Luxembourg Centre for Systems Biomedicine

Impact of Big Data & AI analytics on healthcare

19-20 November 2026

Campus Belval, Luxembourg

Conference & abstracts booklet

Table of contents

Introduction 2

Sponsors 3

Programme 5

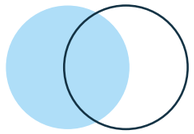
Venue map 7

Speakers 8

Flash talks selected from abstracts 22

Abstract list and poster numbers 24

Abstracts..... 28



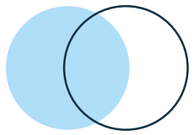
Introduction

The Luxembourg Centre for Systems Biomedicine (LCSB) is an Interdisciplinary Research Centre of the University of Luxembourg. It combines experimental and computational approaches to analyse complex biological systems and disease processes with a strong focus on neurodegeneration. The overarching strategy of the LCSB is to systematically combine interdisciplinary approaches to contribute to the prevention, diagnosis and cure of neurological diseases.

The conference Impact of Big Data & AI Analytics on Healthcare is built around four complementary themes:

- Big Data and Machine Learning
- Target discovery and drug development
- Clinical decision support systems
- Technological and industrial advances

The event offers added value through its integrated coverage of the entire translational pipeline and emphasises European research infrastructure and priorities. It serves as a platform for interdisciplinary exchange, fostering joint strategies and collaborative initiatives within the international bioinformatics community.



The scientific committee

Venkata Satagopam, Senior Research Scientist, Head of the Clinical & Translational Informatics group, Luxembourg Centre for Systems Biomedicine.

Anne Lorant, Computational Biologist and Data Scientist, Luxembourg Centre for Systems Biomedicine.

Georgios Pavlopoulos, , Head of the Computational Genomics Group, Institute of Fundamental Biological Research, Biomedical Science Research Center “Alexandre Fleming”.

Maria Secrier, Associate Professor in Computational Cancer Biology, UCL Genetics Institute, University College London

Reinhard Schneider, former Professor in Bioinformatics, Head of the Bioinformatics Core group, Luxembourg Centre for Systems Biomedicine. Following his recent and sudden passing, the last session of the conference will be dedicated to honouring his memory and achievements.

Sponsors

ITTM
Information Technology for Translational Medicine

LNDS
LUXEMBOURG NATIONAL DATA SERVICE

elixir LUXEMBOURG

uni.lu Luxembourg Centre for Systems Biomedicine

19–20 NOV 2026

Université du Luxembourg
Campus de Belval
Maison du Savoir
2, place de l'Université
L-4365 Esch-sur-Alzette
www.bigdata.uni.lu

Impact of Big Data & AI analytics on healthcare

LCSB

ITTM Information Technology for Translational Medicine · LNDS LUXEMBOURG NATIONAL DATA SERVICE · elixir LUXEMBOURG

Thursday 19 November

13:45 Welcome words - **Venkata Satagopam**, University of Luxembourg
Simone Niclou, Vice-rector for research, University of Luxembourg

14:00 **Keynote:** Early disease risk prediction from longitudinal clinical records using AI - **Chris Sander**, Harvard Medical School

Session 1 - Big data, AI and Machine Learning: current trends and future perspective

Chair: Alexander Skupin, U. of Luxembourg

14:45 From trajectory-based prediction of disease and outcomes to real-time implementation - **Søren Brunak**, University of Copenhagen

15:15 ELIXIR and European health data spaces: infrastructure for AI-Driven genomic medicine - **Tim Hubbard**, King's College London & ELIXIR

15:45 Protein Function as the Bridge Between Individual Genotypes and Human Disease - **Michal Linial**, The Hebrew University of Jerusalem

16:15 2 Flash talks (from abstract selection):
Tsz Pan Tong, University of Luxembourg
Maciej Skorski, University of Luxembourg

16:30 ***Coffee Break, industry exhibition and posters (n° 1 to 4)***

Session 2 - Target discovery and drug development

Chair: Maria Secrier, UCL

17:00 Sponsor pitch: Making Data Actionable: Cutting Time to Health Data Through Trust - **Andreas Kremer,** Information Technology for Translational Medicine (ITTM)

17:10 Systems biomedicine approach to Parkinson's disease **Alexander Skupin,** University of Luxembourg

17:40 Enabling drug discovery through efficient utilisation of disease biology insights and high-quality open data resources - **Philip Gribbon**, EU Openscreen

18:10 AI transforms molecular sciencesn - **Alexandre Tkatchenko**, University of Luxembourg

18:40 2 Flash talks (from abstract selection):
Soumyabrata Ghosh, University of Luxembourg
Martin Hoffmann-Apitius, Fraunhofer Institute for Algorithms and Scientific Computing SCAI

18:55 ***Networking dinner, industry exhibition and posters (n° 5 to 16)***

Friday 20 November

9:00 **Keynote:** Interpreting genetic variation with AI: from sequence to clinical relevance - **Yana Bromberg**, Emory University

Session 3 - Computational models for precision medicine

Chair: Enrico Glaab, U. of Luxembourg

9:45 Of Big Data then and now - **Alfonso Valencia**, Barcelona Supercomputer Centre

10:15 Exploring the planetary microbial proteome at scale **Georgios Pavlopoulos**, Biomedical Science Research Center "Alexander Fleming"

10:45 ***Coffee Break, industry exhibition and posters (n° 17 to 20)***

11:15 Computational and AI models of tumour evolution for clinical decision support - **Maria Secrier,** University College London

11:45 2 Flash talks (from abstract selection):
Kavya Singh, University of Liège
Ulrike Weirauch, Fraunhofer Institute for Cell Therapy and Immunology IZI

Session 4 - Technology and innovation: AI applications in pharma and medicine

Chair: Shekoufeh Gorgi Zadeh, U. Luxembourg

12:00 Building data enabled health innovations through public-private partnerships - **Christian Muehlendyck**, Johnson&Johnson

12:30 ***Lunch, industry exhibition and posters (n° 21 to 29)***

14.00 From bioinformatics to individualised cancer therapy: lessons from Biotech innovation - **Georg Casari**, YGION Biomedical Gmbh

14:30 Precision medicine partnerships: integrating genomics, AI and industry innovation to Build the 360° Patient Beyond Disease Labels - **'Matladi N. Ndlovu**, UCB

15:00 Luxembourg will host the future WHO bureau of AI in health **Charles Betz,** Luxembourg Ministry of Health

15:15 Poster prize ceremony

Session 5 - Advances in bioinformatics: honouring the contributions of Prof. Reinhard Schneider

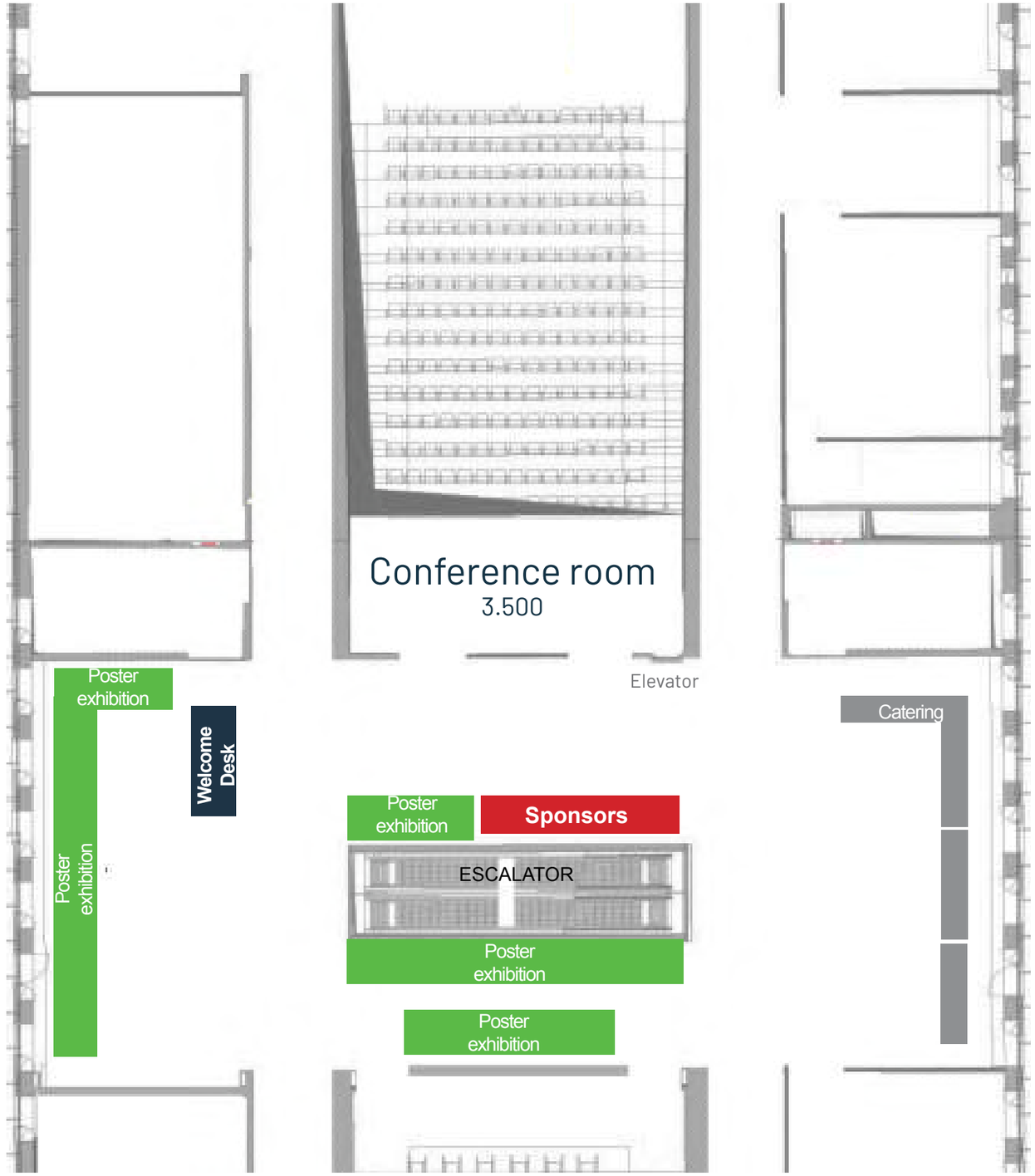
Chair: Venkata Satagopam, U. of Luxembourg

15:30 Chris Sander (10'), Burkhard Rost (10'), Rudi Balling (10')

16:00 Panel: Regina Becker, Alfonso Valencia, Georg Casari, Evangelos Pafilis

16:30 ***Wrap up & coffee***

Venue


Conference room
3.500
Elevator
Poster exhibition
Welcome Desk
Poster exhibition
Poster exhibition
Sponsors
ESCALATOR
Poster exhibition
Poster exhibition
Catering

Speakers

Yana Bromberg

Professor of Bioinformatics
Emory University



Talk title

Interpreting genetic variation with AI: from sequence to clinical relevance

Biosketch

Dr Yana Bromberg is Professor in the Departments of Biology and Computer Science at Emory University, a Principal Fellow of the Emory Center for AI Learning, and an Adjunct Professor in the Department of Biomedical Engineering at Georgia Tech. She was recently elected a Fellow of the International Society for Computational Biology in 2025 and serves on its Board of Directors. Dr Bromberg earned her Ph.D. in Biomedical Informatics from Columbia University in 2007 and is a Hans Fischer Fellow at TU Munich's Institute for Advanced Study.

The primary focus of Dr Bromberg's research is the concept of 'function' in biology: how biological functionality is encoded in genomic data, whether by a single gene, a genome, or a metagenome. Since before it was "cool" her lab has been building machine-learning and AI methods to make sense of genome and metagenome data, with work in precision medicine, environmental microbiome analysis, and origins of life. Tools developed in her lab span variant-effect prediction and whole genome analysis for disease predisposition, microbiome functional annotation, and, more recently, pretrained DNA and protein language models for reference-free function discovery. She also recently testified before the U.S. National Academy of Sciences on AI and biosecurity and co-let the Asilomar AI + Biotech working group.

Link to her lab

Søren Brunak

Professor of Disease Systems Biology
University of Copenhagen



Talk title

From trajectory-based prediction of disease and outcomes to real-time implementation

Biosketch

Søren Brunak is professor of Disease Systems Biology at the University of Copenhagen at the Dept. for Public Health. His research group has a strong life-course perspective on multimorbidities and the associated polypharmacy and combines multiomics data with analysis of electronic patient records, disease registry information and biobank cohort questionnaires. The overall aim to understand disease progression and disease interdependencies in patients who suffer from several diseases simultaneously. He initiated population-scale disease trajectory analysis using data from millions of patients to define subgroups with shared disease progression patterns and their underlying driver mechanisms. Søren Brunak has worked in the machine learning domain since the 1980s and use currently deep learning transformer architectures to analyse population-wide multimodal disease trajectories and predict risk and time to event of future diseases. He is a member of the Danish Royal Academy of Sciences and Letters as well as the Royal Swedish Academy of Sciences.

Link to his lab

Georg Casari

Co-founder, CEO & COO
YGION Biomedicl GmbH



Talk title

From bioinformatics to individualised cancer therapy: lessons from Biotech innovation

Biosketch

Dr. Georg Casari is an Austrian biochemist and biotechnology entrepreneur with extensive leadership experience in life science companies and research organisations. He obtained his MSc in Biochemistry from the University of Vienna in 1988 and completed his PhD at IMP Vienna/University of Vienna in 1991. Following postdoctoral work in protein structure analysis at the University of Salzburg and in sequence analysis and genome annotation at EMBL Heidelberg and the European Bioinformatics Institute, he transitioned to roles at the interface of bioinformatics and pharmaceutical research

In 1997 he co-founded LION bioscience AG in Heidelberg, where he built and co-led the Life Science Informatics business unit as Vice President, later serving as Senior Vice President Life Science Informatics Research and establishing a high-performance bioinformatics research organization to support drug discovery. He subsequently held executive positions as Vice President Informatics at Cellzome AG, Chief Information Officer at Oridis-biomed GmbH, and Administrative Director at the CeMM Research Center for Molecular Medicine of the Austrian Academy of Sciences, with responsibility for finance, IP and technology transfer, and research infrastructure.

Since 2010 Dr. Casari has been co-founder and Managing Director of Haplogen GmbH and, since 2017, Managing Director of Haplogen Bioscience GmbH, overseeing company strategy, partnerships, licensing, financing, and operational teams.

In 2021 he co-founded YGION Biomedical GmbH, a biotechnology company developing an individualised immune-stimulation cancer vaccine based on patient-specific tumor neoepitopes, which is currently being tested in clinical phase.

Dr. Casari habilitated in Bioinformatics at the University of Salzburg in 2002 and has served on several scientific and industry boards, including the Max Planck Institute for Cybernetics and BIOTECH AUSTRIA.

Link to YGION

Philip Gribbon

Director General
EU Openscreen



Talk title

Enabling drug discovery through efficient utilisation of disease biology insights and high-quality open data resources

Biosketch

Dr Philip Gribbon is Head of Discovery Research at the Fraunhofer Institute for Translational Medicine and Pharmacology in Hamburg, Germany and Director General of the European Infrastructure for Chemical Biology, EU-OPENSCREEN. Currently Philip coordinates the IDERHA project which is building an EHDS aligned secure processing environment for analysis of heath data including RWD. Previously, Philip was Chief Scientific Officer of the European ScreeningPort. Previously, Philip has worked at GSK developing new drug discovery and development workflows and at Pfizer covering multiple therapeutic areas. He has a PhD and undergraduate degrees from Imperial College London and Post-doctoral experience from the University of Manchester. Philip is highly engaged in open science initiatives and the application of advanced data analysis methods and has > 120 publications. He has served on the Board of the Society for Laboratory Automation and Screening, including as President in 2025.

Link to YGION

Tim Hubbard

Professor of Bioinformatics
King's College London



Talk title

ELIXIR and European health data spaces: infrastructure for AI-Driven genomic medicine

Biosketch

Tim Hubbard is Professor of Bioinformatics at King's College London. From 2024 to 2026 he served as Director of ELIXIR, the European Research Infrastructure for biological data. ELIXIR coordinates bioinformatics projects across its 23 member states and the European Bioinformatics Institute (EMBL-EBI), including the EU funded Genome Data Infrastructure (GDI) for the Beyond 1 Million Genomes (B1MG) project. He is also a member of the Technical Advisory Group for Genomics (TAG-G) at the World Health Organisation. Between 2013 and 2024 he was seconded to Genomics England for the 100,000 genomes project, mainstreaming whole genome sequencing in the National Health Service (NHS). He has been an advocate of "Trusted Research Environments" (TREs) to enable research on health data without data distribution to protect privacy, building on experience developing and operating the Genomics England TRE. The planned Secure Processing Environments (SPEs) of the European Health Data Space (EHDS) expand this model to support federated data analysis across EU member states as does GDI for B1MG.

Until 2013 he was Head of Informatics at the Wellcome Sanger Institute where he was one of the organisers of the sequencing of the human genome. In 1999 he co-founded the Ensembl project to analysis, organise and provide access to the human genome and from 2007 led the GENCODE project to annotate the structure of all human genes. Between 1996 and 2007 he was an organiser of the biannual Critical Assessment of Structure Prediction (CASP), evaluating methods to solve the protein folding problem, which was substantially solved by the AI Alphafold in 2020.

Link to list of publications

Michal Linial

Professor of Biochemistry & Bioinformatics
The Hebrew University of Jerusalem



Talk title

Protein Function as the Bridge Between Individual Genotypes and Human Disease

Biosketch

Michal Linial is a Professor of Biochemistry and Bioinformatics of the Hebrew University of Jerusalem (HUJI). She served as the Director of the Israeli Institute for Advance Studies (IIAS) in Jerusalem (from 2012-2018) and as a Director of the Sudarsky Center for Computational Biology at the Hebrew University (2001-2011).

Prof. Linial obtained her Ph.D. in Molecular biology from HUJI. She then completed her post-doctoral training at Stanford (CA, USA) on the field of cellular neurochemistry. On 1989 she joined the faculty of HUJI. She is a founder and the Chair (initated in 1999) of the educational program for Computational Biology in HUJI. She is the Israeli representative of pan-European ELIXIR Israeli Node. ML serves in the Board of Directors and as a Vice-President of the International Society of Computational Biology (ISCB) and was selected as a society fellow. She served as the chair of the European conference series in computational Biology (ECCB, 2005-2012). From 2005 she serves at the steering committee of ISMB, RECOMB and ECCB the international conferences of the Bioinformatics and Computational Biology communities and of CAFA competition for automatic functional annotation initiative. Michal Linial was a visiting professor in University of Washington in Seattle and MCR in Cambridge, MA-USA and a visiting scholar in Microsoft Research Institute (MRI, Cambridge Boston).

Linial lab developed and maintains large-scale databases and web-tools that are open to the community including ProtoNet for protein classification and annotations. miRrorSuite for combinatorial regulation by miRNAs and targets, and PANDORA for integration of protein annotations. Recently she developed PWAS-Hub as a source for genotype to phenotype for hundreds of diseases exposing sex-dependent genetics.

Linial has authored over 320 peer-reviewed papers, and a collection of book chapters, and edited books. She contributed to the development of bioinformatics databases and websites that are open to the large communities of the biomedical and Life science researchers. Her current research interests are on molecular evolution, protein family evolution, microRNA mode of action, methodological development of phenotype to genotype, molecular basis of aging and metabolic diseases and methods for merging proteomics and genomics toward human health. She applies large-scale technologies including NGS, proteomics (Mass spectrometry) for revealing the regulation levels in health and disease.

Link to her lab

Christian Muehlendyck

Scientific Partnerships Lead
Johnson&Johnson MedTech



Talk title

Building data enabled health innovations through public–private partnerships.

Biosketch

Christian Muehlendyck, MD MBA, is Scientific Partnerships Lead at Johnson&Johnson MedTech EMEA. In this role, he drives the public private partnership (PPP) activities of J&J MedTech with focus on the Innovative Health Initiative opportunities, including the industry leadership of IDERHA.

He assumes the role of Chair of the MedTech Europe Research & Innovation Committee and represents MedTech Europe on the IHI Governing Board, helping to shape cross-industry research and innovation priorities across Europe.

He began his career as a physician in Orthopedic and General Surgery and obtained his MD/PhD while co-leading the development of an innovative Medical Device. He subsequently transitioned into the MedTech industry to further his passion for medical innovation. Within Johnson & Johnson he has held a series of local, regional and global roles of increasing responsibility spanning from Medical Training as well as commercial positions to Health Economics & Market Access across the Surgical and Orthopedic spectrum. Within those roles he worked closely with physicians, patients & health authorities and acquired a broad expertise in health care innovation, and research incl. value-based healthcare and healthcare digitalization

Link to Johnson&Johnson MedTech

‘Matladi N. Ndlovu

Global Lead Innovative Precision
Medicine Partnerships, UCB



Talk title

Precision medicine partnerships: integrating genomics, AI and industry innovation

Biosketch

Dr. ‘Matladi N. Ndlovu, MSc, PhD, is Global Lead External Innovation at UCB in Brussels. An epigeneticist and systems thinker with 15+ years in immunology, she bridges cutting-edge science and business strategy, bringing diverse experts together around a shared mission: turning complex data into insight, and insight into lasting impact for patients.

‘Matladi has built public–private and pre-competitive digital partnerships from the ground up. Together with Novartis, she drove the creation of IMI HIPPOCRATES, aligning competing companies on a shared research topic within three months. The resulting 27-partner, €23M European consortium uses machine learning to predict and diagnose psoriatic arthritis earlier, with UCB as industry co-lead. She also brought UCB into IMI BIOMAP (2019–2025), which studied data and samples from 50,000 people with atopic dermatitis or psoriasis; ‘Matladi led UCB’s contribution to big-data integration and interpretation. In the US, she helped lay the foundation for FNIH AMP-AIM, a 20-organisation pre-competitive partnership mapping autoimmune disease at single-cell level. ‘Matladi also took Psorcast from idea to a smartphone app with clinically validated digital measures, within a pre-competitive digital biomarker consortium.

She then applied this data-and-partnership approach to hidradenitis suppurativa (HS), a chronic inflammatory skin disease, establishing the “Think HS” programme at UCB. Analysis of US insurance claims showed that HS symptoms affect men and women equally, not mostly women (3:1) as reported in the literature, doubling the estimated global patient population. ‘Matladi is now extending this work with consumer data and a generative AI foundation model covering more than 300 million patients, to find hidden patient populations at scale.

In Luxembourg, she will show how partnership-built data platforms and AI move precision medicine from consortium insight to patient impact.

Link to UCB

Georgios Pavlopoulos

Research Director, Biomedical Science
Research Center “Alexander Fleming”



Talk title

Exploring the planetary microbial proteome at scale

Biosketch

Georgios A. Pavlopoulos earned a BSc/Diploma in Computer Science from the National and Kapodistrian University of Athens (NKUA) in Greece; an MRes in Bioinformatics from the University of Leeds (Department of Biochemistry) in the UK; and a PhD in Bioinformatics from EMBL-Heidelberg in Germany. He also holds a joint MBA from the Athens University of Economics and Business (AUEB) and the National Technical University of Athens (NTUA).

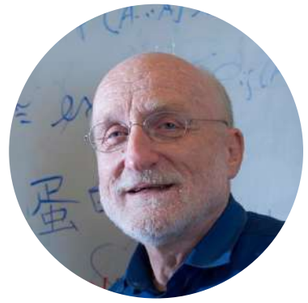
His research career includes postdoctoral work at KU Leuven’s Department of Electrical Engineering (ESAT) in Belgium, and roles as a bioinformatics instructor and postdoctoral researcher at the University of Crete’s Faculty of Medicine in Greece, and at the DOE Joint Genome Institute at Lawrence Berkeley National Laboratory in California. Since 2018, he has consulted in bioinformatics for several U.S. companies and led the bioinformatics program at the Biomedical Sciences Research Center “Alexander Fleming” (BSRC) in Greece, serving as Director of Bioinformatics (Research Associate Professor) and as an elected member of the Scientific Board at the Institute of Fundamental Biomedical Research (IFBR). In parallel, he served as Head of Bioinformatics at NKUA’s Center of New Biotechnologies and Precision Medicine for five years. Appointed in 2024 as Director of Research (Research Full Professor) at BSRC “Alexander Fleming,” he is also a co-founder and board member of BioInnovation Greece; founder and CEO of GC2 Data Science; a bioinformatics data scientist at Ancilia Biosciences (New York); and a Visiting Professor at MBZUAI in the UAE (2026).

Pavlopoulos’s lab focuses on genomics and metagenomics, biological networks, text mining, data visualization, big data analysis, high-performance computing, data integration, machine learning, knowledge management and representation, and personalized medicine.

Link to his lab

Chris Sander

Professor of Systems Biology
Harvard Medical School



Talk title

Early disease risk prediction from longitudinal clinical records using AI

Biosketch

Chris Sander started his career as a theoretical physicist and then switched to theoretical biology, in part inspired by the first completely sequenced genome. He founded two departments of computational biology – at the EMBL in Heidelberg and Memorial Sloan Kettering Cancer Center in New York – and co-founded the research branch of the European Bioinformatics Institute in Cambridge and a biotech startup with Millennium in Boston. Chris joined the Harvard community in 2016 as faculty in Cell Biology and then Systems Biology. Special Advisor for Quantitative Biology to the Ludwig Center at Harvard, and Associate Member of the Broad Institute. He is creating new connections between scientists at Dana-Farber and Harvard Medical School, including building translational collaborative bridges for scientists using quantitative sciences to solve biological problems.

With his group and collaborators, Chris aims to beat drug resistance in cancer using systems biology methods to develop combination therapies. They are also developing the next generation cBioPortal for cancer research and therapy, obtaining biomolecular structures and functional interactions on a large scale using evolutionary information, and adapting machine learning methods to mine millions of genomes. He is collaborating with groups in Denmark and the US to apply AI to longitudinal health records to identify patients at high risk for pancreatic and ovarian cancer and collaborates with clinicians in the design of effective affordable surveillance program aiming at the early detection of cancer.

Link to his lab

Maria Secrier

Associate Professor in Computational Biology
University College London



Talk title

Computational and AI models of tumour evolution for clinical decision support

Biosketch

Maria Secrier is an Associate Professor in Computational Cancer Biology at the UCL Genetics Institute. Her group focuses on understanding cancer evolution in its environmental niche by developing machine learning and data integration methods for complex multi-omics and imaging datasets. Maria trained in Bioinformatics at Jacobs University in Bremen, Germany, and obtained her PhD in Computational Biology from the European Molecular Biology Laboratory in Heidelberg under the supervision of Prof Reinhard Schneider. Afterwards, she worked with Prof Simon Tavaré and Prof Rebecca Fitzgerald on the genomics of oesophageal cancer as a postdoc at the CRUK Cambridge Institute and Bye-Fellow at the University of Cambridge. She also acquired experience in immuno-oncology and new target discovery as a Senior Scientist at AstraZeneca, before joining UCL in 2017 where she started her independent research group.

Link to her lab

Alexander Skupin

Head of the Integrative Cell Signalling Group
University of Luxembourg



Talk title

Systems biomedicine approach to Parkinson's disease

Biosketch

Alexander Skupin is Professor for Integration and Modelling of Biomedical Data at the Luxembourg Center for Systems Biomedicine (LCSB), an Affiliated Professor for Integrative Biophysics at Department of Physics and Material Science at the University of Luxembourg and a Project Scientist at the Department of Neuroscience at University of California San Diego (USA). He studied physics at the Humboldt University (HU) in Berlin where he worked on noise-induced effects in neuronal networks and an application to deep brain stimulation as treatment of Parkinson's disease patients with Lutz Schimansky-Geier. He obtained his PhD in biophysics from the Humboldt University Berlin for his work on calcium oscillation with Martin Falcke. After a first postdoc with Oliver Ebenhöh at the Max-Planck Institute for Molecular Plant Physiology in Potsdam focusing on metabolism including genome-scale modelling, he subsequently joined the Institute for Systems Biology in Seattle (USA) and worked with Aimee Dudley on Systems Genetics, with Sui Huang on cell fate dynamics at single cell resolution and with David Galas on information flow in biological systems. In 2014, he started his interdisciplinary Integrative Cell Signalling research group at the LCSB where the team combines microscopy and single cell analysis techniques with bioinformatics analyses and theoretical concepts from physics to investigate underlying principles of brain disorders with a focus on neurodegeneration.

Link to his lab

Alexandre Tkatchenko

Professor in professor in Theoretical Condensed Matter Physics University of Luxembourg



Talk title

AI transforms molecular sciences

Biosketch

Alexandre Tkatchenko is a professor and former head of the Department of Physics and Materials Science at the University of Luxembourg. Tkatchenko also holds a distinguished visiting professor position at the Technical University of Berlin and has co-founded two deeptech startups. He has delivered more than 500 invited talks, seminars and colloquia worldwide and serves or served on the editorial boards of four society journals: Science Advances (AAAS), AI for Science (IOP), Chemical Science (RSC), and Digital Discovery (RSC). Tkatchenko has received a number of awards, including APS Fellow, Fellow of the Royal Society of Chemistry, Gerhard Ertl Young Investigator Award of the German Physical Society, Dirac Medal from WATOC, Feynman prize from the Foresight Institute, van der Waals prize of the ICNI, and six flagship grants from the European Research Council: a Starting Grant in 2011, a Consolidator Grant in 2017, an Advanced Grant in 2022, and Proof-of-Concept Grants in 2020, 2023, and 2026.

Link to his lab

Alfonso Valencia

Director of the Life Sciences Department Barcelona Supercomputer Center



Talk title

Of Big Data then and now

Biosketch

Prof. Alfonso Valencia is ICREA research Professor, Director of the Life Sciences Department of the Barcelona Supercomputing Center, Director of the Spanish National Bioinformatics Institute INB/ ELIXIR-ES and coordinator of the data pillar of the Spanish Personalised Medicine intiative, IMPaCT. His research interest is the development of Computational Biology methods and their application to biomedical problems. Some of the computational methods he developed are considered pioneering work in areas such as biological text mining, protein coevolution, disease networks and more recently modelling cellular systems (digital twins). He participates in some of the key cancer related international consortia. In terms of community services, he is one of the initial promoters of the ELIXIR infrastructure, founder of the Spanish and International Bioinformatics networks and former president of ISCB, the international professional association of Bioinformaticians. He is Executive Editor of the main journal in the field (Bioinformatics OUP).

Link to the research groups of the Life Sciences Department

Flash talks selected from abstracts



Soumyabrata Ghosh

University of Luxembourg

Abstract

Adept: a secure, no-code platform for exploring governed biomedical datasets across the translational pipeline

Presenting on 19 Nov at 18:30



Martin Hoffmann-Apitius

Fraunhofer Institute for Algorithms and Scientific Computing SCAI

Abstract

Evidence Reloaded

Presenting on 19 Nov at 18:38



Kavya Singh

University of Liège

Abstract

Individual-specific network diffusion reveals biological and therapeutic heterogeneity in mixed Alzheimer's disease and LATE pathology

Presenting on 20 Nov at 11:45

Flash talks selected from abstracts



Maciej Skorski

University of Luxembourg

Abstract

Uncertainty-Aware AI Clinical Narratives for Cardiovascular Risk

Presenting on 19 Nov at 16:23



Tsz Pan Tong

University of Luxembourg

Abstract

FlowGRN: Scalable and Dropout-Robust Gene Regulatory Network Inference via Flow Matching-Based Trajectory Reconstruction

Presenting on 19 Nov at 16:15



Ulrike Weirauch

Fraunhofer Institute for Cell Therapy and Immunology IZI

Abstract

CERTAINTY: A Cellular Immunotherapy Virtual Twin for Personalised Cancer Treatment

Presenting on 20 Nov at 11:53

List of posters

Presenter	Submitter company	Title	Poster n° and session
Marcio Acencio		LLM-assisted reconstruction of a mechanistic disease map from primary literature: a case study	Poster n°1 19 Nov. 16:30-17:00
Esther Ademola	Luxembourg Institute of Health	Uncovering multimodal phenotypic structure in individuals at risk for parkinson's disease using similarity network fusion	Poster n°2 19 Nov. 16:30-17:00
Basel Alshaikhdeeb	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Two-tier biomedical relation extraction pipeline	Poster n°3 19 Nov. 16:30-17:00
Beatriz Garcia Santa Cruz	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Building A Responsible Personalised Digital Therapeutic Solution For Sleep Disturbances: BRAINPATH	Poster n°4 19 Nov. 16:30-17:00
Hanna Ćwiek-Kupczyńska	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	The IDERHA Platform: Integrating federated data governance and analysis for health research	Poster n°5 19 Nov. Networking dinner
Siavash Damari	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Human-in-the-Loop AI Framework for Automated Quantification of Cell Populations in Patient-Derived Midbrain Organoids	Poster n°6 19 Nov. Networking dinner
Beatriz Garcia Santa Cruz	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	MitoSegAI: An Integrated Deep Learning Pipeline for the Segmentation and Quantification of Mitochondrial Ultrastructure in 3D Electron Microscopy	Poster n°7 19 Nov. Networking dinner
Aditya Krishna	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Synthetic Tabular Data Generation Framework for Clinical Decision Support Systems	Poster n°8 19 Nov. Networking dinner
Anne Lorant	Luxembourg National Data Service	AID-Care: Artificial Intelligence for neuroDegenerative diseases Care	Poster n°9 19 Nov. Networking dinner
Taras Lukashiv	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Patient-Specific Digital Twins and Semi-Markov Control for Personalized Extracorporeal Liver Support: A Clinical Decision-Support Framework	Poster n°10 19 Nov. Networking dinner
Tyler McInnes	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	The ELIXIR-LUXEMBOURG training platform	Poster n°11 19 Nov. Networking dinner
Shaman Narayanasamy	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Benchmarking hosted and local System One models on regulated and sensitive text	Poster n°12 19 Nov. Networking dinner
Francesco Nasta	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Characterizing GBA-Associated Parkinson's Disease through Transcriptomics and Morphological Profiling in Midbrain Organoids	Poster n°13 19 Nov. Networking dinner
Marek Ostaszewski	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	The MINERVA Platform: fifteen years of support for disease maps	Poster n°14 19 Nov. Networking dinner
Priyanshu Kumar Panda	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Identifying Genetically Driven Gut Microbiome Components Causally Linked to Parkinson's Disease	Poster n°15 19 Nov. Networking dinner
Sandesh Patil	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	LCSB Data Space: governed management of clinical and biomedical research data at the Luxembourg Centre for Systems Biomedicine	Poster n°16 19 Nov. Networking dinner
Abdul Rafay Pirzada	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	MIGRAPH: a heterogeneous graph framework for multi-omics microbiome integration	Poster n°17 20 Nov. 10:45-11:15
Kavita REGE	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	A sovereign, GDPR-compliant tool assembly for clinical and translational research data: architecture and use in European multi-centre studies	Poster n°18 20 Nov. 10:45-11:15
Carlos Rodrigues	in depth	AI in Radiation Oncology: A five-minute map of where it actually works	Poster n°19 20 Nov. 10:45-11:15
Sara Salesi	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Molecular dynamics supports differential structural engagement of suzetrigine across human voltage-gated sodium channel paralogs	Poster n°20 20 Nov. 10:45-11:15
Eléonore Schiltz	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Decoding memory from thousands of neurons: perspectives on generalization tests for brain–computer interfaces	Poster n°21 20 Nov. 12:30 - 13:30

Presenter	Submitter company	Title	Poster n° and session
Maciej Skorski	University of Luxembourg	Uncertainty-Aware AI Clinical Narratives for Cardiovascular Risk	Poster n°22 Flash talk 19 Nov. 16:23
Gleb Svinin	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	A Putative Compensatory EGR1-Centered Module in Mild Cognitive Impairment along the Alzheimer's Disease Trajectory Is Attenuated in APOE ε4 Carriers	Poster n°23 20 Nov. 12:30 - 13:30
Dilara Uzuner Odongo	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Cell-type and disease-stage dependence of sex differences in the Alzheimer's disease cortex	Poster n°24 20 Nov. 12:30 - 13:30
Astrid Vincze	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Exploring the co-selection of antibiotic, metal, and biocide resistances across environmental and human microbiomes	Poster n°25 20 Nov. 12:30 - 13:30
Prasan Yapa	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	From free text to standardized records: bioc2omop for clinical decision support	Poster n°26 20 Nov. 12:30 - 13:30
Alberto Zancanaro	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	When Models Recognize Patients, Not Pathology: Quantifying Data Leakage in Federated Alzheimer's Disease Detection	Poster n°27 20 Nov. 12:30 - 13:30
Qian Zhang	Luxembourg National Data Service	From Shared Challenges to Shared Practices: Building Luxembourg's Data Management Community	Poster n°28 20 Nov. 12:30 - 13:30
Ulrike Weirauch	Fraunhofer Institute for Cell Therapy and Immunology IZI	CERTAINTY: A Cellular Immunotherapy Virtual Twin for Personalised Cancer Treatment	Poster n°29 Flash talk 20 Nov. 11:53
Kavya Singh	University of Liège	Individual-specific network diffusion reveals biological and therapeutic heterogeneity in mixed Alzheimer's disease and LATE pathology	Poster n°30 Flash talk 20 Nov. 11:45
Tsz Pan Tong	University of Luxembourg	FlowGRN: Scalable and Dropout-Robust Gene Regulatory Network Inference via Flow Matching-Based Trajectory Reconstruction	Poster n°31 Flash talk 19 Nov. 16:15
Soumyabrata Ghosh	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	Adept: a secure, no-code platform for exploring governed biomedical datasets across the translational pipeline	Poster n°32 Flash talk 19 Nov. 18:30
Irina Balaur	Luxembourg National Data Service	Supporting Value Creation from Secondary Use of Data with the LNDS Service Portfolio	Poster n°33
Aquila Farahmand	Luxemboug Centre for Systems Biomedicine, University of Luxembourg	From scattered data to smarter care: AI for neurodegenerative disease	Poster n°34

Abstracts

Poster n° 1

LLM-assisted reconstruction of a mechanistic disease map from primary literature: a case study

Marcio Luis Acencio

No affiliation

Disease maps are diagrammatic and computer-readable representations of disease-specific molecular mechanisms. By integrating individual pieces of mechanistic evidence into a pathway-level overview, these maps can facilitate the investigation of molecular mechanisms regulating disease progression. Disease maps are usually built incrementally through manual biocuration of scientific articles describing experimentally observed molecular interactions associated with the disease of interest. Two crucial steps in disease map construction are (i) the extraction of relevant mechanistic knowledge and (ii) its representation using systems biology computational and graphical notations.

While the extraction of relevant interactions is largely performed manually and relies on the biocurator's domain knowledge, their representation can be assisted by computational tools such as CellDesigner (CD), which help translate acquired knowledge into standardized notations. Nevertheless, biocurators must still learn how to represent biological knowledge using graphical and computational formats such as Systems Biology Graphical Notation (SBGN) and CD Extensible Markup Language (CD XML). Both the extraction and representation of biological knowledge are laborious and time-consuming, and some degree of automation is therefore desirable to accelerate the production of highly curated maps. In this context, Large Language Models (LLMs), such as Generative Pre-trained Transformers (GPTs), could serve as valuable assistants for biocuration and disease map construction.

I present here a case study evaluating the potential of an LLM, GPT-5.6 Luna, to accelerate disease map construction. Keratinocyte and fibroblast submaps of the Inflammatory Skin Disease Map (ISD map,) were used as training and test cases, respectively. Using the keratinocyte submap and its supporting literature, the LLM derived biocuration rules for extracting and representing mechanistic knowledge, which were formalized in YAML, while CD XML representation rules were captured separately. These rules were then applied to the literature supporting the fibroblast submap, without access to the corresponding actual curator-generated map. The LLM reconstructed the principal causal architecture linking CGRP signalling through the RAMP1/CALCRL receptor complex to ANPEP/CD13 expression, together with IL-4/IL-4R autocrine regulation, and the generated CD XML was successfully validated. The actual and reconstructed maps agreed on the main mechanisms, while differences primarily concerned biological granularity, including generic versus CALCA/CALCB-specific representation of CGRP and a more compact representation of IL-4 biosynthesis by the LLM. In addition, the LLM captured the experimentally supported effect of retinoic acid on CD13 activity, which was absent from the curator-generated map. Although automated annotation of entities with official gene identifiers and linkage of supporting PMIDs to individual reactions remain to be implemented, this case study supports the potential of LLMs as assistants to biocurators in disease-map construction.

Poster n° 2

UNCOVERING MULTIMODAL PHENOTYPIC STRUCTURE IN INDIVIDUALS AT RISK FOR PARKINSON'S DISEASE USING SIMILARITY NETWORK FUSION

Esther Ademola (1,2), Ahmed Hemedan (1), Mahdi Bohlouli (5), Rejko Krüger (1,3,4)

1) Luxembourg Research Clinic, Transversal Translational Medicine, Luxembourg Institute of Health (LIH), Strassen, Luxembourg
2) Faculty of Science, Technology and Medicine (FSTM), University of Luxembourg, Esch-sur-Alzette, Luxembourg
3) Parkinson Research Clinic, Centre Hospitalier de Luxembourg (CHL), Luxembourg City, Luxembourg
4) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg
5) Research and Innovation, Petanux GmbH, Bonn Germany

Parkinson's disease (PD) is characterised by substantial heterogeneity in clinical presentation, biological features and disease progression [1]. Individuals at increased risk of PD may therefore exhibit diverse profiles across clinical and biological domains. Characterising this multimodal heterogeneity may provide a complementary perspective to conventional outcome-based approaches to risk assessment. We investigated whether unsupervised integration of multimodal baseline data could characterise participant similarity and reveal structure associated with subsequent PD phenoconversion.

We analysed 752 participants from the Parkinson's Progression Markers Initiative (PPMI) who had no PD diagnosis at baseline and for whom data were available across four modality blocks: clinical measures, dopamine transporter (DAT) imaging, MRI and genetic measures. Of these participants, 55 subsequently converted to PD during follow-up. Within each modality, participant affinity networks were constructed independently and integrated using Similarity Network Fusion (SNF) [2]. The resulting fused network was explored using spectral embedding and clustering without incorporating subsequent phenoconversion status. Cluster separation was quantified using silhouette scores across two to six candidate clusters, with a sensitivity analysis repeated after excluding motor clinical measures. Finally, conversion status was overlaid onto the outcome-independent spectral representation, and local enrichment of future converters was assessed using a 20-nearest-neighbour permutation analysis.

The fused multimodal representation showed heterogeneous participant structure without strong separation into discrete groups. Spectral clustering across two to six clusters yielded low silhouette values (0.018–0.022), indicating limited separation between candidate participant groups. Excluding motor clinical measures resulted in even weaker separation, with silhouette values close to zero. Despite the absence of clearly separated subgroups, future converters showed local enrichment within the spectral representation: 11.4% of their 20 nearest neighbours were also converters, compared with 7.2% under random permutation of conversion labels (permutation p=0.0008).

These findings indicate that multimodal similarity in this cohort was not well represented by a small number of strongly separated participant subgroups. Instead, the outcome-independent fused representation captured heterogeneous participant structure across clinical, imaging and genetic domains, within which future converters showed non-random local organisation. This provides an exploratory indication that multimodal similarity structure may capture information relevant to subsequent phenoconversion beyond discrete subgroup assignment. Ongoing work will determine whether network-derived representations are associated with time to conversion and whether they can contribute to longitudinal risk modelling.

References

1) *Berg D, Borghammer P, Fereshtehnejad SM, et al. Prodromal Parkinson disease subtypes—key to understanding heterogeneity. Nat Rev Neurol. 2021;17:349–361. doi:10.1038/s41582-021-00486-9.*

2) *Wang B, Mezlini AM, Demir F, et al. Similarity network fusion for aggregating data types on a genomic scale. Nature Methods. 2014;11:333–337. doi:10.1038/nmeth.2810.*

TWO-TIER BIOMEDICAL RELATION EXTRACTION PIPELINE

Basel Alshaikhdeeb (1), Prasan Yapa (1), Jaleh Shoshtarian Malak (1), Soumyabrata Ghosh (1), Anne Lorant (1), Venkata Satagopam (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 6, avenue du Swing, Esch-sur-Alzette, L-4367, Luxembourg

In medical Natural Language Processing (NLP), the main task is to identify medical/biomedical entities, including diseases/symptoms, medications/drugs, lab test/results, and others. Such an identification task is known as Named Entity Recognition (NER). However, an additional and consequential task is Relation Extraction (RE), in which the potential relationships among the extracted entities are identified. The importance of such a task is represented through the detection of Drug-Drug Interactions (DDIs), which play a vital role in patient safety, where these interactions may harm the patient and boost the cost of healthcare. Similarly, investigating the Protein-Protein Interactions (PPIs) would lead to a better understanding of the biological processes and disease trajectories that facilitate drug discovery. Lastly, recognizing clinical relations (*e.g., lab results, tests, medications/drugs,* etc.) contributes toward automating the process of structuring and standardizing unstructured clinical documents, which facilitates clinical decision-making [1].

Although RE might seem a straightforward classification task, it has different aspects that increase the challenges. Because of document complexity, RE can be handled at the document level rather than the sentence level, where the document level has longer input (*i.e., a paragraph or passage*) than the sentence level (*i.e., a sentence*). Within the document-level, the task will be much more complicated, involving implicit relations such as 'it activates inflammatory response' where 'it' refers to an entity that has been mentioned in earlier sentences. This indeed would require longer memorization of the sequences [2]. On the other hand, the RE task can be treated as a binary classification task (*i.e., Relation vs. No Relation*) or a multi-class classification task (*i.e., Activate, Inhibit, Upregulate, Downregulate, etc.).* Similarly, the RE can be treated as a separate pipeline method or a joint extraction method that is integrated with the NER task itself [3, 4].

The aforementioned challenges will be addressed in this paper within the context of the FAIRClinical project (https://fairclinical.github.io/), with the aim of supporting the transformation of supplementary data files and published case reports into more structured, machine-actionable, and FAIR data. For this purpose, a systematic and rigorous benchmarking for the RE task will take place to identify the most accurate and efficient technique. The benchmarking will include various models and publicly available gold-standard datasets that cover a wide range of medical/biomedical relations. Lastly, the identification of the best method for RE will contribute toward deploying and running a two-tier RE pipeline on the FAIRClinical corpus of supplementary data files.

References

1) *Yadav, S., et al., Relation extraction from biomedical and clinical text: Unified multitask learning framework. IEEE/ACM transactions on computational biology and bioinformatics, 2020. 19(2): p. 1105-1116.*

2) *Chaturvedi, R., et al. Temporal relation extraction in clinical texts: a span-based graph transformer approach. in Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers). 2025.*

3) *Fei, H., et al., A span-graph neural model for overlapping entity relation extraction in biomedical texts. Bioinformatics, 2021. 37(11): p. 1581-1589.*

4) *Ye, Q., et al., Subsequence and distant supervision based active learning for relation extraction of Chinese medical texts. BMC Medical Informatics and Decision Making, 2023. 23(1): p. 34.*

Building A Responsible Personalised Digital Therapeutic Solution For Sleep Disturbances: BRAINPATH

Beatriz García Santa Cruz (1), Soumyabrata Ghosh (1), Artur Jorge Oliveira de Carvalho (1), Sandrine Medves (1), Philip Santangelo (2), Venkata Satagopam (1)

[1] Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 avenue du Swing, L-4367 Esch-sur-Alzette, Luxembourg.
[2] Department of Behavioural and Cognitive Sciences, Faculty of Humanities, Education and Social Sciences (FHSE), University of Luxembourg, Maison des Sciences Humaines, 11 Porte des Sciences, L-4366 Esch-sur-Alzette, Luxembourg

Sleep is often profoundly affected by traumatic experiences, and sleep disturbances can persist long after other post-traumatic symptoms have improved. People who have experienced combat, emergency work, violence, accidents, forced displacement, or repeated stressful events may experience recurrent nightmares, fearful awakenings, fragmented sleep, and non-restorative nights. These symptoms can contribute to daytime exhaustion, distress, and avoidance, further affecting quality of life and recovery. Effective non-pharmacological interventions for nightmares and sleep disturbances are available, yet many people affected by these symptoms do not receive appropriate treatment. Access to trained specialists can be limited, waiting lists can be long, and support is often least accessible precisely when it is most needed—for example, during the night following a distressing dream [1].

BRAINPATH aims to address this gap by developing a responsible and personalised digital therapeutic (DTx) based on established therapeutic approaches [2]. The proposed intervention will be delivered through a structured programme over several weeks and will be designed for use at home, either independently or with clinical supervision. It will combine psychoeducation, techniques to reduce distress before sleep, strategies for managing night-time awakenings, and guided therapeutic work targeting recurrent nightmares. Our solution will include two interconnected components: **a mobile application for patients and an interactive dashboard for clinicians.** Through the mobile application, patients will access the therapeutic programme and track relevant aspects of their sleep and progress. The clinician dashboard will allow treating professionals to monitor progress and adapt the intervention when appropriate. The aim is therefore not to replace professional care, but to **extend its reach and provide support between clinical encounters**.

The development of BRAINPATH will follow an evidence-oriented framework for digital therapeutics, treating design, development, testing, and post-deployment monitoring as interconnected stages [3]. Evidence requirements will be considered from the beginning of development to ensure that the resulting intervention can be evaluated rigorously [4]. A central principle of BRAINPATH is **co-design**. Prospective users, clinicians, and researchers in behavioural and cognitive sciences will contribute to the development process through structured co-design activities. Their input will inform the cultural and linguistic adaptation of therapeutic content, the wording of emotionally sensitive material, the usability of the interfaces, and the pace and sequencing of the therapeutic programme.

Attention will be given to **safety, clinical oversight, privacy, and appropriate escalation pathways**, recognising that users engaging with trauma-related material may experience distress. Through the integration of evidence-based therapeutic approaches, digital technology, clinical supervision, and user-centred design, BRAINPATH aims to provide a **safe, accessible, and personalised digital pathway for people experiencing trauma-related sleep disturbances**.

References

1. Morgenthaler TI, Auerbach S, Casey KR, et al. Position Paper for the Treatment of Nightmare Disorder in Adults: An American Academy of Sleep Medicine Position Paper. *Journal of Clinical Sleep Medicine*. 2018;14(6):1041–1055.

2. Krakow B, Zadra A. Clinical Management of Chronic Nightmares: Imagery Rehearsal Therapy. *Behavioral Sleep Medicine*. 2006;4(1):45–70.
3. Kim M, Patrick K, Nebeker C, et al. The Digital Therapeutics Real-World Evidence Framework Approach for Guiding Evidence-Based Digital Therapeutics Design, Development, Testing, and Monitoring. *Journal of Medical Internet Research*. 2024;26:e49208.
4. Digital Therapeutics Alliance. *Digital Therapeutics: Clinical Evidence Basics*. 2023.

The IDERHA Platform: Integrating federated data governance and analysis for health research

Hanna Ćwiek-Kupczyńska (1), François Ancien (1), Dennis Asamoah (1), Valentin Grouès (1), Sascha Herzinger (1), François Xavier Owona (1), Nirmeen Sallam (1), Abetare Shabani (1), Jaleh Shoshtarian Malak (2), Venkata Satagopam (1) & the IDERHA Consortium

1) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg
2) Luxembourg Institute of Health (LIH)

Health data research increasingly requires diverse multi-source datasets, while data privacy demands that sensitive information always remains protected under institutional control. A federated approach to data analysis answers this challenge by moving the computation to the data, and is a promising route to cross-institutional studies.

We present the IDERHA Platform [1] — a federated infrastructure that integrates data governance and federated learning in a single standards-based framework, ensuring security, sovereignty and control over data for data providers, while offering a unified interface to data for data users. The platform supports health-related research by enabling multiple steps of the data lifecycle: from dataset discovery, to access request and permit issuance, to federated analysis and learning. It aligns with the FAIR principles, GDPR, and the spirit of the EHDS regulation [2].

The platform's architecture integrates established technologies and standards [3] to enforce data sovereignty, so that data providers retain control at every stage — each institution advertises its data via a HealthDCAT-AP-based federated catalogue, implements local policy-based access control (ODRL policies), issues permits (GA4GH Visas) [4], which enable access authorization, and approves submitted analysis code before its execution. Federated analysis is orchestrated via NVFlare [5] extended with custom governance plugins, and runs in secure isolated environments from which only derived outputs are released, and returned to the user after central aggregation. All platform events are logged for audit compliance.

The platform is being validated through a lung cancer study integrating clinical and imaging data (OMOP CDM, DICOM) from European hospitals for risk prediction and malignancy diagnosis use cases.

The IDERHA project is supported by the IHI JU and its members, under grant agreement No 101112135.

References

1) *Boutsma E, Herzog K, Rancourt RC, et al. A sustainable platform for federated health data access, AI innovation, and regulatory acceptance in alignment with the European Health Data Space principles. Open Res Europe 2026;6:181. doi:10.12688/openreseurope.23459.3*
2) *Hussein R, Balaur I, Burmann A, Ćwiek-Kupczyńska H, et al. Getting ready for the European Health Data Space (EHDS): IDERHA's plan to align with the latest EHDS requirements for the secondary use of health data. Open Res Europe 2024;4:160. doi:10.12688/openreseurope.18179.1*
3) *Shoshtarian Malak J, von Berlin L, Balaur I, et al. Bridging Gaps in Standards for the Secondary Use of Health Data: Exploratory Standard, Tool Assessment, and Feasibility Study. JMIR Form Res 2026;10:e88619. doi:10.2196/88619*
4) *Voisin C, Linden M, Dyke SOM, et al. GA4GH Passport standard for digital identity and access permissions. Cell Genomics 2021;1(2):100030. doi:10.1016/j.xgen.2021.100030*
5) *Roth HR, Cheng Y, Wen Y, et al. NVIDIA FLARE: Federated Learning from Simulation to Real-World. arXiv:2210.13291 (2022).*

Human-in-the-Loop AI Framework for Automated Quantification of Cell Populations in Patient-Derived Midbrain Organoids

Shweta Vikas Lokhande (1), Siavash Damari (2), Holger Voos (3), Noemi Asfogo (2), Javier Jarazo (4), Jens C. Schwamborn (2,4), Shekoufeh Gorgi Zadeh (2)

1) Faculty of Science, Technology and Medicine, University of Luxembourg
2) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg
3) Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg
4) OrganoTherapeutics SARL, Esch-sur-Alzette, Luxembourg

Reliable quantification of cellular populations in patient-derived midbrain organoids is important for studying disease-associated changes, but manual analysis of fluorescence microscopy stacks is time-consuming and difficult to scale. Therefore, we developed a human-in-the-loop AI framework to quantify total cells and tyrosine hydroxylase-positive (TH+) dopaminergic neurons. Because TH staining labels not only cell bodies but also axons and dendrites, counting neurons directly in the TH channel is unreliable, especially in dense regions. The framework therefore first segments all nuclei and then classifies each cell as TH+ or TH−, so that each neuron is counted exactly once.

For development and evaluation, we used 60 patient-derived midbrain organoids: 30 embedded wild-type, 15 non-embedded wild-type and 15 non-embedded knockout. Nuclei were segmented from individual two-dimensional planes of the nuclear-fluorescence Z-stacks using Cellpose-SAM [1], stitched across adjacent planes to preserve cell identities, and refined with a binary nuclear mask (NucMask) obtained by difference-of-Gaussians filtering with fixed intensity and size thresholds [2]. In a dedicated interface, experts removed false detections, added missed nuclei, corrected inaccurate or fused instances and assigned TH labels; these corrections were kept as reusable reference data for evaluation and training. Of the 60 organoids, 19 were reviewed across three adjacent Z-planes (172 image regions, 18,514 nuclei) and split at the organoid level into training, validation and held-out test sets. TH status was then predicted for each segmented cell by a pretrained ResNet18 [3] fine-tuned on the expert labels; input representation, target-cell masking, fine-tuning depth and Z-plane context were selected on validation data only.

Permissive Cellpose-SAM inference improved nucleus recovery but could produce boundaries extending beyond the visible nuclear signal; tightening them with NucMask gave the best balance between instance-level agreement and cell-count accuracy. On the held-out segmentation test set, the validation-selected Cellpose-SAM + NucMask pipeline achieved a mean AP@0.5 of 0.715 and a relative cell-count error of 3.3%, versus 0.654 – 0.683 and 4.8 – 7.4% for single-stage fine-tuned Cellpose-SAM models. On the held-out test set, the TH classifier achieved an F1 score of 0.669 ± 0.015 and an average precision of 0.734 ± 0.005 across three random seeds.

A preliminary post hoc comparison of automated and expert-corrected masks found that 86.4% of matched nuclei were essentially unchanged (IoU > 0.95); substantial corrections concentrated in denser regions, mainly merged or under-separated nuclei. This suggests that expert effort can focus on a minority of difficult instances, and that future improvements should refine these selectively rather than replace the pipeline. Overall, the framework provides a practical basis for scalable, expert-guided organoid image analysis.

References

1) *Pachitariu M, Rariden M, Stringer C. Cellpose-SAM: superhuman generalization for cellular segmentation. bioRxiv 2025. doi:10.1101/2025.04.28.651001*

2) Jarazo J, Santos da Silva E, Glaab E, Perez-Bercoff D, Schwamborn JC. SARS-CoV-2 infection induces dopaminergic neuronal loss in midbrain organoids. J. Neurochem. 2026, 170: e70486.

3) *He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc. IEEE CVPR 2016, pp. 770–778.*

MitoSegAI: An Integrated Deep Learning Pipeline for the Segmentation and Quantification of Mitochondrial Ultrastructure in 3D Electron Microscopy

Beatriz García Santa Cruz (1), Carmen Lahr (1), Alexander Skupin (1), Venkata Satagopam (1)

[1] Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 avenue du Swing, L-4367 Esch-sur-Alzette, Luxembourg.

Mitochondrial morphology and internal membrane architecture are closely linked to organelle function. Alterations in cristae number and shape, membrane organisation, and matrix structure have been associated with ageing, neurodegeneration, and metabolic disease [1]. Three-dimensional electron microscopy (3D EM) resolves these structures at nanometre scale, yet converting images into quantitative measurements still relies largely on manual delineation. This process is slow and subjective, and it does not scale to large tomogram collections. Thin, densely packed structures such as cristae are particularly difficult to trace consistently. General-purpose segmentation tools typically yield whole-organelle masks. They neither distinguish mitochondrial sub-compartments nor provide the per-organelle measurements that biological interpretation requires. As a result, ultrastructural datasets that are costly to acquire are often analysed only qualitatively, or on a small subset of organelles.

MitoSegAI addresses this gap with a modular, reproducible pipeline that takes raw EM volumes to quantitative, per-mitochondrion morphometrics. It combines two components. The first is a two-stage segmentation cascade based on the U-Net architecture [2]. A localisation network identifies each mitochondrion within a region of interest, and a class-conditioned network then delineates three sub-compartments: the intermembrane space, cristae, and matrix. The first-stage prediction is passed to the second stage as a continuous confidence map rather than a thresholded mask, so the inputs seen during training match those encountered at inference. The second component is a quantification module that reports morphometric features in physical units. These include compartment volumes and volume ratios, cristae number and dimensions, intermembrane-space thickness, and organelle width. Predictions can also be exported to IMOD [4] for expert inspection and correction within established ultrastructural workflows.

MitoSegAI was developed around a validation strategy designed for small but richly annotated datasets. Fully hand-annotated mitochondria are scarce, so performance is assessed by leave-one-mitochondrion-out cross-validation, in which every annotated sample contributes to both training and held-out evaluation. Each sub-compartment is compared with manual annotations using standard overlap metrics. These metrics under-represent thin structures, so cristae are additionally evaluated with the topology-aware centerline Dice (clDice) metric [3]. This metric assesses whether their structure is captured rather than whether their boundaries coincide exactly. Preliminary results show reliable delineation of the intermembrane space and matrix, and good preservation of cristae topology. Derived measurements fall within biologically plausible ranges, and round-trip conversion to and from IMOD reproduces the original segmentations exactly.

Particular attention has been paid to transparency about limitations. These include sensitivity to mitochondrial orientation within the cropped region, the variance inherent to small evaluation sets, and the current reliance on prior organelle localisation. MitoSegAI is implemented in Python/PyTorch and includes scripts for high-performance computing environments. By integrating cascade segmentation, physical-unit quantification, and interoperability with expert tools, MitoSegAI aims to make quantitative analysis of mitochondrial ultrastructure faster, more reproducible, and scalable.

References

1. *Jenkins BC, Neikirk K, Katti P, et al. Mitochondria in disease: changes in shapes and dynamics. Trends in Biochemical Sciences. 2024;49(4):346–360.*

2. *Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015:234–241.*

3. *Shit S, Paetzold JC, Sekuboyina A, et al. clDice – a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2021:16560–16569.*

4. *Kremer JR, Mastronarde DN, McIntosh JR. Computer Visualization of Three-Dimensional Image Data Using IMOD. Journal of Structural Biology. 1996;116(1):71–76.*

SYNTHETIC TABULAR DATA GENERATION FRAMEWORK FOR CLINICAL DECISION SUPPORT SYSTEMS

Aditya Krishna, Anne Lorant, Venkata Satagopam

Luxembourg Centre for Systems Biomedicine

Clinical datasets are useful for building clinical decision support systems. To build robust and more generalisable decision-making models, such datasets must be centralised to a single location. However, there are risks involved in transporting such datasets from their original location concerning security, privacy and legitimacy. The use of synthetic data, as a replacement of real clinical data, is a potential solution to this problem, where the modelling utility of the original dataset for downstream tasks and the privacy of the patients is preserved.

This project presents an extensible framework which trains a generative deep learning model, produces a synthetic dataset, evaluates the synthetic dataset against the original dataset to calculate a quality score, and uses this score to search for optimal hyperparameters for model training and for generating the highest quality synthetic data. The framework consists of three customisable components. In the first component, a CTGAN [1] or TVAE [1] model architecture must be selected depending on its suitability to the original dataset. This model is trained and used to generate a synthetic dataset. The second component calculates the overall quality score of the synthetic dataset by evaluating a customisable set of metrics relating to statistical fidelity, downstream modelling utility and privacy risk. The results are consolidated into the quality score. The third component uses this as an objective score for hyperparameter search. The score resulting from the set of hyperparameter values used in model training is used as a trial to search for the optimal set of hyperparameters for a higher score. This hyperparameter search is accomplished using Optuna [2]. The search can be modified by selecting from a predetermined set of hyperparameters for which hyperparameters must be searched and which must remain fixed. The number of search trials performed may also be set. This framework is suited for clinical datasets which are organised in tabular format, contain numerical and categorical variables, and may contain missing values. This framework can be easily extended to train other types of model architectures, calculate additional evaluation metrics for the quality score and consider additional hyperparameters for the hyperparameter search.

References

1) *Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante & Kalyan Veeramachaneni. 2019, Modeling Tabular Data using Conditional GAN. Advances in Neural Information Processing Systems.*

2) *Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta & Masanori Koyama. 2019, Optuna: A Next-generation Hyperparameter Optimization Framework. In KDD.*

AID-Care: Artificial Intelligence for neuroDegenerative diseases Care

Anne Lorant (1), Soumyabrata Ghosh (1), Sandrine Medves (1), AID-CARE consortium, Venkata Satagopam (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg, 7 avenue des Hauts-Fourneaux, Belval Campus, L-4362 Esch-sur-Alzette, Luxembourg

Clinical data offer major opportunities for artificial intelligence (AI), but their use is often limited by heterogeneous formats, unstructured information, and complex governance requirements. AID-Care is a Luxembourg–Japan collaborative project developing an interoperable, AI-enabled infrastructure for multimodal data in dementia and related neurodegenerative disorders.

The project integrates clinical narratives, laboratory data, imaging and associated reports. AI-supported pipelines are used to extract clinically relevant information from unstructured records and transform it into standardized, machine-readable representations. Data harmonization builds on established standards, including OMOP Common Data Model and HL7 FHIR, within a hybrid data-space architecture designed to support secure and compliant data use under European and Japanese regulatory frameworks.

AID-Care further connects this harmonized data layer with scalable computational resources and user-facing AI applications, including analytical dashboards and conversational interfaces.

By combining AI, interoperability, data governance and high-performance computing, AID-Care provides a reusable framework for transforming heterogeneous clinical data into analysis-ready resources and supporting future AI-driven research in neurodegenerative diseases.

Patient-Specific Digital Twins and Semi-Markov Control for Personalized Extracorporeal Liver Support: A Clinical Decision-Support Framework

Taras Lukashiv (1), Soumyabrata Ghosh (1), Igor V. Malyk (2), Ahmed A. Hemedan (3), Venkata P. Satagopam (1)

1) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Belvaux, Luxembourg
2) Department of Mathematical Problems of Control and Cybernetics, Yuriy Fedkovych Chernivtsi National University, Chernivtsi, Ukraine
3) Transversal Translational Medicine, Luxembourg Institute of Health, Strassen, Luxembourg

Background. Intermittent extracorporeal liver support (ELS) is used in acute and acute-on-chronic liver failure to remove albumin-bound toxins and bridge patients to transplantation or hepatic recovery, with sessions conventionally prescribed on fixed schedules. Its effect is complicated by marked inter-patient variability in bilirubin kinetics and by post-treatment rebound caused by redistribution between physiological compartments. Standardized schedules may therefore be unnecessarily intensive for some patients and insufficiently adapted for others. We developed a patient-specific digital-twin framework that translates semi-Markov stochastic stabilization theory into clinically interpretable decision support for personalized ELS.

Methods. Digital twins were constructed from published patient-specific pharmacokinetic parameters for eight patients using a two-compartment model of conjugated bilirubin [1]. For each patient, clinically feasible five-day treatment protocols were evaluated by varying treatment start offset, session duration, inter-session interval, and effective clearance intensity. We ranked the protocols using a hierarchical multi-objective criterion that balanced bilirubin control, physiological safety, post-treatment rebound, cumulative treatment time, integrated clearance effort, and treatment smoothness. We accepted personalization only when the improvement over the original clinical protocol exceeded a predefined clinically meaningful threshold; otherwise, we retained the standard schedule. After selecting the protocol, we applied an adaptive semi-Markov feedback controller [2] to provide bounded local corrections for stochastic physiological deviations around the personalized operating point while preserving the selected clinical schedule.

Results. The optimization revealed substantial physiological heterogeneity despite all patients starting from the same standard ELS protocol. Personalized schedules were selected for seven of eight patients, while the original protocol was retained for one patient because the predicted benefit was marginal. Among personalized cases, the composite objective improved by 5.6–96.6%. Selected schedules ranged from 4 to 6 h per session, 22 to 28 h between sessions, and 0.75 to 1.00 in relative clearance intensity. Safety time-in-range remained 99.8–100%, while preferred therapeutic time-in-range reached approximately 80% in the best-performing patients. Importantly, improved performance did not require uniform escalation of treatment intensity: several patients benefited from shorter sessions, altered timing, or reduced clearance. Adaptive semi-Markov correction further reduced transient deviations while preserving the optimized protocol. Pareto analysis demonstrated that no single schedule simultaneously maximized therapeutic benefit and minimized treatment burden across all patients.

Conclusions. Patient-specific digital twins combined with multi-objective protocol optimization and adaptive semi-Markov regulation provide a clinically transparent framework for personalized ELS. The approach separates long-term protocol planning from short-term stochastic correction, supports conservative personalization only when meaningful benefit is predicted, and exposes explicit trade-offs between efficacy, safety, and treatment burden. This creates a practical pathway toward data-driven clinical decision-support systems for intermittent extracorporeal therapies.

References

1) *Jung A, Krisper P, Haditsch B, Stauber RE, Trauner M, Holzer H, Schneditz D. Bilirubin kinetic modeling for quantification of extracorporeal liver support. Blood Purif. 2006;24(4):413–422. doi:10.1159/000094575.*
2) *Lukashiv T, Malyk IV, Hemedan AA, Satagopam VP. Optimal Control of Stochastic Dynamic Systems with Semi-Markov Parameters. Symmetry. 2025;17(4):498. doi:10.3390/sym17040498.*

The ELIXIR-LUXEMBOURG training platform

Tyler McInnes (1), Milena Zizovic (2) Eléonore Schiltz (1), Vilem Ded (1), Hana Marčetić (1), Nima Gozalpour (1), Natalia Martinez Panadero (1), Pinar Alper (2), Quian Zhang (2), Zahra Hashemi (2), Nene Djenaba Barry (2), Xavier Adam (2), Roland Krause (1)

1) LCSB, Université du Luxembourg
2) Luxembourg National Data Service

Sitting at the interface of biology and computer science, methods in bioinformatics progress at an extremely rapid pace. Additionally, researchers who have traditional wet lab skills may choose to acquire computational data analysis skills at some point in their career. High quality, targeted workshops are an efficient and effective method of making bioinformatics skills available to all levels of the research community[1].

Since 2016 ELIXIR-Luxembourg has been providing high-quality data science training to research staff and students in the form of workshops. Workshops delivered via ELIXIR-LU are typically between two and three days in length, and are comprised of a mixture of theory, demonstration, and practical work. ELIXIR-LU workshops can be broadly grouped into three categories: *fundamental programming* skills (*e.g.,* core aspects of various computer programming languages, how to think about programming, how to practically manage working in a digital environment), *advanced workflows and data handling* (specific workflows that address a known task *e.g.,* spatial transcriptomics, appropriate storage and treatment of data), and *training* (to improve the overall teaching and training capacity of staff, *e.g.,* the ELIXIR-Goblet *Train-the-trainer* system).

Data science is undergoing a fundamental shift as researchers of all levels are using LLMs to write code. Despite the increasingly high quality of this code, without fundamental knowledge knew users are vulnerable to making critical mistakes. There is widespread concern, and growing evidence, that relying on LLMs reduces learning outcomes of individuals[2][3]. We aim to acknowledge the benefit of LLMs for coding support, while striving to ensure researchers gain and retain key competencies in data science.

This work presents an overview of the existing ELIXIR-Luxembourg training portfolio and argues that our focus on fundamental skills is critical for users who are adopting LLMs as coding assistants.

References

[1] Jordan et al., "Short-Format Workshops Build Skills and Confidence for Researchers to Work with Data."
[2] Bastani et al., "Generative AI without Guardrails Can Harm Learning."
[3] Stromberg et al., "The Generative AI Learning Penalty."

Benchmarking hosted and local System One models on regulated and sensitive text

Shaman Narayanasamy (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg (shaman.narayanasamy@uni.lu)

System One models (SOMs) became popular in September 2026, after the company TypeSafe launched Jev, their hosted model. SOMs answer questions with either a yes/no, or a probability. The main advantages of these models are speed and cost compared to frontier LLMs. SOMs could be an effective strategy for texts from regulated industries and settings, such as clinical, legal, policy, finance and cybersecurity. However, hosted services are unsuitable for the aforementioned sensitive records as they may not leave the premises. To that end, I aimed to test whether open models run locally can replace highly capable hosted frontier SOMs.

I evaluated Jev and seven open models of 0.3 billion (B) to 27B parameters on identical items with frozen question wording. A total of 14 data sets with 28,063 scored records were evaluated, covering FDA recalls (1), pharmacovigilance reports (2), manually annotated peer-reviewed biomedical literature (3), privacy policies (4), insurance policies (4), contracts (5), tax statutes (6), central-bank communication (7) and prompt injection (8). In addition, I controlled for memorisation using FDA recalls and manually annotated biomedical literature that post-dated model releases. Each open SOM was compared with Jev by paired AUROC differences with cluster-bootstrap intervals, pooled by random effects against a margin of ±0.05 fixed in advance.

Two open 27B models matched Jev's performance: the pooled AUROC difference was +0.000 (95% CI −0.010 to +0.011) for Decider and −0.004 (−0.014 to +0.005) for InnerJev. Open 9B models trailed by 0.018 to 0.025, and the two models under 1B parameters reached a mean AUROC of 0.64 and 0.69. On PubMedQA every open model of 9B or more reached an AUROC of 0.95 to 0.97 (Jev 0.96), and on FDA recall classification 0.77 to 0.84 (Jev 0.81). Jev changed its label on 314 of 29,738 identical repeat requests (1.06%), while the local SOMs altered none. On records published after the model releases, Jev and all five open models of 9B or more ranked every yes/no pair correctly on 66 biomedical articles, and reached an AUROC of 0.75 to 0.79 on 70 FDA recalls. Importantly, a yes/no cut-off on the probability, tuned on one data set, lost accuracy on the others for every model.

To conclude, I show that an open 27B model running on a single A100 GPU, at about 0.1 s per item, can stand in for the hosted model, and it returned the same answer on every repeat. Overall, replacing hosted models with local GPU-hosted implementations is a viable strategy for running SOMs on sensitive text.

References

1) *U.S. Food and Drug Administration. openFDA enforcement reports. openFDA https://open.fda.gov/ (2026).*
2) *Sun, Z. et al. PHEE: A Dataset for Pharmacovigilance Event Extraction from Text. Preprint at https://doi.org/10.48550/arXiv.2210.12560 (2022).*
3) *Jin, Q., Dhingra, B., Liu, Z., Cohen, W. W. & Lu, X. PubMedQA: A Dataset for Biomedical Research Question Answering. Preprint at https://doi.org/10.48550/arXiv.1909.06146 (2019).*
4) *Guha, N. et al. LegalBench: A Collaboratively Built Benchmark for Measuring Legal Reasoning in Large Language Models. Preprint at https://doi.org/10.48550/arXiv.2308.11462 (2023).*
5) *Koreeda, Y. & Manning, C. D. ContractNLI: A Dataset for Document-level Natural Language Inference for Contracts. Preprint at https://doi.org/10.48550/arXiv.2110.01799 (2021).*
6) *Holzenberger, N., Blair-Stanek, A. & Van Durme, B. A Dataset for Statutory Reasoning in Tax Law Entailment and Question Answering. Preprint at https://doi.org/10.48550/arXiv.2005.05257 (2020).*
7) *Shah, A. et al. Words That Unite The World: A Unified Framework for Deciphering Central Bank Communications Globally. Preprint at https://doi.org/10.48550/arXiv.2505.17048 (2025).*
8) *deepset. prompt-injections. Hugging Face (2023).*

Characterizing GBA-Associated Parkinson's Disease through Transcriptomics and Morphological Profiling in Midbrain Organoids

Francesco Nasta (1,2), Enrico Glaab (1), Jens Schwamborn (2)

1) Biomedical Data Science, Luxembourg Centre for Systems Biomedicine, University of Luxembourg
2) Developmental and Cellular Biology, Luxembourg Centre for Systems Biomedicine, University of Luxembourg

High-content cellular imaging has been shown to capture disease-relevant phenotypes in PD models. However, the reliability of deep learning (DL) predictions depends on biological validation and interpretability. This study aims to identify converging biological trends across modalities by integrating transcriptomic profiling and exploratory image-based analysis to characterize GBA-associated molecular and morphological signatures from donor-matched organoid models. We performed differential expression (DE) analysis on bulk RNA-sequencing data from two datasets of six and four patient-derived midbrain organoids, respectively, using linear mixed models to account for pseudo-replication and confounding effects. Subsequently, the DE results were used for network perturbation analysis to identify dysregulated gene interactions in GBA carriers and highlight any overlap. In parallel, a dataset of 16,000 high-content single-cell organoid images from the same donors with matched RNA-seq data was used to train DL architectures, such as DenseNet and Vision Transformer (ViT); given the limited number of donors, these image-based analyses are considered exploratory. Gradient-based attribution and transformer attention analysis were applied to assess which image regions drive model decisions. Network perturbation analyses of the GBA-specific networks identified dysregulation of receptor tyrosine kinase (RTK) signaling and converged on Wnt signaling dysregulation under proteostatic stress. Exploratory image analysis suggested that the ViT model, compared to convolutional architectures, may focus on more spatially localized and biologically coherent organoid regions, particularly near dense organoid areas and boundaries. Explainability analyses are being extended to align image-derived spatial features with the transcriptomic pathways previously identified. Confounder-aware transcriptomic analysis of patient-derived midbrain organoids confirms known GBA-associated dysregulation in two datasets, while highlighting candidate regulators for further investigation. DL explainability analyses suggest that spatial image features near organoid boundaries and dense regions may carry genotype-relevant information, providing a preliminary basis for cross-modal alignment with transcriptomic findings. Further validation work is ongoing to substantiate these observations.

The MINERVA Platform: fifteen years of support for disease maps

Marek Ostaszewski(1), Venkata P. Satagopam(1), Rudi Balling(2), and Reinhard Schneider(1)

1.LCSB, University of Luxembourg
2.University of Bonn

Abstract

In 2011 a project was kickstarted at Luxembourg Centre for Systems Biomedicine (LCSB) to develop the first graphical and computational resource of pathways perturbed in Parkinson's disease (PD) – the PD map [1]. The goal of the map was to enable application of systems biology methods to study complex relationships of multiple pathophysiological factors contributing to PD-related neurodegeneration. Sharing the PD map with a broad research community was challenging, because the map was encoded into large-scale systems biology diagrams, which were challenging to parse, distribute, and maintain. To address this challenge, the MINERVA (Molecular Interaction NEtwoRk VisuAlization) Platform was developed [2]. MINERVA allowed online exploration of complex diagrams of the PD map, browsing molecular annotations and literature supporting pathway interactions, and visualisation of omics data expression.

Since then, MINERVA was further developed to support the research community building increasingly complex disease maps, covering areas of chronic immune and allergic diseases [3], [4], mechanisms of inflammation and viral infection [5], [6], physiology of liver metabolism [7] or adverse outcome pathways [8]. This broad range of applications guided MINERVA development towards FAIR (Findable, Accessible, Interoperable, Reproducible) principles [9], to enable authors of the maps and their end users for transparent content sharing and use in building research hypothesis. Another area of development was handling different community standards for diagram representation [10], including CellDesigner, Systems Biology Markup Language (GPML), Systems Biology Graphical Notation Modelling Language (SBGN-ML) and Graphical Pathway Markup Language (GPML). Finally, MINERVA Application Programming Interface (API) and plugin architecture support reproducible analysis [11], enabling translation of any manual browsing activity into a workflow, or a part of a larger computational biology pipeline, including modelling [12].

In this poster, we will illustrate the fifteen years of continuous, community-driven development of the MINERVA Platform, highlighting its evolution to respond to the needs of the researchers developing and using systems biology diagrams. We will focus on key components supporting exploration of knowledge encoded in complex diagrams, and on functionalities needed to turn manual content browsing into reproducible workflows linking thousands of literature-supported interactions with tens of thousands of potential candidate molecules emerging from omics pipelines. Finally, we will illustrate next steps for MINERVA development, including integration with the rapidly developing area of Large Language Models.

The MINERVA Platform is an open access software available at minerva.uni.lu, and is a part of the service portfolio of ELIXIR Luxembourg (elixir-luxembourg.org).

References

[1] K. A. Fujita et al., 'Integrating pathways of Parkinson's disease in a molecular interaction map', Mol Neurobiol, vol. 49, no. 1, pp. 88–102, Feb. 2014, doi: 10.1007/s12035-013-8489-4.

[2] P. Gawron et al., 'MINERVA-a platform for visualization and curation of molecular interaction networks', NPJ Syst Biol Appl, vol. 2, p. 16020, 2016, doi: 10.1038/npjsba.2016.20.

[3] M. L. Acencio et al., 'The SYSCID map: a graphical and computational resource of molecular mech-

anisms across rheumatoid arthritis, systemic lupus erythematosus and inflammatory bowel disease', Front Immunol, vol. 14, p. 1257321, 2023, doi: 10.3389/fimmu.2023.1257321.

[4] A. Mazein et al., 'AsthmaMap: An interactive knowledge repository for mechanisms of asthma', Journal of Allergy and Clinical Immunology, vol. 147, no. 3, pp. 853–856, Mar. 2021, doi: 10.1016/j.jaci.2020.11.032.

[5] C. N. Serhan et al., 'The Atlas of Inflammation Resolution (AIR)', Mol Aspects Med, vol. 74, p. 100894, Aug. 2020, doi: 10.1016/j.mam.2020.100894.

[6] M. Ostaszewski et al., 'COVID19 Disease Map, a computational knowledge repository of virus-host interaction mechanisms', Mol Syst Biol, vol. 17, no. 10, p. e10387, Oct. 2021, doi: 10.15252/msb.202110387.

[7] L. Ladeira et al., 'Unlocking liver physiology: comprehensive pathway maps for mechanistic understanding', Front Toxicol, vol. 7, p. 1619651, 2025, doi: 10.3389/ftox.2025.1619651.

[8] A. Mazein et al., 'An explorable model of an adverse outcome pathway of cytokine release syndrome related to the administration of immunomodulatory biotherapeutics and cellular therapies', Front Immunol, vol. 16, p. 1601670, 2025, doi: 10.3389/fimmu.2025.1601670.

[9] I. Balaur et al., 'FAIR assessment of Disease Maps fosters open science and scientific crowdsourcing in systems biomedicine', Sci Data, vol. 12, no. 1, p. 851, May 2025, doi: 10.1038/s41597-025-05147-w.

[10] D. Hoksza, P. Gawron, M. Ostaszewski, J. Hasenauer, and R. Schneider, 'Closing the gap between formats for storing layout information in systems biology', Brief Bioinform, vol. 22, no. 1, p. 608, Jan. 2021, doi: 10.1093/bib/bbaa030.

[11] P. Gawron, E. Smula, R. Schneider, and M. Ostaszewski, 'Exploration and comparison of molecular mechanisms across diseases using MINERVA Net', Protein Sci, vol. 32, no. 2, p. e4565, Feb. 2023, doi: 10.1002/pro.4565.

[12] A. A. Hemedan, V. Satagopam, R. Schneider, and M. Ostaszewski, 'Cohort-specific boolean models highlight different regulatory modules during Parkinson's disease progression', iScience, vol. 27, no. 10, p. 110956, Oct. 2024, doi: 10.1016/j.isci.2024.110956.

Identifying Genetically Driven Gut Microbiome Components Causally Linked to Parkinson's Disease

Priyanshu Kumar Panda (1), Paul Wilmes (1), Patrick May (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg.

Historically viewed as an isolated central nervous system disorder, Parkinson's disease (PD) has now been gaining traction as a complex, systemic pathological state greatly influenced by the peripheral systems one of them being the gut brain axis-this paradigm shift catalyzed by the Braak's hypothesis, which states that α-synuclein pathology begins from the enteric nervous system. Indeed, there is growing observational evidence linking microbial alteration to PD, but distinguishing these signatures from their confounding factors and them, not being a secondary artifact acts as a major roadblock. To resolve this directional ambiguity, this study employs a three-arm Mendelian Randomization (MR) framework, which leverages its ability to use SNPs as instruments for microbial traits to perform an in-silco randomized controlled trial (RCT).

We present a MR framework built on the Luxembourg PD(LUXPARK) cohort, pairing genotyped data with 16s microbiome dataset, which is harmonized across five cohorts (1 internal + 4 external). To identify if the causal signal is best captured on a community level or single taxon level resolution, MR is performed in parallel across three exposure strata: co-abundance ecological modules, genera, and ASVs. Microbial-QTL GWAS was performed in PLINK2 adjusting for age, sex, ten genetic principal components and disease status. Genetic instruments underwent proper linkage-disequilibrium clumping, LD proxy substitution and allele harmonization before two-sample MR using five complementary estimators. The pipeline integrates outlier and pleiotropy sensitivity analyses, per-variant directionality checks, bidirectional MR using PD lead SNPs as reverse-direction exposure, colocalization at each microbial-QTL lead locus, and cross-cohort observational validation in three independent PD cohorts.

We systematically evaluate genetically predicted gut microbial taxa and functional metabolic pathways to PD risk. In doing so, our multi-tiered genetic design can help better understand pathways that might be essential in understanding the disease pathology better.

References

1) *Jiang L, Li JC, Tang BS, Guo JF. Associations between gut microbiota and Parkinson disease: A bidirectional Mendelian randomization analysis.* Eur J Neurol. *2023;30(11):3471-3477. doi:10.1111/ene.15848*

2) *Braak H, Rüb U, Gai WP, Del Tredici K. Idiopathic Parkinson's disease: possible routes by which vulnerable neuronal types may be subject to neuroinvasion by an unknown pathogen.* J Neural Transm (Vienna). *2003;110(5):517-536. doi:10.1007/s00702-002-0808-2*

3) *Nalls MA, Blauwendraat C, Vallerga CL, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. Lancet Neurol. 2019;18(12):1091-1102. doi:10.1016/S1474-4422(19)30320-5*

Poster n° 16

LCSB Data Space: governed management of clinical and biomedical research data at the Luxembourg Centre for Systems Biomedicine

Sandesh Patil (1), Vilem Ded (1), Venkata Satagopam (1)

1. Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg

Researchers at the Luxembourg Centre for Systems Biomedicine generate large and heterogeneous datasets across clinical studies, instrument platforms and computational environments. LCSB Data Space is now in production as the centre's research data management system. Its purpose is to capture these datasets early and keep the conditions under which they were collected attached to them, so the data can be used and passed downstream without re-checking what is permitted. We describe the architecture, the coupling to institutional identity and governance, and the data ingestion routes from submission systems, institutional platforms and researchers.

The system is connected to DAISY, the centre's data governance registry, which records each research project, its datasets, the people responsible for them and the legal basis under which the data were collected (1). Each project and dataset in the registry gets its own storage area, created automatically with its metadata and access rules, and linked to the legal agreements that cover it. Access follows project membership as recorded in the registry, rather than a separately maintained list. Authentication uses the institutional identity provider and single sign-on, so one identity works across every access route: the web portal, the command line and desktop applications.

Data arrive from external collaborators through the institutional submission system, from instrument platforms inside the centre, or directly from researchers. Submitted raw data go through validation and automated quality control, and are preserved for further use. A desktop application presents project data as an ordinary drive, so researchers can work in the tools they already use. Files can be corrected and republished without overwriting the original submission. The web portal is equipped with an AI assistant that answers questions against the collected metadata and the system documentation using retrieval-augmented generation.

References

1. Becker R, Alper P, Grouès V, Munoz S, Jarosz Y, Lebioda J, et al. DAISY: A Data Information System for accountability under the General Data Protection Regulation. GigaScience. 2019;8(12):giz140.

Poster n° 17

MIGRAPH: A HETEROGENEOUS GRAPH FRAMEWORK FOR MULTI-OMICS MICROBIOME INTEGRATION

Abdul Rafay Pirzada (1), Viacheslav Petrov (2), Paul Wilmes (3), Patrick May (4)

1-4) Luxembourg Centre for Systems Biomedicine, University of Luxembourg

Integrating multi-omics microbiome data presents a major structural challenge. Analytical workflows depend on flat feature matrices that do not represent relationships across biological scales, so the connections between genes, genomes, functions, and clinical measurements remain implicit and unavailable to the models trained on them. An even more fundamental obstacle precedes this: because gene prediction runs independently on each metagenomic assembly, biologically equivalent genes receive different identifiers in different samples. Without a shared coordinate system, genes cannot be compared across patients, across cohorts, or between local and public resources, which constrains both integration and the volume of data available for training.

To address this, we are developing a structured, multi-scale framework that represents the microbiome as a heterogeneous graph rather than a set of isolated tables. Working across locally processed cohorts [3] and large public microbiome resources [2], homologous microbial genes are grouped by sequence similarity [1] into a namespace of Unique Genetic Elements (UGEs). This provides consistent identifiers for homologous sequences across individuals while retaining both characterized and uncharacterized genes.

In the resulting architecture, MiGraph, the UGE serves as the central biological observation. Patient-specific measurements from multiple omics layers are mapped onto this shared reference, and UGE nodes are linked to a broader biological context including taxonomy, functional annotation, genomic context, and associated metabolites. The same gene identity therefore carries a different biological state in every patient, and that difference becomes part of the structure the model can learn from.

This graph representation is processed by a self-supervised algorithm to capture the underlying biological context. The aim is to learn embeddings for genes, genomes, and whole samples independently of any disease label. Learning microbiome structure before introducing clinical outcomes is intended to yield representations that transfer across cohorts and that support testable biological hypotheses rather than opaque predictions.

References

1) *Steinegger M, Söding J. MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat Biotechnol. 2017;35(11):1026–1028. doi:10.1038/nbt.3988*

2) *Richardson L, Allen B, Baldi G, et al. MGnify: the microbiome sequence data analysis resource in 2023. Nucleic Acids Res. 2023;51(D1):D753–D759. doi:10.1093/nar/gkac1080*

3) *Narayanasamy S, Jarosz Y, Muller EEL, et al. IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. Genome Biol. 2016;17:260. doi:10.1186/s13059-016-1116-8*

A sovereign, GDPR-compliant tool assembly for clinical and translational research data: architecture and use in European multi-centre studies

Kavita Rege (1) *, David Fitala (1), Sandrine Medves (1), Soumyabrata Ghosh (1), Venkata P Satagopam (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg
**kavita.rege@uni.lu*

Multi-centre European health studies increasingly rely on collecting, harmonising and sharing sensitive data across countries with different legal frameworks, technical capacities and clinical workflows. A single monolithic platform rarely fits all of these needs, while hyperscaler cloud solutions raise concerns about data sovereignty and GDPR compliance. We present the clinical and translational research data infrastructure developed at the Luxembourg Centre for Systems Biomedicine: not one system, but a governed assembly of interoperable tools.

The assembly combines open-source components, academically licensed software such as REDCap for electronic data capture, commercial software-as-a-service where it adds value (e.g. LamaPoll for online surveys, hosted in an ISO 27001-certified EU cloud), and in-house developments: custom plugins and API handlers, ETL and curation pipelines, and the Adept no-code analytics platform [1]. Core services run on university-hosted infrastructure, including a Kubernetes (OKD) cluster, and are connected through a common data model and shared data dictionaries mapped to standard ontologies such as OMOP, SNOMED CT, MeSH and MedDRA. Governance is built into the stack: pseudonymisation at source, federated authentication, role-based access, audit logging, GDPR accountability records in DAISY [2] and dataset descriptions in HealthDCAT-AP specifications [3]. Data are managed in levels, from raw to quality-checked to aggregated, supporting FAIR reuse. The same curated, harmonised layer makes the data AI-ready: it feeds machine-learning analyses, and AI assistants can explore metadata and aggregate statistics through a Model Context Protocol (MCP) interface without access to record-level data.

In the Horizon Europe-funded LEOPARD project, which develops AI-based models to reduce mortality among liver transplant candidates and support consistent organ allocation across Europe, the infrastructure provides a comprehensive data capture and management system with over 1,000 data points per patient, followed longitudinally pre- and post-transplant. The LEOPARD Glossary serves as common data model, and live recruitment and data-quality dashboards support over 40 clinical researchers. Partners using their own systems integrate via the shared dictionary. Currently more than 750 patients are recruited in the project across Europe.

In HeBA, the Healthy Brain Ageing study co-funded by the Michael J. Fox Foundation, the assembly supports population-based screening for Parkinson's disease risk in Kassel, Innsbruck, Barcelona and Luxembourg [4]. Multilingual online surveys in six languages, smell testing and site-specific REDCap projects for clinical follow-up feed common curation pipelines that processed 42,606 responses into a harmonised dataset of 29,709 eligible participants, including 8,589 smell tests and 1,207 clinically followed-up participants. Through governed transfers, curated data are also shared with international initiatives such as the Parkinson's Progression Markers Initiative (PPMI) and the Global Parkinson's Genetics Program (GP2).

This modular design lets each project choose the best component per task without losing control, provenance or compliance. By catalogued datasets, approval-based access and harmonised data models, the approach anticipates the secondary-use requirements of the European Health Data Space (EHDS) and offers a reusable, sovereign blueprint for European clinical research infrastructures.

References

1) *Adept: https://getadept.eu; LCSB instance: https://adept.lcsb.uni.lu*
2) *Becker R et al. DAISY: A Data Information System for accountability under the General Data Protection Regulation. GigaScience 8(12):giz140 (2019).*
3) *HealthDCAT-AP specification: https://healthdcat-ap.github.io/*
4) *Schade S, Ghosh S, Garrido A et al. Identifying individuals at-risk of developing Parkinson's disease using a population-based recruitment strategy: The Healthy Brain Ageing Kassel Study. npj Parkinson's Dis. 11, 216 (2025).*

AI in Radiation Oncology: A five-minute map of where it actually works

Carlos Rodrigues (1), Annelisa d'Angelo (1), Nordine Oural (2)

1) In Depth, Walferdange, Luxembourg
2) Laboratoire National de Santé, Department of Microbiology, Dudelange, Luxembourg

Artificial intelligence is now touching nearly every corner of radiation oncology (radiation therapy and nuclear medicine/theranostics alike) but adoption is uneven, and the headlines rarely match clinical reality.

This flash talk offers a fast, practitioner's map across four areas where AI is measurably changing practice today

- In radiation therapy preparation, deep-learning auto-contouring is commercially available for common sites, with roughly two-thirds of centres now reporting routine use and time savings of 40–50% for some structures, though benefit varies sharply by region of interest and requires local validation [1, 2]. In delivery, AI-driven registration and synthetic-CT generation make online adaptive replanning feasible within minutes, while machine-learning models are starting to flag abnormal setup errors and predict per-plan pass/fail outcomes before treatment [3, 4, 5, 6]

- In nuclear medicine and theranostics, deep learning is denoising PET/SPECT reconstructions to cut acquisition time and tracer dose without losing lesion contrast, while AI is also standardising tumour-burden quantification and approximating patient-specific dosimetry for radiopharmaceutical therapy [7, 8].

- Across both domains, AI-based risk and toxicity models are moving from retrospective research into adaptive treatment monitoring, though clinical translation remains inconsistent, and validation, bias and governance questions stay unresolved [9, 10].

AI in radiation oncology is not one technology but several, spanning two distinct modalities and multiple maturity levels — and knowing which is which is what separates hype from deployable practice.

References

1) *Pang, E.P.P., Tan, H.Q., Wang, F. et al. Multicentre evaluation of deep learning CT autosegmentation of the head and neck region for radiotherapy. npj Digit. Med. 8 (2025).*
2) *Vorbach SM, Putz F, Ganswindt U, Janssen S, Grohmann M, Knippen S, Heinemann F, El Shafie RA, Peeken JC. Contouring in transition: perceptions of AI-based autocontouring by radiation oncologists and medical physicists in German-speaking countries. Strahlenther Onkol. 201 (2025)*
3) *Chen ZF, Larocque B, Wang J, Chen D, Liu W, Choi HHF, Raman S. Artificial intelligence applications in adaptive radiotherapy—a narrative review. Transl Cancer Res 15(3) (2026).*
4) *Mastella E, Calderoni F, Manco L, Ferioli M, Medoro S, Turra A, Giganti M, Stefanelli A. A systematic review of the role of artificial intelligence in automating computed tomography-based adaptive radiotherapy for head and neck cancer. Phys Imaging Radiat Oncol. 33 (2025)..*
5) *Lin W, Xiao C, Xiao L, Fang J, Xu X, Fang Y. Research on real-time detection of radiotherapy setup errors and intelligent quality control methods based on artificial intelligence and big data. Front Oncol. 16 (2026).*
6) *Deng J, Zhao Y, Huang D, Zhang Q, Hong Y, Wu X. AI-enhanced cancer radiotherapy quality assessment: utilizing daily linac performance, radiomics, dosimetrics, and planning complexity. Front Oncol. 15 (2025).*
7) *Miwa K, Yamao T, Hashimoto F, et al. Innovations in clinical PET image reconstruction: advances in Bayesian penalized likelihood algorithm and deep learning. Ann Nucl Med. 39(9) (2025)*
8) *Banala M, Ezuddin S, Foley M, Kuker R. Artificial intelligence across oncologic theranostics: evidence for patient stratification, dosimetry, and adaptive radiopharmaceutical therapy. Front Nucl Med. 6 (2026).*
9) *Pallavi YC. Artificial intelligence for prognostic modelling and adaptive treatment monitoring in radiation oncology. Cureus. 18(6) (2026).*
10) *Zhu M, Gou L, Zhang C, Zheng D. Trustworthy artificial intelligence in radiation oncology: cross-industry lessons for development, validation, and deployment. Front Oncol. 16 (2026).*

Molecular dynamics supports differential structural engagement of suzetrigine across human voltage-gated sodium channel paralogs

Sara Salesi (1,2), Alexander Skupin (1), Vahid Shariati (2) and Patrick May (1)

1) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 4362 Belvaux, Luxembourg
2) National Institute of Genetic Engineering and Biotechnology, P.O. Box 14965/161, Tehran, Iran

Suzetrigine (VX-548; Journavx) is a first-in-class non-opioid analgesic that selectively inhibits the voltage-gated sodium channel $Na_v1.8$ and is approved for the treatment of moderate-to-severe acute pain. Although its pronounced selectivity for $Na_v1.8$ over other human paralogs has been demonstrated experimentally, the extent to which this selectivity can be reproduced from structural and dynamic properties remains unclear. Here, we computationally assessed suzetrigine interactions across a panel of human Na_v channel paralogs representing cardiac, central nervous system, and peripheral nociceptive channels. We combined structural modeling, ligand-binding pocket identification, blind molecular docking, and replicate membrane-embedded molecular dynamics simulations in explicit lipid bilayers. Docking scores alone did not reproduce the experimentally established selectivity of suzetrigine for $Na_v1.8$. Several paralogs showed similarly favorable predicted binding, with $Na_v1.5$ and $Na_v1.6$ ranking above $Na_v1.8$, highlighting the limitations of static docking scores for predicting paralog selectivity. In contrast, membrane-embedded molecular dynamics simulations revealed distinct patterns of ligand-channel interaction. Suzetrigine showed stable and persistent engagement with $Na_v1.8$, accompanied by reduced structural fluctuations and increased compactness of the channel. The $Na_v1.6$ complex displayed greater structural fluctuations and less persistent ligand engagement. $Na_v1.5$ showed an intermediate interaction profile, characterized by modest reductions in flexibility together with channel compaction and persistent ligand-protein hydrogen bonding despite its favorable docking score. Thus, the $Na_v1.5$ response differed from $Na_v1.8$ primarily in the magnitude of the structural response rather than its overall direction. Together, these results show that static docking scores alone are insufficient to reproduce the experimentally observed selectivity of suzetrigine, whereas membrane-embedded molecular dynamics simulations provide additional information that supports preferential engagement of $Na_v1.8$ relative to the other paralogs examined. The residual structural interaction observed for $Na_v1.5$ further illustrates that favorable molecular engagement does not necessarily imply functional channel inhibition. Integrating molecular docking with dynamic simulations may therefore provide a useful framework for assessing paralog-specific ion channel interactions and prioritizing potential off-target effects for subsequent experimental validation (1,2).

References

1) *Huang X, Jin X, Huang G, Huang J, Wu T, Li Z, et al. Structural basis for high-voltage activation and subtype-specific inhibition of human Nav1.8. Proceedings of the National Academy of Sciences. 2022 Jul 26;119(30):e2208211119. doi:10.1073/pnas.2208211119*

2) *Medhat RM, Kotb OA, Baecker D. Suzetrigine, a NaV1.8 Inhibitor as a Novel Approach for Pain Therapy—A Medicinal and Chemical Drug Profile. Molecules. 2026 Jan 20;31(2):358. doi:10.3390/molecules31020358 PubMed PMID: 41599406; PubMed Central PMCID: PMC12843740.*

Decoding memory from thousands of neurons: perspectives on generalization tests for brain–computer interfaces

Eléonore Schiltz (1,2), Lara Cardinaels (3), Martijn Broux (2), Sebastian Haesler (2)

1) LCSB, Luxembourg
2) KU Leuven - Leuven Research Institute for Neuroscience and Disease (LIND)
3) KU Leuven

Brain–computer interfaces (BCIs) are moving from the lab to patients. At their core, these systems are machine-learning decoders trained on high-dimensional neural population activity. Like any supervised model, a neural decoder can reach high accuracy on held-out test data and still fail when population drift or rely on confounding features rather than the variable of interest. BCI decoders are typically calibrated in controlled sessions, then used during natural, self-initiated behavior, and they often need frequent recalibration. How much behavioral context alone changes what can be decoded from a given brain region is rarely tested directly. Here, we combined two independent large-scale datasets to compare decoding in a passive, controlled setting with decoding during active decision-making. We studied the mouse olfactory–hippocampal circuit, a pathway that links sensory information to memory and is affected early in Parkinson's and Alzheimer's disease.

Using Neuropixels 1.0 and 2.0 probes, we recorded about 2,000 single neurons in seven brain regions, from primary olfactory cortex to the entorhinal cortex and hippocampus. Mice encountered novel and familiar odors in two similar settings: a passive exposure while head-restrained [1] and a freely moving task in which the choice of reward port is determined by odor novelty [2].

We trained linear decoders (support vector machines and logistic regression) on neural population activity and tested them with generalization checks. In the passive task, cross-condition generalization was done by leaving held-out odors out of training, allowing to detect odor-identity shortcuts. In the active task, we conducted cross-epoch generalization by training during odor sampling and testing during navigation.

In head-restrained mice, novelty was decodable along the whole pathway. It was encoded independently of odor identity in olfactory cortex, but in an odor-specific way in entorhinal–hippocampal regions.

In freely moving mice, movement and the upcoming decision took over. Population activity in every region mainly tracked the animal's position and goal, and hippocampal novelty decoding dropped to near chance. Only posterior piriform cortex kept a stable novelty code from sampling to reward. Correlated decoder weights showed that the same neurons carried the information throughout.

So behavioral context alone, with similar stimuli, changes which brain regions can be decoded reliably. Accuracy measured in one context says little about accuracy in another. For BCIs and AI in health care, decoders should be validated under the conditions in which they will be used. Generalization tests and interpretable model weights can help identify the regions and neurons whose signals stay stable, which could reduce the need for recalibration.

References

1) Eleonore Schiltz, Martijn Broux, Cagatay Aydin, Pedro J Goncalves, Haesler Sebastian (2026) Experience shapes the transformation of olfactory representations along the cortico-hippocampal pathway eLife 13:RP103373

2) Eleonore Schiltz, Lara Cardinaels, Karl Farrow, Sebastian Haesler (2026) Encoding of Novelty and Spatial Choice in the Olfactory Hippocampal Pathway During Goal-Directed Navigation. Available at SSRN: https://ssrn.com/abstract=7485301 or http://dx.doi.org/10.2139/ssrn.7485301

Uncertainty-Aware AI Clinical Narratives for Cardiovascular Risk

Maciej Skorski (1), Samuel Desire Kom Sande (1), Martin Theobald (1)

(1) Department of Computer Science, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Clinicians increasingly rely on AI-based risk assessment, including for cardiovascular disease (CVD) — the leading cause of mortality worldwide — drawn from patient data supplied in free-form text: clinical notes, discharge summaries, laboratory test results, or whatever other documentation is available.

Trusting these AI-based assessments needs understanding their reasoning: recent explainable-AI (XAI) research aims to pair statistically validated, model-based impact attributions with LLM-generated clinical narratives for trustworthy explanations [1,3]. While such narratives improve user comprehension [4], the open question that remains is how *faithful* they are — *measured quantitatively* rather than judged by fluency [2] — and whether the findings *generalize from benchmark datasets to real clinical cohorts*, as confirmed by domain experts. LLMs are known to be fluent in narration, but poor in forecasting their own uncertainty; verbalized and inferred confidence is not well aligned with accuracy, whereas attribution divergence against a trusted model is a better reliability measure [6].

We present **DoctorSHAP**, a model-agnostic pipeline rather than a fixed model pairing. It has three main components: a **risk assessor** (a medical-domain LLM scoring risk, e.g. MedGemma); a decoupled, exchangeable LLM **narrator** (e.g., Qwen) that turns free-form clinical text into concepts on the way in and the scored result into a clinician-facing summary on the way out; and an **uncertainty layer** that masks a concept (e.g., the clause “patient has marker X at level Y”) and resamples the risk assessor’s score, yielding the uncertainty signal. The result is a per-patient reliability measure, grounded in how consistently the risk assessor’s attribution behaves across resamples — never presented as unconditionally reliable. Our reference instantiation pairs MedGemma, Google’s medical LLM, as risk assessor with Qwen as narrator, on a dual-GPU local deployment. DoctorSHAP closes a real gap: a substrate-agnostic risk assessor validated against a live LLM, with its uncertainty layer rendering within-model variability as narrative rather than a number.

We validate this pipeline on LURIC, a real-world cardiovascular cohort used in our companion study [5], assessing narrative faithfulness (agreement with ground-truth attributions in sign, rank, and relative magnitude) and uncertainty calibration — bridging the domain-expert, real-cohort evaluation gap identified in prior work [4].

References

1) *F. Musicco, D. Danese, G. Fasano, A. Lombardi, A. C. M. Mancino, and T. Di Noia. “Who Are You Explaining To? A Multi-Agent System for Audience-Aware XAI Narratives.” arXiv:2608.11033, 2026.*
2) *T. Ichmoukhamedov, J. Hinns, and D. Martens. “How Good Is My Story? Towards Quantitative Metrics for Evaluating LLM-Generated XAI Narratives.” arXiv:2412.10220, 2024.*
3) *A. Zytek, S. Pido, S. Alnegheimish, L. Berti-Équille, and K. Veeramachaneni. “Explingo: Explaining AI Predictions Using Large Language Models.” IEEE International Conference on Big Data (IEEE BigData), 2024. arXiv:2412.05145.*
4) *D. Martens, J. Hinns, C. Dams, M. Vergouwen, and T. Evgeniou. “Tell Me a Story! Narrative-Driven XAI with Large Language Models.” Decision Support Systems, vol. 191, 114402, 2025. DOI: 10.1016/j.dss.2025.114402.*
5) *S. D. Kom Sande, M. Skorski, M. Theobald, J. Schneider, and W. März. “Regression vs. Medical LLMs: A Comprehensive Study for CVD and Mortality Risk Prediction.” Artificial Intelligence in Medicine (AIME 2026), LNCS vol. 16748, pp. 109–114, Springer, 2027. DOI: 10.1007/978-3-032-30710-1_13.*
6) *A. Dasula, P. Desikan, and J. Srivastava. “LLM Doesn’t Know What It Doesn’t Know: Detecting Epistemic Blind Spots via Cross-Model Attribution Divergence on Clinical Tabular Data.” arXiv:2606.19509, 2026.*

Poster n° 23

A Putative Compensatory *EGR1*-Centered Module in Mild Cognitive Impairment along the Alzheimer's Disease Trajectory Is Attenuated in *APOE* ε4 Carriers

Gleb Svinin (1), Deborah Powers (1), Enrico Glaab (1)

1) Biomedical Data Science Group, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Mild cognitive impairment (MCI) is a symptomatic predementia state along the Alzheimer's disease (AD) trajectory, but the cell-type-specific cortical expression changes that accompany it, and their dependence on *APOE* genotype, are only partly characterized. To address this, clinically defined MCI was studied after exclusion of recorded alternative causes of cognitive impairment. We analyzed prefrontal cortex single-nucleus RNA-sequencing data from ROSMAP and PsychAD as discovery cohorts and used SEA-AD for targeted validation. Cell-type annotations provided by the individual studies were harmonized using the MapMyCells portal, and donor-level pseudobulk differential expression models were adjusted for sex, age at death, and *postmortem* interval.

Since AD risk differs by *APOE* genotype, with the ε3/ε4 genotype associated with approximately 2- to 4.5-fold higher odds of AD than ε3/ε3 across population groups [1], we first analyzed non-ε4 carriers (ε2/ε2, ε2/ε3, and ε3/ε3) to characterize MCI-associated transcriptional changes within a more comparable, lower-risk group. We then fitted an MCI-by-ε4-carrier interaction model to estimate the ε4-associated difference in the MCI transcriptional response. Differentially expressed genes were integrated using CARNIVAL [2] with OmniPath prior knowledge [3] to infer signed, directed networks, and the resulting modules were additionally assessed by pooled and cohort-specific coexpression analysis. Pathway-level changes were evaluated using preranked gene set enrichment analysis (GSEA).

Among the cell types tested, medial ganglionic eminence (MGE)-derived interneurons, which contribute to cortical excitation-inhibition balance, provided a compact inferred network. In non-ε4 carriers, MCI was associated with concordant upregulation of *EGR1* and of the activity-regulated immediate early genes *ARC*, *FOS*, *GADD45B*, and *FOSB*. Using CARNIVAL with the signed, directed interactions from the OmniPath prior knowledge database, we inferred links from *EGR1* to *ARC*, *FOS*, and *GADD45B*, and from *FOS* to *FOSB*. Separately, all four gene pairs showed significant coexpression in the pooled non-ε4 analysis, providing complementary evidence of coordinated expression but not independent validation of the inferred directionality. In SEA-AD, all five genes displayed concordant effect directions in targeted dementia-versus-no-dementia validation among non-ε4 donors. The MCI-by-ε4-carrier interaction estimate was negative for *EGR1*, *ARC*, *FOS*, and *FOSB*. The Reactome NGF-stimulated transcription gene set showed the same opposing pattern, with positive enrichment in non-ε4 MCI (normalized enrichment score [NES] = 2.28, FDR = 1.3×10^{-4}) and negative enrichment for the MCI-by-ε4-carrier interaction (NES = −2.00, FDR = 0.017); *EGR1*, *ARC*, *FOS*, and *FOSB* were shared leading-edge genes. Overall, in this observational study, the activation in non-ε4 MCI together with the negative interaction estimates supports the hypothesis that the *EGR1*-centered program reflects a transient compensatory response, whereas its attenuation in ε4 carriers may accompany lower neuronal resilience.

References

1) Belloy ME, Andrews SJ, Le Guen Y, et al. APOE genotype and Alzheimer disease risk across age, sex, and population ancestry. JAMA Neurol. 2023;80:1284–1294. doi:10.1001/jamaneurol.2023.3599.

2) Liu A, Trairatphisan P, Gjerga E, et al. From expression footprints to causal pathways: contextualizing large signaling networks with CARNIVAL. npj Syst Biol Appl. 2019;5:40. doi:10.1038/s41540-019-0118-z.

3) Türei D, Valdeolivas A, Gul L, et al. Integrated intra- and intercellular signaling knowledge for multicellular omics analysis. Mol Syst Biol. 2021;17:e9923. doi:10.15252/msb.20209923.

Poster n° 24

CELL-TYPE AND DISEASE-STAGE DEPENDENCE OF SEX DIFFERENCES IN THE ALZHEIMER'S DISEASE CORTEX

Dilara Uzuner Odongo (1), Enrico Glaab (1)

1) Biomedical Data Science Group, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, 6 Avenue du Swing, L-4367 Belvaux, Luxembourg

Women account for approximately two thirds of Alzheimer's disease (AD) cases and show faster cognitive decline after diagnosis than men[1,2]. The cell types that carry the underlying molecular changes, and the dependence of these changes on disease stage, are still incompletely defined. We therefore tested whether sex-dependent transcriptional signatures reproduce across independent human cohorts and whether they correspond to those of a longitudinal AD mouse model. Sex-stratified differential expression and sex-pattern classification were applied to single-nucleus RNA-sequencing (snRNA-seq) data from two non-overlapping human cortical cohorts (ROSMAP[3] and PsychAD[4]; >700 donors), with donors staged by Braak stage, CERAD score and cognitive status, and with category assignments replicated across cohorts and checked by sensitivity and down-sampling analyses. The same framework was applied to two longitudinal bulk RNA-sequencing datasets from the 5xFAD mouse cortex[5,6] to add temporal resolution, followed by pathway enrichment, protein interaction, regulatory network and AD genetics analyses across datasets and species. In both human cohorts, sex-specific changes were clustered in a few cell types, most often neurons and oligodendrocytes in men and neurons, astrocytes or vascular cells in women, with the leading cell type differing between cohorts. Seventy-nine genes, including the AD risk gene *APOE*, received the same sex-dependent category in both cohorts. In the mouse datasets, female-specific changes predominated, and mevalonate and cholesterol biosynthesis activity increased with age in three complementary analyses. The same pathway carried a female-specific signal in the human cortex through the mevalonate enzymes *MVD* and *FDPS*, with decreased expression at end-stage disease. Immune and inflammatory processes were shared between sexes and species. Cell-type distribution and gene categorization replicated across the two human cohorts, whereas cross-species agreement was strongest at the pathway level. Sex- and cell-type-resolved pathway signatures may therefore transfer between species more reliably than individual gene orthologs, and biomarker cutoffs and trial designs in AD should be stratified by sex.

References

1) *2026 Alzheimer's disease facts and figures. Alzheimer's & Dementia. 2026;22(4):e71345. doi:10.1002/alz.71345*
2) *Li A, Klinger HM, Seto M, et al. Elevated temporal tau PET predicts faster cognitive decline in women than men: A meta-analysis. Alzheimer's & Dementia. 2026;22(2):e71031. doi:10.1002/alz.71031*
3) *Mathys H, Peng Z, Boix CA, et al. Single-cell atlas reveals correlates of high cognitive function, dementia, and resilience to Alzheimer's disease pathology. Cell. 2023;186(20):4365-4385.e27. doi:10.1016/j.cell.2023.08.039*
4) *Fullard JF, Nm P, Lee D, et al. Population-scale cross-disorder atlas of the human prefrontal cortex at single-cell resolution. Sci Data. 2025;12(1):954. doi:10.1038/s41597-025-04687-5*
5) *Lee CYD, Daggett A, Gu X, et al. Elevated TREM2 Gene Dosage Reprograms Microglia Responsivity and Ameliorates Pathological Phenotypes in Alzheimer's Disease Models. Neuron. 2018;97(5):1032-1048.e5. doi:10.1016/j.neuron.2018.02.002*
6) *Forner S, Kawauchi S, Balderrama-Gutierrez G, et al. Systematic phenotyping and characterization of the 5xFAD mouse model of Alzheimer's disease. Scientific Data. 2021;8(1):270. doi:10.1038/s41597-021-01054-y*

Exploring the co-selection of antibiotic, metal, and biocide resistances across environmental and human microbiomes

Astrid Vincze (1), Viacheslav A. Petrov (1), Tuesday Lowndes (1), Alissa Caroline Muller (1), Delphine Collard (2), Henry-Michel Cauchie (2), Emma Schymanski (1), Patrick May (1), Paul Wilmes (1)

Affiliations

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Esch-sur-Alzette, Luxembourg

2) Luxembourg Institute of Science and Technology, Esch-sur-Alzette, Luxembourg

The microbial resistome includes the complete set of resistance genes in microbial communities and is essential for understanding antimicrobial resistance and co-selection. Although several tools detect resistance genes, few integrate antibiotics, metals, and biocides within a unified framework that also explores their interactions. Many existing methods rely primarily on sequence-based identification, and although ML and deep learning approaches are emerging, they are not yet widely integrated into standard pipelines, which limits the discovery of novel resistance genes. This project aims to integrate machine learning models into the PathoFact microbiome pipeline to enable pattern-based and reduced-dependence on strict sequence similarity for identifying resistance genes and to apply this approach to environmental and human microbiome datasets. The workflow begins with the construction of a database incorporating metal and biocide resistance genes from BacMet, UniProt, CARD, ResFinder, and AMRFinderPlus. Machine learning models are trained and benchmarked against existing tools, and the best models integrated into PathoFact. The improved pipeline will be applied to publicly available human gut microbiome datasets and to seasonal river samples from Luxembourg. Environmental samples of 20 liters will be concentrated, filtered, and processed for metagenomics and metatranscriptomics. Expected outcomes include a tool capable of detecting known and novel resistance genes and revealing interactions within the resistome. This work will enhance resistome profiling in environmental and clinical contexts and support advances in antimicrobial resistance research.

References

1. *Narayanasamy S. et al. (2016) IMP: a pipeline for reproducible reference-independent integrated metagenomic and metatranscriptomic analyses. Genome Biology 16;17(1):260.*
2. *de Nies L. et al. (2021) PathoFact: a pipeline for the prediction of virulence factors and antimicrobial resistance genes in metagenomic data Microbiome. 17;9(1):49.*
3. *Delgado L. et al. (2026) PathoFact 2.0: an integrative pipeline for the prediction of antimicrobial resistance genes, virulence factors, toxins and toxin-associated proteins, and biosynthetic gene clusters in metagenomes. GigaScience 22;15:giag062.*

FROM FREE TEXT TO STANDARDIZED RECORDS: *bioc2omop* FOR CLINICAL DECISION SUPPORT

Prasan Yapa, Basel Alshaikhdeeb, Jaleh Shoshtarian Malak, Soumyabrata Ghosh, Anne Lorant, Venkata P. Satagopam

Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg

Clinical decision support (CDS) and observational research increasingly rely on data harmonized to the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM), a standard maintained by the OHDSI collaborative and adopted across hundreds of institutions worldwide (1). Yet much clinically relevant knowledge remains embedded in unstructured text, and converting natural language processing (NLP)-extracted mentions, often represented in interchange formats such as BioC, into standardized OMOP records remains under-studied (2). A particular risk for downstream decision-making arises when NLP annotation schemes assign the same label to genuine drug mentions and unrelated chemical or laboratory-value mentions, potentially resulting in their incorrect mapping to OMOP tables. Distinguishing drug from non-drug chemical entities remains a persistent challenge, with substantial disagreement even in dedicated shared tasks (3).

We present *bioc2omop*, a configurable pipeline that maps BioC-format NLP annotations to OMOP CDM tables. The pipeline supports five interchangeable concept-matching strategies, spanning lexical and embedding-based approaches, enabling the most appropriate method to be selected for each use case. To address the drug-chemical ambiguity described above, we introduce a relation-aware mechanism that classifies entities based on the relations in which they participate, rather than relying solely on their entity-type labels. We evaluated the pipeline across four heterogeneous corpora, including public literature benchmarks and an in-house clinical case-report corpus, using document-level bootstrap confidence intervals, paired significance tests, and targeted ablation studies. The pipeline and evaluation code are openly available on Zenodo (https://zenodo.org/records/22010477).

Embedding-based matching achieved an F1 score of up to 0.97 for drug and chemical entities and significantly outperformed the external baseline in 6 of 8 tested corpus-table combinations, although its advantage was less consistent for disease and condition entities. The relation-aware mechanism correctly routed 12.2% of chemical mentions to drug exposure records, with no false positives among the validation terms and no effect on corpora lacking the underlying tagging problem. This directly reduces a class of misclassification with potentially important downstream consequences for CDS. Ablation studies confirmed these embedding gains were driven specifically by fine-tuning objective, with F1 differences of 0.06-0.25 over an untuned base model, and further surfaced two cautionary findings that a single benchmark run would have missed entirely.

These results further demonstrate that conclusions about model performance were sensitive to modelling and evaluation choices. BioMistral, a biomedical-tuned model, performed significantly worse than its general-purpose counterpart in three of four comparisons, driven by a calibration difference rather than a knowledge gap, and an apparent gain from a larger candidate pool proved smaller than ordinary sampling variance across five seeds. No single concept-matching strategy was universally best; we argue rigorous, multi-corpus, ablation-tested evaluation, not a single benchmark result, should guide how such pipelines are built and deployed in clinical settings. A key limitation is that none of the four evaluation corpora are genuine electronic health record (EHR) or clinical-practice data, three are published literature and the fourth is drawn from case reports, so findings may not directly transfer to true EHR-derived pipelines. Validating *bioc2omop* against real clinical-practice notes, extending coverage to additional OMOP tables, and benchmarking further concept-matching tools and language-model families are natural next steps toward deployment-ready CDS systems.

References

1) *Hripcsak G, et al. Observational Health Data Sciences and Informatics (OHDSI): Opportunities for Observational Researchers. Stud Health Technol Inform. 2015.*
2) *Vipina K, et al. Representing and utilizing clinical textual data for real world studies: An OHDSI approach. Journal of biomedical informatics 142:104343. 2023.*
3) *Krallinger M, et al. CHEMDNER: The drugs and chemical names extraction challenge. Journal of cheminformatics 7, no. Suppl 1. 2015.*

When Models Recognize Patients, Not Pathology: Quantifying Data Leakage in Federated Alzheimer's Disease Detection

Alberto Zancanaro (1), Venkata Satagopam (2)

1) LCSB, alberto.zancanaro@uni.lu
2) LCSB, venkata.satagopam@uni.lu

Federated Learning [1] is an increasingly popular paradigm for privacy-preserving medical Deep Learning, and the automated detection of Alzheimer's disease from structural MRI is one of its most natural applications, since neuroimaging data are highly sensitive and remain isolated within individual clinical sites.

While the literature reports near-ceiling classification accuracies [2], these results are often obtained under evaluation protocols that may be affected by *data leakage* [3], i.e., the inadvertent flow of test information into training data. This effect is well documented in centralized studies, but in the federated learning literature the term leakage has almost exclusively denoted privacy leakage, and the impact of performance-inflating, subject-level leakage remains largely underreported. In this work we investigate this gap directly. We train two CNN architectures, Demnet and VGG, on the ADNI dataset using the FedAvg algorithm, and we compare two data-partitioning strategies: a *random* split and a clinically realistic *subject-aware* split, in which all scans of a given patient are confined to a single subset. Under random partitioning our federated models reproduce the high accuracies reported in the literature (test accuracy around **0.9**), but the same models collapse to roughly **0.45** under subject-aware partitioning. Similar results were obtained for the Cohen's kappa score.

This discrepancy indicates that the models primarily rely on a misleading shortcut: they identify individual subjects and retrieve their corresponding labels, instead of learning the pathological patterns that differentiate Alzheimer's patients from healthy individuals. We argue that this behavior simultaneously constitutes a methodological flaw that artificially boosts reported accuracy and a breach of the privacy principles underlying federated learning. Consequently, we assert that stringent, subject-aware evaluation protocols are crucial for achieving trustworthy and broadly generalizable federated Alzheimer's detection.

References

1) *B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in Artificial intelligence and statistics, pp. 1273–1282, PMLR, 2017*

2) *A. G. Vrahatis, K. Skolariki, M. G. Krokidis, K. Lazaros, T. P. Exarchos, and P. Vlamos, "Detection of Alzheimer's disease using deep learning models: A systematic literature review," Computer Science Review, 2024. Systematic review reporting predominantly 90–99% accuracies across DL-for-AD studies.*

3) *V. M. Young, S. Gates, L. Y. Garcia, and A. Salardini, "Data leakage in deep learning for alzheimer's disease diagnosis: A scoping review of methodological rigor and performance inflation," Diagnostics, vol. 15, no. 18, p. 2348, 2025.*

From Shared Challenges to Shared Practices: Building Luxembourg's Data Management Community

Qian Zhang, Nene Djenaba

Luxembourg National Data Service (LNDS), Belval, Luxembourg

The growing adoption of Artificial Intelligence (AI) in healthcare has created promising opportunities to improve healthcare quality and efficiency, facilitate research, and assist in decision-making in patient care. However, the potential of these healthcare AI systems depends not only on algorithms, but also on the availability of high-quality, well-managed, and reusable data. Challenges such as fragmented data management practices, heterogenous data sources, insufficient data stewardship, and inconsistent metadata and documentation can limit data interoperability and reuse, ultimately constraining the development, deployment and trustworthiness of healthcare AI solutions.

To address these challenges, the Data Management National Community of Practice (DM-CoP), was established to connect professionals, share expertise, and facilitate common approaches to data management and stewardship. With more than 180 members across Luxembourg and beyond, the DM-CoP fosters a community for knowledge exchange, capacity building and cross-organisational learning.

By sharing data management practices and strengthening Luxembourg's data stewardship capabilities, the DM-CoP contributes to creating FAIR-R[1] and AI-ready healthcare data. It helps build the robust data foundations required for responsible, trustworthy, and sustainable AI adoption in healthcare, while strengthening Luxembourg's national data ecosystem.

References

1) *Verhulst, Stefaan and Zahuranec, Andrew and Chafetz, Hannah, Moving Toward the FAIR-R principles: Advancing AI-Ready Data (March 04, 2025). DOI: http://dx.doi.org/10.2139/ssrn.5164337*

CERTAINTY: A Cellular Immunotherapy Virtual Twin for Personalised Cancer Treatment

Ulrike Weirauch (1,2), Michael Rade (1), Markus Kreuz (1), Maximilian Ferle (3,4), Maria Quaranta (5), Andreas Kremer (5), Stefan Franke (3), Holger Fröhlich (6), Yasamin Salimi (6), Sebastian Schwick (6), , Regina Demlova (7), Carl Steinbeisser (8), Maximilian Merz (9), Kristin Reiche (1,10) on behalf of the CERTAINTY consortium

1) Fraunhofer IZI, Leipzig, Germany
2) Institute for Clinical Immunology, University of Leipzig, Leipzig, Germany
3) Innovation Center Computer Assisted Surgery (ICCAS), University of Leipzig, Leipzig, Germany
4) Department of Hematology, Hemostaseology, Cellular Therapy and Infectiology, University Hospital of Leipzig, Leipzig, Germany
5) Information Technology for Translational Medicine, Esch-sur-Alzette, Luxembourg
6) Fraunhofer SCAI, Sankt Augustin, Germany
7) Faculty of Medicine, Masaryk University, Brno, Czechia
8) Collaborate Project Management, Munich, Germany
9) Myeloma Service, Memorial Sloan Kettering Cancer Center, New York, USA
10) Interdisciplinary Transformation University (IT:U), Linz, Austria

Adoptive cellular immunotherapies (ACIs) like chimeric antigen receptor (CAR) T cells have the potential to revolutionize the treatment of cancer, autoimmune disease, and other diseases, but substantial challenges persist: complex manufacturing, significant toxicities, resistance mechanisms, and high costs. Additionally, these therapies are heterogenous and complex, with a stochastic biological nature. AI-based tools enable the modeling of complex physiological and pathological processes and treatment outcomes by integrating multimodal real-world data. The knowledge and models generated allow for tailored decision-making for individual patients in their current status. A key concept is the virtual twin (VT), a digital patient representation used to simulate patient-specific trajectories for precision medicine.

The CERTAINTY project [1] pioneers a VT ecosystem to support individualized CAR T-cell treatment strategies for relapsed/refractory multiple myeloma (MM) throughout the entire patient journey. Here, we have assembled an interdisciplinary team of AI specialists, computer scientists, data managers, clinicians, biologists, and legal experts, dedicated to the development of the ACI-VT.

We defined minimum design specifications for adapting VTs to "living drugs" like CAR T cell therapy, pioneering a foundational framework for ACI-VTs that simulate individual patient pathophysiology [2]. As a basis for both, VT development and predictive modelling, we are compiling patient-level data encompassing ~10,000 MM patients, with over 1,000 of them receiving anti-B-cell maturation antigen (BCMA) CAR T therapy in real-world settings. Analysis of 61 patients revealed superior complete response rates with cilta-cel over ide-cel (78% vs. 38%) and longer progression-free survival. A longitudinal single-cell multi-omics atlas of 135 samples demonstrated CD4+ cytotoxic T cell expansion associated with response and toxicity, while non-responder CD8+ T cells exhibited impaired effector function. Soluble BCMA reduction correlated with CAR T expansion and systemic inflammation, revealing mechanisms driving differential therapeutic outcomes [3]. Furthermore, we developed a model predicting individual MM progression events from routine blood work, trained on CoMMpass (N = 1,186) and externally validated in GMMG-MM5 (N = 504) [4].

Within CERTAINTY, we curated extensive MM datasets to develop initial predictive models for clinical progression and side effects using clinical annotations and molecular data. Currently, we are advancing multi-scale mechanistic models and implementing the VT prototype. As ACI therapy expands, we envision ACI-VTs evolving in parallel to enable true precision health.

References
1. Reiche K, et al., Lancet Haematol. 2025 Jul;12(7):e490-e491.
2. Weirauch U, et al., NPJ Digit Med. 2025 Aug 1;8(1):493.
3. Rade M, et al., Cancer Cell. 2025 Dec 4:S1535-6108(25)00490-8
4. Ferle M, et al., NPJ Digit Med. 2025 Apr 30;8(1):231

Acknowledgements
The CERTAINTY project is funded by the European Union (Grant Agreement 101136379).

Individual-specific network diffusion reveals biological and therapeutic heterogeneity in mixed Alzheimer's disease and LATE pathology

[1]Kavya Singh, [2,3]David W. Fardo, [1]Kristel Van Steen

[1]BIO3-GIGA-R Molecular and Computational Biology, University of Liège, Place du 20 Août 7, B-4000 Liège, Belgium

[2]Sanders-Brown Center on Aging, University of Kentucky, Lexington, KY, United States

[3]Department of Biostatistics, University of Kentucky, Lexington, KY, United States

Abstract

Mixed Alzheimer's disease (AD) and limbic-predominant age-related TDP-43 encephalopathy (LATE) pathology is common in older adults, yet conventional diagnostic categories provide limited information about the molecular heterogeneity underlying these overlapping conditions. This heterogeneity complicates patient stratification and may obscure opportunities for targeted therapeutic development.

We developed an individual-specific network framework to characterize disease-related molecular programs and their therapeutic implications in 347 patients from the Religious Orders Study and Rush Memory and Aging Project [1]. Transcriptomic and epigenomic features were integrated into patient-specific molecular networks, after which network diffusion was performed using AD- and LATE-associated gene sets as alternative disease priors. For each patient, the resulting diffusion profiles were summarized as gene-level contrasts quantifying, for each gene, its relative network proximity to AD- versus LATE-associated programs.

Two distinct but connected layers of unsupervised discovery were performed: first at the patient level (biological states), then at the gene level (conserved gene modules within those states). This yielded two stable biological states defined by molecular proximity to AD- versus LATE-associated programs (AD-predominant and Non-AD-predominant; clustering robust under bootstrap resampling [2], mean Adjusted Rand Index ≈ 0.85 [3]). Independent network-community analyses identified conserved gene modules within both states. The two largest gene modules were highly similar across states (Jaccard similarities of 0.90 and 0.89 [4]), indicating a shared molecular scaffold rather than two completely distinct disease processes. Nevertheless, the second-largest AD-predominant module (143 genes) showed additional enrichment for mitochondrial energy metabolism (oxidative phosphorylation), a signal absent in the corresponding Non-AD-predominant module (154 genes), despite otherwise similar nucleotide-metabolism enrichment, providing independent support for its biological interpretation.

Finally, to investigate therapeutic implications, gene-level diffusion scores were aggregated across each drug's known target genes to generate patient-specific drug-ranking profiles. Unsupervised clustering of these drug profiles identified two therapeutic states. The biological and therapeutic classifications were strongly associated for AD-predominant patients, 98.9% of whom belonged to the same therapeutic state. In contrast, Non-AD-predominant patients displayed substantially greater therapeutic heterogeneity, with an approximately even distribution between the two therapeutic states. This association was highly significant (Fisher's exact test [5], odds ratio = 76.2, $p < 10^{-6}$). Drug-target-derived gene modules reinforced the strong alignment between the AD-predominant biological and therapeutic states, while uncovering additional functional signals among Non-AD-predominant patients that were not captured by biology-based stratification alone.

These results show that individual-specific network diffusion can recover stable molecular states across a spectrum of AD and LATE pathology burden, including individuals without LATE pathology, without relying on conventional diagnostic labels. They further demonstrate that molecular similarity does not translate uniformly into therapeutic similarity: patients with AD-leaning molecular profiles exhibit highly coherent drug-target-based therapeutic profiles, whereas patients with less AD-leaning profiles show substantially greater therapeutic heterogeneity despite sharing broad disease-related features. By integrating personalized multi-omics networks with drug-target information, this framework provides a data-driven strategy for prioritizing patient subgroups, generating drug-repurposing hypotheses, and informing future precision-treatment decision-support systems for individuals with mixed or overlapping AD/LATE pathology.

References

[1] *Bennett DA, Buchman AS, Boyle PA, Barnes LL, Wilson RS, Schneider JA (2018). Religious Orders Study and Rush Memory and Aging Project. Journal of Alzheimer's Disease, 64(s1), S161–S189.*

[2] Efron B (1979). Bootstrap methods: another look at the jackknife. The Annals of Statistics, 7(1), 1–26.

[3] Hubert L, Arabie P (1985). Comparing partitions. Journal of Classification, 2(1), 193–218.

[4] Jaccard P (1912). The distribution of the flora in the alpine zone. New Phytologist, 11(2), 37–50.

[5] Fisher RA (1935). The Design of Experiments. Oliver and Boyd, Edinburgh.

FlowGRN: Scalable and Dropout-Robust Gene Regulatory Network Inference via Flow Matching-Based Trajectory Reconstruction

Tsz Pan TONG (1,2), Jun PANG (1,2)

1) Department of Computer Science, University of Luxembourg, Esch-sur-Alzette, Luxembourg
2) Institute for Advanced Studies, University of Luxembourg, Esch-sur-Alzette, Luxembourg

Inferring directed gene regulatory networks (GRNs) from single-cell RNA sequencing data is essential for understanding cellular regulation, yet sequencing can only capture destructive snapshots rather than continuous cell trajectories. Sparse temporal sampling, high-dimensional expression profiles, and dropout further obscure regulatory relationships. GRN inference methods without using temporal information scale well but discard dynamic information, while conventional neural ordinary differential equation models incur substantial computational cost and scalability issues. We present FlowGRN, a framework that applies conditional flow matching and score matching to reconstruct cellular trajectories and uses tree-based dynamic network inference to infer directed, signed regulatory networks from snapshot data.

FlowGRN first introduces a new dropout-robust cell similarity measure. Since dropout turns gene expressions into zeros, zero-expressed genes are contaminated by dropout, and only non-zero expressions are reliable. To mitigate dropout, we mask out zero-expressed genes during pairwise cell distance computation, which we use to define optimal transport costs between snapshots. Using pseudotime-ordered snapshots and inferred lineages, conditional flow matching learns a nonlinear vector field directly in gene expression space, with score matching regularizing the learned flow. Trajectories are subsequently reconstructed by integrating the learned vector field forward and backward from observed cells. These trajectories provide time-series inputs to dynGENIE3 (Huynh-Thu & Geurts, 2018), whose random forests estimate directed regulatory strengths from expression dynamics. Regulatory signs are determined by the Jacobian of the learned flow.

We evaluated FlowGRN against eight baselines on the BEELINE benchmark (Pratapa et al., 2020), covering ten simulated datasets (six synthetic, four curated) and fourteen experimental datasets. Experimental evaluation used reference networks from DoRothEA and CollecTRI. Across the ten simulated datasets, FlowGRN achieved the best average ranks for AUPRC and early precision ratio, at 1.5 and 2.0, respectively, compared with 3.9 and 3.1 for GENIE3 (Huynh-Thu et al., 2010). It also achieved the best signed-network inference performance on seven of ten simulated datasets, with advantages in long cascades and cyclic dynamics. On experimental data, FlowGRN ranked first or second in ten of fourteen tasks for each metric and scaled to 1,783 genes.

Ablation studies suggest that trajectory reconstruction contributes to FlowGRN's performance. Using reconstructed trajectories with dynGENIE3 generally outperformed extracting edges directly from the Jacobian of the learned flow. This may reflect the difficulty of distinguishing direct from indirect regulation. Augmenting static GENIE3 with reconstructed trajectories did not consistently improve performance. Overall, FlowGRN offers a scalable approach to directed GRN inference by integrating dropout-aware trajectory reconstruction with dynamic network modeling. Future work will explore incorporating biological constraints into flow matching to jointly learn cellular dynamics and regulatory network structure within a unified framework.

References

1) *Vân Anh Huynh-Thu and Pierre Geurts. 2018. dynGENIE3: dynamical GENIE3 for the inference of gene networks from time series expression data. Scientific Reports 8, 1 (2018), 338*

2) *Vân Anh Huynh-Thu, Alexandre Irrthum, Louis Wehenkel, and Pierre Geurts. 2010. Inferring regulatory networks from expression data using tree-based methods. PLOS ONE 5, 9 (2010), e12776.*

3) *Aditya Pratapa, Amogh P Jalihal, Jeffrey N Law, Aditya Bharadwaj, and TM Murali. 2020. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. Nature Methods 17, 2 (2020), 147–154*

Poster n° 32

Adept: a secure, no-code platform for exploring governed biomedical datasets across the translational pipeline

Soumyabrata Ghosh (1)*, Artur Carvalho (1), Adept Development Team (1), Venkata P Satagopam (1)

1) Luxembourg Centre for Systems Biomedicine, University of Luxembourg

**soumyabrata.ghosh@uni.lu*

Biomedical research generates large, heterogeneous datasets to be shared between various stakeholders. In practice, clinicians, data stewards and data scientists often exchange spreadsheets, ad-hoc scripts (including R-shiny, Steamlit apps) and static reports. These workflows are slow to share, hard to reproduce and risky for access-controlled data. General-purpose tools such as Power BI carry administrative overhead, lack integration with research data-access approvals and rely on commercial cloud hosting, complicating GDPR compliance for sensitive data. We present Adept, a web-based visual analytics platform that brings import, organisation, filtering, visualisation and export of biomedical research data into one secure environment.

Adept is in operational use at the LCSB [1]. It supports data access through a GDPR compliant data governance mechanism. Data providers import tabular data files, confirm field types, and organise datasets into hierarchical collections. Each root collection is linked to an institutional dataset ID, so every data access can be traced back to the approval it was granted under. On the live LCSB Adept instance, approvals come through the ELIXIR Luxembourg Data Catalog [2] or an internal dataset registry through DS-Pack Tool assembly [3]. Role-based access control extends to row- and column-level permissions. Without writing any code, users build multiple dashboards per dataset from table, chart, metric and markdown note widgets, apply and save filters, and export filtered subsets for downstream analysis.

Adept also exposes its functionality through a Model Context Protocol (MCP) server [4], allowing AI assistants such as Claude to interact with the platform in natural language. Through it, users can browse collections, retrieve dataset summary statistics, create and organise collections and datasets, build dashboards and widgets, and monitor imports and exports, all within their existing Adept access permissions. Only metadata and aggregate statistics are exposed; record-level data is never shared with the AI model.

LCSB Adept instance supports projects spanning clinical trial monitoring (recruitment by centre, data completeness, dropout causes), cohort studies and drug-screening results. Adept shows how European research infrastructures can offer reproducible, governed data exploration that bridges domain experts and technical teams.

References

1) *Adept Website: https://getadept.eu, LCSB Adept instance: https://adept.lcsb.uni.lu*
2) *Elixir Data Catalogue: https://datacatalogue.elixir-luxembourg.org*
3) *DS-pack Tool Assembly: https://elixir-luxembourg.org/tools/ds-pack*
4) *Adept MCP URL: https://adept.lcsb.uni.lu/api/mcp*

Poster n° 33

Supporting Value Creation from Secondary Use of Data with the LNDS Service Portfolio

Pinar ALPER (1), Wei GU (1), Christophe TREFOIS (1),

Regina BECKER (1), Bert VERDONCK(1)

1) Luxembourg National Data Service, 6, avenue des Hauts-Fourneaux, L-4362 Esch-sur-Alzette

Realising the potential of big data and artificial intelligence in healthcare depends not only on access to data, but on the capacity to make data discoverable, accessible, interoperable, trustworthy and usable within appropriate governance and security frameworks. This challenge is particularly significant for secondary data use, where valuable insights may depend on bringing together heterogeneous data held across public administrations, research organisations and other sectors.

The Luxembourg National Data Service (LNDS) addresses this challenge through an integrated service portfolio designed to enable value creation from secondary use of data across the data lifecycle [1]. Its services span four complementary areas: Data Services, Technology Services, Ethical, Legal and Societal Implications (ELSI) Services, and Training and Community Services. Together, they support data holders and users from data discovery and preparation through access, analysis and eventual reuse or archiving.

At the data layer, the portfolio provides capabilities for cataloguing; access request and review; data management and stewardship; extraction, enrichment and merging; identifier and pseudonym management; data quality and curation; synthetic data generation; and analysis and visualisation. These are complemented by technology capabilities including a Secure Processing Environment and research data archiving, enabling sensitive data to be processed within controlled infrastructures. ELSI services address data protection, governance and responsible data and AI use, while training and national communities of practice strengthen organisational capabilities and promote shared approaches across the ecosystem.

Rather than treating data access as an isolated technical transaction, the LNDS approach addresses the end-to-end conditions required for responsible data reuse, connecting data holders, data users, governance expertise, stewardship and technical infrastructure through data projects. This is particularly relevant to healthcare and AI, where data quality, provenance, semantic interoperability, privacy protection and trustworthy processing directly affect the feasibility and reliability of downstream analytics.

By combining these capabilities within a reusable national service portfolio, LNDS provides an operational foundation for translating data governance principles into practical data reuse. The approach illustrates how national data service infrastructures can reduce organisational and technical barriers to secondary use while enabling sensitive data to contribute responsibly to research, evidence generation and data-driven innovation, including AI-enabled applications in healthcare.

References

1) *LNDS Service Team. (2026, September 30). Services to Enable Secondary Use of Data. https://www.lnds.lu/services/*

FROM SCATTERED DATA TO SMARTER CARE: AI FOR NEURODEGENERATIVE DISEASE

Aqila Farahmand (1), Soumyabrata Ghosh (1), Venkata P. Satagopam (1),

1) Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Esch-sur-Alzette, Luxembourg

Neurodegenerative diseases such as Parkinson's disease (PD) and dementia are a major and growing burden for ageing societies: about 50 million people live with dementia worldwide, a number projected to triple to 152 million by 2050 (2). Large volumes of relevant data are already generated in research cohorts and routine care, yet they remain hard to use. Key clinical information is recorded in unstructured, multilingual notes; cohort data are spread across sites and software systems; study data dictionaries are not interoperable; and cross-border regulations (GDPR, the European Health Data Space and Japan's APPI) constrain data reuse.

Unlike clinical NLP pipelines that stop at variable extraction, we aim to develop an end-to-end artificial intelligence (AI)-driven clinical data science pipeline that addresses these barriers in five steps: (i) secure data collection with ETL and pseudonymisation; (ii) extraction of clinical variables from free text using multilingual natural language processing and large language models; (iii) harmonisation to the OMOP Common Data Model and HL7 FHIR, using SNOMED CT and LOINC terminologies; (iv) integration in a hybrid federated dataspace coordinated by an agentic AI orchestrator; and (v) prediction and clinician-facing decision support. Its originality lies in combining LLM-based harmonisation of heterogeneous cohort and routine-care data with compliance-aware orchestration across European and Japanese regulatory frameworks, with GDPR, EHDS, EU AI Act and FAIR requirements built in by design.

In the Healthy Brain Ageing (HeBA) study, a four-country European hub that uses stepwise, population-based screening to identify people at risk of PD (1), a pseudonymised multi-site data platform integrates online surveys, smell tests and clinical visits; preliminary counts reach 36,868 surveys, 7,067 UPSIT smell tests and 609 in-person visits, with eligible participants referred to the Parkinson's Progression Markers Initiative. In ORCHESTRA–CONVINCE, which links COVID-19 cohorts with PD assessments (MDS-UPDRS), 902 variables from four REDCap data dictionaries were integrated into one structure covering 19 domains; only 31 variable names were shared across sources, motivating concept-level mapping. In AID-CARE, a planned Luxembourg–Japan collaboration, unstructured dementia electronic health records will be transformed into OMOP/FHIR datasets that feed a clinician dashboard and chatbot flagging patients at risk of rapid decline.

Together, these projects show how compliant, AI-ready data integration can turn scattered clinical data into actionable insight, enabling earlier detection of prodromal PD and dementia, reproducible cross-cohort research and more personalised care. Next steps include machine-learning risk stratification in HeBA, harmonisation of ORCHESTRA–CONVINCE, and clinician-in-the-loop validation of the AID-CARE tools.

References

1) *Schade S, Ghosh S, Garrido A, et al. Identifying individuals at-risk of developing Parkinson's disease using a population-based recruitment strategy: the Healthy Brain Ageing Kassel Study. npj Parkinsons Dis. 2025;11:216. doi:10.1038/s41531-025-01008-w*

2) *World Health Organization. Dementia: number of people affected to triple in next 30 years [news release]. Geneva: WHO; 7 December 2017. Available from: https://www.who.int/news/item/07-12-2017-dementia-number-of-people-affected-to-triple-in-next-30-years*

Evidence Reloaded

Pegah Shakibnia (1), Rojin Sadat Yasini (2), Negin Babaiha (2) and Martin Hofmann-Apitius (1) (2),

1) University of Bonn, Bonn-Aachen International Center for Information Technology (B-IT), 53115 Bonn
2) Fraunhofer Institute for Algorithms and Scientific Computing (SCAI), Schloss Birlinghoven, 53757 Sankt Augustin

Scientific evidence is typically encoded in natural language ... at least we believe it is evidence when we read assertions backed up by references in science publications . On the way to become evidence, some signal in a data set was previously identified and found to be reproducible in other data. Although epistemology knows more ways to establish insights that are informative, generalizable and ultimately robust against variation of observational boundary conditions, in biomedicine we usually deal with evidence that has its roots in well-designed experiments delivering (hopefully) high-quality data.

When we analyze evidence in the scientific literature, however, we find that data are often not linked to the evidence itself . Au contraire! Evidence seems to be somehow “uncoupled” from the experimental design and the data and the data analytics that are underlying the evidence. This makes non-trivial to trace back biomedical evidence to the original data where it stems from. Nanopublications addressed this issue for the first time, but fail to establish full provenance of data and signals in data used to establish evidence. Even in Evidence-Based Medicine (EBM), in most cases it will not be possible to trace-back the evidence that comes from Randomized Controlled Trials (RCTs). Meta-analysis of RCTs – a technique used to establish new guidelines based on a consensus of experts-in-the-field about evidence – does not require to reproduce the signal in the original data. But at least we know that there was a study with a certain design and we can build trust in the process of evidence synthesis.

Pre-clinical R&D by and large, however, does not yet have any systematic approach for the generation of new evidence. We extract assertions of the type “pathway A is upregulated under disease conditions” from the scientific literature and aggregate these statements in “disease maps”. We even know the exact position of the statement made in a referenced paper. However, this is typically where it stops. We do not know whether the referenced publication was the original one (the one with the underlying data such as clinical study data, multi-omics data, epidemiological / observational data) claiming the new evidence and we do not know whether this original one provided access to the data underlying the evidence-generating analytics. This can have substantial impact on the quality of pathway data, disease maps and other knowledge graphs, as un-rooted claims may propagate errors or inflated evidence not really based on signals in the original data. That this problem is a real problem becomes evident from the work of John Ioannidis at Stanford.

We have now developed an LLM-based claim-evidence matching approach that traces back evidence to the underlying data set. We call this process “evidence rooting” and can measure how large the discrepancy is between the number of claims made (“pathway A is upregulated in disease B”) and the claims that can be directly linked to underlying data. The process is fully machine-actionable and establishes a full evidence lineage layer. To our knowledge this is the first automated way to measure how many published claims are actually backed by accessible data.

The numbers we found in one of our application use cases (Kynurenic Acid as an anti-inflammatory compound) are frightening: the vast majority of assertions extracted from topic-specific corpora do not link back to underlying data and patterns in those data. We face a substantial disconnect between what we perceive as “evidence” and real evidence that can be tracked back to the underlying data.

This work was done in the context of EU-project COMMUTE (https://www.commute-project.eu/en/about.html), a project that has seen Reinhard Schneider as the PI for COMMUTE partner University of Luxembourg.

References

1. Baethge C, Jergas H. Systematic review and meta-analysis of quotation inaccuracy in medicine. Res Integr Peer Rev. 2025;10:13. doi:10.1186/s41073-025-00173-z

2. Hamilton, Daniel G., et al. "Prevalence and predictors of data and code sharing in the medical and health sciences: systematic review with meta-analysis of individual participant data." Bmj 382 (2023).

3. Kuhn, Tobias, et al. "Nanopublications: a growing resource of provenance-centric scientific linked data." 2018 IEEE 14th International Conference on e-Science (e-Science). IEEE, 2018.

4. Ioannidis, J. P. (2005). Why most published research findings are false. PLoS medicine, 2(8), e124.

5. Ioannidis, J. P. (2008). Why most discovered true associations are inflated. Epidemiology, 19(5), 640-648.