

---

## Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** Securing Drinking Water Supplies: the role of organic matter on water treatability

**Creator:** Samuel Yeboah Nyarko

**Principal Investigator:** Samuel Yeboah Nyarko

**Affiliation:** Cranfield University

**Funder:** Engineering and Physical Sciences Research Council (EPSRC)

**Template:** EPSRC Data Management Plan

**ORCID iD:** 0000-0002-7064-806X

### Project abstract:

Water transfer from water surplus regions to water deficit regions is being considered in the Anglian Water region of the UK and is also being debated in Ghana to improve the resilience of drinking water supplies. However, we do not know how this will influence water quality (e.g., natural organic matter (NOM), which changes temporally and spatially) and its treatability. These water transfers will invariably result in blends of water containing NOM of different characteristics, resulting in variable removal potentials by normal coagulation and sorption processes. NOM character also influences the downstream formation of potentially carcinogenic disinfection by-products (DBPs). Therefore, understanding the characteristics of organic matter will be of paramount importance in effectively managing and treating new water resources and blends in both the UK and Ghana.

This project will employ new techniques to understand and link NOM profiles and fingerprints to removal by coagulation and other processes (e.g., ion exchange, adsorption). This includes measuring molecular weight, charge and hydrophobicity profiles and the molecular characterisation of the different sized fractions, which can give a cohesive understanding of NOM profiles when combined with advanced characterisation techniques (e.g., fluorescence).

More specifically, the project seeks to:

1. Characterise and assess the spatial and temporal variability of NOM in both the UK and Ghana lowland surface waters;
2. Determine which features of NOM control their removal by water treatment processes (such as coagulation, ion exchange, adsorption);
3. Determine the fraction of NOM that is removed and not removed by the treatment processes and their influence on DBPs formation;
4. Identify the differences in NOM between the UK and Ghana source waters and understand how this might influence water resource management plans.

The Methodology will involve a literature review to compare and analyse studies on NOM characterisation and removal techniques to determine a laboratory experimental work baseline. The laboratory work will include jar tests, bench experiments, and a pilot plant to replicate treatment systems. Here, we will test the influence of changing

**ID:** 96041

**Start date:** 06-09-2021

**End date:** 27-02-2026

**Last modified:** 06-04-2023

**Grant number / URL:** EP/S023666/1

**Copyright information:**

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# Securing Drinking Water Supplies: the role of organic matter on water treatability

---

## Data Collection

### What data will you collect or create?

Most of the data obtained in this research will come from experimental measurements, such as UV254, charge density, molecular weight, hydrophobicity, fluorescence and other water quality parameters, such as DOC, alkalinity, pH, sulphate, nitrate, etc.

This research will also use analytical data from the sponsor, Anglian Water. This data will be in the .xlsx format, which is the company's standard for storing data. Because some of this information may be sensitive, it will be treated following Anglian Water's rules and regulations. None of the information will be kept on a personal computer. The raw data will have to be kept on the sponsor's server. In addition, a literature review will be conducted, with data from published literature being collected in a docx document.

During the project, reports will be written as .docx documents, numeric data will be processed as .xlsx documents for versatility, and convenience. JPEG format will be used to record photographs from project sites, experimental set-ups, and other inanimate objects. At the end of the project (or project sub-sections), any long-term data of value in proprietary formats will be converted into standard non-proprietary formats such as PDF/A, .csv, .txt, .JPEG. to ensure easy access and sharing across different systems and software

The volume of all generated data is expected to be under 20 GB.

Samuel Yeboah Nyarko, the research student, is in charge of data collection. Analytical tools will be calibrated regularly per standard operating procedures, and repetitions of the experiments will be run to confirm the quality of the data collected. Professor Peter Jarvis and Dr Irene Carra will provide further quality control by reviewing the generated data.

### How will the data be collected or created?

Data will be collected by the lead investigator, Samuel Yeboah Nyarko. Where applicable, industry-standard methods (e.g., Standard Methods for the Examination of Water and wastewater, APHA) and units (SI units) will be used. The quality of experimental data will be ensured by instrument calibration, repeatability and validation and will be reviewed by the project supervisors, Professor Peter Jarvis and Dr Irene Carra. Test data files shall use the naming convention 'dd-mm-yy-variable-test number', and version control tables shall be used for vital documents.

## Documentation and Metadata

### What documentation and metadata will accompany the data?

A README text file will accompany the data to be easily read and interpreted in the future, even by secondary users.

The README file will include basic details, like who created or contributed to the data, its

title, date of creation and under what conditions it can be accessed. Furthermore, details on the methodology used, analytical and procedural information, definitions of variables, vocabularies, units of measurement, any assumptions made, and the format and file type of the data will be highlighted too.

Data will use CORD metadata.

## **Ethics and Legal Compliance**

### **How will you manage any ethical issues?**

The work has been approved through Cranfield University's Ethics System (CURES) and shall adhere to Cranfield University's ethics framework. There will be no personal data collected as part of this research.

### **How will you manage copyright and Intellectual Property Rights (IPR) issues?**

Data generated by this project shall be owned by the company (Anglian Water and the Ghana Water Company). Per the intellectual property agreement between the data owners, reports shall be given to Anglian Water and the Ghana Water Company for written approval before publication. Anglian Water and the Ghana Water Company may request information be kept confidential. Confidential data shall not be shared with third parties before three years from the research approval date if the data pertains to the report or before five years for any other confidential data. The data will be licenced for re-use under the CC-BY terms.

## **Storage and Backup**

### **How will the data be stored and backed up during the research?**

The experimental data will be named by describing the analysis, sample information (e.g. parameter, water type, and date of sampling), the measurement date, and the abbreviation of the researcher taking the measurement, e.g., DOC\_RivTrent\_15/03/22\_21/03/22\_SYN. Data will be stored on Cranfield University's network drive, which is automatically backed up by Cranfield IT daily to multiple data centres to minimise risks from data loss. Any essential data recorded in lab books or independent electronic devices will be promptly transferred to the network drive to minimise accumulating risks from data loss (e.g. at the end of each day). Lab books shall be kept securely when not in use and will be handed to the library at the end of the project for preservation and digitisation according to Cranfield university's lab book policy.

### **How will you manage access and security?**

The data will be kept on the university network drive which has restricted access, with access given to the sponsor. The principal investigator (Samuel Yeboah Nyarko) will use a secure personal password

which will be updated yearly to access the data. Additional security measures such as data encryption will be used.

## **Selection and Preservation**

### **Which data are of long-term value and should be retained, shared, and/or preserved?**

The selection of data considered to be of long term value for retention, sharing and preservation shall be at the collective discretion of the principal investigator, supervisors, Anglian Water and the Ghana Water Company Limited. The criteria used to decide what data to be preserved will be based on:

- How the data fall within the repository's scope and other relevant legal requirements
- Uniqueness and potential for data redistribution
- Ease of data replicability

### **What is the long-term preservation plan for the dataset?**

Data of long term value shall be kept on the Cranfield Online Research Data (CORD) repository, which uses the figshare platform and preserves data for at least ten years after the project ends, with datasets assigned a DOI for long-term accessibility, per Cranfield's Management of Research Data Policy. No costs are associated with this activity.

## **Data Sharing**

### **How will you share the data?**

Data will be shared publically through open access scientific publications and Cranfield University's CORD repository according to the university's guidelines (except where confidentiality restrictions apply). Reference to the supporting data and meta-data will be included in reports (with links where appropriate). Contact details of the principal investigator will also be provided in reports.

### **Are any restrictions on data sharing required?**

Data on the report deemed confidential by Anglian Water and Ghana Water Company Limited may be restricted from public view for up to five years from project approval. Here, published metadata will contain the confidentiality reason and any condition(s) which need(s) to be met before access is granted.

Non-confidential data will be made accessible to the public following the publication of a research paper and at the end of the project.

## **Responsibilities and Resources**

### **Who will be responsible for data management?**

The principal investigator, Samuel Yeboah Nyarko, will be responsible for data management including updating and reviewing under the supervision of Professor Peter Jarvis and Dr Irene Carra of Cranfield University. Cranfield University's IT Department will be responsible for managing the network storage facility, backups and availability of long-term data.

### **What resources will you require to deliver your plan?**

RDM costs are included in the project grant, and no extra resources are expected at this stage.