

Quentin-Gabriel Thurier

quentin.thurier@gmail.com

+64 29 127 2170

Profiles:  

129 Hobsonville road, West Harbour, Auckland, New Zealand

France & New Zealand citizen

SKILLS

Science: Machine Learning, Deep Learning, Statistics, Natural Language Processing

Languages: Python, Bash, SQL, R, SAS, Java

ML Engineering: TensorFlow, scikit-learn, spaCy, Git, Github, Prodigy, Argilla, DVC, Prefect, Docker

Data Engineering: Snowflake, Redis, ArangoDB, Elasticsearch, Hadoop, Hive, Spark, Big Query, Redshift

Cloud Platforms: Amazon Web Services, Google Cloud Platform

Visualisation: Dash, Shiny, Tableau, Snowsight, MicroStrategy, Superset

Project Management: Confluence, Jira, Jira automation, Miro

Development Environments: VS Code, PyCharm, JupyterLab, JupyterHub, RStudio

EXPERIENCE

- **Decoded Health** Auckland, New Zealand
Senior Machine Learning Engineer June 2023 – Present
 - **Develop and maintain Decoded Health's AI medical resident deployed in urgent care clinics.**
 - Prompt engineering, Few-shot learning, Retrieval-augmented generation, R&D on *named entity linking* and clinical *text generation* using large language models - Establishing standards for data versioning, ML-based services experimentation and monitoring - Taking part in the on-call schedule
- **Xero - Science, Evaluation and Measurement team** Auckland, New Zealand
Senior Applied Scientist - ML Products November 2020 – May 2023
 - **Worked as a senior member of the cross-functional team in charge of Xero's ML-powered bank reconciliation production system.**
 - R&D on *learning to rank* and *named entity recognition* architectures applied to bank statements records - ML engineering - Collaborating with product and intellectual property teams - Mentoring applied scientists
- **PredictHQ - Data Intelligence team** Auckland, New Zealand
NLP Data Scientist May 2019 – November 2020
 - **Developed and maintained PredictHQ's API ML-driven features together with data and software engineers.**
 - Leading ML projects for events labelling/deduping/ranking from data collection to productionisation - R&D on *named entity linking* systems with open source knowledge graphs - Mentoring data analysts
- **Orion Health - Research team** Auckland, New Zealand
Data Scientist February 2017 – April 2019
 - **Shaped Orion Health's ML-based products features with clinicians and software engineers.**
 - Implementing ML *classifiers* for health risk calculation - Principal investigator for a multi-disciplinary project on *AI interpretability* - Deep learning for clinical NLP - Mentoring interns and junior data scientists
- **Qrious - Spark Ventures** Auckland, New Zealand
Data Scientist September 2015 - January 2017
 - **Worked with data engineers to deliver solutions leveraging Spark New Zealand mobile devices activity data.**
 - Bespoke *geospatial studies* for the tourism and transport industries - Automating insights discovery with *big data* technologies - R&D on population flows *clustering*, transport mode and work/home locations inference
- **NetBooster - Data and Analytics team** Paris, France
Senior Data Manager July 2014 – August 2015
 - **Implemented quantitative digital marketing prototypes with server logs data in collaboration with web analytics experts.**
 - Clickstream *data mining*, Quantifying the efficiency of multi-channel marketing campaigns - R&D on users online journeys *clustering* - Data engineering

- **Société Générale Corporate and Investment Banking - Anti-Fraud team** Paris, France
Data Scientist October 2010 – October 2013
 - Worked as part of the cross-functional team responsible for implementing a data-driven strategy to prevent unauthorized trading.
 - Bespoke analyses on daily financial transactions - *Unsupervised ML* to segment/rank portfolios - Writing/auditing the scripts at the core of the fraud prevention software
- **Société Générale Corporate and Investment Banking - Anti-Fraud team** Paris, France
Intern - Master's thesis focused on statistical methods for preventing unauthorized trading April 2010 – Sept. 2010
- **Colorado State University - Statistics Department** Fort Collins, USA
Intern - Time series analyses of carbon emissions June – July 2009
- **Euler Hermes SFAC (credit insurance) - Economic Studies department** Paris, France
Intern - Implementating dynamic economic dashboards July – August 2008

EDUCATION

- **Telecom ParisTech** Paris, France
Post Master's degree in Machine Learning and Big Data ("Mastère spécialisé") October 2013 – June 2014
- **University of Washington** Seattle, USA
Certificate in Data Science - Distance training January – June 2013
- **ENSAI, National School for Statistics and Information Analysis** Rennes, France
Master of Science in Statistics equivalent ("Diplôme d'ingénieur") 2005 – 2010

PATENTS

- | | |
|--|------------------------|
| Method, Program, and Apparatus for Processing Sensitive Data | Published - May 2023 |
| Methods and Systems for Training Attribute Prediction Models | Published - March 2023 |
| Transaction Data Processing Systems and Methods | Published - June 2022 |

RESEARCH

- | | |
|--|----------------|
| Physician understanding, explainability, and trust in a hypothetical ML risk calculator | JAMIA 2019 |
| Smart MedRec: Using ML for reading dose instructions and incorporating this in clinical software | HINZ 2019 |
| Inspecting a ML-based clinical risk calculator: a practical perspective | IEEE CBMS 2019 |
| Interpretable ML for healthcare | HINZ 2018 |
| Improving clinical named entity recognition with transfer learning | HIC 2018 |
| HOPE (Health Outcomes Prediction Engine) for stroke | HINZ 2017 |
| New Zealand health data review | HINZ 2017 |

CERTIFICATIONS

- | | |
|---------------------------------|--------------------|
| Google Cloud Platform Developer | Google |
| Developer for Apache Hadoop | Cloudera |
| Apache Spark Developer | Databricks |
| SAS Base programming | SAS Institute inc. |
| Data Science Associate | EMC |

INTERESTS

- Judo (2nd degree black belt, coaching, former club secretary)
 Surf
 French Cinema