USING DATA SCIENCE TO PROTECT TAP WATER QUALITY

REVISED Proposal to the Lucy Family Institute for Data and Society

May 30, 2022

This proposal is a revision of the one submitted to the Lucy Institute's 2022 call for proposals. In response to reviewer comments, we have reduced the project duration to one year (from two years) and the budget to $78K (from $175K). This was mainly achieved by removing the comprehensive field survey of homes in the South Bend community. We have added a social scientist, Dr. Danielle Wood, to advise on future external proposals addressing field surveys.

The principal investigators are:

- **Rob Nerenberg**, Professor, CEEES
- **Mike Lemmon**, Professor, EE
- **Matt Sisk**, Research Professor, Lucy Family Institute

The team also includes **Dr. Danielle Wood**, a social scientist from the Center for Civic Innovation (CCI), as an advisor.

Dr. Sisk will lead the GIS modeling effort for Study 1. Dr. Nerenberg will lead the field data collection effort for Study 1, and the modeling and lab experiments for Study 2. Dr. Lemmon will lead the machine learning and data assessment for Study 2. Dr. Wood help devise suitable sampling strategies to be used in subsequent proposals for community sampling.

We propose two environmental engineering graduate students and one electrical engineering graduate student for this work. For Study 1, the lead will be grad student **Isaiah Murrell-Thomas**. Isaiah completed his BS in Environmental Engineering at Notre Dame in May of 2022, and he worked as an intern for the City of South Bend the previous summer. During the last academic year, he has been working as an undergraduate research assistant in the Nerenberg lab on the microbial quality of drinking water. He is an underrepresented minority, first generation college student. For Study 2, the leads will be **Emily Clements**, a PhD student in the Nerenberg lab who is studying biofilms and opportunistic pathogens in drinking water systems, and **Yuying Duan**, an ongoing female PhD student in electrical engineering who is studying fairness issues in Machine Learning under Dr. Lemmon.

ABSTRACT

The proposed research uses data science to identify homes at risk for unhealthy tap water and develop data science-based strategies to mitigate these risks. While tap water is safe in the majority of homes in the US, there are notable exceptions, as seen recently in Flint, MI and Benton Harbor, MI. Unfortunately, water quality problems often impact low-income families, which are the least prepared to manage them.

In most cases, water supplied to homes from the public distribution networks meets stringent EPA drinking water standards. The problems occur *within* homes. Pipes may leach toxic metals, such as lead and copper, or promote growth of pathogenic bacteria, such as *Legionella pneumophila*. In both cases, the decay in water quality correlates to the "water age" or stagnation within the plumbing system. Longer water ages provide more time for metals to leach and for bacteria to grow.

Given the complex nature of plumbing systems that 1) may have been built over 100 years ago and lack documentation, 2) may have been modified over the years without documentation, 3) have pipes that are often hidden beneath the ground or behind walls, and 4) have highly variable water demands whose characteristics change depending on the type of fixtures and appliances, and the number and type of occupants, home plumbing systems often behave as black boxes. *Data science can be used to identify which homes are at risk, and to develop strategies to mitigate the risks*. For example, actively controlling water age in premise plumbing systems could help improve water quality and reduce health risks.

The specific objectives of this project are to (1) develop a GIS-based model to predict homes with high water ages, and (2) develop machine-learning-based strategies to reduce water age.

The Lucy funding would be used to develop *preliminary data* that then could support proposals to external funding agencies. We propose a novel data-based approach to identify homes with potential water quality concerns:, to identify homes likely to have high water ages. We also propose field sampling of two or three homes, belonging to the PIs or colleagues, to develop our methodology and obtain preliminary data for an external proposal.

The proposed research uses data science to address a key societal problem, the safety of public water supplies. This is an especially important problem for the most vulnerable urban populations, low-income and older community members. It addresses one of the themes identified at the Lucy Institute's inaugural symposium on October 27, 2021: "Data Science for Tackling Health Inequities & Disparities."

PROJECT DESCRIPTION

GOALS AND OBJECTIVES

The **goals** of the proposed research are to use data science to

1. Identify homes at risk for unhealthy tap water
2. Develop strategies to mitigate these risks

The **objectives** of this project are to *develop methods and preliminary data* to

1. Develop a GIS-based model that predicts tap water risk based on easily available data on homes and water use, and develop a plan and methodology to validate this model by collecting limited field data from homes in South Bend
2. Develop strategies for mitigating water quality issues, and develop methodologies to test these strategies

The basic motivation is to use big data methodologies at a community wide level to both monitor and improve community residential water quality. The project integrates environmental engineering, data science, modeling, and field studies with PIs from Environmental Engineering, Electrical Engineering, and the Lucy Family Institute. It includes the Center for Civic Innovation (CCI) and the City of South Bend as partners. It involves Notre Dame undergraduate and graduate students, and the South Bend community.

Water utilities capture raw water, treat it to EPA standards, and distribute it to users via a piped network (Figure 1a). Utilities must comply with EPA standards up to the user's connection. But once water enters a private building's plumbing system, water quality is typically the owners' responsibility. The conditions differ significantly from the public distribution system (Figure 1a), and there often is much greater degradation of water quality.(Rhoads et al., 2016)

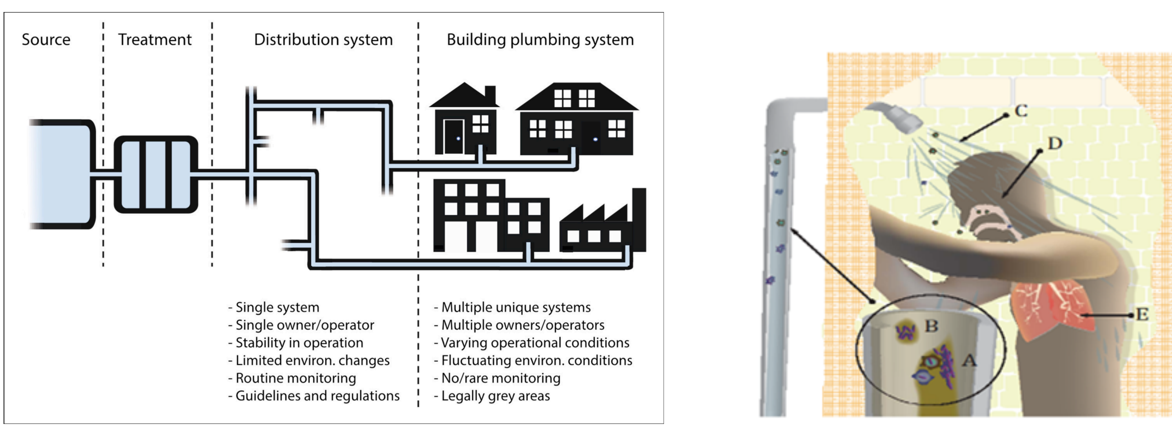


Figure 1. Left: Schematic of potable water system, indicating water source, treatment, distribution, and premise plumbing systems. Right: inhalation of opportunistic pathogens in shower water(Ley et al., 2020; Neu and Hammes, 2020; Proctor et al., 2016)

Several recent, high-profile cases show how water quality can be compromised within homes. In Flint, MI, the water supply met EPA standards at the treatment plant, but was severely compromised at the tap. Both lead and Legionella were linked to building plumbing systems. More recently, Benton Harbor, MI detected lead concentrations at the tap greatly exceeding EPA limits. As in the cases of Flint and Benton Harbor, these problems more commonly occur in poorer areas of cities with more vulnerable populations (Allaire et al., 2018; Marshall et al., 2020; Stillo and Gibson, 2017; Weisner et al., 2020).

Microbial pathogens are also a concern in plumbing systems(Leslie et al., 2021). While bacteria in well-managed drinking water systems are typically innocuous, some may be opportunistic pathogens, i.e., causing infections in immunocompromised individuals. For example, a key opportunistic pathogen is *Legionella pneumophila* (Lau and Ashbolt, 2009). The incidence of Legionella infections has significantly increased over the last 20 years, even though only 11% of Legionnaires' Disease cases are correctly diagnosed, suggesting the infection rate is even higher (Cassell et al., 2019). Legionella is associated with biofilms growing on pipe walls. Legionella can survive adverse conditions, such as chlorine disinfectants, by growth in biofilms and by association with protozoa. When they are inhaled or ingested by humans, Legionella bacteria can cause infection (Figure 1b) (Cullom et al., 2020).

Water quality degrades in piped networks, often in correlation with the "water age" or water residence time.(Ley et al., 2020) This is because **water age correlates with chlorine dissipation, disinfection by-product formation, leaching of toxic metals from pipes, leaching of organic chemicals from plastic pipes or fittings, and growth of microbial biofilms with opportunistic pathogens on pipe walls**. Public water utilities design and operate their networks to limit water age. But there typically is no control in homes.

Homeowners often lack knowledge about their plumbing systems (pipe materials, pipe sizes, pipe layout), especially in older homes and poorer neighborhoods. Thus, **the plumbing network is essentially a black box**. Also, unlike public water networks with thousands of users, water demands in homes are highly variable and cannot be controlled by the utility. A house may have no demands overnight, over a weekend, or several weeks when the occupants are on vacation. A guest bathroom may go unused for several weeks. In many cases, older apartment buildings in poorer neighborhoods are only partly occupied. Or houses built for large families may be occupied by one or two elderly residents. These can lead to very high ages and poor quality, posing a health risk for populations that tend to be more vulnerable.

High water age is known by utilities to degrade water quality, and public networks are actively managed to avoid high age However, the effects water age tend to be much more significant in premise plumbing systems. This is due to the (1) higher water temperatures compared to the public distribution system, where pipes are buried in the ground; (2) greater "surface effects" (high surface to volume ratios) for the much smaller pipes used in premise plumbing systems; and (3) low or non-existent chlorine residuals due to their much faster degradation with warm temperature and high surface effect .(Leslie et al., 2021). Thus, the greater water age significantly increases the risk for opportunistic pathogens. Controlling water age in premise plumbing systems could help improve water quality and reduce health risks.

Note that the above effects are likely to become even more significant in the future. Population growth and climate change are likely to increase water scarcity, leading to more conservation measures and lower water usage. Climate change will lead to higher water temperatures in many cases. And both will likely increase the use of marginal water sources, which are more likely to contain organic matter and promote the growth of biofilms.

PROPOSED APPROACH AND METHODS

Based on the above, tools are needed to identify houses with risks for water quality problems. Houses with greater risk could be prioritized for improvements, such as federal programs to replace lead pipes. Home age alone is insufficient as a predictor of risk. **Water age is also an important factor**. In addition, strategies are also needed to mitigate risks.

We propose two studies. One uses data science to help identify homes with risks of unhealthy water quality. The other uses data science to develop strategies to mitigate the risks. For future proposals based on these studies, we will propose using **the City of South Bend as a test bed**. This has been discussed with Eric Horvath, City Engineer. It also has been discussed with Jay Brockman of the Center for Civic Innovation (CCI). Both enthusiastically endorsed the proposed research and are interested in participating in the future field research stage.

Study 1 - Identifying Potential Risk

We propose using GIS records of home age and size, and water use records from water utilities, to identify homes likely to have high water ages. We also propose a very limited field sampling of homes to confirm the correlation.

Task 1. **Data aggregation and analysis.** Water usage data, at the individual property level, are available from the City of South Bend. These data, however, require significant cleaning, aggregation and geocoding to be useful. Once this process is automated, we will combine them with publicly available property records to give a structure size and construction date. This work will primarily be carried out by Dr. Sisk, a GIS expert affiliated with the Lucy Institute.

Task 2. Risk modeling. When the data have been aggregated at the household level, we plan to use a combination of geospatial and machine learning techniques to model the risk of individual households. This work will primarily be carried out by Dr. Sisk.

Task 3. Model validation using volunteer households. In this task, we propose using up to three homes in South Bend, owned by the PIs or colleagues, to test the hypothesis that greater water age can lead to impaired water quality. The methodology and results from this sample study can be used to then propose a more comprehensive study across the city. This sampling and analysis will be carried out by Isaiah Murrell-Thomas, a former undergraduate student from CEEES and an incoming MS student. He will be supported by doctoral student Emily Clements. We expect that households with low water usage, large homes, and older construction would have a higher water age.

Task 4. Planning future proposal. The PIs and Dr. Danielle Woods will help develop a sampling strategy for the proposed research program. Using a sampling frame defined by tasks discussed above, we will cluster sample candidate households meeting the selection criteria by neighborhood. This study targets areas with lower socioeconomic status, which can have residents with multiple participation barriers. Recruitment strategies will include information sessions through neighborhood association meetings, flier information through community partners (e.g. WIC, REAL Services), and door-to-door. Furthermore, participant households will be provided with gift card incentives.

Study 2 - Mitigating Risk

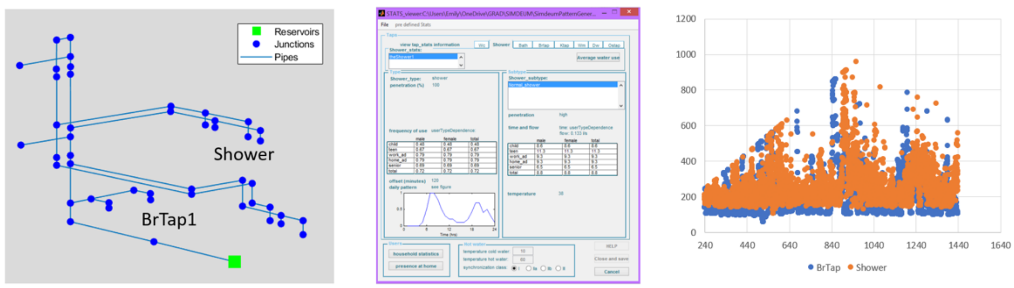
Flushing residential plumbing systems after periods of stagnancy will improve water quality and mitigate the associated health risks to residents. Determining an optimal flushing schedule, however, is complicated by the fact that water age is a function of plumbing configuration and water usage patterns. Providing a single flushing schedule to be followed by all homes in a community will be suboptimal and lead to excessive flushing that is undesirable in areas with water scarcity. Another approach would install sensors in each home to measure a home's water age and then initiate flushing when the water age exceeds a given threshold. This approach would be optimal, but it is economically impractical for poorer households and so the health risks would fall disproportionally on less affluent neighborhoods in the community. A more practical and effective plan for risk mitigation can be achieved using federated learning concepts (Li et al., 2020). In particular, this study will develop a risk mitigation plan that first trains models predicting water age for a statistical sampling of community residences and then uses federated learning to identify aggregated models from which neighborhood-specific purging schedules can be obtained. By using federated learning, we can statistically sample a small number of households while developing models that are tailored to each neighborhood. The proposed work would develop the proposed federated learning framework and then use simulation studies to assess its ability to reduce residential water age in a cost effective and neighborhood agnostic manner.

This proposed work can be broken down into the three tasks described below. The first task will develop deep learning methods for **training local models** that predict a given residence's water age based on water usage patterns and the home profile (age, number of bathrooms, number of occupants, etc.). The second task will develop a **federated learning framework** that takes local models from task 1 and aggregates them to obtain a set of neighborhood-specific models that predict neighborhood water age as a function of neighborhood and home profile. These neighborhood-specific models would then be used to generate purging schedules that can be voluntarily followed by the neighborhood residents. The third task will be a **simulation study** assessing the cost effectiveness of the proposed method for developing neighborhood purging schedules. This study will be used to develop a plan that can be followed to minimize health risks due to stagnant residential water in South Bend.

Task 1 - Training Local Models for Residential Water Age: Federated Learning is a machine learning technique that trains models using data from multiple edge devices. In our case, data from individual homes (edge device) is used to train a local model that predicts that home's water age as a function of the home's profile and its water usage pattern. This local model would then be forwarded to a cloud server where a global model of the community's water usage can be trained using the data from multiple households. This task will develop the deep learning methods needed to learn these ``local models'' for residential water age.

These local models will be obtained using standard backpropagation training of deep neural networks. We will use existing Machine Learning software such as TensorFlow for training these deep networks. The data to be used in training these models will come from computer simulation models of a home's water age. The computer simulation models will be developed using plumbing modeling software (EPANET) and the stochastic water demand generator (SIMDEUM). Preliminary model data is shown in Figure 2. Water quality modeling will be added as part of this task by doctoral student Emily Clements with support from PI Nerenberg.

This task will also build a **lab-scale test bed** to validate model predictions. The test bed will consist of two single pipes, one for hot water and one for cold, where the water demands can be controlled by valves controlled by a computer. The computer will implement a stochastic demand pattern, with and without a simulated purging device. Sensors for flow and temperature will continuously monitor the water in the pipe, as well as the water coming out. This work will be carried out by doctoral student Emily Clements, with support from master's student Isaiah Murrell-Thomas.



EPANET plumbing network model

SIMDEUM stochastic water demand generator

Simulated water ages for tap and shower, as a function of time

Figure 2. Preliminary data from EPANET plumbing network model

The simulation data from selected home profiles as well as the lab-scale test bed will then be used to train and evaluate deep neural networks (TensorFlow) predicting water age as a function of water usage and home profile. This work will be done by doctoral student Yuying Duan with support from PI Lemmon.   The main deliverables from this task will be 1) the neural network architecture to be used for local modeling of water quality, 2) evaluation results characterizing the accuracy of these models, 3) a list of in-situ sensors that would gather the data used to train these local models, and 4) an initial design of the software and hardware to be later used in training local water age models for instrumented residences.

Task 2 - Federated Learning of Neighborhood Residential Water Age: This task studies the use of federated learning techniques in training neighborhood-specific models for water age. Figure 3 illustrates the proposed system architecture in which individual homes are treated as edge devices.   Each home would use the results from Task 1 to train a local model for that home. The local model would be sent to a cloud server over the world-wide web using the home's WiFi hub as an access point. The cloud server would first cluster these models into distinct groups that we hypothesize will represent the distribution of home types and neighborhoods. The models in each cluster would then be used to train a ``neighborhood'' model for water usage and from that model an optimal purging schedule would be generated that minimizes water age subject to a constraint on the

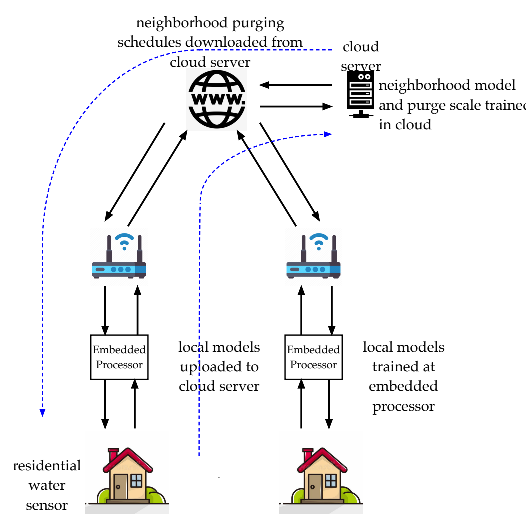


Figure 3. Federated Learning Framework

total amount and frequency of purging. This work will be conducted by doctoral student Yuying Duan with support from PI Lemmon using the local models generated in task 1.

There are two major challenges facing the development of the proposed federated learning algorithms. The first challenge concerns the fact that a statistical sampling of neighborhoods may not be independent and identically distributed (Konevcny et al., 2016). We propose addressing this challenge using generative adversarial networks (GANs) (Goodfellow et al. 2014), that learn how to generate random training data with the same distribution as our small sample set. This approach also addresses the second issue regarding the privacy of individual home models used for training. Since GANs independently sample the neighborhood distribution, the actual training samples are not representative of any particular home, thereby preserving the privacy of the individual. The use of GANs, therefore, will play a major role in our approach to federated learning. This work will be done by doctoral student Yuying Duan with support from PI Lemmon. The main deliverables from this task will be 1) federated learning algorithms using GANs to train neighborhood water age models, 2) assessment of the GAN's ability to ensure the privacy of homeowner data and neighborhood fairness, and 3) initial design of software objects for the federated learning framework.

Task 3 - Simulation Study of Water Purging Schedules: This task uses the federated neighborhood models from task 2 to synthesize purging (flushing) schedules for neighborhood households. The proposed schedules would minimize water age subject to inequality constraints on volume and frequency of flushed water discharges. This work will be done by doctoral student Yuying Duan with support from PI Lemmon. The main deliverables from this task will be 1) optimization algorithm used to generate flushing schedules, 2) simulated assessment of the flushing schedule's optimality, and 3) a study to assess the method's risk mitigation using neighborhood data from South Bend.

Targeted Societal Impact

This research will have the greatest impact on the most vulnerable populations, including poorer urban communities and older community members. There is extensive quantitative and qualitative data demonstrating that environmental toxins disproportionately affect low-income and marginalized groups.(Egendorf et al., 2021; Marshall et al., 2020) These populations are most likely to live in older homes with greater stagnation times, and also have less resources to purchase bottled water or have access to advanced water treatment systems.

By working directly with the municipality (the City of South Bend) and institutions with a proven track-record of outreach and collaboration with local communities (CCI and ND-LIT) we are confident that we can build upon these relationships and reach the populations most at risk from environmental contaminants.

Expected Long-Term Benefits

The proposed research will develop a new strategy to identify households and businesses that are at greater risk. We will measure success in two main ways. First, the successful creation and testing of models, technologies and sensors for minimizing risk of water-borne toxins will be a key indicator of success. Second, and equally important, will be using these tools and technologies to inform community members of the potential risk and effective mitigations.

Using our affiliation with CCI, we will carefully craft messaging and outreach that allow us to inform community members, businesses and municipal entities about these topics. As both aspects of this are novel investigations, we expect that a series of publications, additional grant proposals and conference presentations will follow. Additionally, sensor and purging tools may have translational potential.

Alignment with Mission of the Institute

The proposed research uses data science to address a key societal problem, the safety of public water supplies. This is an especially important problem for the most vulnerable urban populations, low-income and older community members. It addresses one of the themes identified at the Lucy Institute's inaugural symposium on October 27, 2021: "Data Science for Tackling Health Inequities & Disparities." With a proposed methodology that combines data science modeling techniques with a societal problem, this project is a natural fit for the Lucy Family Institute.

Rationale

Plumbing systems can significantly degrade water quality, especially when the water age is high. But most plumbing systems behave as a **black box**, and data science can be used to shine light on their behavior. It is a highly interdisciplinary project.

Sustainability

Following successful completion of this pilot work, we anticipate applying for additional, longer-term, funding from various federal and private funding agencies. Potential targets include:

(1) Water Research Foundation (WRF), which is comprised of water utilities, consultants, and water supply vendors. Each year the WRF has a call for unsolicited proposals.

(2) National Science Foundation (NSF), in the Engineering Division and the Directorate of Chemical, Bioengineering, Environment, and Transport (CBET) in the Environmental Engineering group.

(3) Environmental Protection Agency (EPA)

(4) Housing and Urban Development (HUD),

(5) Sloan Valve Company.

The proposed research complements other existing Notre Dame research (e.g. the ND Lead Innovation Team) with a history of private and grant funding. We anticipate sharing resources with this team for some of our activities.

PROJECT TIMELINE AND MILESTONES

The proposed funding period is one year: from July 1, 2022 to Jun 30, 2023.

STUDY 1

June – September. **Task 1**: Data aggregation and analysis. **Milestone 1**: Completion of data aggregation and analysis.

September - November 2022. **Task 2**: Risk modeling. **Milestone 2**: Completion of preliminary model

June 2022 – December 2022. **Task 3**:Model validation using volunteer households. **Milestone 3**: Testing of real homes complete

September 2023 - April 2024. **Task 4**: Write proposals for external funding to expand testing broadly throughout South Bend, in collaboration with the City of South Bend and the CCI. **Milestone 4**: Proposals submitted.

STUDY 2

June 2022 – October 2022. **Task 1**: Training Local Models for Residential Water Age. **Milestone 1**: complete evaluation of local models and initial hardware/software design. **Milestone 2:** identification and preliminary testing of proposed flow and water quality sensors.

October 2022 – February 2023. **Task 2:** Federated Learning of Neighborhood Residential Water Age. **Milestone 3:** complete evaluation of federate learning algorithms and initial software design.

March 2022- May 2023. **Task 3:** Simulation Study of Water Purging Schedule: **Milestone 4:** complete development of purging schedules and preliminary simulation results characterizing the effectiveness of purging.

BUDGET AND BUDGET JUSTIFICATION

We propose a one-year project, including 1 year of support for Isaiah Murrell-Thomas, 0.5 years for Emily Clements, and 0.5 years for Yuying Duan) at rate of $30,000 per student-year (without fringe benefits), 100 hours of undergraduate support, and $10,000 for materials, supplies, and analytical instrument time. Drs. Nerenberg, Sisk, and Lemmon will provide their time as an in-kind contribution to the project. The computational work will use existing PI and campus resources.

Graduate student stipends	$60,000
Fringe benefits for graduate student	$7,000
Undergraduate students for 100 hours	$1,000
Materials, supplies, instrument time	$10,000
TOTAL	$78,000

ADDITIONAL MATERIALS

1. **Previous support from the Lucy Family Institute**

Over the past year, Matthew Sisk has worked with the Lucy Family Institute on the dataMichiana (civic data portal) project, and as of April 1, 2022 is a full-time as a Professor of the Practice in the Lucy Family Institute. The other co-PIs have never had any Lucy Family Institute support.

2. **Biosketch for all members of the proposing team (using the NSF Biosketch Template).**

See below.

3. **Current and Pending support for all senior members of the proposing team (using the NSF Current & Pending Template).**

See below.

NAME: Robert Nerenberg

POSITION TITLE & INSTITUTION: Professor, University of Notre Dame

A. PROFESSIONAL PREPARATION
(see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
University of Buenos Aires	Bs. As., Argentina	Civil Engineering	BS	1990
Wayne State University	Detroit, MI	Civil Engineering	MS	1992
Northwestern University	Evanston, IL	Environmental Engineering	Ph.D.	2003

B. APPOINTMENTS
(see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
2019 - present	Full Professor, University of Notre Dame
2017- present	Kellogg faculty fellow, Kellogg Institute for International Studies
2011- 2012	Visiting Professor, Dept. of Hydr. and Env. Eng. Pontifical Catholic University of Chile
5/2011 - 6/2011	Visiting Professor, Swiss Federal Institute for Aquatic Science and Technology (EAWAG),
2010 – 2015	Director of Graduate Studies, Dept. of Civil and Env. Eng. and Earth Sci., U. of Notre Dame
2010 - 2019	Associate Professor, University of Notre Dame
2004 - 2010	Assistant Professor, University of Notre Dame
6/2003 - 12/2003	Post-Doctoral Research Fellow, Northwestern University
1998 -2003	Research Assistant, Northwestern University
1992-1997	Environmental Engineer, Harza Engineering Company (now Stantec, Inc.), Chicago, IL
1990 - 1992	Environmental Engineer, Ayres Lewis, Norris & May, Ann Arbor, MI

C. PRODUCTS
(see PAPPG Chapter II.C.2.f.(i)(c))
Products Most Closely Related to the Proposed Project

1. Pavissich, JP; Li, MF; Nerenberg, R (2021) Spatial distribution of mechanical properties in Pseudomonas aeruginosa biofilms, and their potential impacts on biofilm deformation. Biotechnology and Bioengineering. 118:4: 1564-1575. DOI: 10.1002/bit.27671
2. Kim, B.; Perez-Calleja, P.; Li, M.; Nerenberg, R (2020). Effect of predation on the mechanical properties and detachment of MABR biofilms. Water Research Vol. 186, Article 116289 DOI: 10.1016/j.watres.2020.116289
3. Jiang, MM; Zheng, JJ; Perez-Calleja, P; Picioreanu, C; Lin, H; Zhang, XH; Zhang, YY; Li, HX; Nerenberg, R (2020). New insight into CO2-mediated denitrification process in H2-based membrane biofilm reactor: An experimental and modeling study. Water Research Vol. 184, DOI: 10.1016/j.watres.2020.116177
4. Li, Mengfei; Matous, Karel; Nerenberg, Robert (2020). Predicting biofilm deformation with a viscoelastic phase-field model: Modeling and experimental studies. Biotechnology and Bioengineering 117:11:3486-3498, DOI: 10.1016/j.watres.2020.116177
5. T. Seviour; N. Derlon, M. Dueholm, H-C. Flemming, E. Girbal-Neuhauser, H. Horn, S. Kjelleberg, M.C. van Loosdrecht, T. Lotti, F.M. Malpei, R. Nerenberg, T.R. Neu, E. Paul, Y. Han-qing, Y. Lin (2019). Extracellular polymeric substances of biofilms: suffering from an identity crisis. Water Research. doi.org/10.1016/j.watres.2018.11.020
6. G. Tierra, J. P. Pavissich, R. Nerenberg, Z. Xu, M. S. Alber (2015). Mechanical role of EPS in biofilm deformation and detachment: a modeling study. Journal of the Royal Society Interface. 12: 20150045. DOI: 10.1098/rsif.2015.0045.

Other Significant Products, Whether or Not Related to the Proposed Project

1. Kazemifar, F; Blois, G; Aybar, M; Calleja, PP; Nerenberg, R; Sinha, S; Hardy, RJ; Best, J; Smith, GHS; Christensen, KT. (2021). The Effect of Biofilms on Turbulent Flow Over Permeable Beds. Water Resources Research 57:2 Art 2019WR026032. DOI: 10.1029/2019WR026032
2. M. Aybar, P. Perez-Calleja, M. Li, J.P. Pavissich, R. Nerenberg (2018). Predation Creates Unique Void Layer in Membrane-Aerated Biofilms. Water Research. https://doi.org/10.1016/j.watres.2018.10.084
3. F. Sabba*, A. Terada, G. Wells, Barth Smets, R. Nerenberg (2018). Nitrous Oxide Emissions from Biofilm Processes for Wastewater Treatment. Applied Microbiology and Biotechnology. doi: 10.1007/s00253-018-9332-7
4. F. Sabba*, C. Picioreanu, J. Pérez, R. Nerenberg (2015). Hydroxylamine Diffusion Can Enhance N2O Emissions in Nitrifying Biofilms: A Modeling Study. Environmental Science & Technology. DOI: 10.1021/es504

D. SYNERGISTIC ACTIVITIES
(see PAPPG Chapter II.C.2.f.(i)(d))

• Chair of International Water Association (IWA) Biofilms Specialist Group (11/2019 - present).
• Organizer of IWA Biofilms 2020 virtual conference. Currently organizing IWA Biofilms 2021 virtual conferenc
• Associate Editor of npj Biofilms and Microbiomes (2021 - present)
• Organized NSF-funded workshop on Biofilm Mechanical Properties (April 22-24, 2018). U. of Notre Dame
• Prepared and delivered hands-on course on biofilm research techniques at Notre Dame. Attended by 14 PhD students, post-docs, and faculty from across the US (2019).
• Board member of Awards Committee for Association of Environmental Engineering and Science Professors (AEESP) (2019-present)

NAME: Michael Lemmon

POSITION TITLE & INSTITUTION: Professor, University of Notre Dame

A. PROFESSIONAL PREPARATION - (see PAPPG Chapter II.C.2.f.(i)(a))

INSTITUTION	LOCATION	MAJOR/AREA OF STUDY	DEGREE (if applicable)	YEAR (YYYY)
Stanford University	Stanford, CA	Electrical Engineering	B.S.	1979
Carnegie-Mellon Univ.	Pittsburgh, PA	Electrical & Computer Eng.	M.S.	1987
Carnegie-Mellon Univ.	Pittsburgh, PA	Electrical & Computer Eng.	Ph.D.	1990

B. APPOINTMENTS - (see PAPPG Chapter II.C.2.f.(i)(b))

From - To	Position Title, Organization and Location
2004-present	Professor of Electrical Engineering, University of Notre Dame, Notre Dame, IN
1996-2003	Associate Prof. of Electrical Eng., University of Notre Dame, Notre Dame, IN
1990-1996	Assistant Prof. of Electrical Eng., University of Notre Dame, Notre Dame, IN
1986-1990	Research Assistant, Carnegie-Mellon Univ. Pittsburgh, PA
1983-1986	Control/Signal Processing Engineer, General Electric, Valley Forge, PA
1981-1983	Research Engineer, Lockheed Missiles and Space Co., Sunnyvale, CA
1979-1981	Member of Technical Staff, TRW, Redondo Beach, CA

C. PRODUCTS - (see PAPPG Chapter II.C.2.f.(i)(c)) Products Most Closely Related to the Proposed Project

1. T. Gong, T. Zhang, XS Hu, Q. Deng, M.D. Lemmon, and S. Han, ``Reliable Dynamic Packet Scheduling over Lossy Real-time Wireless Networks'', 31st Euromicro Conference on Real-Time Systems (ECRTS 2019)
2. B. Wu, M.D. Lemmon, H. Lin, ``Local methods for stability analysis of networked control systems with IEEE 802.15.4 protocol'', IEEE Transactions on Control Systems Technology, 26(5), 1635-1645, 2017
3. L. Li, X. Wang, M.ZD. Lemmon, ``Efficiently attentive event-triggered systems with limited bandwidth'', IEEE Transactions on Automatic Control, 62(3), 1491-1497, 2016
4. B. Hu, M.D. Lemmon, ``Distributed switching control to achieve almost sure safety for leader-follower vehicular networked systems'', IEEE Transactions on Automatical Control, 62(12), 3195-3209, 2016.
5. X. Wang, and M.D. Lemmon, ``Event-=triggering in distributed networked control systems'', IEEE Transactions on Automatic Control, 56(3), 586-601, 2010.

Other Significant Products, Whether or Not Related to the Proposed Project

1. X. Wang, M.D. Lemmon, ``Self-triggered Feedback Control Systems with Finite-Gain L2-stability'', IEEE Transactions on Automatic Control, 54(3), 452-467, 2009.
2. Sergiuu Rafiliu, Petru Eles, Zebo Peng, M.D. Lemmon, ``Stability of on-line resource managers for distributed systems under execution time variations'', ACM Transactions on Embedded Computing Systems, 14(2),1-29, 2015
3. A.M. Nightingale, B. Goodwine, M.D. Lemmon, E.J. Jumper, ``Phase-locked-loop adaptive-optic controller and simulated shear layer correction'', AIAA journal, 51(11), 2714-2726, 2013.
4. X. Wang, M.D. Lemmon, ``On event design in event-triggered feedback systems'', Automatica, 47 (10):2319-2322,2012.
5. T. Chantem, X.S. Hu, M.D. Lemmon, ``Generalized Elastic Scheduling for Real-time Tasks'', IEEE Transactions on Computers, 58(4):480-495, 2009.

D. SYNERGISTIC ACTIVITIES - (see PAPPG Chapter II.C.2.f.(i)(d))

1. Associate Chair and Director of Graduate Studies, Electrical Engineering, University of Notre Dame, 2021-present.
2. Lead investigator in project forming a partnership between Notre Dame, city of South Bend, Indiana, and a small business (EMNET LLC) to commercialize sensor-actuator network technologies used in controlling combined sewer overflows, 2005-2009.
3. Served on executive committee for Notre Dame's GLOBES interdisciplinary training program, 2010-2013.
4. Associate Editor, IEEE Transactions on Control Systems Technology, 1999-2004
5. Associate Editor, IEEE Transactions on Neural Networks, 1990-1994

NSF BIOGRAPHICAL SKETCH

NAME: Sisk, Matthew L

POSITION TITLE & INSTITUTION: GIS Librarian, University of Notre Dame

(a) PROFESSIONAL PREPARATION -(see PAPPG Chapter II.C.2.f.(a))

INSTITUTION	LOCATION	MAJOR / AREA OF STUDY	DEGREE (if applicable)	YEAR YYYY
University of South Carolina	Columbia, SC	Anthropology	AB	2001
University of South Carolina	Columbia, SC	Marine Science	BS	2001
Stony Brook University	Stony Brook, NY	Anthropology	MS	2006
Stony Brook University	Stony Brook, NY	Anthropology (Archaeology)	PHD	2011

(b) APPOINTMENTS -(see PAPPG Chapter II.C.2.f.(b))

2015 - present GIS Librarian, University of Notre Dame, Notre Dame, IN
2013 - 2015 Postdoctoral Research Fellow, University of Notre Dame, Notre Dame, IN
2012 - 2012 Chercheur invité (Invited researcher), University of Toulouse, Toulouse
2008 - 2013 Research Scientist, Applied Biomathematics, Setauket, NY

(c) PRODUCTS -(see PAPPG Chapter II.C.2.f.(c))

Products Most Closely Related to the Proposed Project

1. Ringwald P, Chapin C, Iceman C, Tighe ME, Sisk ML, Peaslee G, Peller J, Wells EM. Characterization and within-site variation of environmental metal concentrations around a contaminated site using a community-engaged approach. Chemosphere. 2021 January. Other: 129915
2. Tighe M, Knaub C, Sisk M, Ngai M, Lieberman M, Peaslee G, Beidinger H. Validation of a screening kit to identify environmental lead hazards. Environ Res. 2020 Feb;181:108892. PubMed PMID: 31735346.
3. Beidinger-Burnett H, Ahern L, Ngai M, Filippelli G, Sisk M. Inconsistent screening for lead endangers vulnerable children: policy lessons from South Bend and Saint Joseph County, Indiana, USA. J Public Health Policy. 2019 Mar;40(1):103-113. PubMed PMID: 30559451.

Other Significant Products, Whether or Not Related to the Proposed Project

1. White R, Mensan R, Bourrillon R, Cretin C, Higham TF, Clark AE, Sisk ML, Tartar E, Gardère P, Goldberg P, Pelegrin J, Valladas H, Tisnérat-Laborde N, de Sanoit J, Chambellan D, Chiotti L. Context and dating of Aurignacian vulvar representations from Abri Castanet, France. Proc Natl Acad Sci U S A. 2012 May 29;109(22):8450-5. PubMed Central PMCID: PMC3365179.
2. Molik DC, Tomlinson D, Davitt S, Morgan EL, Sisk M, Roche B, Meyers N, Pfrender ME. Combining natural language processing and metabarcoding to reveal pathogen-environment associations. PLoS Negl Trop Dis. 2021 Apr;15(4):e0008755. PubMed Central PMCID: PMC8055023.

3. Maillie L, Lazarev S, Simone CB 2nd, Sisk M. Geospatial Disparities in Access to Proton Therapy in the Continental United States. Cancer Invest. 2021 Jul-Aug;39(6-7):582-588. PubMed PMID: 34152235.

(d) SYNERGISTIC ACTIVITIES -(see PAPPG Chapter II.C.2.f.(d))

1. Serves on Digitizing Hidden Collections Review Panel for the Digital Library Federation
2. Maintaining spatial components of large digital collections like Building South Bend and Cities in Text (https://hue.nd.edu/)
3. Works with the Palliative Care Association of Uganda to develop and assess mHealth applications and train researchers / practitioners in their use
4. Session chair and conference planning committee, Digital Library Federation 2018-2020
5. Collaboration with local municipality and public library system to assess the data needs of regional non-profits.

Projects/Proposals

1.*Project/Proposal Title : GOALI: Effect of Hydroxylamine on the Structure and Function of Nitrifying Biofilms

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available): CBET-1805406_A1

*Source of Support: National Science Foundation (NSF)

*Primary Place of Performance : University of Notre Dame

Project/Proposal Start Date (MM/YYYY) (if available) : 10/2018

Project/Proposal End Date (MM/YYYY) (if available) : 09/2021

*Total Award Amount (including Indirect Costs): $ 369,492

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1. 2019	0.50	4.	
2. 2020	0.50	5.	
3. 2021	0.50		

2.*Project/Proposal Title : Wastewater Management Studies for ECOM Wet Mills

*Status of Support : (•) Current ( ) Pending ( ) Submission Planned ( ) Transfer of Support

Proposal/Award Number (if available):

*Source of Support: ECOM Agroindustrial Corp. Ltd

*Primary Place of Performance : University of Notre Dame

Project/Proposal Start Date (MM/YYYY) (if available) : 04/2021

Project/Proposal End Date (MM/YYYY) (if available) : 03/2022

*Total Award Amount (including Indirect Costs): $ 67,421

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1. 2022	0.30	4.	
2.		5.	
3.			

Projects/Proposals

3. *Project/Proposal Title : Preliminary ECO-CBET: Unraveling Biofilm Mechanical Behavior: A Convergent Approach (This Pre-proposal)

*Status of Support : ○ Current ◉ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available): Cayuse-22-0324

*Source of Support: National Science Foundation (NSF)

*Primary Place of Performance : University of Notre Dame

Project/Proposal Start Date (MM/YYYY) (if available) : 01/2022

Project/Proposal End Date (MM/YYYY) (if available) : 12/2025

*Total Award Amount (including Indirect Costs): $ 1,522,795

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1. 2022	1.00	4. 2025	1.00
2. 2023	1.00	5.	
3. 2024	1.00		

4. *Project/Proposal Title :

*Status of Support : ○ Current ○ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support:

*Primary Place of Performance :

Project/Proposal Start Date (MM/YYYY) (if available) :

Project/Proposal End Date (MM/YYYY) (if available) :

*Total Award Amount (including Indirect Costs): $

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)	Year (YYYY)	Person Months (##.##)
1.		4.	
2.		5.	
3.			

*PI/co-PI/Senior Personnel Name: Michael D. Lemmon

***Required fields**

Note: NSF has provided 15 project/proposal and 10 in-kind contribution entries for users to populate. Please leave any unused entries blank.

Project/Proposal Section:

Current and Pending Support includes all resources made available to an individual in support of and/or related to all of his/her research efforts, regardless of whether or not they have monetary value.[1] Information must be provided about all current and pending support, including this project, for ongoing projects, and for any proposals currently under consideration from whatever source, irrespective of whether such support is provided through the proposing organization or is provided directly to the individual. This includes, for example, Federal, State, local, foreign, public or private foundations, non-profit organizations, industrial or other commercial organizations, or internal funds allocated toward specific projects. Concurrent submission of a proposal to other organizations will not prejudice its review by NSF, if disclosed.[2]

[1] If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.
[2] The Biological Sciences Directorate exception to this policy is delineated in PAPPG Chapter II.D.2.

Projects/Proposals

1. *Project/Proposal Title : CPS:Small: Learning How to Control: A Meta-Learning Approach for the Adaptive Control of Cyber-Physical Systems (THIS PROPOSAL)

*Status of Support : ○ Current ◉ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support: National Science Foundation (NSF)

*Primary Place of Performance : University of Notre Dame

Project/Proposal Start Date (MM/YYYY) (if available) : 01/2023

Project/Proposal End Date (MM/YYYY) (if available) : 01/2026

*Total Award Amount (including Indirect Costs): $ 500,000

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

	*Year (YYYY)	*Person Months (##.##)
1.	2023	0.75
2.	2024	0.75
3.	2025	0.75

	Year (YYYY)	Person Months (##.##)
4.		
5.		

*Overall Objectives : This project develops a new meta-learning framework for the control of complex CPS found in IoT-enabled manufacturing systems.

*Statement of Potential Overlap : 100% overlap. This is the project

Projects/Proposals

2. *Project/Proposal Title : Collaborative Research: PPoSS: LARGE: Towards Scalable Design of Resilient Mission-Critical IoT Systems

*Status of Support : ○ Current ◉ Pending ○ Submission Planned ○ Transfer of Support

Proposal/Award Number (if available):

*Source of Support: National Science Foundation (NSF)

*Primary Place of Performance : University of Notre Dame

Project/Proposal Start Date (MM/YYYY) (if available) : 10/2022

Project/Proposal End Date (MM/YYYY) (if available) : 09/2027

*Total Award Amount (including Indirect Costs): $ 1,665,041

*Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project

*Year (YYYY)	*Person Months (##.##)
1. 2023	1.25
2. 2024	1.25
3. 2025	1.25
4. 2026	1.25
5. 2027	1.25

*Overall Objectives : Development of scalable methods for the design and management of mission-critical IoT systems

*Statement of Potential Overlap : 20% overlap

In Kind Contributions

*Required fields

In-Kind Contribution Section:

Current and Pending Support also includes in-kind contributions (such as office/laboratory space, equipment, supplies, employees, students). If the in-kind contributions are intended for use on the project being proposed to NSF, the information must be included as part of the Facilities, Equipment and Other Resources section of the proposal and need not be replicated in the individual's Current and Pending Support submission. In-kind contributions not intended for use on the project/proposal being proposed that have associated time obligations must be reported below. If the time commitment or dollar value is not readily ascertainable, reasonable estimates should be provided.

Please enter your support entries so they are grouped together based on the "Status of Support" and are in the order of Current to Pending from top to bottom

In Kind Contributions

1.*Status of Support : (•) Current ( ) Pending

*Source of Support : University of Notre Dame

*Primary Place of Performance : University of Notre Dame

*Summary of In-Kind Contributions : 1 month of summer support for serving as director of graduate studies in ND department of electrical engineering

Time Commitment - Month(s) (or Partial Person-Months) Committed Per Year

If the time commitment is not readily ascertainable, reasonable estimates should be provided.

	*Year (YYYY)	*Person Months (##.##)
1.	2022	1.00
2.	2023	1.00
3.	2024	1.00

	Year (YYYY)	Person Months (##.##)
4.	2025	1.00
5.	2026	1.00

*Dollar Value of In-Kind Contribution: $ 15,000

*Overall Objectives : summer support for serving as director of graduate studies for ND-EE

*Statement of Potential Overlap : no overlap with this project.

PI/co-PI/Senior Personnel: Sisk, Matthew

PROJECT/PROPOSAL CURRENT SUPPORT

1. Project/Proposal Title: A New Paradigm for Environmental Lead Reduction: Community Lead Screening

   Proposal/Award Number (if available): INLTS0017-20

   Source of Support: Department of Housing and Urban Development (HUD)

   Primary Place of Performance: University of Notre Dame

   Project/Proposal Support Start Date (if available): 11/2020

   Project/Proposal Support End Date (if available): 09/2022

   Total Award Amount (including Indirect Costs): $700,000

   Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2021	0.6
2022	0.5

   Overall Objectives: Scale up of screening kit for environmental lead in households

   Statement of Potential Overlap: none

PROJECT/PROPOSAL PENDING SUPPORT

1. Project/Proposal Title: Frameworks: PeerPresQT: Improving Preservation Quality and Data Sharing in a Peer-to-Peer Model using FAIR principles (This Proposal)

   Proposal/Award Number (if available):

   Source of Support: UIUC - NSF

   Primary Place of Performance: University of Notre Dame

   Project/Proposal Support Start Date (if available): 07/2022

   Project/Proposal Support End Date (if available): 06/2025

   Total Award Amount (including Indirect Costs): $608,139

Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2022	0.5
2023	1
2024	1
2025	0.5

Overall Objectives: The goal of PeerPresQT is to enhance the existing PresQT (Preservation Quality Tool) services with Peer-to-Peer technology for data sharing and preservation under consideration of FAIR principles.

Statement of Potential Overlap: This proposal

2. Project/Proposal Title: Partnership to address environmental contamination and brain health

Proposal/Award Number (if available):

Source of Support: National Institutes of Health (NIH) / Purdue University

Primary Place of Performance: University of Notre Dame

Project/Proposal Support Start Date (if available): 11/2021

Project/Proposal Support End Date (if available): 12/2026

Total Award Amount (including Indirect Costs): $972,026

Person-Month(s) (or Partial Person-Months) Per Year Committed to the Project:

Year	Person-months per year committed
2021	1
2022	1
2023	1

Overall Objectives: Address environmental toxins in local communities

Statement of Potential Overlap: None

REFERENCES

Allaire, M., Wu, H. and Lall, U. 2018. National trends in drinking water quality violations. Proceedings of the National Academy of Sciences of the United States of America 115(9), 2078-2083.

Cassell, K., Gacek, P., Rabatsky-Ehr, T., Petit, S., Cartter, M. and Weinberger, D. 2019. Estimating the True Burden of Legionnaires' Disease. American Journal of Epidemiology 188(9), 1686-1694.

Cullom, A., Martin, R., Song, Y., Williams, K., Williams, A., Pruden, A. and Edwards, M. 2020. Critical Review: Propensity of Premise Plumbing Pipe Materials to Enhance or Diminish Growth of Legionella and Other Opportunistic Pathogens. Pathogens 9(11).

Egendorf, S.P., Mielke, H.W., Castorena-Gonzalez, J.A., Powell, E.T. and Gonzales, C.R. 2021. Soil Lead (Pb) in New Orleans: A Spatiotemporal and Racial Analysis. International Journal of Environmental Research and Public Health 18(3).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. 2014. Generative adversarial nets, Advances in neural information processing systems, 27.

Konevcny, J., McMahan, H.B., Yu, F.X., Richtarik, P., Suresh, A.T. and Bacon, D. 2016. Federated learning: strategies for improvin communication efficiency, ArXiv preprint arXiv:1610.05492.

Lau, H. and Ashbolt, N. 2009. The role of biofilms and protozoa in Legionella pathogenesis: implications for drinking water. Journal of Applied Microbiology 107(2), 368-378.

Leslie, E., Hinds, J. and Hai, F. 2021. Causes, Factors, and Control Measures of Opportunistic Premise Plumbing Pathogens-A Critical Review. Applied Sciences-Basel 11(10).

Ley, C.J., Proctor, C.R., Singh, G., Ra, K., Noh, Y., Odimayomi, T., Salehi, M., Julien, R., Mitchell, J., Nejadhashemi, A.P., Whelton, A.J. and Aw, T.G. 2020. Drinking water microbiology in a water-efficient building: stagnation, seasonality, and physicochemical effects on opportunistic pathogen and total bacteria proliferation. Environmental Science-Water Research & Technology 6(10), 2902-2913.

Li, Tian, Sahu, A.K., Talwalkar, A., and Smith, V. 2020. Federated learning: challenges, methods, and future directions, IEEE Signal Processing Magazine. 37(3):50-60.

Marshall, A.T., Betts, S., Kan, E.C., McConnell, R., Lanphear, B.P. and Sowell, E.R. 2020. Association of lead-exposure risk and family income with childhood brain outcomes. Nature Medicine 26(1), 91-+.

Neu, L. and Hammes, F. 2020. Feeding the Building Plumbing Microbiome: The Importance of Synthetic Polymeric Materials for Biofilm Formation and Management. Water 12(6).

Proctor, C., Gachter, M., Kotzsch, S., Rolli, F., Sigrist, R., Walser, J. and Hammes, F. 2016. Biofilms in shower hoses - choice of pipe material influences bacterial growth and communities. Environmental Science-Water Research & Technology 2(4), 670-682.

Rhoads, W.J., Pruden, A. and Edwards, M.A. 2016. Survey of green building water systems reveals elevated water age and water quality concerns. Environmental Science-Water Research & Technology 2(1), 164-173.

Stillo, F. and Gibson, J. 2017. Exposure to Contaminated Drinking Water and Health Disparities in North Carolina. American Journal of Public Health 107(1), 180-185.

Weisner, M., Root, T., Harris, M., Mitsova, D. and Liu, W. 2020. Tap water perceptions and socioeconomics: Assessing the dissatisfaction of the poor. Sustainable Production and Consumption 21, 269-278.