

Thomas Charlon, MEng, PhD

Boston, MA, USA — LinkedIn — GitHub — Website — thomas_charlon@hms.harvard.edu

EXPERIENCE

Harvard Medical School, Department of Biomedical Informatics <i>Research Associate - Senior postdoc, CELEHS laboratory</i>	Jan 2024 - Present <i>Boston, MA, USA</i>
<ul style="list-style-type: none">GenAI LLM agents with Azure/Ollama for endoscopy scores, heart failure and suicidality chart reviews, ontology creation (GPT-5, Llama, fine-tuned Phi4, R, Pytorch)Semantic search ontology webapp using LLM embeddings and vector database Elasticsearch, BGE fine-tuning, evaluation of multiple BERT models (R/Shiny, Pytorch)Suicide risk prediction models based on hospital data (EHR codified data and clinicians' notes), de-identification of notes, embedding clustering of suicide-related publications using GlovePatient history dashboard for rheumatology treatment switches, deployed at MGB hospital, VA-fundedOntology mappings creation used by multiple institutions (MGB, CHA, Duke, Pittsburgh, Mayo Clinic)Insurance claims co-occurrence analysis and knowledge graphs of 210M patients (Inovalon)Mentorship of 4 research assistants and 1 PhD student, lead of group meetings, collaborations with 2 postdocsProject quality processes, grant writing, onboarded at MGB and CHA hospitals, part of 4 IRBs	
Independent Researcher	Feb 2018 - Dec 2023
<ul style="list-style-type: none">Development of a real-estate market estimation application using R/ShinySocial network analysis on Twitter, identifying withheld content across Europe, chronological topic modelingAlgorithmic cryptocurrency trading, fully automated market-neutral system using APIs	
Quartz Bio, part of Precision for Medicine <i>Bioinformatician & PhD Student</i>	Feb 2014 – Jan 2018 <i>Geneva, Switzerland</i>
<ul style="list-style-type: none">Led genetic analyses for the EU-funded PreciseSADs, developed genome-wide clustering algs for autoimmunityVisiting researcher at Genyo Institute (Pfizer / University of Granada, Spain)	

SKILLS

Data Science & AI: Embeddings Semantic Search, LLM Fine-tuning, Unsupervised Clustering

Software & Web Development: R/Shiny, Pytorch, Langchain, FastAPI, JavaScript, SAS

Server Administration & Databases: SQL, PostGIS, Elasticsearch, Git, Docker Compose, AWS

LONG-TERM PROJECTS

Center for Suicide Research and Prevention — <i>Survival analysis, Risk prediction, EHR, LLM</i>	2024 – Present
<ul style="list-style-type: none">Analysis of patient data at MGB and CHA, cohorts and outcomes definitions using 2M patientsInnovative methods core (knowledge graphs) and demographics stratified modelingLLM agent chart reviews, feature interpretation with NLP, integration of results in R/Shiny web apps	
Real-Estate Price Estimation App — <i>Market pricing, Public-facing R/Shiny & JavaScript app</i>	2021 – 2023
<ul style="list-style-type: none">Visualization and neighborhood analysis of 15M French real-estate transactions over 5 yearsPublic facing app, server admin (PostGIS, OpenStreetMap), business model and marketing	
PreciseSADs, genetics of autoimmune diseases — <i>Unsupervised density-based clustering</i>	2014 - 2018
<ul style="list-style-type: none">Analysis of genome-wide SNP data of 10k patients, density-based clustering, haplotype summarization	

SELECTED TALKS ON YOUTUBE

- R/Medicine 2025:** Co-occurrence analysis and knowledge graphs for suicide risk prediction
- SoCal Linux 2025:** R/Python pipelines for biomedical semantic search
- R/Pharma 2024:** Knowledge graphs for drug discovery, **R/Medicine 2024:** Word embeddings in mental health

LEADERSHIP

- Open-Source Projects:** Published 6 R packages on CRAN, multiple citations (Nature), 50k downloads
- Mentorship:** Supervision of junior researchers, lectures at Harvard DBMI, collaborations with postdocs
- Project Management:** Led projects from inception to delivery, developing specifications with clients

EDUCATION

University of Geneva <i>PhD in Computer Science / Bioinformatics</i>	2014 - 2019 <i>Geneva, Switzerland</i>
Clustering and sparse coding of genome-wide data and proteomics in systemic autoimmune diseases	

EPITA <i>Master of Engineering in Computer Science</i>	2010 - 2013 <i>Le Kremlin-Bicêtre, France</i>
Artificial intelligence and cognitive sciences major, C++, Java, Matlab	