Keynote Speaker (Plenary 1): Dr Demner-Fushman, MD

Title: Getting AI Generated Results into Decision Support Workflows: Research, Clinical and Policy Perspectives

From both research and clinical perspectives, Dr Demner-Fushman will share insights drawn from her decade-long experience utilising a Clinical Decision Support (CDS) tab within the National Institute of Health (NIH) Electronic Health Record (HER) system. Additionally, Dr Demner-Fushman will present recent research on the application of large language models (LLMs) for predicting patient outcomes and generating progress notes. Dr Demner-Fushman will provide an overview of community efforts on regulating LLM technical cards and reporting LLM results, as well as informal discussions with FDA on regulating LLM as medical device.

Dr Dina Demner-Fushman, MD is an Investigator at the National Library of Medicine, NIH, HHS. She leads research in the areas of Text and Image Processing for Clinical Decision Support and Education. The outgrowths of these projects are the evidence-based decision support system used at the NIH Clinical Center from 2009 to 2020, an image retrieval engine, Open-i, launched in 2012, and an automatic question answering service CHiQA launched in 2018. Dr. Demner-Fushman earned her doctor of medicine degree from Kazan State Medical Institute in 1980, and clinical research Doctorate (PhD) in Medical Science degree from Moscow Medical and Stomatological Institute in 1989. She earned her MS and PhD in Computer Science from the University of Maryland, College Park in 2003 and 2006, respectively. She authored more than 300 articles and book chapters in the fields of information retrieval, natural language processing, and biomedical and clinical informatics. Dr. Demner-Fushman has received sixteen staff recognition and special act NLM awards since 2002. She is a recipient of the 2012, 2022, and 2023 NIH Award of Merit, a 2013 NLM Regents Award for Scholarship or Technical Achievement and a 2014 NIH Office of the Director Honor Award.

Invited Speaker: Dr John Vardakis

Title: Exploring the Frontier: The Future of GenAI Investments and Collaborations in Industry and Academia"

The rapid advancement of Generative Artificial Intelligence (GenAI) marks a pivotal juncture in innovation, particularly within data-driven healthcare. This presentation delves into the multifaceted landscape of GenAI investments and the indispensable role of collaboration between industry and academia in propelling groundbreaking innovations.

The talk provides an overview of GenAI investment trends, highlighting burgeoning areas of growth and potential. Pivotal factors influencing investment decisions, including technological preparedness, regulatory frameworks, and market opportunities, are also discussed. We touch upon successful case studies, spotlighting strategic investments that have expedited the translation of GenAI research into tangible healthcare solutions, underscoring the significance of targeted funding and support. The discussion underscores the critical importance of robust collaborations between industry and academia. This symbiotic relationship creates an environment that fosters the flourishing of cutting-edge research and facilitates its seamless integration into commercial applications. Additionally, the role of

intellectual property is emphasised as a key driver and protector of innovation, ensuring the foundation for sustained advancements in GenAI technologies.

The FinTech industry is an active participant in this revolution, leveraging distinctive financial models and industry insights to facilitate the funded development and implementation of GenAI technologies, poised to revolutionise healthcare delivery, diagnostics, and patient care.

Aligning investment strategies with collaborative endeavours holds the key to unlocking opportunities for innovation, which will ultimately culminate in improved outcomes and the transformation of the healthcare landscape, and beyond

John Vardakis, PhD has a diverse work experience, primarily in the field of scientific research. John is currently serving as the Head of Science Card Investment Group at Science Card, where they are involved in innovative banking. Previously, they worked as a Research Scientist at the University of Glasgow, where they focused on simulation-driven microelectromechanical systems design. John also worked as a Research Scientist at INSERM, investigating tumor mechanics and vascular fractality quantification. At the Center for Computational Imaging & Simulation Technologies in Biomedicine (CISTIB), they conducted research in cardiovascular and cerebrovascular biomechanics. Their work at UCL involved computational modeling of dementia and image-guided neurosurgery for epilepsy. John completed their DPhil in Biomedical Engineering at the University of Oxford, where their research focused on fluid transport in the brain. John also gained experience in the fields of vascular diagnosis and market research at the University of Heidelberg and P2i, respectively. Additionally, they worked as a Mechanical Engineer at Enactus and completed a placement at King's College London's Centre for Robotics Research

---

Keynote Speaker (Plenary 2): Dr Christopher Kelly

Title: Generative AI in healthcare

Dr Kelly will explore the evolution of AI in healthcare, from its early days to the latest advances in generative AI. Recent work in large language models and multimodal AI has unlocked many new opportunities, and we'll explore their potential to impact healthcare. Dr Kelly will conclude by examining the future of AI in health, including the challenges and opportunities that lie ahead.

Dr. Christopher Kelly is a clinician scientist working in Google's artificial intelligence research team focused on healthcare. Prior to joining Google, he completed a PhD in medical imaging at King’s College London. As a clinician, he created an award-winning smartphone app called NeoMate, used worldwide to support doctors and nurses looking after sick newborn infants. He continues to practise as a paediatrician on the neonatal intensive care unit at St Thomas’ Hospital, London. Alongside medicine, Chris previously founded and sold two internet companies and is an active angel investor in technology startups.

STUDENTS

Student: Gregory Kell
Title (Oral Presentation): RealMedQA: A pilot biomedical question answering dataset containing realistic clinical questions
Abstract: Question answering systems have the potential to improve the quality of clinical care by providing health professionals with the latest and most relevant evidence. Despite progress in recent years, biomedical question answering systems have still not been widely adopted. We present a method to create biomedical question answering datasets with realistic clinical questions and actionable answers. We show that the LLM is more cost-effective at generating questions even when accounting for quality. We also show that the question answering task is more difficult for several systems on our dataset than on BioASQ.

Student: Sabaoon Zeb
Title (Oral Presentation): Clinical decision support management of thoracic malignancy through multi-omics data science
Abstract: Thoracic cancer, encompassing various intra-thoracic tumours, like lung, oesophageal, mesothelioma, tracheal, and mediastinal regions, presents a growing health challenge, particularly non-small cell lung cancer (NSCLC) and lung squamous cell carcinoma (LUSC). Hence, urgent enhancements in early detection and prognosis are vital to mitigate this escalating global health burden. Currently, by employing both statistical and machine learning methods, we have constructed two robust analytical pipelines for scRNA-seq NSCLC data analysis (pipeline-I), and multi-omics data integration analysis (pipeline-II) for TCGA-LUSC bulk RNA-seq and methylation data. Pipeline-I employed methods for cell-type identification and marker detection, revealing highly variable genes and a diverse tumour microenvironment. Pipeline-II integrated methylation and RNA-seq data, highlighting DNA methylation's significant role in LUSC. We have currently applied access to relevant data sources for further thorough exploration into thoracic cancers, to unravel more intricate patterns and identify novel insights.

Student: Robert Cobb
Title (Oral Presentation): Improved Classification Learning from Highly Imbalanced Multi-Label Datasets of Inflamed Joints in Maraciclatide Imaging of Arthritic Patients by Natural Image and Diffusion Model Augmentation
Abstract: Gamma camera imaging of the novel radiopharmaceutical 99mTc-maraciclatide can be used to detect inflammation in patients with rheumatoid arthritis. Due to the novelty of this clinical imaging application, data are scarce with only one dataset composed of 48 patients available for development of classification models. In this work we classify inflammation in individual joints in the hands of patients using only this small dataset. Our methodology combines diffusion models to augment the available training data for this classification task from an otherwise small and imbalanced dataset. We also explore the use of augmenting with a publicly available natural image dataset in combination with a diffusion model. We use a DenseNet model to classify the inflammation of individual joints in the hand. Our results show that compared to non-augmented baseline classification accuracy, sensitivity, and specificity metrics of 0.79 ± 0.05, 0.50 ± 0.04, and 0.85 ± 0.05 respectively our method improves model performance for these metrics to 0.91 ± 0.02, 0.79 ± 0.11, 0.93 ± 0.02 respectively, improving the sensitivity by 58%. When we use an ensemble model and combine natural image augmentation with m* augmentation we see performance increase to 0.92 ± 0.02, 0.80 ± 0.09, 0.95 ± 0.02 for accuracy, sensitivity, and specificity respectively.

Student: Davide Ferrari
Title (Demo): Navigating multi-objective training for Symbolic Regression to harness the potential of evolutionary algorithms for Human-in-the-loop Machine Learning
Abstract: Multi-objective Symbolic Regression (MOSR) is proving to be an exceptional alternative to off-the-shelf machine learning algorithms for high-stakes domain and difficult predictive challenges. Particularly in datasets with high class imbalance, it consistently proves to outperform state-of-the-art alternatives. Moreover, it provides the computing framework supporting the input of human knowledge and domain specific expertise to profoundly change for better how traditional ML models train and behave. In this demonstration it will be shown how to leverage a novel open-source implementation of MOSR to improve predictive models with non-standard objective functions.

Student: Dina Farran
Presentation (Oral Presentation): Stroke prevention in patients with atrial fibrillation and co-morbid physical and mental health problems
Abstract: This presentation will look at how to improve patient safety by screening for stroke risk among people with Atrial Fibrillation (AF). Methods: (a) Conducted a systematic review on prevalence, management, and outcomes of AF in people with Serious Mental Illnesses (SMI). (b) Evaluated oral anticoagulation (OAC) prescription trends in people with AF at KCH. (c) Identified the recorded rates of OAC prescription among people with AF in SLaM. (d) Implemented an electronic clinical decision support system (eCDSS) and assessed its acceptability by clinicians. Results: (a) AF patients with SMI were less likely to receive OAC than the general population. SMI was not associated with an increased risk of stroke or major bleeding. (b) Among AF patients having a high stroke risk, those with co-morbid SMI were less likely to be prescribed any OAC. (c) Among AF patients with co-morbid SMI, warfarin was less likely to be prescribed to those having substance dependency, serious self-injury, hallucinations or delusions and activities of daily living impairment. (d) The eCDSS prompts clinicians to clinically assess stroke and bleeding risks and refer patients at high risk of stroke to OAC clinics. Conclusions: This study presents a unique opportunity to prompt best practice for AF patients.

Student: Leo Xinyue Zhang
Title (Oral Presentation): Rethinking Text Generation Evaluation: A Unified Evaluation Theory for Reflective and Open-Ended Generation Tasks
Abstract: Natural language generation (NLG) has progressed significantly in the last decade. There are two types of generation tasks. One is reflective generation task (RGT), where the model output is a reflection of the information of the input such as machine translation. On the other hand, Open-ended generation task (OGT) requires to generate new information that does not exist in the input such as report generation. However, very different approaches are used in evaluating these two tasks. Reference-based measurements are used in RGT whereas in OGT evaluation, due to the challenge of coming up with references, there are few studies available and most of them are on distribution level evaluations. In this work, a theoretical view of unified generative text evaluation is proposed to bridge the gap. Based on this theoretical framework, the problem of the most commonly used pre-written reference-based measurement is pointed out in RGT evaluation and an automated language quality evaluation for OGT is attempted

Student: Mary Chinwe Abichi
Title (Oral Presentation): Organisation level Outcomes and inequalities for those with Long-Term conditions accessing NHS Talking Therapies
Abstract: Approximately fifteen million people in England live with one or more medical physical long-term conditions (LTCs). Of those with LTC, thirty percent have a comorbid mental health condition such as anxiety and depression. Compounding evidence elucidates mental healthcare inequalities for those with LTCs compared to those without an LTC, as they tend to have worse treatment outcomes. A recent area of interest is the organisational impact on the outcome of those with LTC versus those without. NHS Talking Therapies delivers care using the 'stepped-care' model; however, there is no robust analysis investigating the types of care pathways utilized in a real-world context, specifically for those with LTCs. We used Process Mining Algorithms and Clustering algorithms to extract themes from large complex pathways for the 15 sites in our study. Our findings show variance in outcomes between the different services and therapeutic inequality within services for those with LTCs compared to those without. Our study highlights the importance of barriers and inequalities in healthcare that need to be addressed, especially for those with LTCs, and different forms of therapeutic intervention delivery may be required rather than the more generalized stepped-care model recommended by the National Institute for Health and Care Excellence (NICE).

Student: Miquel Serna Pascual
Title: SPARKS vital sign analysis platform
Abstract (Demo): High-fidelity biomedical waveform time-series data such as ECG, pulse oximetry, or resting respiratory pattern recordings are currently being recorded on a regular basis in clinics and physiological laboratories. At present, this data is used in a limited way when making diagnoses and research assessments. Usually, single-point measurements, such as the maxima and minima, serve as the basis for those decisions, disregarding most of the information contained in these signals. Despite substantial research highlighting the relevance of waveform morphology and variability, these more nuanced features are not conventionally utilised. One reason for this is the complexity of their assessment and measurement. The Symmetric Projection Attractor Reconstruction (SPAR) was developed to facilitate the qualitative evaluation and quantitative measurement of these features. At the start of my PhD project, it was recognised that SPAR attractor generation and subsequent analysis of its metrics was a time-consuming process, prone to error and necessitating of programming and data science expertise. This was deemed impractical for clinicians, physiologists or biomedical scientists, representing a barrier to uptake of the SPAR method. The work presented in this demo aims to bring together this novel mathematical analysis with biomedical research by engineering a SPAR graphical user interface able to simplify SPAR attractor generation, metric computation and its subsequent analysis and interpretation.

WORKSHOPS

Workshop 1: Regulatory and Quality Management Considerations for Medical Devices
Facilitator: Hannah Leadbetter

In the past few years, the number of software and AI-based apps and health management systems have increased substantially. But are these medical devices, and do the manufacturers know if what they have created is a medical device? The reality is that very few software products have CE or UKCA marking and as such there is little quality control plus significant privacy concerns.

The high standards of quality and safety for medical products are set out in regulation; protecting the health of patients and users by ensuring all products meet established safety and performance requirements. This workshop will introduce regulatory requirements (and how to navigate them) for device manufacturers, and processes for quality management used in the development of clinically usable products. With particular regard to software and AI, we will address the qualification and classification of medical devices. A high-level summary of the conformity assessment process for medical device manufacturers will show the lifecycle approach to device manufacture, and how following a Quality Management System is a key function of this.

Additionally in an ever-changing landscape where innovation must live in balance with regulatory conformity, the presentation will discuss regulations across the UK and EU as they stand, plus insight on what we can expect from the future regulations in the UK in general, and with regard to software as a medical device, UKCA marking and Approved Bodies. Recently, the Medicines and Healthcare products Regulatory Agency (MHRA) have announced other access pathways for devices to be placed on the Great Britain market, and this workshop will address this. In all, the presentation will provide an overview of how creating SaMD and AIaMD (software as a medical device / AI as a medical device) that meets the regulations contributes to patient safety and reduces healthcare burden.

Workshop 2: Code rusts, Models rot, how do we fight that entropy?
Facilitator: Dr Michal Grochmal

Please note: This is a hands-on workshop that requires a laptop and good internet access. Come prepared to code if possible.

You have trained a model and you have thoroughly tested that it generalises quite well, and that most of its metrics are where you'd like it to be. Now what? What's next? The model will not magically appear as a service that a user can interact with. You need to find a way to either deploy the model or pass a good deal of data through the model and make the data available to a user. OK, Mike, which one it is? Deploy the model or pass lots of data through it? We are going to discuss that. Moreover, when the model is "alive" (AI safety pun intended), many complications come into play. How quickly it predicts? Does it predict this quickly on the deployed machine? How much data is passing though? Do we have any accuracy metrics on the prediction? We will try to address a few of these. Metrics during model training are well understood, metrics during prediction are less studied. Finally comes the data. The world does not stay in place, and eventually we are going to retrain the rotten model. Models rot because the world they have modeled has changed and we need to address that in order for a model to be useful in the long term. To pass data around for training purposes but to also allow for experimentation with the model is a task tantamount to the impossible: everyone will have a different idea on how to do it. We will explore a few of the ideas. Most ideas involve some form of automation and orchestration over how we pass the data around. And what happens if your AI is a "modern" AI workflow? By which i mean that

you do not even have a model in hands. You just call the plethora of the GTP APIs as your AI. Then all the things above you go away, right Mike? Wrong. Then all the above becomes harder because you need more metrics to see whether things may be going awry. The workshop requires reasonable knowledge of python programming, in an IDE or editor, not only notebooks. Some knowledge of huggingface's environment will be beneficial, and so will be some awareness of OpenAI's models. We will use several tools to interact with both environments which we will explain. Yet the usage of huggingface and OpenAI will be left mostly to the attendee to understand.