STARS Project Determination

**Title:** Evaluating the use of computer vision approaches for environmental health issues

**Description:** This is a STARS project to evaluate the use of computer vision in regard to environmental health issues. As a field, computer vision uses traditional and machine learning or artificial intelligence approaches to process and extract information from digital imagery, including object detection and classification. Computer vision approaches are being used across a wide variety of domains, including classification of objects that may be relevant to environmental health issues. The long-term goal of this technology is to automate the identification of environmental health conditions to relieve the burden on state, tribal, local, and territorial health agencies to undertake on-site evaluations of environments. This project builds on prior pilot efforts at CDC to automate the identification of water cooling towers in metropolitan areas with the Tower Scout project. This pilot effort used machine learning approaches to train a model that evaluated satellite imagery to identify water cooling towers. This program achieved sufficient accuracy in identifying these structures and is now being developed for STLT use by CDC. This STARS project will leverage the lessons learned from this pilot effort to evaluate other environmental health conditions that include, but are not limited to, onsite wastewater systems, pools, harmful algal blooms, potential mosquito breeding sites, and other conditions identified by STLT environmental public health agencies. No new data collection will occur; all data will be from existing data sources. No data will be collected from human subjects.

**Goals:** This activity will help the Division of Environmental Health Science and Practice (DEHSP) evaluate the utility of computer vision to identify environmental health conditions. It directly aligns with division-wide efforts to improve public health practice using technology. This activity directly relates to the following DEHSP Focus Areas and objectives: Focus Area 2: Build the environmental public health evidence base • Objective 4: Evaluate public health actions and practices Focus Area 3: Develop and implement environmental public health best practices • Objective 5: Develop evidence-based environmental public health strategies and interventions • Objective 6: Promote the adoption and implementation of environmental public health best practice.

**Objective:** This comprehensive and coordinated approach will bring together programmatic and research-related efforts to improve public health practice via technology. This work will include, but not be limited to, how technology is being used to assess the health of populations, assure and inform populations, and develop policies that improve population health; all of which incorporate a health equity approach.

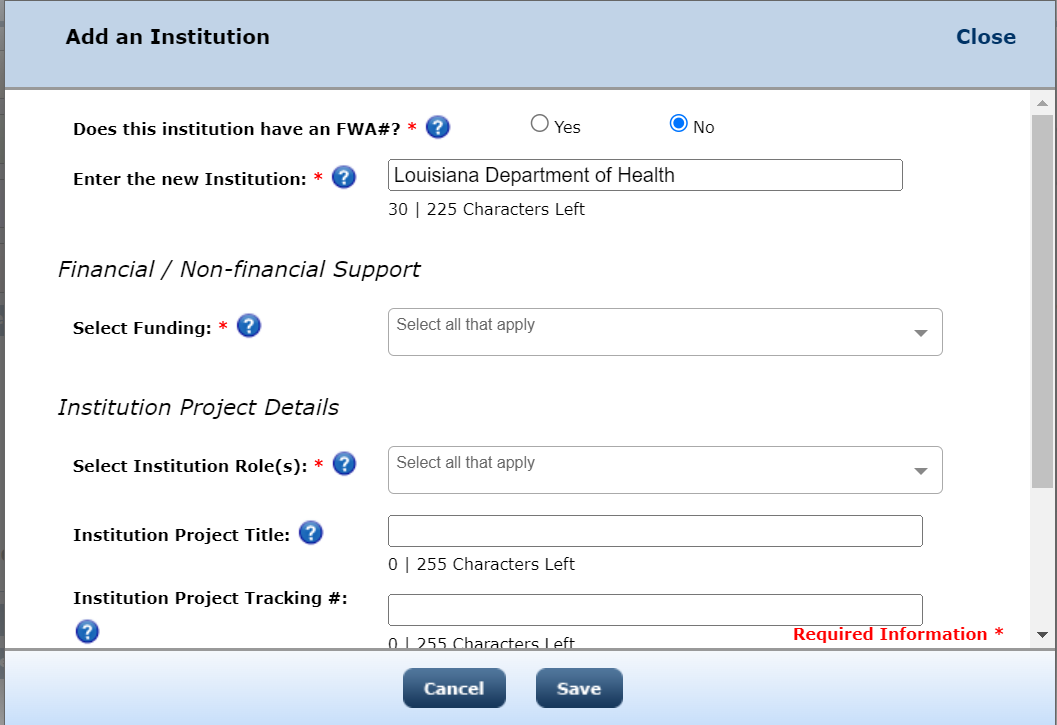
**Methods:** This work will focus on evaluating the use of computer vision approaches to enhance public health practice in STLT environmental public health agencies. Aerial and satellite imagery will be used to identify or detect specific extractable features that could indicate concerning public health situations, such as man-made structures, species habitat extent, evidence of sewage or chemical spills, algal blooms, or other features identified by STLT partners. This activity will develop both presentations and journal articles to communicate the findings and advance work in this field.

**Data collection:** Geospatial information will be used from the repository of systems available to CDC in ATSDR’s Geospatial Research, Analysis, and Services Program (GRASP). Investigators will discuss potential topics of interest with STLT jurisdictions that would be suitable for identification using computer vision. Investigators will then work with these STLT partners to identify the most appropriate data sources to accurately identify the environmental health issue being studied. This approach is solely intended to identify structures and environmental features that have been labeled by researchers according to the location identifiers available from either publicly-available or state/local government-owned datasets. These features may be geocodes or addresses that are converted to geocodes. There is no intent to identify individuals using this approach.

**Expected use:** This extension of pilot efforts will help advance understanding of using these approaches to more efficiently identify environmental health conditions. STLT partners will benefit from this work through the dissemination of these methods for their jurisdiction to use as well as the potential that CDC host these services for jurisdictions to use.

Institutional partners:

* Louisiana DoH



**Public access justification:** Proposing both public and non-public

* This data is owned by the partnering jurisdictions and collection of the data is not funded by CDC. The participating jurisdictions have agreed to allow CDC to analyze their data. Data availability is left to each jusrisdiction’s discretion and is governed by their policies regarding data availability. CDC will make any data tables from publications available to the public in an open, machine-readable, non-proprietary format. This project is exploratory in nature and will use commercial datasets that CDC has purchased access to but does not own. Public access would likely not be of use until the project is proven to be of utility to the agency.

**How access will be provided:** CDC will make any data tables from publications available to the public in an open, machine-readable, non-proprietary format.

**Preservation:** Dataset will be archived in a shared location only available to CDC staff. The dataset will be maintained for the period of time specified by CDC guidelines.