

Vinayak Mehta

+91 9868 525 444
vmehta94@gmail.com
www.vinayakmehta.com

Experience

- Feb 2018 **Grofers – Data Engineer** – Bangalore, India.
-present
- Working with the data team to develop tools and services that help the organization become more data-driven.
 - Created a CLI tool which enables anyone to create an Airflow DAG using a YAML config. Set up a Jenkins pipeline which translates a YAML pushed to GitHub into a DAG and deploys it to the production Airflow instance.
 - Created a Python library which can be used by data scientists and analysts inside Jupyter notebooks to perform tasks such as sending an email, creating database connections, and loading data into S3 and Redshift.
 - Started the monthly **Grofers Tech Talks** [meetup](#) in Bangalore office.
- Jun 2016 **SocialCops – Data Engineer** – Delhi, India.
-Jul 2018
- Designed, built and deployed a part of the internal data platform responsible for data ingestion (REST/SOAP, Google Sheets) and transformation (R/Python scripts), which enabled anyone within the company to self serve their ETL needs. Created a web interface which was used to set up data pipelines for the [DISHA dashboard](#), among other company-wide projects. Created a 1-click process to build existing data pipelines into a debian package for on-premises deployment.
 - Conducted technical interviews and helped with hiring. Mentored an intern to develop features in existing tools.

Projects

Camelot

A Python library to extract tabular data from PDFs. ([code](#)) ([docs](#))
~2200 stars on GitHub.
#2 open-source project of the month - [Mybridge](#) (October 2018)

Excalibur

A web interface to extract tabular data from PDFs. ([code](#)) ([docs](#))
~400 stars on GitHub.
#1 research tool of the week - [Labworm.com](#)
Featured in [PyCoder's Weekly](#) (November 2018)

Skills

Languages Python, Javascript, HTML, CSS, C, C++

Tools Git, Flask, Celery, Airflow, Docker, Ansible, PostgreSQL, \LaTeX

Talks

- May 2019 **PyCon US** – Cleveland, Ohio, US.
Extracting tabular data from PDFs with Camelot & Excalibur ([video](#)) ([slides](#))
- Jun 2018 **PyData Delhi meetup #25** – Delhi, India.
Introduction to Apache Airflow ([slides](#))

Articles

- Oct 2018 [Announcing Camelot, a Python Library to Extract Tabular Data from PDFs](#)
Top post on Hacker News upon release.
- Aug 2018 [Airflow, Meta Data Engineering, and a Data Platform for the World's Largest Democracy](#)
Top 10 post on Hacker News upon release.

Education

- May 2016 **Bharati Vidyapeeth's College of Engineering** – Delhi, India.
Bachelor of Technology in Computer Science and Engineering – Grade: 71.6%
- May 2012 **St. Cecilia's Public School** – Delhi, India.

Interests Running (ran 1094 km in 2018), playing guitar, amateur astronomy.