

Ding CHENG

+86 17521619529 | mkbo@me.com | github.com/dipcarbon

EDUCATION

Fudan University <i>Bachelor degree of Chemistry, Minor in Data Science</i>	Shanghai, Full-time 2014 – 2018
Fudan University <i>Master Candidate of Software Engineering</i>	Shanghai, Part-time 2021 – Present

EXPERIENCE

Data Engineer - Data Manager <i>Roche, PD Data Science(under transformation)</i>	Jun 2018 – Present <i>Shanghai</i>
<ul style="list-style-type: none">• Use data surveillance tools and strategies to provide aggregate level reviews designed to identify patterns in data to ensure high-quality results• Build analytic tools that utilize the data pipeline to provide actionable insights into operational efficiency or other key business performance metrics• Work with stakeholders to assist with data-related technical issues and explore ways of automation and insights generation	
Clinical Programmer <i>Roche, PD Biometrics</i>	<i>Shanghai</i>
<ul style="list-style-type: none">• Responsible for 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• 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 <i>Groningen Research Institute of Pharmacy</i>	May 2017 – Sep 2017 <i>The Netherland</i>
<ul style="list-style-type: none">• Developed and support for the development of the proteomics pipeline with Python• Conducted data visualization, including volcano plots, violin plots and interactive user interface and realized the data upload and analysis	

RELATED PROJECTS

PD Study Dashboard <i>R, Linux Server, Docker, Google Workspace, Google Data Studio</i>
<ul style="list-style-type: none">• Developed an E2E dataflow solution and using with R serving as a background and Data Studio as the front-end• Implemented data integration and analysis from RAVE, IVRS, and CTMS• Collaborated with Study Data Manager, CRA, and Global Study Managers to suggest features and get feedback
Monthly Portfolio Report Platform <i>R, SQL, Shiny, Javascript/D3, Spotfire</i>
<ul style="list-style-type: none">• Developed a SQL server and web collection interface for 20-people groups to ensure high-quality information• Optimized and standardized workflow and delivered an automated warning methodology• Visualized portfolio results(e.g. China Launch Chart) to PD leadership teams and other stakeholders with Spotfire
Data Curation and Integration <i>R, R Package, Shiny, SQL, Python, REST API</i>
<ul style="list-style-type: none">• Supported R packages(<i>CurateR</i>, <i>Scrubj</i>) development and a full-stack web app(<i>Curation Workshop</i>) using R Shiny• Implemented continuous delivery using Jenkin and Docker to build the plugin upon new a release

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