# Barry Shapira principal software engineer

https://github.com/originalgremlin/portfolio
barry.shapira@gmail.com

## Experience

**Humane [https://humane.com]**
**Machine Learning Infrastructure Lead, September 2021 to Present**
- Responsible for building the entire machine learning infrastructure at Humane.
- Developed systems to manage thousands of artifacts such that the provenance of every model and dataset is known.
- Trained models on hundreds of V100 and A100 GPUs on AWS and Azure. Deployed these models in production. Built tooling for other engineers and data scientists to do the same.
- As a feature engineer was responsible for running speech models both on-device and in the cloud, building the image processing pipeline, and Lora fine-tuning LLMs for custom natural language tasks.

**Facebook [https://www.facebook.com]**
**E5, January 2020 to October 2020**
- On the FBLearner Flow team, an established library and framework to enable distributed function execution at scale, particularly for AI training workloads. Migrated all projects on the existing custom build system to use the company-wide standard systems. Developed an improved DAG API for visualizing workflow execution in real-time.
- Stopped working at Facebook to focus on raising my 5-year old daughter and 7-year old son until school restarts in September 2021.

**Whiterabbit.ai [https://www.whiterabbit.ai]**
**Principal Software Engineer, June 2018 to November 2019**
- Led a team of six developers tasked with deploying AI models to health-care clinics to provide real-time inference results to doctors and staff. Hired four full-time engineers and three interns. Time was split evenly between management and personal code contributions.
- Wrote two libraries in Python to perform inference. One used microservices communicating via ZMQ sockets and was deployed on-prem. The other used asyncio-enabled kafka clients and was deployed in the cloud.
- Designed and developed a hybrid-cloud docker swarm to treat cloud and on-prem nodes as a single compute resource. Work was performed using terraform, packer, kvm, and custom python services.

**Redbooth [http://www.redbooth.com]**
**Principal Software Engineer, March 2015 to February 2018**
- Developed a front-end chat application that used a single codebase to provide native apps for desktop and mobile using Electron, React, and React Native.
- Designed and built a fault-tolerant, scalable, distributed docker swarm on AWS.
- Created a real-time voice-to-text and translation app using a custom-built stream processing framework in Go and commercial AI offerings from AWS, GCS, and IBM Bluemix.

**StoryCloud [http://www.storycloud.co]**
**Vice President of Engineering, December 2013 to December 2014**
- Designed and built the foundation of a highly distributed backend-as-a-service platform using Hadoop, HBase, Kafka, elasticsearch, and Java Jersey.
- Assembled an engineering team of a dozen fantastic individuals, including a UX designer, a sysadmin, a data scientist, several backend experts, and several mobile developers.

**EGG-energy [http://www.egg-energy.com]**
**Director of Information Systems, August 2012 to August 2013**
- Created an internal business management website for an off-grid energy startup in Tanzania.
- Managed robust data synchronization at remote stations where power and Internet connectivity failures were common.
- Programmed Android phones and tablets using Apache Cordova.
- Coded all aspects of the site using NodeJS, CouchDB, elasticsearch, and Backbone.

**Go Go Gremlin**
**Lead Engineer, September 2011 to April 2012**
- Designed and developed a daily deal search and alert system using MVC design, RESTful APIs, and advanced JavaScript interactions.
- Architected an extensible platform for growth with a high level of abstraction.
- Coded all aspects of the site using Python/Django, JavaScript/YUI, and HTML5.

**Analog Analytics**
**Engineer, February 2011 to September 2011**
- Developed a model to estimate performance for a syndication network of advertisers.
- Unified all web analytics code under Google Analytics.
- Wrote a server to generate custom reports in xml, json, and delimited formats.

**Independent Programming and Study**
**June 2009 to February 2011**
- Wrote an object-oriented, Ajax-friendly JavaScript widget library based on Prototype and Scriptaculous to enable rapid creation of highly interactive web applications. [http://sourceforge.net/projects/scriptwidgets/]
- Developed agent-based models in Java using the Mason library to explore the effects of network topology on trading and scheduling of trains in the Paris Metro.
- Studied economics at UCSD, attending classes on AI, game theory, econometrics (statistics), microeconomics, and computational macroeconomics, which focused on methods for interpolation, optimization, matrix manipulation, and other important numerical tasks.

**United States Peace Corps**
**Education Volunteer in Cameroon, June 2007 to June 2009**
- Taught English and Computer Literacy to francophone high school students.
- Managed the school computer center of 40 desktop computers and 3 servers.
- Wrote a grade and discipline management program in Java.

**Snap**
**Developer, December 2004 to June 2006**
- Mined data sets of 100 billion records to improve the search results ranking algorithm.
- Wrote the Ajax library powering the Snap.com search interface.
- Dealt with cross-browser issues, memory management, and object-oriented JavaScript.

**Global Human Fulfillment**
**Developer, August 2001 to September 2002**
- Designed a media search engine of over 2,000,000 images and 100,000 movies.
- Developed a multi-threaded, distributed Python robot to crawl the web.
- Generated dynamic HTML pages using PHP and MySQL.

## Education

California Institute of Technology
Bachelor of Science in Organic Chemistry, June 1998

## Programming Skills

- Languages: PHP, Python, Perl, Java, bash, SQL, HTML, CSS, JavaScript, NodeJS, Go
- Machine Learning: scikit-learn, numpy, pandas, pytorch
- JavaScript Libraries: Prototype, YUI, Backbone/Marionette, jQuery, Underscore, React
- Devops: AWS, Terraform, Packer, Docker, Docker Swarm
- Databases: MySQL, CouchDB, HBase, TinkerPop
- Other: Hadoop, Kafka, Redis, elasticsearch, Solr, Git

## Social Links

- Github [https://github.com/originalgremlin]
- Bitbucket [https://bitbucket.org/originalgremlin]
- LinkedIn [https://www.linkedin.com/in/originalgremlin]

## Patents

- Modulators of Protein Tyrosine Phosphatases (PTPases)
  [United States Patent Numbers 6,410,556; 6,613,903; and 7,019,026]
- Method of inhibiting protein tyrosine phosphatase 1B and/or T-cell protein tyrosine phosphatase 4 and/or other PTPases with an Asp residue at position 48
  [United States Patent Number 7,115,624]
- Method and apparatus for network controlled access to physical spaces
  [United States Patent Number 9,508,207]
- Automatic generation and registration of alter-ego web service accounts
  [United States Patent Number 9,634,997]

## Foreign Languages

- Native English speaker
- Fluent in written and spoken French
- Intermediate in written and spoken Spanish