

David Ravensborg - Developer/Data Scientist

<https://www.linkedin.com/in/david-ravnsborg-b624b658/> | <https://stackoverflow.com/users/724403/david>

Key Skills/Features

- 6 FTE years of software development and FTE 3 years of data science experience.
- Experience with data science, product analytics, product development, data wrangling, visualization, data pipelines, hypothesis testing, predictive modeling, and cloud/local deployments.
- Love of science and mathematics. Constantly curious and exploring how the world works.
- Currently reading The Shape of Data, which is about underlying geometries in data and their applications in data analysis and ML.
- Proficient with task tracking software (Jira, Shortcut, MS Planner, ad-hoc spreadsheets, etc.) and project planning. Not too attached to any one approach, so long as it's consistent and facilitates communication.
- Lots of experience with remote and hybrid teams, in multiple time zones covering North America, Europe, and India.
- Primary languages: Python, R, SQL; plus background with JavaScript, C#, Java, PHP, Matlab, and Bash.

Projects & Experience

Data Scientist - Deployment Specialist

04/2022 to Present

Purpose Analytics

Edmonton, AB

- Focused on building data pipelines and automation solutions to integrate data from a variety of sources, with an emphasis on idempotent design to ease debugging and reproducibility.
- Platform engineering: Data pipelines and tooling support are used on many dashboarding projects, a CKAN open data portal we are preparing to publish, and a GIS web app. Solutions are developed in R and Python, and integrate with or are deployed across a variety of Azure infrastructure.
- On my largest project, I work closely with 2 other data scientists and coach them on how to use various development tools like git, logging and debugging practices, validating processed data, etc; while collaborating on data modeling and aggregation strategies based on the raw datasets they explore.

Data Scientist (personal finance project)

09/2017 to Present

Petrichor

Edmonton, AB

- Utilizing web scraping and financial data APIs to train machine learning models, which generate stock market insights for my own personal curiosity and investing. This has been a long-term personal project which has made use of the following technologies: Python, Matplotlib, Pandas, Vaex, SHAP, SQL, Sklearn, Keras, XGBoost, and MLFlow.
- I currently use an on-demand ELT (Extract, Load, Transform) approach to pull and transform most of my data from 3rd party APIs. However, I have a new and rather large options dataset for which I am planning idempotent pre-processing jobs. Being idempotent just means it can be run multiple times (even concurrently) and not cause unexpected additional changes to the target dataset in the data warehouse (i.e. the operation is repeatable while giving a consistent result). This is a good design principle for most types of concurrent systems, and APIs/data pipelines, especially when writing data. That way I can also run the processing job manually to get the latest data on-demand if I need to, and not worry about negatively impacting already processed data in my data warehouse target, if it was run recently.
- Hypothesis-testing pipeline which allows me to test many hypotheses (combinations of features, models, etc) in an automated way against a particular goal (alpha/profit relative to a particular index like the S&P 500) after writing experiment and data transformation parameters in their respective YAML files. This also lets me easily re-visit old experiments/datasets.

David Ravensborg - Developer/Data Scientist

<https://www.linkedin.com/in/david-ravnsborg-b624b658/> | <https://stackoverflow.com/users/724403/david>

- For automated data downloading I created a bot using Google App Engine with Flask, which I schedule with Cron jobs, storing data as unstructured files in Google Cloud Storage.

Technical Lead/Co-owner (startup)

04/2021 to 12/2021

TechUp

Edmonton, AB

- Created a proof of concept for a QR-code driven digital inspection app, from ideation. I was approached by a personal contact about a perceived gap we could fill in his industry in digital inspections, and so we explored: current inspection practices, ideal UX workflows for both workers and managers, identified competitors, created wireframes, and began creating a pair of multi-tenanted apps to test the concept.
- We eventually found another competitor already offering a similar suite of features, with years of history in compliance and safety inspections, a large team, venture capital financing, and a product at a much lower price point than we had anticipated. This cast doubt on the viability of the chosen business model. I advised against continuing with this project to avoid wasting time and resources, and we put the project on an indefinite hold.

Senior Software Developer (startup)

06/2021 to 11/2021

Trust Science

Edmonton, AB

- I worked on JavaScript web app development across multiple disparately structured apps, to support our clients with loan origination. Key challenges included being the first hire on the apps team in over 2 years and applying for a Data Developer role but joining as a Software Developer, with many more front end app related responsibilities than data pipeline/API related responsibilities. I focused on learning their applications, infrastructure, fixing bugs, delivering new features, and improving the on-boarding process for subsequent hires.
- I also supported our data modeling initiative with the goal of getting new developers and data scientists up to speed quickly with our data pipelines. I updated the documentation and parameterized their Python pipeline to make it config driven, in order to reduce the amount of manual edits and copy+pasting needed to begin modeling a new data source. The motivation behind this came from some data workshops where we tried getting additional developers, analysts, and data scientists - myself included - exposed to the company's primary data modeling pipeline. I listened for the challenges that some developers were having, documented other challenges I had encountered, solicited feedback, developed a setup & debugging guide, and shared it.

Software Developer

03/2019 to 06/2021

ATB Financial

Edmonton, AB

- Using JavaScript with React, Express, Enzyme, GraphQL, as well as Selenium, Pytest, GCP/Kubernetes, SAP, and our internal enterprise services to create great web applications for our internal ATB users. I worked on a cross-functional scrum team on all parts of the application, including development, testing, and analytics.
- On the data science and machine learning side, I worked on: a loan default propensity modeling project (short term), an internal hackathon to build a model estimating a client's propensity to seek new credit (short term), and exploratory analysis of user analytics data from the main web application I helped develop (long term). In these I primarily used SQL, Python, Pandas, Numpy, Sklearn, Keras, and GCP Big Query.
- Key Achievements:
 - Designed a solution to make our UI layer work seamlessly with our back-end authorization systems, to improve the UX for our users who have varying levels of back-end permissions. I identified it as a missing gap in our application, since presenting users with functionality they are unable to use, and showing an error when

David Ravensborg - Developer/Data Scientist

<https://www.linkedin.com/in/david-ravnsborg-b624b658/> | <https://stackoverflow.com/users/724403/david>

they attempt to use that functionality, is a poor user experience. I researched our SAP security framework and identified classes of users by entitlements/authorizations, and led most of the design and implementation of the solution; working closely with product owners, security, and UX designers. This authorization check became a core piece of functionality for responsive UX and analytics across the app.

- Revived our in-app user analytics, fixed key issues, analyzed and created an ETL pipeline to transform data and produce insights (including visualizations), and put that in the hands of stakeholders with the help of our product management team. I began by taking the initiative on early exploration and visualization of our existing analytics data, as the efforts had stalled when we lost the original analytics implementor. This included leading the design and rollout of new data structures for our analytics events across the app, which had two additional benefits: it captured additional relevant information like the lifecycle of a particular feature interaction, and it was more easily consumable by myself and our analysts. This ended up seeing widespread use by our partners/stakeholders within ATB, as well as by our product team, to prove ATB users were making growing use of our application. It also gave insights on what features they were using most, which helped us prioritize bugs and stories.

QA Developer

10/2017 to 12/2018

Questionmark

Edmonton, AB

- Worked remotely on automated testing, environment deployment and configuration, and coordinating releases for 3 teams working on one common set of applications (customer portal). Roughly 70/30 split between time spent on development of our C# automation framework, and other aspects of the role.
- Core areas of focus include: creating a comprehensive test plan, coordinating team releases and testing, monitoring testing to provide feedback to the development teams as a gatekeeper for product quality, and improving test agents/suites to improve their feedback and increase test runner resilience (i.e. reducing false positives).
- Key Achievements:
 - Reduced false positive rate by factor of 10 - the rate at which tests fail without a real underlying issue - of 3 test suites, with 40-200 tests each, from 35% to 3-5%.
 - Initiated the development of better test agent cleanup practices to prevent memory leaks that were severely impacting our ability to run tests, and new extensible methods for element verification that saw use across the QA Automation team.

Undergraduate Researcher

01/2017 to 08/2017

University of Alberta

Edmonton, AB

- Received an 8-month stipend from the university for my undergraduate research project, where I programmed and analyzed machine learning models in Python to predict a participant's subsequent steps from walking data gathered in the lab.
- To explore the hyperparameter space I used a monte carlo search over a wide area, then a grid search procedure within the best performing ranges, to train and compare many versions of an LSTM neural network model for predicting time-series data related to gait analysis.

Research Assistant in Clinical Gait Analysis (4 month practicum)

05/2016 to 08/2016

Alberta Health Services (Glenrose Rehabilitation Hospital)

Edmonton, AB

- Investigated the kinetics and kinematics underlying gait initiation using the Syncrude and Courage in Motion Centre gait labs.

David Ravensborg - Developer/Data Scientist

<https://www.linkedin.com/in/david-ravnsborg-b624b658/> | <https://stackoverflow.com/users/724403/david>

- Programmed models in Matlab for analyzing and anticipating gait initiation on the Computer Assisted Rehabilitation Environment (CAREN), working with high frequency force plate data (~30000hz) measured from a special treadmill.
- Assisted in the Syncrude gait lab with participants undergoing gait assessment.

Hardware Technician (part-time while in school)

06/2014 to 01/2015

GFI Systems

Edmonton, AB

- Hardware configuration and script deployment, software testing of GPS and fleet management systems, being the liaison between customers and the development team on certain feature rollouts (i.e. vehicle on-board diagnostics), testing of new and returned equipment, plus client support and troubleshooting.
- Frequently constructed SQL queries for checking parameters not yet displayed on our front-end systems and modified Java programs for client-tailored equipment configurations.
- Wrote and used extensive process documentation for communicating with other hardware technicians, client service representatives, accounting, and the development team.

Game Developer (startup)

04/2013 to 10/2013

Hungry Moose Games

Edmonton, AB

- Primary project: Developed and supported an Android and iOS game for children and parents called "9Lives: Casey and Sphynx".
- Used Mono C# and the Unity engine, which is an entity-component oriented framework.
- Secondary project: Supported development of a nanotech educational game for grade 4 students sponsored by the National Institute of Nanotechnology in Edmonton.

Software Developer (contract)

03/2012 to 02/2013

General Electric

Edmonton, AB

- Did web application and platform development: focused on C#/jQuery feature implementation for the UX team and integrating apps with new/existing systems using ASP.NET MVC, Windows Azure and WCF (SOAP) for the platform team in an n-tier architecture.
- Created unit/system tests to ensure continued system stability.

Data Analyst / Web Developer (startup)

05/2011 to 01/2012

Amateur Sports Database (asdb.com)

Vancouver, BC

- Worked with the owner and lead developer to analyze key concerns around data-gathering and website presentation.
- Saved the company from hiring multiple manual data entry positions through the use of automated scripts and web crawlers to gather up-to-date statistics on minor and junior hockey.
- Created a data import and cleaning pipeline for handling both automated and manual uploads to the database, using PHP, JavaScript, and SQL.

Independent Game Developer

01/2010 to 03/2011

Eccentric Duck Games

Vancouver, BC

- Created a downloadable game called Gyroball, for the Xbox Live Indie Games store, over the course of 14 months.
- Implemented 3D rendering, collision detection, multiplayer networking, user interfaces, peripheral input, remote debugging (on console), and agent-based AI using Microsoft XNA (DirectX 9c) and C#.
- Supported and tested the game extensively, including with others in multiplayer, to ensure that it was fun and free of bugs.

David Ravensborg - Developer/Data Scientist

<https://www.linkedin.com/in/david-ravnsborg-b624b658/> | <https://stackoverflow.com/users/724403/david>

Education

Data Science Certificate

09/2019 to 12/2019

NAIT

Edmonton, AB

- Completed on weekends while working full time at ATB. It focused on data analysis, data mining techniques, data visualization, and especially machine learning. We typically learned theory on Friday and Saturday in the mornings, and applied it by Saturday afternoon and Sunday morning.
- The final course in the series was a project based capstone where we predicted loan delinquency. We started exploring a time-series LSTM model, but quickly switched to simpler Sklearn and XGBoost models due to their shorter training time, and reduced number of transformations needed to prepare the data. In the end, the XGBoost model performed best.
- Python Libraries covered: NumPy, Pandas, Matplotlib, Seaborn, Sklearn, Keras, XGBoost.

Certified Tester Foundation Level

12/2017

ISTQB/CSTB

Edmonton, AB

BSc in Kinesiology

09/2013 to 04/2017

University of Alberta

Edmonton, AB

- Received my faculty's Undergraduate Research Certificate; granted for taking courses with a focus on research methods, a research practicum (Research Assistant in Clinical Gait Analysis), and presentation of my research at a 2016 student researcher conference.
- Included coursework in computer science, mathematics, and statistics; additional coursework with complex systems and assessment of disorders in physiology and neuroscience; and the application of math/stats to domains like biomechanics and epidemiology.
- I chose to do this program after getting bad whiplash when I was rear-ended, and physiotherapy helped me recover. I realized during my practicum that I like analyzing data, programming, and running experiments more than I like being a clinician/trainer.

Computing Science

09/2010 to 04/2011

Capilano University

Vancouver, BC

- Studied computing science and mathematics, but the second year of the program was discontinued due to major administrative changes at the university, so I pursued an opportunity with a Vancouver-based startup (Amateur Sports Database).