

# Zachary (Zack) Novak

[zack@znovak.io](mailto:zack@znovak.io) | 440-840-8527 (cell)

All developer / social links available at [znovak.io](http://znovak.io).

---

## **TECHNICAL SKILLS:**

- **Development:** Java, Scala, Python, Atlassian Tools (Jira, Confluence, Crucible), Shell, Spark, Kafka, SVN, Git, SQL, Alteryx, Jenkins
- **Cloud:** AWS
- **DevOps:** Terraform, Kubernetes, Docker
- **Orchestration:** Airflow, Prefect
- **Database:** MySQL, Postgres, MSSQL, SSRS, Redshift, Neo4J, Hive
- **Front-End Development:** HTML, CSS, JQuery, JavaScript, Spring, Flask, Thymeleaf, React
- **Analytics:** Elasticsearch, Tableau, Excel, R, Python

## **CERTIFICATIONS:**

- Certified Associate Developer for Apache Spark 2.4 with Scala 2.11, issued by Databricks on January 2020

## **RELEVANT WORK EXPERIENCE:**

### ***Progressive Insurance | Mayfield, Ohio***

March 2020 – Present

Data Engineer within Progressive's Machine Learning Engineering team

- Key technologies: Python, Prefect, Git, Terraform, Docker, Shell (Bash / Zsh), AWS (EC2, VPC, S3, Redshift, DynamoDB, IAM), Flask
- Contribute to development of model serving applications directly plugged into Progressive's quoting flow for commercial and personal vehicles.
- Created an infrastructure as code solution using Terraform to deploy a Prefect orchestration execution module to business groups throughout Progressive.
- Built an open-source python app to migrate entire projects (repository, code, comments, issues, projects) from GitHub Enterprise to GitHub.com using a configuration file and the GitHub API, allowing groups to use GitHub Actions, GitHub apps, open source their work, and keep all projects on one platform.
- Developed and maintain internal Flask website that automates AI/ML team's work intake, reducing communication burdens on team, improving customer experience, and shortening new work intake process by ~900%.
- Shortened code elevation process from ~ 4 weeks to ~ 5 days through shell automation, unit / integration tests, and approved security infrastructure templates.
- Introduced new technologies to Progressive by leading security reviews, evaluating software against currently used technologies, and highlighting why the new software would be beneficial.

### ***Visa Inc. | Mentor, Ohio***

January 2019 – March 2020

Data Engineer within CardinalCommerce's Big Data team

- Key technologies: Spark, Scala, Python, Airflow, SVN, Shell (Bash), Kafka, Elasticsearch, Tableau, MySQL, AWS (EC2, S3, Redshift), Databricks, Neo4J, Spring, Thymeleaf, HTML, CSS, JQuery, PGP/GPG Encryption, SFTP
- Assisted in development of a data dictionary using neo4j that enabled business users to self-service requests as to where data fields were coming from, standard values to expect, and the field's lineage throughout our data system.
- Lead creation of automated parsing system for all internal system logs including > 2 year backfill process. Based on this data, I developed tools and dashboards that analyzed the numerous services within the logs to

provide actionable insights into user behavior, identified business opportunities, and drove several business decisions regarding log analysis and service design.

- Developed and maintained an anomaly detection system that works on a customizable schedule for all of our cloud data sources that checks for abnormalities in input data using historical data and automatically alerts appropriate internal stakeholders of abnormalities.
- Due to widespread adoption of Databricks / Spark within CardinalCommerce, I created an “Introduction to Spark” user guide, a “Spark Optimization Guide” that helps to improve query / notebook speed using both beginner and advanced techniques and lead several lunch and learn sessions geared towards working with Spark.
- Optimized Airflow batch processing jobs by investigating known and suspected bottlenecks to ultimately lead to a 300% speed increase in jobs.
- Assisted in creation of standardized Airflow deployment system that includes testing changes in local Airflow VM instance and automated deployment using SVN, boto3, Jenkins, and bash shell scripting.
- Updated existing data processing jobs to use latest versions of libraries, and latest schemas while also developing an internal job QA that runs to detect schema changes of the data being read (expected schema vs actual schema).
- Worked with program / product managers as a technical lead to design, test, implement, and maintain data solutions projects with 3<sup>rd</sup> party clients while complying with Visa and federal data compliance regulations.
- Automated report / file delivery to 3<sup>rd</sup> party clients using SFTP, pgp/gpg key encryption, boto3, and bash scripting.
- Lead and pushed for use of documentation and knowledge sharing within team for greater visibility when troubleshooting and preventing knowledge siloes.
- Upgraded legacy scripts and projects to use data collected from external APIs where applicable for more accurate / reliable data rather than static files created manually by internal employees.
- Investigated Cardinal’s overall data pipeline and created documentation for others to reference on where a dataset is sourced / derived from, when it can be expected to be populated, and made recommendations to my manager on where lag can be cut down to increase processing speed.
- Investigated database indexes within MSSQL used for our external reporting system and deleted unnecessary indexes improving overall query performance and freeing up system memory.
- Lead backfill projects on Elasticsearch indexes and created a standardized procedure with how to create a new index, how to update an existing indexes’ template, and how to perform a backfill on multiple ES indexes.
- Provided technical support to business users on how to search for transactions within Elasticsearch (ES), how to create dashboards / visualizations, and how to overall increase their experience with ES.

## **OTHER WORK EXPERIENCE:**

***Diebold Nixdorf* | North Canton, Ohio**

Sept 2017 – May 2018

Jira Specialist / Project Management Intern

***CardinalCommerce* | Mentor, Ohio**

May 2016 – Aug 2016, May 2017 – Aug 2017, May 2018 – Aug 2018

Business Intelligence Engineering Intern / Big Data Intern

## **EDUCATION:**

***The University of Akron***, College of Business Administration

Akron, Ohio

Bachelor of Business Administration in Information Systems Management

Graduated December 2018

GPA: 3.65 / 4.00

Dean’s List: 4 semesters

Outstanding 2018 – 2019 Management Information Systems Graduate