pdgroves@gmail.com https://petergroves.com Chicago, Illinois

# PETER GROVES

## ABOUT

Software engineer (18yoe) with a focus on developing applications with machine learning functionality.

Specialties: Machine Learning, DevOps, REST, Data Modeling, Python, Java, Linux, AWS.

## EXPERIENCE

### National Center for Supercomputing Applications — *Senior Software Engineer*

2015 - PRESENT
As a Senior Research Software Engineer, my focus is on the Backend, Devops, and Machine Learning aspects of cloud-based applications designed for scientists by Visual Analytics at the NCSA.

1. **Lead Backend Developer for Nest, a platform for data science and web applications.** This platform is based on common data types shared between frontend, backend, and data science jobs. Nest integrates mature but modern open-source components as docker containers, including CI jobs and one-click deploy to AWS EC2. (2015-present)
   Online Whitepaper
   *Tech: AWS (EC2, Route53, S3, IAM), Docker, Postgres, Python Flask, SQLAlchemy, Jenkins.*

2. **Product Manager and Lead Developer of the Pixsure project** built on Nest. Pixsure is a human-centric web tool for annotating complex medical images and simplifies data management of the resulting ML training data. (2020-present)
   *Tech: Nest (described above), Conda, Scikit-image.*

3. **Cloud engineer for a proof-of-concept API for Mayo Clinic** to store and access a particular subset of genomics data used by clinicians when determining cancer treatments. I evaluated technologies and built the MVP of the data management system and API on Google Cloud while the remaining team of doctors and DBAs created a sophisticated data model in Postgres. (2021)
   *Tech: GCP (Compute, Cloudsql-Postgres), Docker-Compose, Traefik (a reverse proxy), Postgresql, PostgREST, OpenAPI, RapidDoc, Jupyter Notebooks.*

4. **Lead Backend Developer on Phyloflow, a bioinformatics pipeline** to compute phylogenetic trees from tumor mutation data in collaboration with a CS prof and Mayo Clinic. This is an up-and-coming area of cancer research and my role was to pull together a collection of bleeding-edge commandline tools, mainly from University lab Github repos, and package them consistently into docker containers and WDL tasks, and then run the resulting pipeline on both a local HPC system and GCP.
   Phyloflow On Github, Phyloflow On Dockstore
   *Tech: AWS (ECR, Route53), GCP (CloudFiles, terra.bio), Docker, Workflow Description Language (WDL), MiniWDL, Conda, Dockstore*

5. **Lead Backend Developer for Omix** built on Nest. Omix is a web app that visualizes microbiome analytical results built in collaboration with Mayo Clinic's Center for Individualized Medicine. (2015-2017)
   Project Page
   *Tech: Nest (described above), Conda, Numpy/Scipy*

### GROUPON — *Software Engineer*

SEPTEMBER 2013 - MARCH 2015

As part of the Automated Merchandising team, I was responsible for bundling deals (Groupons) into widgets to be displayed anywhere on the Groupon website or mobile apps. Personal areas of focus:

- Machine learning system for adjusting the relative rankings of widgets in different contexts.
- Integration test suite of the core system, written in Python.
- Build automation using Python, Capistrano, Maven, Jenkins, and cron.
- General software engineering of a Java webapp on an 8 person agile software team.
- International rollout of core service to 20 countries in Europe and Latin America.

### INDEPENDENT CONTRACTOR — *Machine Learning Software Developer*

OCTOBER 2006 - SEPTEMBER 2013

As an independent contractor, I worked on the following projects:

1. Nexlp (2013). An e-Discovery startup in Chicago. Their core product uses a natural language processing toolkit from the University of Illinois and a graph database (Neo4j) to analyze sets of millions of emails at a time. The primary deliverable was a pattern detection module that combined frequent item set analysis and anomaly detection to generate patterns of the form "Bob emailed Sally late at night about Chicago 12 times during the week of Dec 4, 2006, but normally this occurred 0.03 times/week."
2. rVibe (2013). A boutique maker of training software for the pharmaceutical industry. Handled devOps for the company for nine months. Designed and implemented build tools and a performance benchmarking suite. Performed weekly deploys to production servers and adjusted agile release schedule and methodology as needed.
3. Fuzzy economics project (2011/2012). Privately funded by a (stealth) organization, this was a bleeding edge project to create a system for creating expansive yet detailed ontologies of hypotheses and their supporting evidence. Built the full stack prototype using Java EE with JSP and the JavaBayes toolkit for Bayesian Networks.

### DESIGNBYROBOTS — *Founder*

SEPTEMBER 2006 - SEPTEMBER 2013

Developed an algorithmic trading application written in OCaml using a genetic algorithm and statistical modeling techniques to find market price patterns on a time scale of less than one hour. Trading strategy optimization was the first application that uses DesignByRobot's data model for machine learning and automated design technology.

https://designbyrobots.com

## EDUCATION

### UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN — *MS Computer Science*

2001 - 2003

### UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN — *BA Agricultural Engineering*

1997 - 2001