

02 October
2025

Thursday

LECTURE**MEET & EAT**

Light lunch provided

11.00 - 12.00 pm**12.00 - 13.30 pm**

Best practices for doing science with data

ABSTRACT

Data scientists and researchers write a lot of code but often without the practices that make it robust and reproducible. That gap can lead to serious consequences, such as the infamous Reinhart-Rogoff Excel error, which influenced global economic policy due to a simple spreadsheet mistake.

This talk focuses on how to close that gap using reproducible analytical pipelines (RAPs), a practical approach built on tools and techniques like Git, functional programming, dependency management, automated testing, and containerisation software. These practices improve two things every research project needs: reliability and traceability.

Research software engineers (RSEs) have played a key role in bringing these tools into research settings by turning fragile scripts into sustainable code. Now, large language models (LLMs) are accelerating that shift. They help RSEs support more teams and make it easier for researchers to adopt best practices independently by explaining tools, generating templates, and catching common issues.

The goal isn't to turn researchers into software engineers, but to build workflows that are as solid and reusable as the results they produce.

High-quality real-world data are essential for both scientific research and science policy. Data-driven and evidence-based decision-making depend on the availability of accurate, comprehensive, and reliable information. Ensuring that research findings are not only scientifically sound but also relevant and actionable is critical for guiding effective disease control strategies and supporting efficient allocation of healthcare resources. To best support its mission, the National Cancer Registry of Luxembourg (Registre National du Cancer) actively promotes initiatives in data quality, interoperability, and data linkage – and supports this lecture as part of that effort.



SPEAKER

Dr. Bruno Rodrigues

Head of Statistics department

Ministry of Research and Higher Education, Luxembourg


 THE GOVERNMENT
OF THE GRAND DUCHY OF LUXEMBOURG
Ministry of Higher Education and Research

HOST:

 Cancer Epidemiology and Prevention (EPI CAN) Group
Department of Precision Health (DoPH)
Luxembourg Institute of Health (LIH)

RESPONSIBLE SCIENTIST:

 Bruno Lima
Postdoctoral Fellow, Cancer Epidemiology and Prevention
(EPICAN) Group

*Please note that in-person attendance is subject to limited availability and requires prior registration. To secure your spot, kindly send an email to epican@lih.lu

Locations:

Lecture:

 1 A-B rue Thomas Edison, 1445 Strassen
Salle Marie S. Curie & Salle Louis Pasteur

To join the Webinar:

JOIN

 Event number: 2788 108 0667
Event password: yZAZW25P2eQ
