



Job title	<i>Data Scientist</i>
Reports to	<i>Director of Software and Data Engineering</i>

Job purpose

This position is in the Kyulux data engineering team. This world-class team is responsible for developing software tools; performing high through virtual screening with quantum chemistry and cheminformatics; using machine learning and data science on theoretical and experimental data with the goal of discovering new materials for OLED displays alongside our colleagues in the lab. Your job is to push forward our science with an approach that leverages scientific intuition, evidence-based reasoning, practicality and wrangling copious amounts data.

Duties and responsibilities

- Use advanced statistical and data science concepts, combined with experimental/theoretical chemistry data, to answer scientific inquiries about materials
- Create and evaluate machine learning models to answer these scientific questions
- Clean messy data from disparate sources
- Design, write, & debug command line tools using the python programming language
- Use version control tools to collaborate on code with colleagues
- Document your work such that colleagues can use your tools unattended or repeat your scientific inquiry, such as in a Jupyter notebook, slide deck, or documentation
- Present work periodically on biweekly evening video calls with team in Japan or in person with local team

Qualifications

- Ph.D. or similar experience in chemistry, engineering or physics; a CS background with exposure to general chemistry concepts would also be considered
- At least 2 years experience writing software in python or similar language
- Some project based work that demonstrates proficiency in data science concepts
- Evidence of effective verbal and written and communication skills, for example, experience giving seminars and writing peer-reviewed scientific papers

Preferred Skills

- Version control with git or similar
- Big picture understanding of neural networks (hands-on experience a plus)
- Database querying using Django or SQL; understanding of relational databases
- Scientific computing packages including numpy, pandas, scipy, matplotlib (or other plotting favorite), jupyter notebook
- Machine learning packages such as scikit-learn

Working conditions

40 hour work week mostly in downtown Boston office excepting 1-2 hours a week from home to participate in weekly phone calls with Japan (approx. 8pm). Travel to Japan is an option but not required.