

Alexandre Yahia

PhD candidate in Biomedical Informatics

*Department of Biomedical Informatics
PH-20, 622 W 168th Street, New York, NY 10032
☎ +1 (718) 737 6678
✉ alexandre.yahia@columbia.edu
📍 AYahia*

EDUCATION

- 2015–present **Doctor of Philosophy**, *Columbia University*, New York, NY.
Department of Biomedical Informatics
- 2015–2018 **Master of Philosophy**, *Columbia University*, New York, NY.
Department of Biomedical Informatics
- 2015–2017 **Master of Arts**, *Columbia University*, New York, NY.
Department of Biomedical Informatics
- 2013–2015 **Master of Science in Electrical and Computer Engineering**, *Georgia Institute of Technology*, Atlanta, GA.
College of Engineering, Department of Electrical and Computer Engineering
- 2011–2014 **Master of Science in Engineering - Diplôme d'ingénieur**, *Supélec*, (now CentraleSupélec), Paris, France.
Top-3 Engineering School in France
- 2011–2012 **Bachelor of Science in Fundamental Physics**, *University Orsay - Paris XI*, Orsay, France.
Double degree obtained through evening courses during my 1st year at CentraleSupélec.
- 2008–2011 **Advanced Physics, Mathematics and Chemistry preparatory classes**, *Lycée Janson de Sully*, Paris, France.
Intensive 3-year preparation for the national competitive entrance examinations to French graduate engineering schools, specializing in mathematics, physics and chemistry: achieved with High Honors

RELATED AND PROFESSIONAL EXPERIENCE

- 2016–2017 **Scientific Advisor**, *Roam Technologies, LLC*, New York, NY.
Sold IP to a smart cities startup working in tech innovation for travel and hospitality, accompanied the development of their business model and the portage of our technology into their product.
- 2014–2015 **WELL**, New York, NY.
Prototyped a crowdsourced health platform connecting researchers and volunteers as part of IE@Columbia Program. Developed business plan, implemented iPhone app using HealthKit and Amazon Web Services, and regularly presented to Columbia University Business School faculty and angel investors.
- 2015 **Innovation and Entrepreneurship @ Columbia**, *Columbia Business School*, New York, NY.
Competed for entry and accepted into cross-University program that develops and leverages ideas for new ventures. Received training from faculty members, alumni, and industry mentors to develop and fundable, high-potential business plan.

Fall 2013 **Engineering consultant**, *Cell Constraint & Cancer*, Paris, France, 4-month part-time mission.

Analyzed experimental data for a biotech start-up on a disruptive innovating treatment in oncology, in partnership with the Curie Institute in Paris (leading research institute in biophysics and oncology)

PUBLICATIONS, PREPRINTS AND MANUSCRIPTS

Peer-reviewed publications

1. F.C.G. Polubriaginof, R. Vanguri, K. Quinnes, G.M. Belbin, **A. Yahi**, H. Salmasian, T. Lorberbaum, V. Nwankwo, L. Li, M.M. Shervey, P. Glowe, I. Ionita-Laza, M. Simmerling, G. Hripcsak, S. Bakken, D. Goldstein, K. Kiryluk, E.E. Kenny, J. Dudley, D.K. Vawdrey, N.P. Tatonetti. (2018) Disease Heritability Inferred from Familial Relationships Reported in Medical Records, *Cell*.
2. **A. Yahi**, R. Vanguri, N. Elhadad, N.P. Tatonetti. (2017) Generative Adversarial Networks for Electronic Health Records: A Framework for Exploring and Evaluating Methods for Predicting Drug-Induced Laboratory Test Trajectories, *Neural Information Processing Systems (NIPS) 2017, Machine Learning for Health Workshop*.
3. P.J. Kennel, A. Saha, D.A. Maldonado, R. Givens, D.L. Brunjes, E. Castellero, X. Zhang, R. Ji, **A. Yahi**, I. George, D.M. Mancini, A. Koller, B. Fine, E. Zorn, P.C. Colombo, N.P. Tatonetti, E.I. Chen, P.C. Schulze. (2017) Serum exosomal protein profiling for the non-invasive detection of cardiac allograft rejection, *The Journal of Heart and Lung Transplantation*.
4. K.W. Johnson, K. Shameer, **A. Yahi**, R. Miotto, L. Li, D. Ricks, J. Jebakaran, P. Kovatch, P. Sengupta, A. Gelijins, A. Moskovitz, B. Darrow, D.L. Reich, A. Kasarskis, N.P. Tatonetti, S. Pinney, J.T. Dudley. (2016) Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai Heart Failure Cohort, *Pacific Symposium on Biocomputing 2016*.
5. K.W. Johnson, K. Shameer, **A. Yahi**, R. Miotto, D. Ricks, J. Jebakaran, P. Kovatch, P. Sengupta, D.L. Reich, A. Kasarskis, S. Pinney, N.P. Tatonetti, J.T. Dudley. (2016) Pharmacological Factors Associated With Congestive Heart Failure Hospital Readmission: A Case-Study Using 15,768 Heart Failure Patients From Two Health Systems, *Circulation*.
6. R. Brossel, **A. Yahi**, L.M. Velasquez, J.M. Guinebrière. (2016) Mechanical signals inhibit growth of a grafted tumor in vivo: Proof of Concept, *PLOS ONE*.
7. **A. Yahi**, A. Chassang, L. Raynaud, H. Duthil, D.H.P. Chau. (2015) Aurigo: An Interactive Tour Planner for Personalized Itineraries, *ACM IUI 2015*.
8. **A. Yahi**, N.P. Tatonetti. (2015) A knowledge-based, automated method for phenotyping in the EHR using only clinical pathology reports, *AMIA Joint Summits 2015*.

Pre-prints (non-refereed)

1. **A. Yahi**, T. Lappalainen, P. Mohammadi, N.P. Tatonetti. (2018) RecNW: A fast pairwise aligner for targeted sequencing, *bioRxiv*

Manuscripts under review

1. **A. Yahi**, T. Lappalainen, P. Mohammadi, N.P. Tatonetti. RecNW: A fast recurrent pairwise aligner for targeted sequencing.
2. **A. Yahi**, P.J. Kennel, Y. Naka, D.M. Mancini, C.C. Marboe, K. Akat, T. Tuschl, E.M. Vasilescu, E. Zorn, N.P. Tatonetti, P.C. Schulze. (2018) Identification of serum miRNA predictors and biomarkers of allograft rejection following heart transplantation using next generation small RNA sequencing.

Manuscripts in preparation

1. **A. Yahi**, P. Hoffman, P. Mohammadi, N.P. Tatonetti, T. Lappalainen. EdiTyper: a high-throughput tool for genotype analysis of CRISPR-Cas9 genome editing experiments.
2. **A. Yahi**, F. Polubriaginof, J.D. Romano, D.K. Vawdrey, N.P. Tatonetti. Open Data: A review of current methods used for discovery in biomedical research.
3. **A. Yahi**, A. Basile, N.P. Tatonetti. Predicting trends in drug-induced laboratory measures: using deep embeddings for irregularly sampled time series.

CONFERENCE PRESENTATIONS

- 2018 **Generative adversarial networks: evaluating drug effect in synthetic time series generated from EHR Data**, *AMIA 2018 Informatics Summit, San Francisco, CA*.
- 2017 **Repurposing drugs for precision medicine: haplotype-specific modular assembly of drug safety subnetworks (hapMADSS)**, *AMIA Joint Summits on Translational Science 2017, San Francisco, CA*.
- 2015 **Improving user engagement and insight through Contextualized Quantified Self**, *AMIA Symposium 2015, San Francisco, CA*, (Student Design Challenge, Second Place).
- 2015 **A knowledge-based, automated method for phenotyping in the EHR using only clinical pathology reports**, *AMIA Joint Summits 2015, San Francisco, CA*.
- 2015 **Aurigo: An Interactive Tour Planner for Personalized Itineraries**, *ACM Intelligent User Interface 2015, Atlanta, GA*.

RESEARCH EXPERIENCE

- 2015–present **PhD Candidate**, *Columbia University, New York, NY*.
Doctoral Advisor: Nicholas P. Tatonetti
Thesis Committee: George Hripcsak, Noémie Elhadad
Using deep generative models to simulate drug effects on laboratory test time series: advancing the field with deep representation of irregularly sampled time series, developing novel quantitative metrics specific to EHR data to evaluate generative models, and building multi-modal generative adversarial networks (GANs) models for EHR data generation and simulation.
- Fall 2015 **Research Rotation**, *New York Genome Center*, with Tuuli Lappalainen.
Developed a novel fast pairwise aligner for next-generation sequencing data and a tool to analyze the results of CRISPR-Cas9 gene editing experiments.

Fall 2015 **Research Rotation**, *Columbia University*, with George Hripcsak.
Helped to implement 46 phenotyping algorithms from the PheKB platform by developing algorithms relying on clinical NLP tools for free text clinical notes, to extract terms and map them to standard terminologies (RxNorm, SNOMED, ICD-9/10) and identify section headers for contextualized named-entity recognition (NER)

2013–2015 **Graduate Student Intern**, *Columbia University - Department of Biomedical Informatics*, New York, NY.

Principal Investigator: Nicholas P. Tatonetti, PhD

- **Full-time:** June-September 2013 / June-December 2014 (9 months)
- **Part-time:** September 2013 - May 2014 / December 2014 - August 2015
- **Projects:** Worked with EHR data (OMOP CDM), explored how to leverage laboratory tests for phenotyping and used medical ontologies and terminologies to augment EHR data-driven methods.

TEACHING EXPERIENCE

Spring 2017 **BINF G4001 - Computational Methods**, *Teaching Assistant*, Department of Biomedical Informatics, Columbia University.

Spring 2017 **BINF G4001 - Computational Methods**, *Guest Lecturer*, Department of Biomedical Informatics, Columbia University.

2 lectures: Decision trees and random forests; Introduction to Deep Learning principles

Fall 2016 **BING G4000 - Acculturation to Programming and Statistics**, *Teaching Assistant*, Department of Biomedical Informatics, Columbia University.

2011–2013 **Advanced Chemistry**, *Examiner - college level*, Lycée Stanislas & Lycée Saint-Louis, Paris, France.

Taught advanced Chemistry through weekly oral examinations for 3 to 4 classes a year

PROFESSIONAL SERVICE

2018 **Finance and Publication co-chair**, *Machine Learning for Health Workshop, NIPS 2018*.

2018 **Reviewer**, *Machine Learning for Health Workshop, NIPS 2018*.

2015-2018 **Reviewer**, *AMIA Symposium and Summits*.

2018 **Reviewer**, *Journal of Biomedical Informatics*.

SELECTED AWARDS AND HONORS

March 2018 **Student Travel Grant**, *Graduate Student Advisory Council*, Columbia University.

December 2017 **Student Travel Award**, *Machine Learning for Health Workshop*, Neural Information Processing Systems (NIPS) 2017.

November 2015 **Finalist, AMIA Student Design Challenge**, *The Human Side of Big Data – Facilitating Human-Data Interaction*, AMIA Symposium, San Francisco, CA.

Awarded 2nd place out of 16 applicants for Contextualized Quantified Self visualization.