support multi-sequence trajectories that can combine three types of temporal data: states, intervals, and events. This allows tracking subjects (patients, users, customers, etc.) over time across multiple dimensions, providing a richer and more complete view of their temporal evolution. The library provides a comprehensive toolkit for multidimensional temporal pattern discovery and analysis (sequence clustering).

URL: <https://tanat.gitlabpages.inria.fr/core/tanat/>

Publication: [hal-05336370](https://hal.archives-ouvertes.fr/hal-05336370)

Contact: Thomas Guyet

7.1.2 Soft-ECM

Name: Soft version of Evidential C-Means

Keywords: Time Series, Clustering, Fuzzing

Functional Description: This software is an implementation of the soft-ECM clustering algorithm. It allows the user to cluster datasets with the choice of its dissimilarity measure, that can be metric or not. A first-citizen application of Soft-ECM is the clustering of time series data, with non-metric dissimilarity measures (e.g. DTW).

URL: <https://gitlab.inria.fr/tguyet/ecm-ts>

Publications: [hal-04929629](https://hal.archives-ouvertes.fr/hal-04929629), [hal-05162452](https://hal.archives-ouvertes.fr/hal-05162452)

Contact: Thomas Guyet

Partner: Université Clermont Auvergne

7.1.3 Synacomp

Keywords: Python, Spike sorting, Library

Functional Description: A Python library and executable for analyzing spike sorting data with different models. The library provides an extensible framework for creating custom user-designed models that can then be integrated into a common processing pipeline. The main pipeline uses the Hydronaut framework to enable easy configuration via YAML files and systematic tracking of results with MLflow.

URL: <https://gitlab.inria.fr/aistrosight/synacomp>

Contact: Arnaud Hubert

7.1.4 MotifLocatorAIstroSight

Keywords: Python, SiRNA

Functional Description: A Python library and command-line tool that searches publicly available databases of genetic and proteic data for targets that meet a number of different user-specified criteria compatible with Theranexus's therapeutic technology.

Contact: Jan-Michael Rye

Participant: an anonymous participant

7.1.5 Pubmed Target Finder

Keywords: Python, Medical applications, NLP

Functional Description: A Python library and web service that uses an LLM to parse medical abstracts and a separate model to identify potential therapeutic targets for THX Pharma in the parsed data. The predictor is trained on carefully curated data provided by Theranexus.

Contact: Jan-Michael Rye

Participant: 3 anonymous participants

8 New results

8.1 Investigating the ultrastructural properties of the endoplasmic reticulum in perisynaptic astrocytic processes and its impact on signaling

Participants: Audrey Denizot.

Astrocytes recently emerged as key regulators of information processing in the brain. Calcium signals in perisynaptic astrocytic processes (PAPs) notably allow astrocytes to fine-tune neurotransmission at tripartite synapses. As most PAPs are below the diffraction limit, their content in calcium stores and the contribution of the latter to astrocytic calcium activity is unclear.

In [2], we reconstruct hippocampal tripartite synapses in 3D from a high resolution electron microscopy (EM) dataset and find that 75% of PAPs contain some endoplasmic reticulum (ER), a major astrocytic calcium store. The ER in PAPs displays strikingly diverse shapes and intracellular spatial distributions. To investigate the causal relationship between each of these geometrical properties and the spatio-temporal characteristics of calcium signals, we implemented an algorithm that generates 3D PAP meshes by altering the distribution of the ER independently from ER and cell shape. Reaction-diffusion simulations in these meshes reveal that astrocyte activity is governed by a complex interplay between the location of calcium channels, ER surface-volume ratio and spatial distribution. In particular, our results suggest that ER-PM contact sites can act as local signal amplifiers if equipped with IP3R clusters but attenuate PAP calcium activity in the absence of clustering. This study sheds new light on the ultrastructural basis of the diverse astrocytic calcium microdomain signals and on the mechanisms that regulate neuron-astrocyte signal transmission at tripartite synapses.

8.2 Evaluating PDE discovery methods for multiscale modeling of biological signals

Participants: Andréa Ducos, Audrey Denizot, Thomas Guyet.

Biological systems are non-linear, include unobserved variables and the physical principles that govern their dynamics are partly unknown. This makes the characterization of their behavior very challenging. Notably, their activity occurs on multiple interdependent spatial and temporal scales that require linking mechanisms across scales. To address the challenge of bridging gaps between scales, we leverage partial differential equations (PDE) discovery. PDE discovery suggests meso-scale dynamics characteristics from micro-scale data.

In [12], we present our framework combining particle-based simulations and PDE discovery and conduct preliminary experiments to assess equation discovery in controlled settings. We evaluate five state-of-the-art PDE discovery methods on particle-based simulations of calcium diffusion in astrocytes. The performances of the methods are evaluated on both the form of the discovered equation and the forecasted temporal variations of calcium concentration. Our results show that several methods accurately recover the diffusion term, highlighting the potential of PDE discovery for capturing macroscopic dynamics in biological systems from microscopic data.

8.3 Enhancing Fluorescence Correlation Spectroscopy with machine learning to infer anomalous molecular motion

Participants: Nathan Quiblier, Jan-Michael Rye, Hugues Berry.

The random motion of molecules in living cells has consistently been reported to deviate from standard Brownian motion, a behavior coined as “anomalous diffusion”. To study this phenomenon in living cells, Fluorescence Correlation Spectroscopy (FCS) and Single-Particle Tracking (SPT) are the two main methods of reference. In opposition to SPT, FCS with its classical analysis methodology cannot consider models of motion for which no analytical expression of the auto-correlation function is known. This excludes for instance anomalous Continuous-Time Random Walks (CTRW) and Random Walk on fractal (RWf). Moreover, the whole acquisition sequence of the classical FCS methodology takes several tens of minutes.

In [9], we proposed a new analysis approach that frees FCS of these limitations. Our approach associates each individual FCS recording with a vector of features based on an estimator of the auto-correlation function and uses machine learning to infer the underlying model of motion and to estimate the values of the motion parameters. Using simulated recordings, we show that this approach endows FCS with the capacity to distinguish between a range of standard and anomalous random motions, including CTRW and RWf. Our approach exhibits performances comparable to the best-in-class state-of-the-art algorithms for SPT and can be used with a range of FCS setup parameters. Since it can be applied on individual recordings of short duration, we show that with our method, FCS can be used to monitor rapid changes of the motion parameters. Finally, we apply our method on experimental FCS recordings of calibrated fluorescent beads in increasing concentrations of glycerol in water. Our results accurately predict that the beads follow Brownian motion with a diffusion coefficient and anomalous exponent which agree with classical predictions from Stokes-Einstein law even at large glycerol concentrations. Taken together, our approach significantly augments the analysis power of FCS to capacities that are similar to state-of-the-art SPT approaches.

This work was carried out in the framework of the ANR project ABC4M (see 10.3).

8.4 Single-cell multi omics data integration for gene regulatory network inference

Participants: Thibaut Peyric, Thomas Guyet.

In [14], we presented a novel three-state gene expression model designed to elucidate the underlying mechanisms of mRNA transcription and its regulation. This is a collaborative work between three Inria’s team (Biotic, Music and AIstroSight) which has been awarded by the best paper of the CMSB conference.

Our model incorporates gene regulatory processes by explicitly including a transcription factor-bound state, thereby capturing the dynamic interplay between transcription activation and chromatin dynamics. We fit the model to paired single-cell ATAC-seq and single-cell RNA-seq data, as these data give us simultaneous information on a gene’s transcriptional state and its accompanying chromatin state. Working at the pseudo-bulk level, we extract biologically meaningful high-level descriptors from homogeneous cell (sub)populations, such as the mean and variance of gene expression as well as the fraction of accessible chromatin. Crucial to the computational feasibility of our approach, these descriptors can be analytically related to our model parameters.

The model parameters reveal a small number of distinct expression strategies among gene clusters, providing data-driven novel insight into context-dependent regulation of gene expression.

8.5 Vascular Segmentation of Functional Ultrasound Images using Deep Learning

Participants: Hana Sebia, Thomas Guyet, Hugues Berry.

fUS is a non invasive imaging method that measures changes in cerebral blood volume (CBV) with high spatio-temporal resolution. It is a recent technique used in premedical studies for instance to analyse the effects of drugs on brain activity. In such purpose, it is then important to distinguish arterial flow from venous flow. However, distinguishing arterioles from venules in fUS is challenging due to opposing blood flow directions within the same pixel.

In [11, 24], we introduce the first deep learning-based application for fUS image segmentation, capable of differentiating signals based on vertical flow direction (upward vs. downward), leveraging annotations derived from Ultrasound Localization Microscopy (ULM) paired images, and enabling dynamic CBV quantification. In the cortical vasculature, this distinction in flow direction provides a proxy for differentiating arteries from veins.

We evaluate various UNet architectures on fUS images of rat brains, achieving competitive segmentation performance, with 90% accuracy, a 71% F1 score, and an IoU of 0.59, using only 100 temporal frames from a fUS stack. These results are comparable to those from tubular structure segmentation in other imaging modalities. Additionally, models trained on resting-state data generalize well to images captured during visual stimulation, highlighting robustness. Although it does not reach the full granularity of ULM, the proposed method provides a practical, non-invasive and cost-effective solution for inferring flow direction—particularly valuable in scenarios where ULM is not available or feasible. Our pipeline shows high linear correlation coefficients between signals from predicted and actual compartments, showcasing its ability to accurately capture blood flow dynamics.

8.6 Methods and frameworks for the analysis of care pathways

Participants: Arnaud Duvermy, Thomas Guyet.

The analysis of care trajectories is essential to assist epidemiologists in their study of the impact of new cares or to stratify patients cohort according to target specific publications for health studies. More specifically, we develop methods and frameworks to cluster patients according to their care trajectoires (represented as sequences of events and as time series).

This year, we had several contributions to this research line:

- The major result of the year is the first release of a new platform, named TanaT (see Software 7.1.1). TanaT is a Python library designed for advanced temporal sequence analysis, with specialized focus on patient care pathways and complex temporal data structures (trajectories). Its purpose is to gather the methodological developments of the team around the analysis of care trajectories and to provide them to the international public health community (and more broadly to all team working on timed sequences). A presentation of this new framework to the national [22] and international data science communities [13] initiated the broadcasting of the software in order to make it used by close collaborators and external teams in a short term.
- The difficulty in clustering care pathways, represented by sequences of timestamped events, lies in defining a semantically appropriate dissimilarity and clustering algorithms. In [8], we adapt two methods developed for time series to the clustering of timed sequences: the drop-DTW metric and the DBA approach for the construction of averaged time sequences. These methods are applied in clustering algorithms to propose original and sound clustering algorithms for timed sequences.
- An alternative for the same challenge as been developed in collaboration with the University of Clermont Ferrand to propose an evidential approach of clustering. In [15, 16], we reformulate the Evidential C-Means (ECM) problem for clustering complex data. We propose a new algorithm, Soft-ECM, which consistently positions the centroids of imprecise clusters requiring only a semi-metric. Our experiments show that Soft-ECM present results comparable to conventional fuzzy clustering approaches on numerical data, and we demonstrate its ability to handle mixed data and its benefits when combining fuzzy clustering with semi-metrics such as DTW for time series data. Soft-ECM is also a software that is available for research (see 7.1.2).

A part of this work is conducted within the SafePaw project.

8.7 Automatic analysis of negation cues and scopes for medical texts in French using language models

Participants: Salim Sadoune, Thomas Guyet, Hugues Berry.

Correct automatic analysis of a medical report requires the identification of negations and their scopes. Since most of available training data comes from medical texts in English, it usually takes additional work to apply to non-English languages. In [10], we introduced a supervised learning method for automatically identifying and determining the scopes and negation cues in French medical reports using language models based on BERT. Using a new private corpus of French-language chest CT scan reports with consistent annotation based on clinical data provided by the Radiology Department of Lyon University Hospital (Hospices Civils de Lyon, HCL), we first fine-tuned five available transformer models on the negation cue and scope identification task. Subsequently, we extended the methodology by modifying the optimal model to encompass a wider range of clinical notes and reports (not limited to radiology reports) and more heterogeneous annotations. Lastly, we tested the generated model on its initial mask-filling task to ensure there is no catastrophic forgetting. On a corpus of thoracic CT scan reports annotated by four annotators within our team, our method reaches a F1-score of 99.4% for cue detection and 94.5% for scope detection, thus equaling or improving state-of-the-art performance. On more generic biomedical reports, annotated with more heterogeneous rules, the quality of the automatic analysis of course decreases, but our best-of-the-class model still delivers very good performance, with F1-scores of 98.2% (cue detection), and 90.9% (scope detection). Moreover, we show that fine-tuning the original model for the negation identification task preserves or even improves its performance on its initial fill-mask task, depending on the lemmatization. Considering the performance of our fine-tuned model for the detection of negation cues and scopes in medical reports in French and its robustness with respect to the diversity of the annotation rules and the type of biomedical data, we conclude that it is suited for use in a real-life clinical context.

9 Bilateral contracts and grants with industry

9.1 Bilateral contracts with industry

Participants: Hugues Berry, Audrey Denizot, Thomas Guyet, Jan-Michael Rye, Benjamin Vidal, Luc Zimmer, Maelle Morange, William Peoc'H, Andrea Ducos, Hana Sebia, Zoe Laffitte, Florian Dupeuble, Lucas Perret.

AIstroSight is a joint project-team with the biotech company [Theranexus](#). A plain “tutelle” of the team, Theranexus brings its research expertise in in vitro cell culture, disease modelling and imaging, both in terms of research workforce and data. The stand-alone “convention d’équipe-projet commune” of AIstroSight lists a group of 30 rare diseases of the central nervous system, that are of direct interest to Theranexus and that are associated with a specific regimen in terms of IP and legal affairs. However, AIstroSight members are allowed to work on pathologies outside this list without any restriction but with a different legal regimen vis-à-vis Theranexus.

10 Partnerships and cooperations

10.1 International initiatives

10.1.1 Participation in other International Programs

AD endfeet

Participants: Audrey Denizot.

Title: Integrating nanostructure, proteomics, and computational modeling to understand the pathological signatures of the astrocytic endfoot in Alzheimer’s disease

Partner Institution(s): • McGill University, Canada (Pr. K. Murai)
• University of Edinburgh, Scotland (Dr. B. Diaz-Castro)

Date/Duration: 10/2025-2026 ; 1 year

Additional info/keywords: This project will resolve structural-molecular-functional relationships regarding remodeling of the endfoot in Alzheimer’s Disease (AD), a goal only possible through interdisciplinary work. This project will facilitate collaborations through focused experimentation, training, and knowledge exchange and allow creation of new open access datasets and models. Preliminary discoveries will be bolstered on how structural and proteomics data relate to (dys)function of endfeet caused by AD-related pathology

10.2 International research visitors

10.2.1 Visits to international teams

Research stays abroad

Hana Sebia

Visited institution: University College London, Center for medical image computing

Country: UK, London

Dates: 11/04/2025 to 11/06/2025

Context of the visit: Collaboration with Daniel Alexander on image generation (super-resolution of functional ultrasound). In the framework of the new results on fUS, see section 8.5.

Mobility program/type of mobility: research stay funded by a grant of the Inria London programme.

10.3 National initiatives

ABC4M

Participants: Nathan Quiblier, Hugues Berry.

Title: Approximate Bayesian computation-driven multimodal microscopy to explore the nuclear mobility of transcription factor

Partner Institution(s): • Inria, Lyon (supervision)
• Institut Langevin, ESPCI, Paris
• Phlam laboratory, Lille

Date/Duration: 2020-2025

Additional info/keywords: Funded by the French National Agency for Research (ANR), Call “AAP2020” (grant ANR-20-CE45-0023-01). We combine computer simulations and Approximate Bayesian computation with simultaneous multiple microscopy methods (FCS and spt-PALM) to quantify the way transcription factors explore the nucleus to find their binding sites.

EngFlea

Participants: Arnaud Hubert, Leonardo Trujillo Lugo, Hugues Berry.

Title: Engram of fast learning in the striatum

Partner Institution(s):

- CIRB, Collège de France, Paris (supervision)
- Inria, Lyon (supervision)

Date/Duration: 2022-2026

Additional info/keywords: Funded by the French National Agency for Research (ANR), Call "AAP2021" (grant ANR-21-CE16-0022-02). We study the link between endocannabinoid-mediated synaptic plasticity and fast learning of rodents thanks to a multidisciplinary approach combining in vitro and in vivo experimental neurophysiology with detailed subcellular biophysical models and large-scale neural network models.

SecNet

Participants: Schayma Ben Marzougui-El Garrai, Hugues Berry, Audrey Denizot.

Title: Spatio-temporal dynamics of second messenger networks

Partner Institution(s):

- Institut de la Vision, Paris (supervision)
- Inria, Lyon

Date/Duration: 2023-2026

Additional info/keywords: Funded by the French National Agency for Research (ANR), Call "AAP2022" (grant 2023-ANR-22-CE16-0034-02). We combine cell biology approaches and mathematical modeling to provide a description of compartmentalized networks of second messengers that specifically regulate axon guidance and cell migration in response to repellent molecules.

SAFEpaw

Participants: Thomas Guyet, Francois-Elie Calvier.

Title: SAFEpaw (PEPR Santé Numérique)

Partner Institution(s):

- CNRS, Paris
- Université de Tours
- Ecole Normale Supérieure Paris-Saclay
- Aix-Marseille Université
- Ecole des Hautes Etudes En Santé Publique
- CHU Grenoble - Université Grenoble Alpes
- CHU Bordeaux - Université de Bordeaux
- Mines Saint-étienne
- Inria, Lyon

Date/Duration: 2023-2027

Additional info/keywords: The SAFEPaw project is a multidisciplinary project to question the improvement or optimization of care organization by distinguishing three points of view: Regulators / Patients / Healthcare professionals that includes doctors, health care institutions and ambulatory care. Our contribution to this project is to develop tools that would support decision making about the organization of care. It requires dually to be able to describe what is actually the current organization of care and to identify changes that may be improved or optimized. For that, we develop innovative visualization, data mining and operational research tools for care pathways analysis, management and planning. Their originality lays in their ability to consider three views of the care pathways: the patient, the regulator and the provider.

InflaMage

Participants: Hugues Berry.

Title: Development and proof of concept of non-invasive MRI-based imaging of brain inflammation

Partner Institution(s):

- Inria, Lyon
- ISC-MJ, CNRS UMR 5229, Lyon (supervision)
- MeLiS, INSERM U1314, Lyon
- Creatis, INSERM U1294, Lyon

Date/Duration: 2025-2029

Additional info/keywords: Funded by the French National Agency for Research (ANR), Call "AAP2024" (grant ANR-25-CE45-4106-02). InflaMage aims to develop a specific dMRI-based method for detecting cerebral inflammation using a machine learning classifier trained with Monte-Carlo simulations of dMRI signals obtained with digital phantoms of the complex microstructure of WM at the cellular level, including in addition to axons the geometries of resting and activated microglia and astrocytes. The classifier will be fine-tuned with histological data and dMRI measurements of the corpus callosum from mouse models of focal brain inflammation (LPS), and more realistic mouse models of human neuroinflammation (neuromyelitis optica (NMO) and multiple sclerosis (MS)).

CELLARD

Participants: Audrey Denizot.

Title: The CELLular Architecture and Reconstructions Database, an open-access web portal of high quality 3D cell reconstructions for structural analysis and simulations

Partner Institution(s):

- Inria, Lyon (supervision)
- Inria Rennes
- McGill University

Date/Duration: 2026-2029

Additional info/keywords: Funded by Inria (Action Exploratoire, AEx 2025). This project aims to build a domain-specific data hub to share 3D cellular reconstructions of volumetric electron microscopy datasets and to develop open-source tools to quantify their structural properties and facilitate their use in computational modeling studies. It will contribute to unraveling the structural determinants of cell function at the nanoscale and provide more accurate meshes for cell digital twins.

10.4 Regional initiatives

QuickRare

Participants: Thomas Guyet.

Title: QuickRare

Partner Institution(s):

- HCL, Lyon
- Univ Lyon 3
- Inria, Lyon

Date/Duration: sept 2024-2026

Additional info/keywords: QuickRare is a multidisciplinary project that aims to develop and analyse AI tools to reduce the length of patients' journey before obtaining a diagnosis for a rare disease. The technical challenge is to conceive a decision support tool that could be used by general practitioners simply by providing past medical reports of patients. The project is focused on the decision to address a patient to a rare disease reference center dedicated to the pediatric nephrology diseases. The team is complemented with social science researchers and philosophers who are investigating questions raised by the use of AI tools in this context. QuickRare is funded by a "Projet d'amorçage" grant of the [SHAPE-Med@Lyon](#) funding program.

BrainChat

Participants: Hugues Berry.

Title: BrainChat

Partner Institution(s):

- Inria, Lyon
- SBRI, Inserm U1208, Lyon
- Univ Claud Bernard Lyon 1

Date/Duration: sept 2024-2026

Additional info/keywords: BrainChat joins forces between teams and platforms working in neurosciences, single cell dataset production and digital sciences to develop a workflow to extract transcriptional signatures related to neurological diseases, while discarding confounding information (i.e. transcriptional differences related to their maturation and/or regional location). Our goal is to produce high resolution spatial RNA-Sequencing datasets that will add important spatial information on the dysregulated processes as well as the proximity of diseased astrocytes to neurons and develop data analysis methods to predict the molecular interactions occurring between astrocytes and with neurons, both in terms of ligand-receptor interactions and in terms of metabolic interactions (cholesterol, lactate. . .). BrainChat is funded by a "Projet d'amorçage" grant of the [SHAPE-Med@Lyon](#) funding program.

11 Dissemination

11.1 Promoting scientific activities

11.1.1 Scientific events: organisation

Member of the organizing committees

- Hugues Berry was part of the organizing committee of the conference “**AI/ML for the analysis of single-cell spatial transcriptomics**” held 15-17 Oct 2025 in Lyon (France).
- Maëlle Moranges co-organized the EXPLAIN’AI workshop on eXplainable Artificial Intelligence (XAI) at the 2025 and 2026 editions of the Knowledge Extraction and Management (EGC) conference.

11.1.2 Scientific events: selection

Chair of conference program committees

- Thomas Guyet was general chair of the **EGC** conference (Extraction et Gestion des Connaissances) in January 2025.

Member of the conference program committees

- Thomas Guyet was part of the program committees of ECML, ECAI, TIME, XAI, MedInfo conferences

11.1.3 Journal

Member of the editorial boards

- Hugues Berry is Section Editor for Neuroscience of PLoS Computational Biology
- Thomas Guyet is adjunct secretary editor for **ROIA** (Revue Ouverte d’Intelligence Artificielle)

Reviewer - reviewing activities

- Audrey Denizot reviewed 1 article for PLOS Computational Biology and 1 article for Cell
- Thomas Guyet reviewed 1 article for Springer CIBM, 1 article for IEEE Transactions on Artificial Intelligence and 1 article for IEEE Transactions on Systems, Man and Cybernetics: Systems.
- Maëlle Moranges reviewed for the EXPLAIN’AI workshop (EGC), Data & Knowledge Engineering, ECML-PKDD, and a Springer Nature book chapter.

11.1.4 Invited talks

We gave the following invited talks in 2025:

- “Adding astrocytes to digital twins for cellular neuroscience: towards computational glioscience”, Seminars of the IGFL (Institut de Genomique Fonctionnelle de Lyon), CNRS UMR5242, Lyon, France, October, Hugues Berry
- “Some examples of the benefit of AI for oncology”, Johns Hopkins Science Diplomacy Summit, Washington DC, USA, April, Hugues Berry
- “Linking astrocyte nano-architecture and function: insights from computational tools”, Shape Analysis Group, McGill University, Montreal, Canada, April, Audrey Denizot
- “Dissecting the Mechanisms Regulating Astrocyte Function at the Nanoscale with Computational Models”, Concordia University, Montreal, Canada, April, Audrey Denizot
- “Insights into the mechanisms regulating astrocyte function at the nanoscale”, “Neural computations without neurons” workshop, 22nd annual Computational and Systems Neuroscience (COSYNE) conference, Mont Tremblant, Canada, April, Audrey Denizot
- “An overview of Project-Team AIstroSight’s research”, Seminars of the Biomedical Imaging and Healthy Aging Laboratory, Concordia University, Montreal, Canada, April, Hugues Berry
- “Striatal endocannabinoid-dependent LTP and one-shot learning”, 5th Synaptic Microenvironment, Mini-symposium and Workshop, Sölden, Austria, March, Hugues Berry

- “Using simulation-based methods to characterize the dance of molecules in living cells”, 13th Manutech-SLEIGHT Graduate School Science Event, Saint-Etienne, France, January, Hugues Berry
- “IA pour la recherche en sciences de la vie” [17], Journée scientifique de l’École doctorale 536, Avignon, Avril, Thomas Guyet
- “Générer un SNIIRAM synthétique sans contrainte de partage”, Meetup Health Data Hub (HdH) - Données synthétiques en santé, Paris, December Thomas Guyet
- “Calcium diffusion inside a mesh of an astrocyte endfoot”, Keith Murai’s team, McGill university, Florian Dupeuble
- “Vascular Segmentation of fUS using Deep Learning”, Machine Learning Interest Group at the Center for Medical Image Computing, Hawkes Institute, London, Hana Sebia

11.1.5 Leadership within the scientific community

- Hugues Berry is a member of the working group on AI of the HIROs organization (Heads of International Research Organizations)
- Hugues Berry was vice-Chairman of the [2025 ANR Call “Thématiques Spécifiques en Intelligence Artificielle \(TSIA\)”](#)
- Chair of the steering committee of the [French Platform of Artificial Intelligence](#), Thomas Guyet

11.1.6 Scientific expertise

- Audrey Denizot was reviewer and member of the scientific committee of the French center for the 3Rs ([FC3R](#)), which aims to reinforce the implementation of the 3Rs (Replace, Reduce, Refine) in France through education, promoting responsible and innovative research, and transparent communication.
- Hugues Berry was member of the scientific evaluation pannel for the [ANR French Korean call for proposals in “Biotechnologies using Artificial Intelligence” 2025](#)
- *Participation to Scientific Advisory Boards (SAB)*
 - SAB of [SMART-HIFU](#), an ANR Labcom on Personalized treatments by High Intensity Focused Ultrasound involving the [LabTau](#) (INSERM U1032) and the company [EDAP-TMS](#), Hugues Berry
 - SAB of [INRAE UMR PRC](#), Tours, France, Hugues Berry
 - SAB of [Inserm U1059 Sainbiose](#), Saint-Etienne, France, Hugues Berry
 - SAB of [JUNON](#), BRGM, Orléans, France, Thomas Guyet

11.1.7 Research administration

- Hugues Berry has been Head of the Office of AI and Digital Sciences of Inserm (Pole IA et numerique) since March 2025 (Mise à Disposition Temporaire) and Chairman of the Strategic Committee of Inserm’s private cloud infrastructure.
- Thomas Guyet is vice-president of the French Society of Artificial Intelligence ([AFIA](#)) (representative of EurAI society)
- *Participations in selection committees:*
 - Hugues Berry has served in two selection committees for assistant professor positions

11.2 Teaching - Supervision - Juries - Educational and pedagogical outreach

11.2.1 Teaching

- Audrey Denizot gave a tutorial entitled “Computational modeling of astrocytes” at the XVII European Meeting on Glial Cells in Health and Disease, Marseille, France, July 2025
- Florian Dupeuble gave 18 hours of practical works to L1 student, in LIFBAP.
- Andrea Ducos gave 24 hours of practical works to L1 student, in LIFBAP.
- ArnaudHubert gave 36 hours of seminal course in algebra to L1 student, in informatics, Lyon1-UCBL

11.2.2 Supervision

PhD students

- Andrea Ducos, “Partial differential equation discovery for spatio-temporal simulations in cells”, since 02/11/2023, supervised by Thomas Guyet, Audrey Denizot and Hugues Berry
- Schayma Ben Marzougui-El Garrai, “Modeling the spatio-temporal dynamics of second messenger networks”, since 01/10/2023, supervised by Audrey Denizot and Hugues Berry
- Florian Dupeuble, “Biophysical modeling of neurovascular coupling at the gliovascular unit”, since 01/09/2023, supervised by Audrey Denizot and Hugues Berry
- Eric Pardoux, “Ethical issues in the use of artificial intelligence in healthcare: the contribution of an epistemological perspective” since 01/10/2021, supervised by Thomas Guyet and M. Laerke (CNRS, MFO)
- Thibaut Peyric, “Single-cell multi-omics data integration for gene regulatory network inference”, since 01/11/2023, supervised by A. Crombach (Inria/Beagle) and Thomas Guyet.
- Hana Sebia, “Deep phenotyping of patients” since 01/11/2022, supervised by Thomas Guyet and Hugues Berry.
- Ismail Bachchar, “Robust counterfactual explanation under distribution changes” since 01/03/2023, supervised by Thomas Guyet; Tassadit Bouadi (IRISA/Lacodam) and Françoise Fessant (Orange Labs).
- Arnaud Hubert, “Endocannabinoid-mediated synaptic plasticity and its implication in fast learning”, supervised by Hugues Berry

Interns

- Zoé Koenig, M1, INSA Lyon, France, 2025.02-06, supervisors: Audrey Denizot (50%) and Florian Dupeuble (50%)

11.2.3 Juries

- HDR juries
 - Y. Le Cunff, Univ. Rennes, May (Hugues Berry, examiner)
 - R. Ureña , Univ. Aix Marseille, February (Thomas Guyet, reviewer)
- PhD juries
 - M. Thomas, Univ. Toulouse, December (Hugues Berry, reviewer)
 - L. Tomy, Univ. Rennes, November (Hugues Berry, reviewer)
 - D-W. Garcia, Univ. Paris-Saclay, November (Audrey Denizot, reviewer)
 - Gioacchino Sterlicchio (Univ. Bary, Italie), December (Thomas Guyet, reviewer)

- Rodrigue Govan, Univ. Nouvelle Calédonie, (Thomas Guyet, reviewer)
- Nada Boudegzame, Univ. Paris Nord, December (Thomas Guyet, reviewer)
- Lilliam Muyama, Univ. Paris Cité, March (Thomas Guyet, reviewer)
- Ali Khudiyev, Univ. Strasbourg, November (Thomas Guyet, examiner)
- Moustafa Saïd Hawchar (Univ. Nantes), December (Thomas Guyet, examiner)
- Thibaut Soullard (Univ. Paris Cité), December (Thomas Guyet, examiner)
- Youssef Oubelmouh (Univ. Tour), June (Thomas Guyet, examiner)
- Armel Soubeiga (Univ. Clermont Ferrand), April (Thomas Guyet, examiner)

11.2.4 Educational and pedagogical outreach

- Audrey Denizot gave a pedagogical talk at [Université Ouverte Lyon 1](#) “Créer des mondes informatiques pour explorer le vivant”, Villeurbanne, France, entitled “Simuler numériquement le fonctionnement de notre cerveau pour mieux le soigner”
- Thomas Guyet taught introduction to artificial intelligence in the Master of Public Health and in the Medical Specialisation in Artificial Intelligence (DU IA).
- Maëlle Moranges taught Technology Watch to ING3 SCIA and ING3 IF students at EPITA and contributed to the design of the educational framework for EPITA’s Bachelor’s program in AI, Biotech, and Health.
- Maëlle Moranges participated in the outreach program "1 scientifique, 1 classe : chiche !" at Simone Weil High School and contributed [an experience-sharing presentation on the numin platform](#).
- Andréa Ducos participated in six events aimed at introducing young people to research and inspiring young girls. These events were: "Sciences : un métier de femmes" for the International Women’s Days ; meeting with high school students for their first-year internship ; meeting with high school students from Annonay as part of "la Fête de la Science" ; discussions with students from Gilberte et Pierre Brossolette High School organized by Pop Sciences ; presentation and tour of the campus with local middle school girls, organized by Lyon 1 University and help with organizing the “Filles et Maths, une équation lumineuse” day at ENS Lyon.

11.3 Popularization

11.3.1 Specific official responsibilities in science outreach structures

Audrey Denizot is a member of the board of directors & editorial committee of the “[Papier-Mâché Sciences](#)” association. The goal of the association is to explain the content of scientific publications in French & to outline the scientific method & publication process.

11.3.2 Productions (articles, videos, podcasts, serious games, ...)

- Thomas Guyet and Maëlle Morange (in connection with Inria Lyon communication services) created a pedagogical game for making discover the design of a medical artificial intelligence (see [details](#)). This game has been played with pupils and adults during the “*Fête de la Science*” in October 2025.
- Maëlle Moranges was interviewed by a journalist about her research conducted within the AISTroSight team for the article ["L'équipe AISTroSight mise sur la coopération avec les praticiennes et praticiens hospitaliers"](#).

11.3.3 Participation in Live events

- Audrey Denizot participated in the “Journée Filles et informatique : une équation lumineuse”, ENS Lyon, France ; speed-meeting with women studying in high school
- Thomas Guyet participated to a round table “AI, health and disabilities” organized by students of the M2 MALIA, Universty Lyon 2, January
- Thomas Guyet participated to a round table “Demondialization” organized by students of the lycée André Paillot, december
- Thomas Guyet made a presentation about artificial intelligence during a seminar on **informational challenges for library and documentation professionals** at the national school of documentalists (ENSSIB, Lyon), January [25].
- Thomas Guyet made an invited popularization presentation about artificial intelligence at the Meyzieu media library for the “*Fête de la Science*” in October 2025.
- Maëlle Moranges participated in a round-table discussion on the use of generative AI in education, organized by Réseau Canopé, entitled “**Vers des critères de choix partagés et pérennes dans les usages des IA génératives en éducation**”.

11.3.4 Others science outreach relevant activities

- Andréa Ducos, Maelle Morange and Thomas Guyet made outreach presentation in college (Chiche! program)

12 Scientific production

12.1 Major publications

- [1] Y. Dembitskaya, C. Piette, S. Perez, H. Berry, P. J. Magistretti and L. Venance. ‘Lactate supply overtakes glucose when neural computational and cognitive loads scale up’. In: *Proceedings of the National Academy of Sciences of the United States of America* 119.47 (14th Nov. 2022). DOI: [10.1073/pnas.2212004119](https://doi.org/10.1073/pnas.2212004119). URL: <https://inria.hal.science/hal-03922367>.
- [2] A. Denizot, M. Fernanda Veloz Castillo, P. Puchenkov, C. Cali and E. de Schutter. ‘The ultrastructural properties of the endoplasmic reticulum govern microdomain signaling in perisynaptic astrocytic processes’. In: *Glia* (16th Oct. 2025). DOI: [10.1002/glia.70091](https://doi.org/10.1002/glia.70091). URL: <https://hal.science/hal-03591633> (cit. on p. 18).
- [3] A. Ducos, A. Denizot, T. Guyet and H. Berry. ‘Evaluating PDE discovery methods for multiscale modeling of biological signals.’ In: *Springer LNBI*. 23rd International Conference on Computational Methods in Systems Biology (CMSB). Vol. 15959. Lyon, France, 10th Sept. 2025. URL: <https://hal.science/hal-05128224>.
- [4] T. Guyet and A. Duvermy. ‘Towards a Library for the Analysis of Temporal Sequences’. In: AALTD 2025 - 10th Workshop on Advanced Analytics and Learning on Temporal Data. Porto (Portugal), Portugal, 19th Sept. 2025. URL: <https://inria.hal.science/hal-05336370>.
- [5] S. Sadoune, A. Richard, F. Talbot, T. Guyet, L. Boussel and H. Berry. ‘Automatic analysis of negation cues and scopes for medical texts in French using language models’. In: *Computers in Biology and Medicine* 197 (2025), p. 110795. DOI: [10.1016/j.combiomed.2025.110795](https://doi.org/10.1016/j.combiomed.2025.110795). URL: <https://hal.science/hal-04564718>.
- [6] H. Sebia, T. Guyet, M. Pereira, M. Valdebenito, H. Berry and B. Vidal. ‘Vascular Segmentation of Functional Ultrasound Images using Deep Learning’. In: *Computers in Biology and Medicine* 194 (May 2025), p. 110377. DOI: [10.1016/j.combiomed.2025.110377](https://doi.org/10.1016/j.combiomed.2025.110377). URL: <https://inria.hal.science/hal-04744271>.

12.2 Publications of the year

International journals

- [7] A. Denizot, M. Fernanda Veloz Castillo, P. Puchenkov, C. Cali and E. de Schutter. ‘The ultrastructural properties of the endoplasmic reticulum govern microdomain signaling in perisynaptic astrocytic processes’. In: *Glia* (16th Oct. 2025). DOI: [10.1002/glia.70091](https://doi.org/10.1002/glia.70091). URL: <https://hal.science/hal-03591633>.
- [8] T. Guyet, P. Pinson and E. Gesny. ‘Clustering of timed sequences – Application to the analysis of care pathways’. In: *Data and Knowledge Engineering* 156 (Mar. 2025), p. 102401. DOI: [10.1016/j.datak.2024.102401](https://doi.org/10.1016/j.datak.2024.102401). URL: <https://inria.hal.science/hal-04849653> (cit. on p. 20).
- [9] N. Quiblier, J.-M. Rye, P. Leclerc, H. Truong, A. Hannou, L. Héliot and H. Berry. ‘Enhancing Fluorescence Correlation Spectroscopy with machine learning to infer anomalous molecular motion’. In: *Biophysical Journal* 124.5 (8th Jan. 2025), pp. 844–856. DOI: [10.1016/j.bpj.2025.01.026](https://doi.org/10.1016/j.bpj.2025.01.026). URL: <https://inria.hal.science/hal-04650578> (cit. on p. 19).
- [10] S. Sadoune, A. Richard, F. Talbot, T. Guyet, L. Boussel and H. Berry. ‘Automatic analysis of negation cues and scopes for medical texts in French using language models’. In: *Computers in Biology and Medicine* 197 (2025), p. 110795. DOI: [10.1016/j.combiomed.2025.110795](https://doi.org/10.1016/j.combiomed.2025.110795). URL: <https://hal.science/hal-04564718> (cit. on p. 21).
- [11] H. Sebia, T. Guyet, M. Pereira, M. Valdebenito, H. Berry and B. Vidal. ‘Vascular Segmentation of Functional Ultrasound Images using Deep Learning’. In: *Computers in Biology and Medicine* 194 (May 2025), p. 110377. DOI: [10.1016/j.combiomed.2025.110377](https://doi.org/10.1016/j.combiomed.2025.110377). URL: <https://inria.hal.science/hal-04744271> (cit. on p. 20).

International peer-reviewed conferences

- [12] A. Ducos, A. Denizot, T. Guyet and H. Berry. ‘Evaluating PDE discovery methods for multiscale modeling of biological signals.’ In: *Springer LNBI*. 23rd International Conference on Computational Methods in Systems Biology (CMSB). Vol. 15959. Lyon, France, 10th Sept. 2025. URL: <https://hal.science/hal-05128224> (cit. on p. 18).
- [13] T. Guyet and A. Duvermy. ‘Towards a Library for the Analysis of Temporal Sequences’. In: AALTD 2025 - 10th Workshop on Advanced Analytics and Learning on Temporal Data. Porto (Portugal), Portugal, 19th Sept. 2025. URL: <https://inria.hal.science/hal-05336370> (cit. on p. 20).
- [14] T. Peyric, T. Lepoutre, A. Crombach and T. Guyet. ‘Three-State Gene Expression Model Parameterized for Single-Cell Multi-Omics Data’. In: 23rd International Conference on Computational Methods in Systems Biology (CMSB 2025). Lyon, France, 19th July 2025. DOI: [10.1101/2025.07.16.665109](https://doi.org/10.1101/2025.07.16.665109). URL: <https://hal.science/hal-05180519> (cit. on p. 19).
- [15] A. Soubeiga, T. Guyet and V. Antoine. ‘Soft-ECM : une extension de l’algorithme Evidential C-Means pour des données complexes’. In: *Extraction et Gestion des Connaissances, EGC’2025*. Strasbourg, France, 29th Jan. 2025. URL: <https://inria.hal.science/hal-04929629> (cit. on p. 20).
- [16] A. Soubeiga, T. Guyet and V. Antoine. ‘Soft-ECM: An extension of Evidential C-Means for complex data’. In: *FUZZ-IEEE 2025 - International Conference on Fuzzy Systems*. Reims, France, July 2025. URL: <https://inria.hal.science/hal-05162452> (cit. on p. 20).

Conferences without proceedings

- [17] T. Guyet. ‘IA pour la recherche en sciences de la vie’. In: *Journée scientifique de l’École doctorale 536 – Numérique et intelligence artificielle pour les agrosociétés*. Avignon, France, 24th Apr. 2025. URL: <https://inria.hal.science/hal-05046080> (cit. on p. 27).

Scientific books

- [18] N. Abadie, G. Ateazing, G. Bonnet, T. Cazenave, A. Cornuéjols, V. Guigue, J.-G. Mailly, F. Mougin, P. Pr ea, F. Schwarzentruher, D. Symeonidou, H. Verhaeghe, A. Wilczynski, T. Guyet, B. Le Blanc, D. Longin, F. Sa s and A. Samet, eds. *Conf erence Nationale d’Intelligence Artificielle Ann ee 2025*. Association Fran aise pour l’Intelligence Artificielle, Sept. 2025. URL: <https://hal.science/hal-05409313>.

Edition (books, proceedings, special issue of a journal)

- [19] *Actes de la conf erence Extraction et Gestion des Connaissances*. Conf erence Extraction et Gestion des Connaissances - EGC’2025. Vol. E.41. Editions RNTI, Jan. 2025. URL: <https://inria.hal.science/hal-04929614>.

Reports & preprints

- [20] H. Berry, P. Gabriel, T. Lepoutre and N. Quiblier. *Subdiffusive fractional limit of a jump-renewal equation*. 13th Jan. 2026. URL: <https://hal.science/hal-05456904>.
- [21] H. Berry, J.-M. Rye and L. Trujillo. *Particle-Based Framework for Continuous Fields of Coupled Phase Oscillators: Exploring Spontaneous Local Synchronization*. 3rd July 2025. URL: <https://hal.science/hal-05143109>.

Other scientific publications

- [22] T. Guyet and A. Duvermy. ‘TanaT : A library for the analysis of temporal sequences’. In: PEPR-SanteNum 2025 - Journ es Annuelles du PEPR Sant  Num rique, Oct 2025, Lille, France. Lille, France, 14th Oct. 2025. URL: <https://inria.hal.science/hal-05351453> (cit. on p. 20).
- [23] T. Guyet, M. Guyomard and R. Urena. ‘Privacy-preserving generation of a realistic synthetic SNIIRAM database’. In: PEPR-SanteNum 2025 - Journ es Annuelles du PEPR Sant  Num rique, Oct 2025, Lille, France. Lille, France, 14th Oct. 2025. URL: <https://inria.hal.science/hal-05350989>.
- [24] H. Sebia, T. Guyet, H. Berry and B. Vidal. ‘Vascular Segmentation of fUS Images using Deep Learning’. In: Intelligence Artificielle en Imagerie Biom dicale (IABM). Nice, France, 17th Mar. 2025. URL: <https://inria.hal.science/hal-04964498> (cit. on p. 20).

Scientific popularization

- [25] T. Guyet. ‘L’IA aujourd’hui Comment  a marche ? Quelles limites ?’ In: Initiation   l’IA et   ses enjeux informationnels pour les professionnels des biblioth ques et de la documentation. Lyon, France, 12th Feb. 2025. URL: <https://inria.hal.science/hal-04946830> (cit. on p. 30).

12.3 Cited publications

- [26] S. Azam, M. Haque, M. Jakaria, S.-H. Jo, I.-S. Kim and D.-K. Choi. ‘G-Protein-Coupled Receptors in CNS: A Potential Therapeutic Target for Intervention in Neurodegenerative Disorders and Associated Cognitive Deficits’. In: *Cells* 9.2 (2020), p. 506 (cit. on p. 8).
- [27] A. Badoual, M. Arizono, A. Denizot, M. Ducros, H. Berry, U. Valentin N gerl and C. Kervrann. ‘Simulation of Astrocytic Calcium Dynamics in Lattice Light Sheet Microscopy Images’. In: *IEEE International Symposium on Biomedical Imaging, ISBI*. 2021, pp. 135–139 (cit. on p. 9).
- [28] N. Bazargani and A. Attwel. ‘Astrocyte calcium signaling: the third wave’. In: *Nature Neuroscience* 19 (2016), pp. 182–189 (cit. on p. 9).
- [29] A. Bespalov, T. Steckler, B. Altevogt, E. Koustova, P. Skolnick, D. Deaver, M. Millan, J. Bastlund, D. Doller, J. Witkin, P. Moser, P. O’Donnell, U. Ebert, M. Geyer, E. Prinssen, T. Ballard and M. Macleod. ‘Failed trials for central nervous system disorders do not necessarily invalidate preclinical models and drug targets’. In: *Nature Reviews Drug Discovery* 15 (2016), p. 516 (cit. on p. 6).

- [30] T. Bie. ‘Subjective interestingness in exploratory data mining’. In: *International Symposium on Intelligent Data Analysis*. 2013, pp. 19–31 (cit. on p. 11).
- [31] M. Bienvenu. ‘Ontology-mediated query answering: harnessing knowledge to get more from data’. en. In: *Proceedings IJCAI’16*. 2016, pp. 4058–4061 (cit. on p. 11).
- [32] S. Brunton, J. Proctor and J. Kutz. ‘Discovering governing equations from data by sparse identification of nonlinear dynamical systems’. In: *Proc Natl Acad Sci USA* 113 (2016), p. 3932 (cit. on p. 10).
- [33] A. Callahan. ‘ACE: the Advanced Cohort Engine for searching longitudinal patient records’. In: *J Am Med Inform Assoc* 28.7 (2021), pp. 1468–1479 (cit. on p. 11).
- [34] E. Chin and E. Goh. ‘MeCP2 Dysfunction in Rett Syndrome and Neuropsychiatric Disorders’. In: *Psychiatric Disorders: Methods and Protocols, Methods in Molecular Biology*. Ed. by F. Kobeissy. Vol. 2011. Humana, New York, NY, 2019, pp. 573–592 (cit. on p. 12).
- [35] A. Cougnoux, J. Yerger, M. Fellmeth, J. Serra Vinardell, K. Martin, F. Navid, J. Iben, C. Wassif, N. Cawley and F. Porter. ‘Single Cell Transcriptome Analysis of Niemann-Pick Disease, Type C1 Cerebella’. In: *Int J Mol Sci* 21 (2020), p. 5368 (cit. on p. 12).
- [36] Y. Cui, V. Paille, H. Xu, S. Genet, B. Delord, E. Fino, H. Berry and L. Venance. ‘Endocannabinoids mediate bidirectional striatal spike-timing dependent plasticity’. In: *J Physiol* 593.13 (2015), pp. 2833–2849 (cit. on p. 9).
- [37] Y. Cui, I. Prokin, A. Mendes, H. Berry and L. Venance. ‘Robustness of STDP to spike timing jitter’. In: *Scientific Reports* 8 (2018), p. 8139 (cit. on p. 9).
- [38] Y. Cui, I. Prokin, H. Xu, B. Delord, S. Genet, L. Venance and H. Berry. ‘Endocannabinoid dynamics gate spike-timing dependent depression and potentiation’. In: *eLife* 5 (2016), e13185 (cit. on p. 9).
- [39] Y. Cui, Y. Yang, Z. Ni, Y. Dong, G.-H. Cai, A. Foncelle, S. Ma, K. Sang, S. Tang, Y. Li, Y. Shen, H. Berry, S.-X. Wu and H. Hu. ‘Astroglial-Kir4.1 in lateral habenula drives neuronal bursts in depression’. In: *Nature* 554 (2018), pp. 323–327 (cit. on p. 9).
- [40] A. Denizot, M. Arizono, V. U. Nagerl, H. Soula and H. Berry. ‘Simulation of calcium signaling in fine astrocytic processes: effect of spatial properties on spontaneous activity’. In: *PLoS Comput Biol* 15.8 (2019), p. 1006795 (cit. on p. 9).
- [41] A. Denizot, C. Cali, H. Berry and E. De Schutter. ‘Stochastic spatially-extended simulations predict the effect of ER distribution on astrocytic microdomain Ca²⁺ activity’. In: *ACM NanoCom* 20 (2021), pp. 1–5 (cit. on p. 9).
- [42] S. Erikainen and S. Chan. ‘Contested futures: envisioning "Personalized", Stratified," and "Precision" medicine’. In: *New Genet Soc* 38.3 (2019), pp. 308–330 (cit. on p. 11).
- [43] T. Filzen, P. Kutchukian, J. Hermes, J. Li and M. Tudor. ‘Representing high throughput expression profiles via perturbation barcodes reveals compound targets’. In: *PLoS Comput Biol* 13.2 (2017), p. 1005335 (cit. on p. 15).
- [44] S. Finkbeiner. ‘Bridging the Valley of Death of therapeutics for neurodegeneration’. In: *Nature Medicine* 16 (2010), pp. 1227–1232 (cit. on p. 6).
- [45] A. Foncelle, A. Mendes, J. Jedrzejska-Szmeck, S. Valtcheva, H. Berry, K. Blackwell and L. Venance. ‘Modulation of spike-timing dependent plasticity: towards the inclusion of a third factor in computational models’. In: *Frontiers in Computational Neuroscience* 12 (2018), p. 49 (cit. on p. 9).
- [46] J. Freitas, K. Johnson, E. Golden, G. Nadkarni, E. Bottinger, B. Glicksberg and R. Miotto. ‘Phe2vec: Automated disease phenotyping based on unsupervised embeddings from electronic health records’. In: *Patterns* 2.9 (2021), p. 100337 (cit. on p. 12).
- [47] K. Friston, A. Mechelli, R. Turner and C. Price. ‘Nonlinear Responses in fMRI: The Balloon Model, Volterra Kernels, and Other Hemodynamics’. In: *NeuroImage* 12 (2000), pp. 466–477 (cit. on p. 9).
- [48] N. J. Gamo, M. R. Birknow, D. Sullivan, M. A. Kondo, Y. Horiuchi, T. Sakurai, B. S. Slusher and A. Sawa. ‘Valley of death: A proposal to build a "translational bridge" for the next generation’. In: *Neurosci Res* 115 (2017), pp. 1–4 (cit. on p. 6).

- [49] G. Gangarossa, S. Perez, Y. Dembitskaya, I. Prokin, H. Berry and L. Venance. ‘BDNF controls bidirectional endocannabinoid-plasticity at corticostriatal synapses’. In: *Cerebral Cortex* 1.30 (2020), pp. 197–217 (cit. on p. 9).
- [50] R. Gupta, D. Srivastava, M. Sahu, S. Tiwari, R. Ambasta and P. Kumar. ‘Artificial intelligence to deep learning: machine intelligence approach for drug discovery’. In: *Molecular Diversity* 25 (2021), pp. 1315–1360 (cit. on p. 6).
- [51] S. Hodavance, C. Gareri, R. Torok and H. Rockman. ‘G Protein-coupled Receptor Biased Agonism’. In: *J Cardiovasc Pharmacol* 67.3 (2016), pp. 193–202 (cit. on p. 9).
- [52] J. Ip, N. Mellios and M. Sur. ‘Rett syndrome: insights into genetic, molecular and circuit mechanisms’. In: *Nat Rev Neurosci* 19 (2018), pp. 368–382 (cit. on p. 12).
- [53] R. Jolivet, J. Coggan, I. Allaman and P. Magistretti. ‘Multi-timescale Modeling of Activity-Dependent Metabolic Coupling in the Neuron-Glia-Vasculature Ensemble’. In: *PLoS Comput Biol* 11 (2015), p. 1004036 (cit. on p. 9).
- [54] M. Kiernan, S. Vucic, K. Talbot, C. McDermott, O. Hardiman, J. Shefner, A. Al-Chalabi, W. Huynh, M. Cudkowicz, P. Talman, L. Berg, T. Dharmadasa, P. Wicks, C. Reilly and M. Turner. ‘Improving clinical trial outcomes in amyotrophic lateral sclerosis’. In: *Nature Reviews Neurology* 17 (Dec. 2020) (cit. on p. 11).
- [55] C. Kloft, M. Trame and C. Ritter. ‘Systems pharmacology in drug development and therapeutic use — A forthcoming paradigm shift’. In: *Eur J Pharm Sci* 94 (2016), pp. 4–14 (cit. on p. 6).
- [56] N. Lim and P. Pavlidis. ‘Evaluation of connectivity map shows limited reproducibility in drug repositioning’. In: *Sci Rep* 11.1 (2013), p. 17624 (cit. on p. 15).
- [57] B. Lőrincz, E. Jury, M. Vrablik, M. Ramanathan and T. Uher. ‘The role of cholesterol metabolism in multiple sclerosis: From molecular pathophysiology to radiological and clinical disease activity’. In: *Autoimmunity Rev* 21.6 (2022), p. 103088 (cit. on p. 15).
- [58] P. Monnin, C. Raïssi, A. Napoli and A. Coulet. ‘Discovering alignment relations with Graph Convolutional Networks: a biomedical case study’. In: 2021, p. 03452182 (cit. on p. 15).
- [59] A. Newman-Tancredi. ‘Translating biased agonists from molecules to medications: Serotonin 5-HT_{1A} receptor functional selectivity for CNS disorders’. In: *Pharmacology Therapeutics* 107937 (2021) (cit. on p. 9).
- [60] P. O’Donnell, L. B. Rosen, R. Alexander, V. N. Murthy, C. H. Davies and E. Ratti. ‘Strategies to Address Challenges in Neuroscience Drug Discovery and Development’. In: *International Journal of Neuropsychopharmacology* 22 (2019), pp. 445–448 (cit. on pp. 6, 11).
- [61] V. Pérez-Nueno. ‘Using quantitative systems pharmacology for novel drug discovery’. In: *Expert Opin Drug Discov* 10.12 (2015), pp. 1315–1331 (cit. on p. 6).
- [62] S. Pfeffer. ‘NPC intracellular cholesterol transporter 1 (NPC1)-mediated cholesterol export from lysosomes’. In: *J Biol Chem* 294.5 (2019), pp. 1706–1709 (cit. on p. 12).
- [63] Z. Qian, K. Kacprzyk and M. Schaar. ‘D-code: discovering closed-form ODEs from observed trajectories’. In: *Proceedings ICLR 2021*. 2021 (cit. on p. 10).
- [64] M. Raissi, P. Perdikaris and G. Karniadakis. ‘Physics-informed neural networks : A deep-learning framework for solving forward and inverse problem involving nonlinear partial differential equations’. In: *J Comput Phys* 378 (2019), pp. 686–707 (cit. on p. 10).
- [65] C. Shivade, P. Raghavan, E. Fosler-Lussier, P. Embi, N. Elhadad, S. Johnson and A. Lai. ‘A review of approaches to identifying patient phenotype cohorts using electronic health records’. In: *J Am Med Inform Assoc* 21.2 (2014), pp. 221–230 (cit. on p. 11).
- [66] S. Shovlin and D. Tropea. ‘Transcriptome level analysis in Rett syndrome using human samples from different tissues’. In: *Orphanet J Rare Dis* 13.1 (2018), p. 113 (cit. on p. 12).
- [67] A. Silberschatz and A. Tuzhilin. ‘On subjective measures of interestingness in knowledge discovery’. In: *Proceedings KDD’95*. 1995, pp. 275–281 (cit. on p. 11).

- [68] T. Takeuchi, A. Duzkiewicz and R. Morris. ‘The synaptic plasticity and memory hypothesis: encoding, storage and persistence’. In: *Philos Trans R Soc Lond B* 369 (2014), p. 20130288 (cit. on p. 9).
- [69] J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer and S. Zhao. ‘Applications of machine learning in drug discovery and development’. In: *Nature Reviews Drug Discovery* 18 (2019), pp. 463–477 (cit. on p. 6).
- [70] M. Vanier. ‘Niemann-Pick disease type C’. In: *Orphanet J Rare Dis* 5.16 (2010) (cit. on p. 12).
- [71] A. Verkhratsky and M. Nedergaard. ‘Physiology of Astroglia’. In: *Physiol Rev* 98.1 (2018), pp. 239–289 (cit. on p. 7).
- [72] B. Vidal, S. Fieux, J. Redouté, M. Villien, F. Bonnefoi, D. Bars, A. Newman-Tancredi, N. Costes and L. Zimmer. ‘In vivo biased agonism at 5-HT_{1A} receptors: characterisation by simultaneous PET/MR imaging’. In: *Neuropsychopharmacology* (2018), pp. 2310–2319 (cit. on p. 9).
- [73] R. Weinshilboum. ‘Inheritance and Drug Response’. In: *N. Engl J Med* 348 (2003), pp. 529–537 (cit. on p. 11).
- [74] H. Xu, S. Perez, A. Cornil, B. Detraux, I. Prokin, Y. Cui, B. Degos, H. Berry, A. De Kerchove D’exaerde and L. Venance. ‘Dopamine- endocannabinoid interactions mediate spike-timing dependent potentiation in the striatum’. In: *Nature Communications* 9 (2018), p. 4118 (cit. on p. 9).
- [75] K. Yin, D. Qian, K.-W. Cheung, B. Fung and J. Poon. ‘Learning phenotypes and dynamic patient representations via rnn regularized collective non-negative tensor factorization’. In: *Proceedings of the 2019 AAAI Conference on Artificial Intelligence*. 2019, pp. 1246–1253 (cit. on p. 12).
- [76] M. Zhang, D. Strnatka, C. Donohue, J. Hallows, I. Vincent and R. Erickson. ‘Astrocyte-only Npc1 reduces neuronal cholesterol and triples life span of Npc1^{-/-} mice’. In: *J Neurosci Res* 86 (2008), pp. 2848–2856 (cit. on p. 12).