# Executive Summary

As a Software Engineer, I have spent 6 years of my career focused on delivering high quality software to my company’s customers, with their focus and needs being on sorting through large numbers of documents in a timely fashion. This has meant understanding ingestion, storage, and display of arbitrary data. It has included custom data visualizations. This was primarily done with [Python](https://www.python.org/) and [Ubuntu Linux](https://ubuntu.com/), but has also included work with [Perl](https://www.perl.org/) and [PHP](https://www.php.net/).

# Relevant Technical Skills

* **Big Data:** [Google Big Query](https://cloud.google.com/bigquery), [HDFS](https://hadoop.apache.org/), [Hive](https://hive.apache.org/), [YARN](https://hadoop.apache.org/), [Alluxio](https://www.alluxio.io/), [Impala](https://impala.apache.org/), [Trino](https://trino.io/), [Kafka](https://kafka.apache.org/), [Kubernetes](https://kubernetes.io/), [Dagster](https://dagster.io/)
* **Database Servers:** [MySQL](https://www.mysql.com/), [PostgreSQL](https://www.postgresql.org/), [Microsoft SQL Server](https://www.microsoft.com/en-us/sql-server)
* **Database Skills:** [PostgreSQL](https://www.postgresql.org/) Database Administration, Relational Schema Design, Structured Query Language (SQL)
* **Programming and Scripting Languages:** [Bash](http://www.gnu.org/savannah-checkouts/gnu/bash/manual/bash.html), C/C++, [Java](https://www.java.com/en/), Javascript, [Perl](https://www.perl.org/), [PHP](https://www.php.net/), [Python](https://www.python.org/)
* **Programming Skills:** [Docker](https://www.docker.com/), [Jenkins](https://www.jenkins.io/), [Jira](https://www.atlassian.com/software/jira), [Intellij IDEA](https://www.jetbrains.com/idea/), Object-Oriented Design, Object-Oriented Programming, Refactoring
* **Software Configuration Management Tools:** [Git](https://www.git-scm.com/), [GitHub](https://github.com/), [GitHub Actions](https://github.com/features/actions), [Mercurial](https://www.mercurial-scm.org/), [Subversion](https://subversion.apache.org/)

# Relevant Job History

## EvolutionIQ - Senior Software Engineer

New York City, NY - Feb 2024 - Current

* Orchestrated [Google Cloud](https://cloud.google.com/) infrastructure with [Terraform](https://www.terraform.io/), managing [GKE](https://cloud.google.com/kubernetes-engine), [BigQuery](https://cloud.google.com/bigquery), and developer virtual machines for scalability and efficiency.
* Redesigned EvolutionIQ’s CI/CD pipeline, enhancing efficiency in development workflows.
* Fixed critical bugs in locally developed [GitHub Actions](https://github.com/features/actions), ensuring smooth and reliable environment promotion.
* Built a template repository to guide EvolutionIQ teams in using [GitHub Actions](https://github.com/features/actions) for deployment, as well as setting up data pipelines with [Dagster](https://dagster.io/).
* Developed a Python-based CI/CD library to manage complex actions beyond [Bash](http://www.gnu.org/savannah-checkouts/gnu/bash/manual/bash.html)’s capabilities.
* Implemented [nektos/act](https://github.com/nektos/act) to enable developers to test build and deploy actions locally before committing to [GitHub](https://github.com/).
* Refined [data retention policies](https://cloud.google.com/bigquery/docs/managing-partitioned-tables" \l "partition-expiration) to guarantee compliance and ensure critical data availability for legal requirements.
* Implemented SLA monitoring with [Grafana](https://grafana.com/), [SendGrid](https://sendgrid.com/), and [PagerDuty](https://www.pagerduty.com/) to proactively detect and alert customers and engineers about overnight job failures, ensuring swift issue resolution.
* On-call support on a rotating basis.

## 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.
* Guided the team through splitting our ETL pipelines into multiple repositories.
* Organized the migration of ETL pipelines from Python 2 to Python 3.

### Data Engineer, Mar 2015 - May 2018

* Implemented data duplication between two [Hadoop](https://hadoop.apache.org/) clusters.
* Created ELT jobs to ingest third party data to make it available internally.
* Installed and configured multiple [Hadoop](https://hadoop.apache.org/) clusters.
* Developed new ETL jobs to aggregate data from Pulsepoint’s RTB exchange.
* Converted [Sqoop](https://sqoop.apache.org/) jobs to use [FreeBCP](https://www.freetds.org/) instead.
* 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.
* Added over a hundred automated tests to our ETL pipeline.
* Managed data backfill issues whenever they arose.

## Weight Watchers - Systems Engineering Lead

New York City, NY - Nov 2014 - Feb 2015

* Developed lightweight monitoring tool for use within my group.
* Configured [Vormetric](https://cpl.thalesgroup.com/encryption/vormetric-data-security-manager) products to ensure [HIPAA](https://en.wikipedia.org/wiki/Health_Insurance_Portability_and_Accountability_Act) compliance for customer data.
* Worked to transfer from [Rackspace Cloud](https://www.rackspace.com/cloud) to [Openstack](https://www.openstack.org/) based private cloud.

## OrcaTec, LLC - Developer

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

* Developed advanced search tool using [Python](https://www.python.org/), [TurboGears](https://www.turbogears.org/), and [jQuery](https://jquery.com/).
* Created new document production framework from scratch.
* Added holds and matters framework, allowing customers to state that documents belong to specific cases and should not be deleted while the cases are ongoing.
* Wrote [Python](https://www.python.org/) framework to manage long running background jobs.
* Reduced multi-hour [SQLAlchemy](https://www.sqlalchemy.org/) bulk database jobs to minutes.
* Found major security hole (remote code execution) and closed it.
* Spearheaded conversion from [YUI 2](http://yui.github.io/yui2/) to [jQuery](https://jquery.com/) and [jQueryUI](https://jqueryui.com/).
* Documented internal server API, wrote a [Python](https://www.python.org/) class to standardize it’s use.
* Added tag cloud (using [awesomecloud plugin for jQuery](https://github.com/russelporosky/jQuery.awesomeCloud.plugin)).
* Added support for allowing customers to login using [OpenID](https://www.openid.net/).
* Mentored other developers in the use of [TurboGears](https://www.turbogears.org/), [SQLAlchemy](https://www.sqlalchemy.org/), [Python](https://www.python.org/), 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](https://www.openstack.org/) nodes, and gather billing information.
* Built [Nagios](https://www.nagios.org/) configuration file generator for in-house web interface for [Nagios](https://www.nagios.org/).
* Configured [Bacula](https://www.bacula.org/) backup system as replacement for custom backup scripts.
* Reconfigured [Nagios](https://www.nagios.org/) monitoring, reducing full check from 8 hours to 2 minutes.
* Refactored in-house [Nagios](https://www.nagios.org/) 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](https://www.credly.com/badges/6c91d726-a924-4b8d-994b-80e49e002428/public_url) (Jan 2024)  
Online Course - Google