

# Ding CHENG

+86 17521619529 | [mkbo@me.com](mailto:mkbo@me.com) | [github.com/dipcarbon](https://github.com/dipcarbon)

## EDUCATION

---

### Fudan University

*Bachelor degree of Chemistry, Minor in Data Science*

Shanghai, Full-time

2014 – 2018

### Fudan University

*Master Candidate of Software Engineering, Data Science and Application*

Shanghai, Part-time

2021 – Present

## EXPERIENCE

---

### Data Engineer

*Roche, PD Data Science(under transformation)*

Jun 2019 – Present

Shanghai

- Use data surveillance tools and strategies to provide aggregate level reviews designed to identify patterns in data
- Work with stakeholders to assist with data-related technical issues and explore ways of automation and insights generation

### Clinical Programmer

*Roche, PD Biometrics*

Feb 2018 – Present

Shanghai

- Responsible for clinical data delivery of RG6042(Huntington's Disease), Crovalimab(Hematology), and support Tecentriq HCC, OC Filing
- Develop SAS programs for extracting and processing data with SDTM implementation and CDISC standards and lead the data collection of clinical database development with clinical data EDC programming
- Perform Spotfire programming and documentation for Risk-Based Quality Management, Data Monitoring, Medical Data Review

### Research Assistant

*Groningen Research Institute of Pharmacy*

May 2017 – Sep 2017

The Netherland

- Developed and support analysis of the genomics/proteomics data with clinical data of COPD with Python
- Conducted data visualization, including volcano plots, violin plots and interactive user interface and realized the data upload and analysis

## RELATED PROJECTS

---

### Data Curation and Integrated Analysis | *R Shiny, SQL, Python, REST API*

- Implemented data integration and analysis from Clinical Trial data of EDC, Real World Data and Omics Data
- Supported continuous delivery using Jenkin and Docker to build the plugin upon new a release and R packages(*CurateR*, *Scrubj*) development and a full-stack web app(*Curation Workshop*) using R Shiny

### Crovalimab Phase I/II Trials Study | *SAS, R, Python, Spotfire*

- Ensure timely and quality development and validation of datasets and outputs for study-documents according to specifications
- Designed and reviewed the eCRF and data structures. Comply with company and industry standards (CDISC SDTM/ADaM) and processes, created the review of programming specifications
- Built analytic tools that utilize the data pipeline to provide actionable insights into key business performance metrics and auto-generated clinical outputs

### PD Study Dashboard | *R, Linux Server, Docker, Google Workspace, Google Data Studio*

- Developed an E2E dataflow solution, collaborated with Clinical Scientist, Study Data Manager, CRA, and Global Study Managers to suggest features and get feedback

## PUBLICATION

---

**First Author:** Ding Cheng, Liang Qiao\*, Peter Horvatovich\*, *Towards Spectral Library-free MALDI-TOF MS Bacterial Identification*, J. Proteome Res., 17 (2018), 2124–2130

## TECHNICAL SKILLS

---

**Languages:** English(TOFEL 108), Mandarin

**Programming Languages:** R, Python, C#, SAS, SQL(MySQL), JavaScript

**BI Tools:** Tableau(Certificated), Google Data Studio, TIBCO Spotfire, Microsoft Power BI

**Libraries and Frameworks:** pandas, NumPy, Matplotlib, Shiny, Tidyverse, Roxygen