


APPENDIX A – PROPOSAL COVER SHEET

PROGRAMME NAME: Research Ireland – Wales Innovation Network Research Alliance Award	
PARENT AWARD PROPOSAL ID AND END DATE (RESEARCH IRELAND): 21/PATH-S/9454 (Research Ireland PATHWAY AWARD); END DATE: 31.05.2026	
TITLE OF COLLABORATION PROPOSAL: Machine Learning-Driven Discovery and Assessment of Metal-Organic Adsorbents for Freshwater Cleanup	
FULL NAME OF RESEARCH IRELAND PROJECT LEAD: Soumya Mukherjee	HOST INSTITUTION: University of Limerick
FULL NAME OF WIN CO-LEAD: Mosab Bazargani	HOST INSTITUTION: Bangor University
% RESEARCH IRELAND LEAD COMMITMENT: 5%	% WIN CO-LEAD COMMITMENT: 5%
TOTAL REQUESTED RESEARCH IRELAND BUDGET (€): 30,000	TOTAL REQUESTED WIN BUDGET (€):* 30,000
TOTAL JOINT BUDGET (€): 60,000	
REQUESTED STARTING DATE: 01.04.2025	PROPOSED DURATION (3-12 MONTHS): 12
PROJECT LEAD ELIGIBLE TO APPLY FOR HORIZON EUROPE FUNDING? (Y/N) INCLUDE PIC NUMBER IF APPLICABLE	Y PIC 999809071

<p>Signatures below confirm acceptance and agreement with the Research Ireland grants and awards Terms and Conditions, and that the institution ensures the applicant meets eligibility requirements, and that the project is in full agreement with all legal and regulatory matters governing research in Ireland, and no aspect of this project is already being funded from another source and all details provided are correct.</p>	
<p>INSTITUTIONAL SIGNATORY AUTHORITY (IRISH LEAD INSTITUTION ONLY)</p> <p>Name: Puneet Saidha</p> <p>Position: Director of Research</p> <p>Email: puneet.saidha@ul.ie</p> <div style="text-align: center;">  </div> <p>Signed:</p> <p>Date: 29 January 2025</p>	

Research Collaboration Plan

Title of Collaboration proposal: Machine Learning-Driven Discovery and Assessment of Metal-Organic Adsorbents for Freshwater Cleanup (Acronym: **MARCIA**)

Synopsis: *Thanks to bottom-up design, crystalline metal-organic framework (MOF) adsorbents are primed to elicit bespoke structure-function rationales that enable trace scavenging of drinking water contaminants, per- and polyfluoroalkyl substances (PFAS). Aligned with the United Nations (UN) Sustainable Development Goals SDG 3: Good Health and Well-being and SDG 6: Clean Water and Sanitation, overall objective of this research proposal is to (A) conduct data-driven, high-throughput screening on the computation-ready experimental metal-organic framework (CoRE-MOF-2019) database to identify the top-performing MOFs for perfluorooctanoic acid (PFOA, a prototypical legacy C8 PFAS) removal in lab-scale (led by the Wales PI); (B) experimentally determine the PFOA removal performances and binding sites in these lead MOFs (led by the Irish PI). Synergy between these two iterative workflows will not only introduce a disruptive machine learning-framework for screening PFAS adsorbent libraries, but will also reveal the sweet spots of strong PFAS binding within the benchmark MOF sorbents, validated both computationally and experimentally. Our award findings will be key to drafting two joint grant proposals, to: 1) Marie Skłodowska-Curie Actions (MSCA) Doctoral Networks grant, 2) Joint Translational Call “Water and Health” (Water4all), both in 2025.*

Background: Per- and polyfluoroalkyl substances (PFAS) are a broad group of over 10,000 synthetic, non-biodegradable organics widely used as surface coatings in the chemical, textile and automotive industries worldwide.^[1] PFAS consist of fully (per-) or partially (poly-) fluorinated linear or branched alkyl hydrophobic chains attached to various polar functional groups, primarily carboxylate, or sulphonate. Each fluoroalkyl chain contains several carbon-fluorine (C–F) bonds, regarded as the “strongest bonds in organic chemistry”, each offering a bond energy as high as $\approx 485 \text{ kJ mol}^{-1}$.^[2] The perfluoroalkyl moieties of PFAS confer ultra-low surface tension, resulting in high hydrophobicity and oleophilicity. Such confluence of these features render PFAS particularly suitable for producing water- and stain-resistant, non-stick surfactants and polymers, supporting a wide range of industrial and consumer applications (Figure 1a), such as paper packaging, textiles and leather, and metal plating.

Unsurprisingly, PFAS are prevalent in various consumer products, and industrial commodities, making their presence ubiquitous in soil, water, and air. However, prolonged exposure to PFAS in drinking water, such as perfluorooctanoic acid (PFOA) and perfluorooctane sulfonic acid (PFOS), even at trace levels ($\leq 10,000$ parts per million, ppm), can lead to human bioaccumulation with a host of adverse effects, including hepatotoxicity, tumour induction, developmental toxicity, immunotoxicity, neurotoxicity, and endocrine disruption (Figure 1b).^[3] A 2023 study by the U.S. Geological Survey reveals that at least 45% of the nation's tap water is contaminated by one or more types of PFAS compounds. Additionally, the Royal Society of Chemistry reported in August 2024 that 35% and 37% of watercourses tested in England and Wales contain medium or high-risk levels of PFOS and PFOA, respectively. Stringent regulations are in place—or under development—for a wide range of PFAS through the European Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) regulation. More importantly, the EU Drinking Water Directive limits the total PFAS concentration in drinking water supplies to 500 ng L^{-1} , with a combined limit of 100 ng L^{-1} for 20 specific PFAS. To this end, the US Environmental Protection Agency (EPA) introduced a new national drinking water regulation in April 2024, capping PFOA and PFOS at 4 ng L^{-1} (*i.e.*, 4 parts per trillion, or ppt) each, and three other emerging PFAS at 10 ng L^{-1} .^[4] These limits are expected to become even stricter in the future as advances in analytical techniques

[1] R. C. Buck, *et al.*, *Integrated Environmental Assessment and Management* **2011**, 7, 513. [2] D. O'Hagan, *Chem. Soc. Rev.* **2008**, 37, 308. [3] K. Kannan, *et al.*, *Environ. Sci. Technol.* **2004**, 38, 4489. [4] Safe Drinking Water Act, Per- and Polyfluoroalkyl Substances (PFAS). in <https://www.epa.gov/sdwa/and-polyfluoroalkyl-substances-pfas>, accessed on 31st January **2025**.

enhance the detection and quantification of PFAS.^[5] Simply put, the profound societal importance of PFAS remediation from freshwater cannot be overstated.

Physicochemical properties of PFAS, particularly their chemical inertness complicate their decontamination from drinking water. At present, the only *ready-for-use* drinking water PFAS remediation technology is adsorption by granular activated carbon (GAC). However, *state-of-the-art* (SOTA) GAC and powdered activated carbon (PAC) variants suffer from low affinity for short-chain PFAS, owing to slow kinetics, large interference during adsorption, and energy-intensive regeneration. *These pitfalls foster the search for alternative energy-*

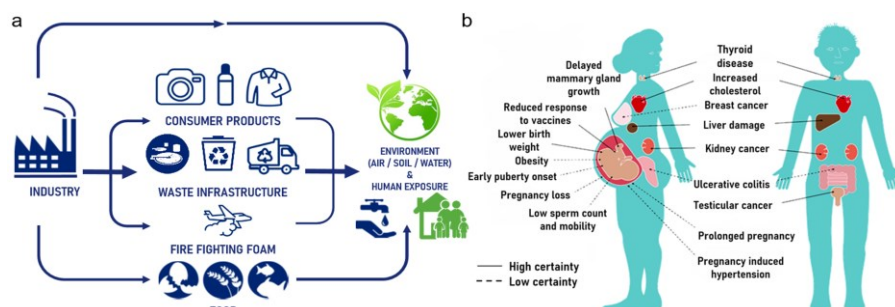


Figure 1: Schematics illustrating a) PFAS exposure pathways; b) diverse impacts on human health caused by prolonged PFAS exposure.

efficient adsorbents that rely upon regenerable adsorption as a technology. Compositionally modular adsorbent solids, especially recyclable physisorbents that offer energy-efficient sorbent recycling is a solution with high upside potential. In this context, reticular metal-organic frameworks (MOFs),

often constructed from transition metals and organic ligands, lend themselves to bottom-up crystal engineering design principles. Synthesised bottom-up, and tailored for recyclable performances, MOFs if designed right, have all the early signs to become the next PFAS remediation solutions, offering at least a few orders of magnitude improvement in PFAS removal efficiencies compared to the activated carbons.

Importance: Although MOFs remain understudied in this context, until very recently, the focus has always been upon harnessing the high surface-area led meso-/microporosity of MOFs to strongly immobilise and bind PFAS molecules within their pore channels, albeit with little and sporadic success.^[6] Breaking away from this known paradigm (that suffers from oft-encountered trade-offs between guest-accessible mesoporosity and water stability)^[7], a recent (2024) study led by the UL principal investigator, PI (applicant) has shown ***how an optimal surface combination of hydrophobicity and electrostatics enables the first observation of trace (environmentally relevant concentrations $\approx 2 \mu\text{g L}^{-1}$, i.e., 2 parts per billion, ppb) removal of PFOA and PFOS in a prototypical water-stable MOF despite inadequate PFAS-accessible surface area, with benchmark removal kinetics and efficiencies.***^[8] Although this study emphasised how ***surface functionalisation alone could elicit strong PFAS binding, as compared to the traditional approach of harnessing pores, the effect of pore size facilitating smaller sized PFAS (lesser than the C8s) capture was also evident from the experimental data.*** The Cambridge Structural Database, CSD^[9] has > 128,000 MOF entries thus far.^[10] Such an *embarras de richesses* makes it inevitable to feel like searching for a needle in a haystack, if relying solely on experimental approaches. To this end, only a tiny subset of *ca.* 20 MOFs are known to demonstrate PFAS capture performances (including experiments that start from both high and low PFAS concentrations).^[6] The vast number of potential MOFs renders purely experimental or molecular simulation-based screening impractical due to the prohibitive computational costs and time constraints. Conventional approaches, such as molecular dynamics (MD) and density functional theory (DFT) simulations, are inherently limited by scalability, typically restricting the screening process to fewer than 200 MOFs, thus far.^[6] This limitation significantly reduces the probability of identifying optimal candidates, as the most promising MOFs are likely embedded within a much larger and largely unexplored dataset of the CSD.^[9] To overcome this challenge, ***we propose a***

[5] V. A. Arias Espana, *et al.*, *Environ. Technol. Innov.* **2015**, 4, 168. [6] R. Li, *et al.*, *Matter* **2022**, 5, 3161. [7] R.-R. Liang, *et al.*, *J. Am. Chem. Soc.* **2024**, 146, 9811. [8] N. Ilić, *et al.*, *Adv. Mater.*, **2025**, 2413120. [9] C. R. Groom, *et al.*, *Acta Crystallogr. B.* **2016**, 72, 171. [10] P. Z. Moghadam, *et al.*, *Chem. Sci.* **2020**, 11, 8373.

machine learning (ML) framework that integrates both labelled and unlabelled data to accelerate the discovery of MOFs with high affinity for PFOA adsorption. Rather than supplanting experimental methodologies, this ML-driven approach strategically narrows the search space, ensuring that only the most viable MOFs proceed to experimental validation, thereby optimising resource allocation and expediting the identification of high-performance adsorbents.

To develop a predictive model for PFOA adsorption, we will extract key physicochemical properties of the MOFs using Zeo++ version 0.3,^[11] which provides high-accuracy descriptors. The selected features, including largest cavity diameter (LCD), pore limiting diameter (PLD), accessible surface area (GSA), void fraction (VF), water contact angle (WCA), and density, serve as fundamental predictors in our model. As a first step, a rule-based expert system is applied to filter the MOF database, eliminating structures that do not meet critical threshold values. This filtering process is guided by prior experimental knowledge, ensuring that only MOFs with a pore limiting diameter greater than 10 Å, a water contact angle exceeding 90°, and an accessible surface area above 500 m²/g are retained. This initial reduction in the dataset significantly decreases the number of MOFs requiring

further evaluation, facilitating the subsequent application of ML models. Despite this streamlining step, thousands of MOFs remain viable candidates, necessitating the use of supervised and semi-supervised learning approaches to further refine the selection. The labelled dataset available for model training is highly imbalanced, with only 20 MOFs experimentally confirmed as effective PFOA adsorbents and 80 identified as ineffective. Given this imbalance, several classification models will be explored, including Random Forest, Support Vector Machines, and XGBoost, as these algorithms have demonstrated robustness in handling structured datasets with limited labelled samples.^[12] To mitigate class imbalance, we employ techniques such as Synthetic Minority Over-sampling (SMOTE),^[13] and class weighting adjustments within the loss

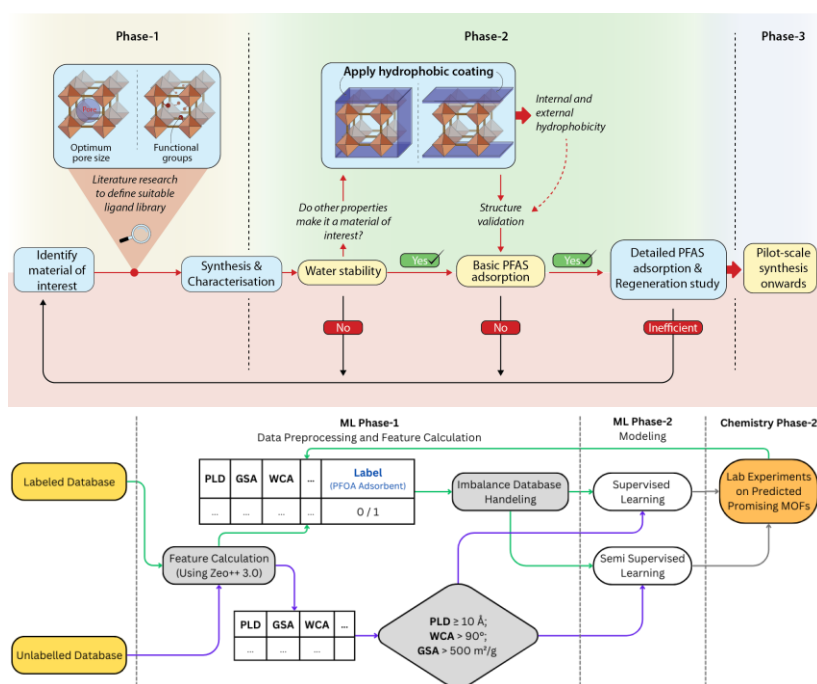


Figure 2: The PFAS adsorbent design paradigm currently underway in the PI's laboratory at UL. Objectives of the phases 1 and 2 are to determine which kind of MOF functionalisation routes (pore vs. surface) elicit optimal surface signatures, just right to exhibit strong affinity to PFAS contaminants (even when present in trace concentrations, such as, ppb) (above). The Machine Learning (ML) design will be conducted by the project Co-Lead at BU. Phase 1 involves data preprocessing, feature calculation, and data filtering, while Phase 3 focuses on ML modelling. Once the most promising MOFs are ranked, they will be sent to the PI's laboratory at UL for the experimental analysis (Chemistry Phase-2). The outcome of the experimental analysis will subsequently be incorporated into the labelled database.

function to ensure that minority class predictions are not systematically underrepresented. These models aim to establish a predictive framework for classifying the remaining MOFs based on their adsorption potential. Since the CSD consist of unlabelled MOFs (their adsorption is unknown), semi-supervised learning plays a crucial role in this project. Semi-supervised approaches, including self-training with pseudo-labelling,^[14] and

[11] R. L. Martin, *et al.*, *J. Chem. Inf. Model.* **2012**, 52, 308. [12] a) T. M. Khoshgoftaar, *et al.*, presented at 19th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2007), 29-31 Oct. **2007**; b) M. Vega Garcia, *et al.*, *Ecol. Inform.* **2020**, 56, 101039; c) M. Khalilia, *et al.*, *BMC Med. Inform. Decis. Mak.* **2011**, 11, 51. [13] R. Blagus, L. Lusa, *BMC Bioinformatics* **2013**, 14, 106. [14] Z. Zhao, *et al.*, *Proc. Int. AAAI Conf.* **2022**, 36, 9208. [15] G. Du, *et al.*, *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 34, 6081.

graph-based label propagation,^[15] will allow the model to iteratively refine its predictions while leveraging the structure of the entire dataset. The self-training process involves training an initial classifier on the labelled data, predicting labels for the unlabelled MOFs, and iteratively retraining the model with high-confidence predictions. Further, graph-based methods such as Label Spreading will be employed to propagate known labels across a network of structurally similar MOFs, enhancing classification accuracy.^[13] Including these semi-supervised techniques ensures that the ML framework can generalise well beyond the small set of experimentally validated MOFs.

If time permits, uncertainty quantification techniques will be integrated into the ML framework to further enhance reliability. Bayesian Neural Networks will be employed to assess prediction confidence, ensuring that the MOFs selected for experimental validation exhibit high certainty in classification outcomes. Furthermore, SHAP (SHapley Additive exPlanations) analysis^[12b] will provide interpretability by identifying the most influential physicochemical features in determining PFOA adsorption potential.^[10] This interpretability is crucial for refining domain knowledge and establishing empirical validation criteria for high-performance MOFs. The anticipated impact of this ML-driven approach is a substantial reduction in the number of MOFs requiring experimental validation, thereby significantly accelerating the identification of high-performance adsorbents for PFOA remediation. By seamlessly integrating domain expertise with advanced data-driven methodologies, this research will establish a scalable and generalisable framework for MOF screening, applicable beyond PFOA removal to a wide spectrum of PFAS contaminants. The primary outcomes of this project will serve as a foundation for securing further funding through the **MSCA Doctoral Networks grant** and the **Joint Translational Call “Water and Health” (Water4All)**. These funding opportunities will enable the expansion and refinement of the proposed methodologies, fostering long-term advancements in ML-assisted materials discovery and water purification technologies.

Project Team: Thanks to the proposed Research Alliance award, our team will rely on strong collaboration ties between two groups: A) the Mukherjee Group led by Dr. Soumya Mukherjee (**SM**), an Assistant Professor of Materials Chemistry from the University of Limerick (UL); B) the Bazargani Group led by Dr. Mosab Bazargani (**MB**), a Lecturer in Data Science from the Bangor University (BU).

A) **SM:** One team member at the Mukherjee Group, Qihao Yin’s PhD project (started from October 2024) is not just limited to PFAS adsorption isotherms guided improvement of structure-function relationships along at least three platforms of water-stable MOFs and organic polymer@MOF hybrids, but to ascertain which pore and surface functionalization routes can lead to top-notch PFAS removal performances and surface wetting (hydrophobic) traits (Figure 2, top). The team is utilising an array of advanced techniques (nuclear magnetic resonance (¹⁹FNMR), liquid chromatography-tandem mass spectrometry (LC-MS/MS), impedance spectroscopy, and high-angle annular dark field-scanning transmission electron microscopy (HAADF-STEM) imaging) to optimise the operating conditions/design of adsorbent columns to maximise PFAS removal efficiencies (vs. the SOTA, GAC). B) **MB:** Dr Mosab Bazargani is a Lecturer in Artificial Intelligence at Bangor University with a background in both academia and industry. He earned his PhD from Queen Mary University of London and has over five years of experience in data science. Before joining Bangor, he worked as a Data Scientist at Tesco and as a Data Analytics contractor for Marsh McLennan, where he gained expertise in applying AI to large-scale real-world challenges. He is an active member of The Operational Research (OR) Society, and contributes to AI standardisation efforts through the British Standards Institute ART/1 committee. He also chaired the UK OR Society’s annual conference, OR66, hosted at Bangor University in September 2024. His research focuses on AI applications in sustainability, with an ongoing project in collaboration with Airbus, using ML and Raman spectroscopy for PFAS identification. This project recently passed Airbus’s Technology Readiness Level 3 (TRL3) review. His team specialises in AI-driven solutions for sustainability and environmental challenges, applying data science methodologies to accelerate material discovery and environmental monitoring.

Goal of the Collaboration: Treating MOF pore/channel-based immobilisation of the C8 PFAS PFOA, to exhibit PFOA-capture from a contaminated freshwater solution, the MOF pore diameter must be higher than the size of PFOA molecule, 1 nm.^[16] Although more than 12,500 three-dimensional (3D) MOFs have been identified

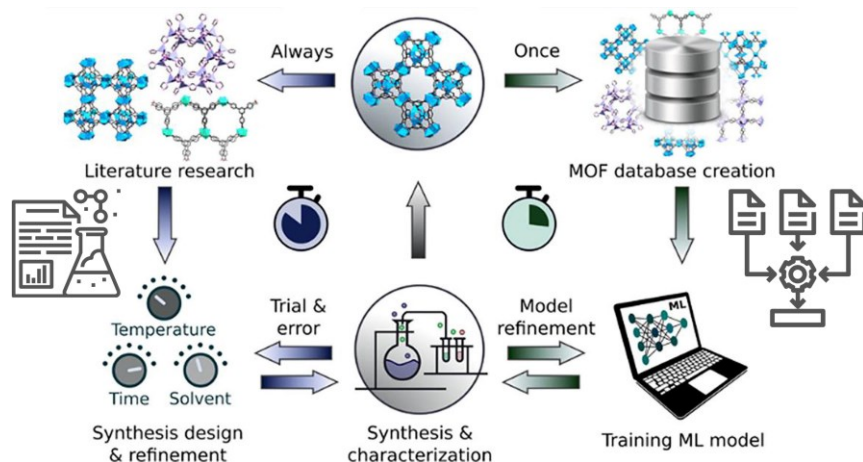


Figure 3: Schematic presenting comparison of traditional (trial-and-error based) (left), and data-driven (ML-based) (right) approach for task-specific MOF sorbents synthesis, such as, PFOA removal from water.

from over 100,000 MOFs in the CSD, which contains more than 1.25 million structures, (1) none have yet met the performance metrics required to challenge the state-of-the-art (SOTA) albeit outperforming activated carbon on kinetics and other key metrics,^[6, 8] and (2) no MOF screening approach currently exists to predictively design MOFs with strong binding affinities for PFAS contaminants, such as the legacy PFAS, PFOA. In this context, predicated upon the

Computation-Ready,

Experimental MOF (CoRE MOF 2019) database (DB)^[17], several of the CoRE MOF 2024 DB entries (40,000+ MOFs),^[18] meet this pore size requirement. Based upon structures obtained by deletion of solvent molecules from the as-synthesised MOF crystal structures, MARCIA will deliver **a previously unreported ML framework to accelerate the discovery of PFOA-adsorbing MOFs. This disruptive ML-driven approach will ensure streamlining of the MOF library to experimental validation, thereby optimising resource allocation and expediting the identification of high-performance adsorbents.**

Objective: Introducing machine learning-driven discovery as a bespoke approach, the vision of this proposal is to predictively screen over 40,000 reported metal-organic adsorbents from a database of crystal structures and subsequently experimentally evaluate the lead sorbent candidates for their PFAS removal performances. This dual computational- experimental strategy to predictively design, synthesise, and critically interrogate MOFs that offer **PFAS-selective strong binding sites** for the remediation of contaminated freshwater will form a cornerstone of the **joint MSCA Doctoral Networks grant proposal in 2025**. In line with the call's eligibility criteria, our submission will include a third legal entity established in a different EU Member State. In this regard, Prof. Dr.-Ing. Jörg E. Drewes, Chair Professor of Urban Water Systems Engineering at the Technical University of Munich, Germany, has already formally agreed to act as a research partner. A 2024 publication by Prof. Drewes's group and the UL PI's group (featured in several popular press releases) highlights the well-established nature of this proposed consortium of PIs in relation to the **MSCA Doctoral Networks grant proposal, call reference: HORIZON-MSCA2025-DN-01-01** (deadline: November 2025). On top of this, the success of the Alliance award will strategically strengthen this tripartite consortium's (UK-IE-DE) chances of securing a joint grant award through the **Horizon Europe partnership-based translational call, Water4All**, specifically to the **2025 JOINT TRANSNATIONAL CALL "Water and Health"**. The applicants to this Horizon EU grant have already received formal (written) support from two of the PI's industrial partners on these proposed tripartite grant submissions, namely from Uisce Éireann (IE) and Suiki TDK (DE). This collaboration aligns directly with overarching themes of two of the 17 UN SDGs, specifically *SDG 3: Good Health and Well-being*, and *SDG 6: Clean Water and Sanitation*.

[16] F. Xiao, et al., *J. Colloid Interface Sci.* **2012**, 368, 505. [17] Y. G. Chung, et al., *J. Chem. Eng. Data* **2019**, 64, 5985. [18] G. B. Zhao, L; Chheda, S; Huang, J; Kim, H; Liu K, et al., **2024**, ChemRxiv®, DOI: 10.26434/chemrxiv-2024-nvmnr.

Proposed Budget

Proposed Budget (Ireland)



Research Alliance Award Budget template (ROI)

Research Ireland Project Lead Prof. Soumya Mukherjee

WIN Project Co-Lead Prof. Mosab Bazargani

Anticipated start date 01.04.2025 Duration 12 months

Cost items	Amount (€)	Total (€)
Materials and Consumables	28,500	28,500
Travel	1,500	1,500
Total (Direct)	30,000	30,000

MATERIALS AND CONSUMABLES

Description	Amount (€)	Total (€)
Material 1	6,200.00	6,200
Material 2	9,000	9,000
Material 3	860	860
Material 4	840	840
Consumable 1	8,000	8,000
Consumable 2	3,600	3,600
	28,500	28,500

TRAVEL & SUBSISTENCE

Description	Amount (€)	Total (€)
Travel cost between PI labs	1,000	1,000
Conference travel	500	500
Total	1,500	1,500

Proposed Budget (Wales)



Research Ireland / WIN - Research Alliance Award 2024

Breakdown of Expenditure - WIN Funding

Item /Activity	Details	Cost £
Staff	PDRA	£15,332
Staff	MB Teaching buyout	£7,954
T&S	Conference	£840
T&S	Research Collaboration Visits	£1,050
Total		£25,176
Overhead 25%		£6,294
Total		£31,470.33

Budget Justification

Ireland:

A) Justifications for all the listed “Materials and Consumables” items are provided below and described in detail subsequently.

Table 1.

Materials and Consumables	Amount (€)
Material 1: Agilent InfinityLab PFC-free HPLC Conversion Kit (accessory) x 2 Pcs	6,200
Material 2: Chemicals (including solvents) for the lead MOF sorbents syntheses	9,000
Material 3: ZORBAX RRHD StableBond C18 Column, 2.1 x 100 mm, 1.8 μ m, 1200 bar; Agilent Part Number: 858700-902 (to couple with the existing Agilent QToF)	860
Material 4: ZORBAX RRHD Eclipse Plus C18 Column, 3.0 x 50 mm, 1.8 μ m, 1200 bar; Agilent Part Number: 959757-302 (to couple with the existing Agilent QToF)	840
Consumable 1: Instrument access charges (characterisation & PFAS quantification) at the University of Limerick (PI's institutional facility that needs to be accessed at ca. 10-40 € / hour access rate)	8,000
Consumable 2: ChemDraw Professional Perpetual licence for two academic users	3,600

Description:

Material 1) For the PFAS removal quantification experiments, the University of Limerick (UL) hosts a top-of-the-line Agilent 6530 Q-TOF mass spectrometer. The proposed procurement of the **Agilent InfinityLab PFC-free HPLC Conversion Kit (accessory)** (two units, one as a potential backup) is intended to replace the existing piping of the Q-TOF with non-PFAS piping material. This is in alignment with the open-access resources available on the Agilent webpage:

<https://www.agilent.com/cs/library/technicaloverviews/public/technicaloverview-pfas-background-reduction-pfc-free-hplc-conversion-kit-5994-2291en-agilent.pdf>. Thanks to this procurement, the accuracy and benchmarking of the analytical measurements will be improved substantially.

Material 2) Chemicals (including solvents) for the lead MOF sorbents syntheses: €9,000 over a year. This project will need specific building blocks to be procured. These will be used for synthesising the data discovery-guided select few top-performing MOFs. Solvents and membrane filters needed for filtration after each PFAS saturation step (for batch adsorption isotherm measurements) have been budgeted herein.

Materials 3 and 4) Stable performance up to 1,200 bar, enabling fast, high-resolution separations of the most complex PFAS contaminants, can only be achieved with the appropriate stationary phase columns suited for C8-PFAS sorbates, such as C8-PFOS and C8-PFOA: for both their detection and removal. Peak shapes across a pH range of 2 to 9 will be recorded, using one of these two C18 ZORBAX fixed-bed columns: A) RRHD StableBond C18 Column, 2.1 x 100 mm, 1.8 μ m, 1200 bar (Agilent Part Number: 858700-902); B) ZORBAX RRHD Eclipse Plus C18 Column, 3.0 x 50 mm, 1.8 μ m, 1200 bar (Agilent Part Number: 959757-302). These columns are to be coupled with the existing Agilent 6530 Q-TOF mass spectrometer.

Consumable 1) Materials characterisation and PFAS quantification studies will be conducted at the Bernal Institute, and the Materials & Surface Science Institute (MSSI). **Instrument access charges** have been budgeted here, in alignment with the hourly access rates of several instruments hosted at UL (<https://ppms.eu/ul/login/?pf=2>, and <https://www.ul.ie/bernal/equipment>) (please see the **Table 2** below):

Table 2:

Instrument	rate/h (in €)	Instrument	rate/h (in €)

Single Crystal XRD	10	FIB-SEM	30
PXRD	20	Ntegra Nanoscope	15
TGA	10	JEOL TEM	30
ATR-FTIR	10	XPS	30
SEM	25	EDS	25
ICP-AES	25	AFM	20
HPLC-MS/MS (Single Quad)	40	LC-MS/Q-TOF	40

Consumable 2) ChemDraw Professional Perpetual academic licence for the UL-based PI and the UL PhD student will be procured from Revvity Signals Software sales team. Upon request, the latter shared a quotation (number Q-104832) in mid-2024 to the UL PI, based on which this estimate is provided.

B) Justifications for all the listed “Travel & Subsistence” items are provided below and described in detail subsequently.

TRAVEL & SUBSISTENCE	
Description	Amount (€)
Travel cost between PI labs: Visits between the collaborating labs, viz., UL and BU (including subsistence), are planned. Each of the two PIs will make one visit to the other’s university, during which pre-arranged research seminar presentations will be delivered. Overnight stays have been accounted for in the budgeting of these visits between Limerick, Munster, and Bangor, Gwynedd.	1,000
Conference travel: Visit of the UL PhD student Qihao Yin to the <u>UKPorMat 2026 RSC conference 2026*</u> and <u>MC18*,**</u> meetings to present MARCIA results	500
Total	1,500

*Venue TBD; **MC18 = 18th International Conference on Materials Chemistry (in UK).

As outlined above, all of these are direct costs, as detailed in the budget spreadsheet. Beyond this, and in line with the grant award guidelines, I request the Research Ireland to make an indirect or overhead contribution to UL, amounting to a percentage of the direct costs excluding equipment (noting that no equipment procurement is included).

Wales:

The budget allocation for the Welsh partner in this collaborative project is strategically structured to align with the objectives of the Research Ireland–Wales Innovation Network (RI-WIN) Research Alliance Award. This funding supports cross-border research collaboration while strengthening the competitiveness of the project for two future Horizon Europe grant applications, as outlined in the preceding Synopsis and Importance sections. Given that the Welsh partner's primary role focuses on the computational and Machine Learning aspects of the research, no budget is allocated for laboratory materials. Instead, the funding is primarily directed toward a Postdoctoral Research Associate (PDRA), whose contributions will be essential in conducting preliminary work, curating and processing datasets, and generating the necessary outputs to support subsequent applications for EU Horizon funding and the MSCA Doctoral Network grant.

Furthermore, Bangor University's in-kind contribution includes access to Supercomputing Wales, a high-performance computing infrastructure that will provide the necessary computational resources for machine learning modelling and large-scale data analysis. This access significantly reduces external computational costs, allowing the project to maximise the efficiency of allocated funds.

Teaching Replacement (15% FTE) – £7,954: To enable Dr Mosab Bazargani to dedicate substantial effort to this project, a portion of his teaching responsibilities will be reallocated. This adjustment ensures his active engagement in collaborative research activities, fostering cross-border knowledge exchange and aligning with the Research Alliance Award's objective to strengthen academic partnerships between Ireland and Wales. His involvement will be critical in supervising the Postdoctoral Research Associate (PDRA), guiding the development of AI-driven methodologies for MOF screening, and ensuring seamless integration between computational modelling and experimental validation. Additionally, his time will be dedicated to developing competitive joint grant applications for Horizon Europe and the MSCA Doctoral Network together with the PI of the project, reinforcing the long-term sustainability and impact of this collaboration. This strategic allocation of time and resources will enhance the project's research output, facilitate high-impact publications, and position the consortium for securing substantial follow-up funding.

Postdoctoral Research Associate (PDRA) – £15,332: The appointment of a PDRA is crucial for advancing the project's research objectives and obtaining preliminary machine learning results. This role will provide the necessary foundation for developing competitive joint grant applications to Horizon Europe. The post is limited to a three-month duration, during which a candidate based at Bangor University's Co-Lead Group will undertake key computational tasks essential for the project's success.

The PDRA will focus on developing code to automate Zeo++ 3.0 for high-throughput feature extraction across all MOFs in the CSD. This automation will enable large-scale computation of essential MOF descriptors, leveraging Supercomputing Wales for efficient execution. Following this phase, the PDRA will implement a rule-based filtering approach to refine the dataset, systematically narrowing the search space to a subset of MOFs with the highest potential for PFAS adsorption.

In the final stage, the PDRA will develop a preliminary machine learning model tailored to this problem, integrating both supervised and semi-supervised learning techniques. The most promising computationally identified MOFs will then be forwarded to the PI's lab in UL for experimental validation, ensuring a seamless transition from AI-driven predictions to real-world testing.

Conference Attendance – £840: This project is inherently multidisciplinary, integrating expertise from artificial intelligence, computational chemistry, and materials science to address the pressing challenge of PFAS remediation. Attending a conference relevant to this research domain is essential for staying informed about the latest advancements, engaging with leading experts, and fostering collaborations. Participation in such events will provide valuable insights into emerging methodologies, experimental validation techniques, and industry applications, ensuring that our approach remains at the forefront of scientific innovation. Moreover, presenting our findings at a reputable conference will facilitate constructive feedback, enhance the project's

visibility within the academic and industrial communities, and strengthen our position for future funding applications.

Research Collaboration Visits – £1,050: This budget allocation supports in-person collaboration with the PI at the University of Limerick, facilitating direct engagement between the Welsh partner and the PDRA. These visits will be instrumental in aligning the computational and experimental aspects of the project, ensuring that AI-driven predictions are effectively integrated into material validation efforts. Additionally, part of this fund will be allocated for meetings with Prof. Dr.-Ing. Jörg E. Drewes, Chair Professor of Urban Water Systems Engineering at the Technical University of Munich, Germany. His expertise in water treatment technologies is highly relevant to this project, and collaboration with him will be pivotal in developing a Horizon Europe funding proposal.

This budget also provides flexibility for supporting Prof. Drewes's visit to Bangor University, should in-person discussions on the grant proposal or project coordination be required. Such strategic in-person engagements will enhance knowledge exchange, strengthen collaborative ties, and improve the competitiveness of future EU funding applications.

Indirect Costs – £6,294: These funds cover essential administrative and operational expenses, ensuring the project's smooth execution. This allocation is consistent with the award's guidelines on funding distribution.

This budget is structured to support the Welsh partner's active involvement in the project, fostering a robust collaboration that aligns with the strategic aims of both governments as outlined in the Research Alliance Award.