

Opportunity Title: EPA Postdoctoral Fellowship on Machine Learning Applications
for Lead Service Line Identification

Opportunity Reference Code: EPA-ORD-CESER-BIL-2024-04

Organization U.S. Environmental Protection Agency (EPA)

Reference Code EPA-ORD-CESER-BIL-2024-04

How to Apply *Connect with ORISE...on the GO!* Download the new ORISE GO mobile app in the [Apple App Store](#) or [Google Play Store](#) to help you stay engaged, connected, and informed during your ORISE experience and beyond!

A complete application consists of:

- An application
- Transcript(s) – For this opportunity, an unofficial transcript or copy of the student academic records printed by the applicant or by academic advisors from internal institution systems may be submitted. All transcripts must be in English or include an official English translation. Click [here](#) for detailed information about acceptable transcripts.
- A current resume/CV, including academic history, employment history, relevant experiences, and publication list
- Two educational or professional recommendations. Click [here](#) for detailed information about recommendations.

All documents must be in English or include an official English translation.

Application Deadline 9/27/2024 3:00:00 PM Eastern Time Zone

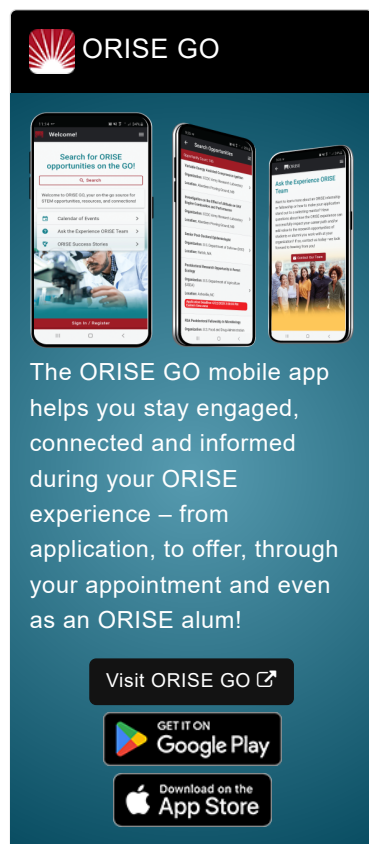
Description ***Applications may be reviewed on a rolling-basis and this posting could close before the deadline.** Click [here](#) for information about the selection process.

EPA Office/Lab and Location: A research opportunity is currently available at the Environmental Protection Agency (EPA), Office of Research and Development (ORD), Center for Environmental Solutions and Emergency Response (CESER) located in Cincinnati, Ohio. If selected for the opportunity, the participant will need to relocate to the appropriate EPA facility. The relocation costs are not reimbursable. The opportunity is not 100% remote, but limited remote participation may be considered at the mentor's discretion.

Research Project: Service lines connect community drinking water distribution pipes to individual buildings. Older buildings may have lead service lines (LSL) that pose significant health risks to consumers. The research participant will gain educational and professional benefits through involvement in a statistical modeling project focusing on machine learning (ML) approaches appropriate for characterizing the likelihood of lead service line presence in community neighborhoods and homes.

Under the guidance of a mentor, research activities may include:

- Reviewing current state of knowledge on machine learning applications for LSL identification
- Investigating approaches to integrate field generated data into ML analyses
- Exploring novel approaches to ML analyses e.g. ensemble modeling
- Integrating recent EPA data science resources e.g. analytical notebooks, cloud computing into LSL research



Opportunity Title: EPA Postdoctoral Fellowship on Machine Learning Applications
for Lead Service Line Identification

Opportunity Reference Code: EPA-ORD-CESER-BIL-2024-04

- Application of machine learning methods to community level decision-making for LSL identification
- Developing guidance on applying and interpreting ML LSL results to aid community/state level decision-making
- Presenting research at professional conferences
- Publishing research results in peer-reviewed journals
- Traveling to professional conferences, research facilities, and field sites

Learning Objectives: Under the guidance of a mentor, the research participant will learn to detect and show patterns in the user data and then make predictions based on these and intricate patterns for answering business questions and solving business problems. The research participant will learn to use machine learning in analyzing the data as well as identifying trends. The research participant will learn to generate manuscripts, presentations and other outputs related to drinking water research projects including lead service line identification projects through collaboration on informational resources in support of clear technical communications and project management. The research participant will learn basic machine learning methods such as supervised machine learning algorithms, unsupervised machine learning algorithms, semi-supervised machine learning algorithms, and reinforcement machine learning algorithms. The participant will also learn to integrate recent EPA data science resources e.g. analytical notebooks, cloud computing into LSL research. The research participant will gain experience in collaborating within a variety of government settings (federal, state and local).

Mentor(s): The mentor for this opportunity is Caleb Buahin (Buahin.Caleb@epa.gov). If you have questions about the nature of the research please contact the mentor.

Anticipated Appointment Start Date: August/September 2024. All start dates are flexible and vary depending on numerous factors. Click [here](#) for detailed information about start dates.

Appointment Length: The appointment will initially be for one year and may be renewed three to four additional years upon EPA recommendation and subject to availability of funding.

Level of Participation: The appointment is full-time.

Participant Stipend: The participant will receive a monthly stipend commensurate with educational level and experience. **The current stipend range for this opportunity is \$71,984 - \$86,279 per year plus a travel/training allowance.** Click [here](#) for detailed information about full-time stipends.

EPA Security Clearance: Completion of a successful background investigation by the Office of Personnel Management (OPM) is required for an applicant to be on-boarded at EPA.

ORISE Information: This program, administered by ORAU through its contract with the U.S. Department of Energy (DOE) to manage the Oak Ridge Institute for Science and Education (ORISE), was established through an interagency agreement between DOE and EPA. Participants do

Opportunity Title: EPA Postdoctoral Fellowship on Machine Learning Applications
for Lead Service Line Identification

Opportunity Reference Code: EPA-ORD-CESER-BIL-2024-04

not become employees of EPA, DOE or the program administrator, and there are no employment-related benefits. Proof of health insurance is required for participation in this program. Health insurance can be obtained through ORISE.

ORISE offers all ORISE EPA graduate students and Postdocs a free 5 year membership to the National Postdoctoral Association (NPA).

The successful applicant(s) will be required to comply with Environmental, Safety and Health (ES&H) requirements of the hosting facility, including but not limited to, COVID-19 requirements (e.g. facial covering, physical distancing, testing, vaccination).

Questions: Please see the [FAQ section](#) of our website. After reading, if you have additional questions about the application process please email ORISE.EPA.ORD@orau.org and include the reference code for this opportunity.

Qualifications The qualified candidate should have received a doctoral degree in one of the relevant fields (e.g. Computer Science, Statistics, Environmental Modeling), or be currently pursuing the degree with completion before the appointment start date.
Degree must have been received within five years of the appointment start date.

Preferred Skills:

- Desired background and/or expertise includes data mining, statistics, machine learning, computer programming, artificial intelligence, and mathematics

Eligibility Requirements

- **Citizenship:** LPR or U.S. Citizen
- **Degree:** Doctoral Degree received within the last 60 months or currently pursuing.
- **Discipline(s):**
 - **Chemistry and Materials Sciences** ([3](#) 👁)
 - **Computer, Information, and Data Sciences** ([17](#) 👁)
 - **Engineering** ([4](#) 👁)
 - **Mathematics and Statistics** ([11](#) 👁)

Affirmation I have lived in the United States for at least 36 out of the past 60 months.
(36 months do not have to be consecutive.)