

**Michael J. Pedersen** : [\(908\) 283-0318](tel:9082830318) : [datacyclist@gmail.com](mailto:datacyclist@gmail.com) : [My Site](#) : [LinkedIn](#)

## Executive Summary

Manager with 5 years of experience guiding teams through challenging conditions, ranging from technical challenges through COVID-19 related challenges. I have kept my team together, and we have thrived on the challenges of taking care of the data for a \$200 million company. Most importantly, all of this has been done while keeping an eye on our budget, making sure to keep our costs down to keep the company profitable right through COVID-19.

## Relevant Technical Skills

- **Big Data:** [HDFS](#), [Hive](#), [YARN](#), [Alluxio](#), [Impala](#), [Trino](#), [Kafka](#), [Kubernetes](#)
- **Database Servers:** [MySQL](#), [PostgreSQL](#), [Microsoft SQL Server](#)
- **Database Skills:** [PostgreSQL](#) Database Administration, Relational Schema Design, Structured Query Language (SQL)
- **Programming and Scripting Languages:** [Bash](#), C/C++, [Java](#), Javascript, [Perl](#), [PHP](#), [Python](#)
- **Programming Skills:** [Docker](#), [Jenkins](#), [Jira](#), [IntelliJ IDEA](#), Object-Oriented Design, Object-Oriented Programming, Refactoring
- **Software Configuration Management Tools:** [Git](#), [GitHub](#), [Mercurial](#), [Subversion](#)

## Job History

### Pulsepoint - Data Engineer and Director of Infrastructure for Data

New York City, NY & Newark, NJ (Telecommute) - Mar 2015 - Nov 2023

#### Director of Infrastructure for Data, May 2018 - Nov 2023

- Architected data streaming that manages 40T of data/day.
- Lead maintainer for ETL pipelines, encompassing over 250 transformations.
- Established new data centers in Europe and in Virginia.
- Migrated data center, moving processing of data pipelines to new data center.
- Split data management team into data platform and data product development.
- Guided the team through splitting our ETL pipelines into multiple repositories.
- Organized the migration of ETL pipelines from Python 2 to Python 3.
- Instituted and formalized processes and procedures for the team.
- Planned capacity to ensure we could handle incoming data throughout the year.
- Replaced [Vertica](#) with [Trino](#).
- Reported on system wide data latency using [ElasticSearch](#), [Kibana](#), and [Grafana](#).
- Conducted interviews for my team and for teams that work closely with my team.
- Automated distribution of incident reports to all affected parties.
- Changed hardware profiles for [Hadoop](#) to remove storage and compute colocation.
- Onboarded new team members, helping them to fully integrate into the team.
- Held weekly 1 on 1 meetings with team members.
- Developed new stories (including estimates) for our [Jira](#) board.
- Prioritized tickets for our [Jira](#) board.
- Passed annual HIPAA training for data protection.
- Developed roadmap for the Data Platform team.

**Michael J. Pedersen** : [\(908\) 283-0318](tel:9082830318) : [datacyclist@gmail.com](mailto:datacyclist@gmail.com) : [My Site](#) : [LinkedIn](#)

### **Data Engineer, Mar 2015 - May 2018**

- Collaborated with other teams to help them use the systems to find the data they need.
- Worked with other teams to define and then implement needed features for our internal ETL pipeline framework.
- Participated in on-call rotation.

### **Weight Watchers - Systems Engineering Lead**

New York City, NY - Nov 2014 - Feb 2015

- Developed lightweight monitoring tool for use within my group.
- Configured [Vormetric](#) products to ensure [HIPAA](#) compliance for customer data.
- Worked to transfer from [Rackspace Cloud](#) to [Openstack](#) based private cloud.

### **OrcaTec, LLC - Developer**

Atlanta, GA (Telecommute) - Jun 2012 - Oct 2014

- Mentored other developers in the use of [TurboGears](#), [SQLAlchemy](#), [Python](#), and JavaScript.
- Organized weekly meetings for members of the frontend (OTGUI) team, providing a chance to discuss (in depth) the issues the team was facing.

### **Choopa.com - Developer**

Sayreville, NJ - Jan 2012 - May 2012

- Developed library to manage [OpenStack](#) nodes, and gather billing information.
- Built [Nagios](#) configuration file generator for in-house web interface for [Nagios](#).
- Configured [Bacula](#) backup system as replacement for custom backup scripts.
- Reconfigured [Nagios](#) monitoring, reducing full check from 8 hours to 2 minutes.
- Refactored in-house [Nagios](#) web interface. This reduced the workload from six files down to one when adding new checks.
- Several smaller bug fixes and features throughout the internal code base.

### **Education**

Bachelor of Science in Computer Science, 2000  
East Stroudsburg University, East Stroudsburg, Pennsylvania

### **Professional Certificates**

[Google Cloud Certified Professional Data Engineer](#) (Jan 2024)  
Online Course - Google