Kassy **Raymond**

BSc, MSc, PhD Candidate

linkedin.com/in/kassy-r-2a477717a github.com/kassyray
613-281-3754 kassyraymond2@gmail.com
https://kassyray.github.io

Skills and Competencies

Programming	Python (pandas, dash, plotly, matplotlib), SQL, Cypher, R (tidyverse, shiny, ggplot2), perl, C, shell, awk, SAS
Database Management	Neo4j, Neo4j Aura, Microsoft Access, RDS, MySQL, PostgreSQL
Dashboard Development	dash plotly, Rshiny, Microsoft PowerBI, Tableau Visualization Software
Operating Systems	Mac OS X, Windows, Linux
Other	docker, kubernetes, git, github, svn, Cloud Platforms (AWS, Azure, GCP), Trivy, GitHub actions, Kubeflow, Kubeflow Pipelines

Experience

Apr 2025 – Present

Data Engineer, Wellington-Dufferin-Guelph Public Health, Hybrid

- Co-designed and maintained automated data pipelines for ISPA and CCEYA immunization reporting using **Python**, YAML, and Typst, improving delivery speed and accuracy across multiple PHUs.
- Developed and modularized an end-to-end reporting pipeline using Bash, **GitHub Actions**, and **Docker** enabling efficient batch processing, reproducibility and performance tracking of the above pipeline.
- Co-design and development of a **data quality** library through incorporation of integrated data validation frameworks (Great Expectations) to enhance data quality and confidence in downstream analyses.
- Contributed to technical **documentation** (WDGPH Development Guide) and public-facing materials, supporting transparency and usability of data products for internal and external stakeholders.
- Contributed to **data governance** efforts through connecting with my network of data governance experts to organize speakers for the Data Governance working group.
- Contributed to **data governance** efforts by implementing column making and schema validation across datasets arriving from other PHUs.
- Submitted, edited, and participated in meetings related to **Privacy Impact Assessments** for the transformation of data pipelines, storage, and infrastructure to support ongoing data and AI transformation efforts.

R | Typst | Python | Git | Kubeflow

Jan 2025 – Current

Data Engineer, Food for Thought Dynamic Dashboard Project, Remote

- Design and implementation a PostgreSQL database to efficiently store user data.
- Leveraged AWS S3 with pre-signed URLs to securely handle file uploads and downloads, ensuring seamless user access to data.
- Development of AWS Lambda functions to automate data processing workflows, including parsing input spreadsheets, triggering model execution, and storing results.
- Integrated cloud-based storage using S3 to manage dynamic datasets, enabling efficient retrieval and interaction with processed results.

AWS | SQL | Python | R | S3

May 2024 – Oct 2024

Research Associate, St. Joseph's Healthcare Hamilton, Hamilton, Ontario

- Development of research protocol for collecting thermal imaging data in hemodialysis unit using a thermal camera (FLIR).
- Development of data pipeline to ingest, process, and deliver sensitive/private image data collected to the university server.
- Communicated protocols via written documentation, ensuring that data documentation was understandable by hospital staff.
- Communicated with collaborators and project manager at hospital, ensuring that protocols were feasible and understandable in the unit.

python | FLIR Research Studio | Sharepoint | C++

Sep 2020 – May 2024

Technical Manager, Global Burden of Animal Diseases (GBADs) Informatics, Remote

- Served as primary subject matter expert for data products such as metadata, datasets, databases, dashboards, and the organization's GitHub.
- Lead the development of dashboards (python dash) displaying livestock population and biomass data at international and national scales. Trained stakeholders and gathered feedback on dashboard progress at an international meeting in Addis Ababa, Ethiopia (in-person).
- Fulfilled data requests from international stakeholders through writing SQL queries to extract data.
- Designed, developed, and maintained a Cloud-Based Graph Database Management System for managing metadata and classifications. Developed an Application Programming Interface in python and R tools.
- Lead efforts on the GBADs Data Portal (named Trusted Animal Information portaL (TAIL - currently in alpha version)).
- Wrote and maintained the GBADs Data Governance Handbook (written in GitHub pages), which outlines procedures, best practices, and organizational data values. Emphasized efforts to make data FAIR (Findable, Accessible, Interoperable, and Reusable).
- Developed webscraping scripts in python to collect Antimicrobrial Use data from javascript rendered webpages.
- Organized and presented (in-person and virtual) a series of workshops for modellers working with data. Topics included : interacting with APIs in R and python, learning GitHub, data governance for GBADs, importance of metadata for data discovery and reuse, creating dashboards and data visualizations, best practices for formatting spreadsheets for data reuse.

RShiny | AWS | GCP | Sharepoint | MS Excel | R | Python | Neo4j | SQL | markdown | RDS

Nov 2019 – Now

Independent Consultant, Self-Employed, Guelph, Ontario

- Global Burden of Animal Diseases Senegal Data Ecosystem Mapping (2023)
  - Collaborated with in-country government stakeholders to assemble a data inventory of available private and open datasets.
  - Created questionnaires in Excel to collect metadata on available data.
  - Cleaned and prepared collected data using R.
  - Developed an interactive data visualization of the data ecosystem in Senegal using kumu.io, allowing the visualization to update in-real time during the in-country stakeholder meeting.
- Global Burden of Animal Diseases Aquaculture Data Management and Governance Strategy (2021)
  - Identified and interviewed key data stakeholders in the salmon industry in three countries.
  - Created a preliminary inventory of relevant data and metadata related to salmon production.
  - Designed the infrastructure for a cloud-based information management system for open and private aquaculture data.
  - Created a preliminary data governance strategy for managing aquaculture data.
  - Disseminated results via a written report. Delivered a Powerpoint to a non-technical audience outlining findings and distilling complex technical details.
- Global Burden of Animal Diseases Ethiopia Pilot Study (2019)
  - Created a comprehensive data inventory of openly available poultry data.
  - Evaluated open data quality for Ethiopia's poultry sector.
  - Identified and communicated with key data stakeholders in Ethiopia.
  - Presented results through reports and voice-annotated data visualizations.

latex | bash | python | C | AWS | Azure | GCP | kumu | R | Rmarkdown | dash | Google Sheets | Excel

Jan 2020 – Apr 2024

Graduate Teaching Assistant, School of Computer Science, University of Guelph

- Served as a Graduate Teaching Assistant for various courses including CIS*2250 Software Design II, CIS*3010 Systems Programming, and CIS*4010 Cloud Computing.
- Created shell scripts to automate grading of assignments.
- Delivered guest lectures, ran labs and seminars, and held office hours.
- Co-authored CIS*2250 course textbook : "Collaborative Design Fundamentals for Software Engineers".

latex | bash | Python | C | perl | AWS | Azure | GCP

Sep 2020 – Dec 2020

Part-time Machine Learning Research Associate, WIL Digital Internship Competition, Remote

- Prepared and cleaned biosignal (heart rate, resp rate, electrodermal activity) data collected from patients exhibiting psychological stress in a lab.
- Created a data pipeline to calculate features from raw signals in preparation to be used in an unsupervised machine learning analysis.
- Used scikit learn to cluster data to determine whether periods of psychological stress could be identified from the signal.

Python | R | scikit learn | matplotlib | ggplot2

May 2020 – Sep 2020 | **Computer Vision for COVID-19 Screening Co-Lead | Developer, iris | University of Guelph, Remote**

- Co-lead of a team responsible for developing software for COVID-19 screening using thermal and visible light images.
- Mentored and managed co-op software development students. Acted as liaison between iris and the team.
- Designed a prototype dashboard to display human biosignals derived from thermal imaging.

Python | git | github | yoloFace | openCV | dash | Rshiny

Education

2025 Doctor of Philosophy - Computational Sciences, College of Engineering and Physical Sciences, University of Guelph, Ontario (4.0 GPA (97%))

2021 Master of Science - Bioinformatics, College of Engineering and Physical Sciences, University of Guelph, Ontario (4.0 GPA (93 %))

2019 Bachelor of Science - Biological Sciences, College of Biological Science, University of Guelph, Ontario