

TRISTAN E. CHONG

San Francisco, CA

tristanchong@gmail.com

[linkedin.com/in/t33chong](https://www.linkedin.com/in/t33chong)

github.com/t33chong

Software engineer with a broad skill set and an affinity for making sense of language data, driven by boundless curiosity and a desire to achieve positive real-world outcomes.

PROFILE

Resourceful problem solver with excellent communication skills and an insatiable appetite for knowledge. Jack of all trades with a record of driving projects from concept to production. Force multiplier who coheres teams and proactively clears obstacles. Formally trained in linguistics and computational linguistics; currently pursuing advanced degrees in computer science and data science. Industry experience spanning a decade in natural language processing, full-stack application development, infrastructure and data engineering.

EDUCATION

Dual M.S., Computer Science and Data Science, 2023-present

University of Pennsylvania

MCIT (Computer and Information Technology): Software development, discrete math, computer architecture, operating systems, data structures, algorithms, databases, networking, cybersecurity

MSE-DS (Data Science): Statistics, linear algebra, big data analytics, machine learning, deep learning, natural language processing, artificial intelligence

M.S., Computational Linguistics, 2011-2015

University of Washington

Classification algorithms, language modeling, part-of-speech tagging, parsing, tokenization, named entity recognition, regular expressions, word sense disambiguation, information retrieval, formal grammars, automatic summarization, dialogue systems, speech recognition & synthesis

B.A., Linguistics and Anthropology, 2004-2009

University of California, Los Angeles

Phonetics, phonology, morphology, syntax, semantics, pragmatics

WORK EXPERIENCE

STATT, Washington, DC (Remote / SF)

Oct 2020 - Oct 2022

Senior Software Engineer

- As engineer #3 on a small team creating an intelligence platform for legislative research, played a variety of roles spanning infrastructure, backend, frontend, and natural language processing
- Developed NLP pipeline to generate summaries and search suggestions for tens of millions of public policy documents
- Built Trends dashboard featuring interactive data visualization modules enabling users to gain insights pertaining to the direction of public policy over time
- Worked behind the scenes to improve operational visibility, developer experience, provisioning and deployment, secrets and state management, and project management practices

PAGERDUTY, San Francisco, CA

Aug 2015 - Aug 2019

Software Engineer, Full-Stack Applications / Data & Machine Learning

I spent the latter half of my employment with PagerDuty in the Product Development org, contributing both as Full-Stack Developer and Data/ML Engineer on the Event Intelligence team.

- Led the architectural planning and development of Similar Incidents, a feature that provides relevant historical context to incident responders, from personal hackday project to general availability
- Built, productionized, and operated the machine learning models, data stores/pipelines, backend services/APIs, and frontend web user interfaces that comprise PagerDuty's various Alert Grouping offerings, reducing noise and repeated interruptions for related issues during triage and remediation
- Provided expertise in chat command interface design and development while on loan to the Platform team, delivering a first-class integration with Atlassian HipChat (RIP) and laying the groundwork for future integrations
- Coached and fostered the growth of interns into full-fledged contributors within an agile development setting

Software Engineer, DevOps & SRE

In my first 2 years at PagerDuty, my time was split evenly between 2 roles in our Infrastructure org: first, in a capacity I would characterize as a Release Engineer on the Developer Tools team, and then as a Site Reliability Engineer on the eponymous team.

- Facilitated the company's transition from a monolithic architecture to microservices by building chat-based tooling around deployment, configuration management, and multi-cloud provisioning
- Drove the organization-wide adoption of modern, immutable infrastructure by educating and supporting teams through the process of containerizing their services
- Enabled the product and engineering organizations to iterate more quickly through the build-measure-learn feedback loop by introducing continuous integration and delivery
- Automated the creation of new production-ready skeleton services compliant with best practices around monitoring, logging, and secrets management, resulting in improved developer happiness and faster time to market

- Built a simplified abstraction layer to allow teams to run applications in a highly available and scalable manner across multiple environments, integrating open-source service discovery, load balancing, and container orchestration technologies into an internal platform
- Led the SRE team's response to various major incidents during regular on-call shifts

WIKIA, San Francisco, CA

Sep 2015 - Oct 2015

Software Engineer

- Accepted a short-term contract role to update the data pipeline and NLP-related services I had built with my former team to run on the company's new containerized infrastructure

BETTERCOMPANY, San Francisco, CA

Jul 2014 - Nov 2014

Software Engineer

- Implemented signup and authentication via Facebook
- Built a database of potentially interested customers by extracting information from user Facebook profiles and mobile phone contacts (if shared) on signup
- Enabled the marketing team to filter prospective users by various criteria with a custom application backed by the aforementioned database
- Automated the sending of email invitations to targeted prospects
- Developed tools to simplify common tasks with Amazon Web Services APIs

WIKIA, San Francisco, CA

Jun 2013 - Jul 2014

Computational Linguist / Software Engineer, NLP

- Built a pipeline utilizing NLTK and the Stanford CoreNLP software suite to parse tens of millions of pages of text
- Extended the capabilities of the parsing pipeline to scale up and down automatically by provisioning and decommissioning AWS EC2 instances in accordance with load
- Implemented a service-oriented architecture designed to extract and cache data on named entities, syntactic heads, coreference chains, dependency relations, and sentiment
- Wrote ETL and load balancing modules in a Python library used for data science research
- Researched various document summarization algorithms, and evaluated n-gram keyword extraction & sentiment analysis for potential business applications
- Developed a heuristic to infer the subject of a wiki using term frequency and weighted scoring
- Trained latent Dirichlet allocation (LDA) models with named entity data and used a distance metric to identify related pages as part of a recommendation system

FLUENTIAL, Sunnyvale, CA

Jan 2011 - Jun 2013

Computational Linguist

- Worked as part of a team of linguists and engineers to develop machine translation software and spoken dialogue systems

- Wrote context-free grammars in a Backus-Naur Form variant for parsing natural language
- Tested and tuned support vector machines (SVMs) and semantic class taggers to achieve higher phrase classification accuracy
- Created training corpora using Python and the Natural Language Toolkit (NLTK) to automate tasks including crowdsourced data collection, text normalization, production of canonical forms via stemming and lemmatization, morphosyntactic operations, and elicitation of use cases for regression testing
- Developed YAML interaction guides to manage dialogue states and conversation flow
- Localized existing applications to different languages, regions, and target markets
- Optimized synthesized speech audio using noise removal, silence trimming, normalization and compression techniques
- Translated text-to-speech pronunciation dictionaries between X-SAMPA and proprietary formats

TECHNOLOGIES

A necessarily incomplete selection of technologies with which I have experience:

Python	JavaScript	Linux	AWS
pandas	HTML	Bash	MySQL
NLTK	CSS	Docker	PostgreSQL
spaCy	Ruby	Nginx	Elasticsearch
Gensim	Rails	Consul	Kafka
scikit-learn	Node.js	Vault	Airflow
Stanford CoreNLP	React	Nomad	Spark
😊 Transformers	Redux	Terraform	Redis
Scala	Ember.js	Chef	Git