

# Shiny in Production 2024

Schedule and Abstracts



The Catalyst, 3 Science Square, Newcastle Helix, Newcastle, NE4 5TG

2024-10-01

# Schedule

## Day 1 - Wednesday 9th October

Time	Title	Room
12:30-13:30	Registration	Foyer
13:30-15:00	Workshop 1	Gorgan
13:30-15:00	Workshop 2	Faraday
13:30-15:00	Workshop 3	Electra
13:30-15:00	Workshop 4	Team Decision Theatre
15:00	Coffee break	Level 1
15:30-17:00	Workshop 1	Gorgan
15:30-17:00	Workshop 2	Faraday
15:30-17:00	Workshop 3	Electra
15:30-17:00	Workshop 4	Team Decision Theatre
17:00	Drinks reception & buffet	Foyer



## Day 2 - Thursday 10th October

Time	Title	Room
08:00	Morning coffee & pastries	Foyer
09:00	Registration opens	Foyer
09:30	Session 1	TED Theatre
11:00	Morning break	
11:30	Session 2	
12:30	Lunch	
13:30	Session 3	
14:30	Afternoon break	
15:00	Session 4	
16:00	Close	



# Abstracts

## Day 1 - Wednesday 9th October

### Workshop 1: Level up your plots: Tips, tricks and resources for crafting compelling visualisations with R and ggplot2 - Time: 13:30 - 17:00

Data visualisations are a great asset in getting people talking about your findings. From making the patterns in the data easy to see, to making a big visual statement and keeping people talking beyond the end of your presentation, transforming your plots from functional to aesthetically pleasing and visually compelling is about so much more than making things pretty.

In this workshop, we'll explore how we can make the most of colours, different plot types, text, and interactivity to maximise the impact of our visualisations. Here's where we're looking to boost your dataviz confidence:

- crafting intuitive dataviz-friendly colour palettes without compromising on accessibility (or creativity!)
- selecting the right type of dataviz for your data and your story
- making the most of typography to optimise text hierarchy and readability
- using annotations wisely to both help interpretation and declutter the visualisations
- turning your ggplot into an interactive plot for additional data exploration
- packaging up your decisions easy reuse across plots (and projects!) This is intended as a hands-on workshop, so bring along a laptop, a plot you're working on or a research question, and some data. Throughout the workshop, I will highlight free resources for each of these aspects of dataviz development. The aim is for you to leave with a plot that you'd be happy to publish, and with some resources you can continue to build on.

### Workshop 2: Building Responsive Shiny Applications - Time: 13:30 - 17:00

The diverse range of devices used for modern web browsing presents challenges when designing an application that works well for all users. Enter responsive design: the practice of building fluid web pages that “work” on huge 4k and 5k monitors, tiny smartphones and all things in between. This course will look at responsive design principles and best practices for Shiny developers, covering page layout, easy-to-add widgets and some simple CSS tricks for when built-in solutions don't quite cut it.

### Workshop 3: Asynchronous Shiny - Time: 13:30 - 17:00

Imagine you couldn't register to attend “Shiny in Production” if someone else was in the process of registering, and you had to wait until they had finished before you could click to “Buy tickets on EventBrite”. This kind of “blocking” shouldn't happen in modern web applications but is surprisingly common in Shiny applications. It happens because a single R process handles all of the server-side processing for multiple



users—one long-running task can prevent any other task from proceeding, hampering interactivity both between and within user-sessions.

Fortunately, Shiny's support for asynchronous programming can alleviate this problem. In the asynchronous approach, you start tasks running without having to wait for them to complete. But, this requires a change in mindset for many programmers and there are a few concepts to understand before you can take advantage of this approach. So, what are you waiting for? Sign up for this workshop!

#### **Workshop 4: Building Apps for Humans - Time: 13:30 - 17:00**

Frameworks like Shiny and Dash can help those with a scientific or mathematical background communicate their research in a way that's interactive and engaging. But while these tools can make constructing a graphical user interface quicker and easier, there's no guarantee that the end product is going to be optimised for human use.

This workshop is aimed at scientists (and the curious) that are interested in learning some basics of human-computer interaction and gaining an understanding for how science itself can assist with the development of better user interfaces that, in turn, lead to improved user experiences.



## Day 2 - Thursday 10th October

### Keynote: Cara Thompson

*Data Visualisation Consultant, Edinburgh, United Kingdom.*



#### **Data-To-Wow: Leveraging Shiny as a no-code solution for high-end parameterised visualisations**

You've created a prototype visualisation, fine-tuned it so it looks amazing and perfectly on-brand, and turned the plot code into a function so that you can run it again on different data and highlight different aspects of the story. Others on the team have seen how good the outputs look and they want in on the magic! But they don't want to learn R.

This talk will offer a behind-the-scenes look at the process of creating a Shiny App that functions as a black box to get straight from the data to high-end parameterised visualisations. We'll start by looking at creating parameterised plot functions using ggplot, before exploring how to bring the data and parameterisation into Shiny to create a seamless no-code data-to-viz workflow for the users.

#### **About the Speaker**

Cara is a freelance data consultant with an academic background, specialising in dataviz and in “enhanced” reproducible outputs. She lives in Edinburgh, Scotland, and is passionate about maximising the impact of other people's expertise.



## Gareth Burns

*Data Scientist at Exploristics, Belfast, Northern Ireland.*



### **Shiny in Secondary Education: Supplementing traditional learning resources to allow students to explore statistical concepts**

The Statisticians in the Pharmaceutical Industry (PSI) Schools Outreach initiative aims at promoting data literacy and statistical concepts to the next generation of Statisticians and Data Scientists. Volunteers attend secondary schools to present from specialised workshops which are designed to be interactive, engaging and aligned to the national curriculum for different age groups.

The PSI Visualisation Special Interest Group (VIS SIG) created a Shiny application to supplement an existing workshop for Asthma. This workshop aims to introduce the students to analysis of continuous data and make them think about concealing treatment assignment and consider false positive and false negative results. The application allowed electronic data capture the ability to dynamically explore their own data, re-enforcing the statistical concepts and making learning more engaging and accessible.

Each school is different in terms of class size, computer resources and student abilities, therefore the application needed to be flexible to account for this and enable independent set up by a volunteer instructor. User experience and accessibility were fundamental in the design concepts to ensuring the application was appropriate for a classroom environment and data visualisation were at an appropriate level for students.

In this presentation we discuss the range in issues required to get a Shiny application being implemented by a team of volunteers into a classroom a classroom setting. This includes flexible project management for a team of volunteers, use of persistent storage to enable multiple simultaneous users and use of Shiny modules to make code flexible and scalable for future Workshops.



## Tan Ho

*Data Scientist at Zelus Analytics.*



### **A minimum viable Shiny infrastructure for serving 95,000 monthly users**

All too often, Shiny gets criticized for scaling poorly - and yet the DynastyProcess Shiny app crushes those expectations by serving over 95,000 unique users each month!

In this talk, we'll deep dive into the motivations, architecture, and design decisions driving the development of a massively popular app and share takeaways for scaling up your own apps.





## Katy Morgan

*Data Analytics Specialist at Government Internal Audit Agency, Bristol, United Kingdom.*



### **More than just a chat bot: Tailoring the use of Generative AI within Government Internal Audit Agency with user-friendly R shiny applications**

Generative AI offers huge potential for driving creativity by suggesting new ideas and perspectives and can also improve efficiency by rapidly processing and extracting insights from large volumes of text. However, using a chatbot-style tool such as ChatGPT can be overwhelming as users have to work out, through trial and error, which questions and instructions give them the outputs they need. The Government Internal Audit Agency's data analytics team has created two R shiny web applications, each of which simplifies the user's experience of using generative AI by providing a user-friendly interface and implementing a set of standardised prompts. The Risk Engine walks the user through a stepwise process to explore and articulate the potential risks that might impact any given business objective. The Writing Engine enables users to analyse and generate text in several ways, including generating a draft audit report from rough notes, and summarising common themes from a set of audit reports. This presentation will cover the process of developing and deploying the web applications and the challenges we faced along the way, describing how we tailored the appearance and functionality of the apps to best meet user needs.



## Vikki Richardson

*Data Scientist at Audit Scotland, Glasgow, Scotland.*



### **Faster than a Speeding Arrow - R Shiny Optimisation In Practice**

The task of optimising your R Shiny apps for great performance can be challenging. Ensuring your code is efficient, using promises where you can, caching resources, and reducing the number of widgets or reactive variables can all help. But datasets can't be squeezed any more – or can they? By storing larger chunks of data in Arrow format and using the Arrow package for manipulation, we were able to speed up some slower computations by at least one order of magnitude - often more.

This presentation will cover a case study of migrating a financial data auditing system to Arrow data storage. Because of Arrow, we were able to drop from two Connect servers to one, making management very happy with the cost savings - and delighting our users with the new, snappier application.



## Pedro Silva

*Data Scientist at Jumping Rivers, Newcastle upon Tyne, UK.*



### Convincing IT that R packages are safe

R packages are essential tools for data science, but how reliable are they? This talk will explore methods for evaluating the quality and potential risks associated with R packages, focusing on key factors like community engagement, development activity, documentation standards, and code quality.

We'll discuss the toolset used as the standard at Jumping Rivers for validating packages with clients, offering advanced scoring methods, customizable risk thresholds, and the ability to incorporate data from external sources to deliver a more comprehensive assessment of package risks.



# Lightning Talks

## David Carayon

*Statistician at INRAE, Bordeaux, FRANCE.*



### **The SK8 project: A scalable institutional architecture for managing and hosting Shiny applications**

Introducing the SK8 Project (Shiny Kubernetes Service), where data scientists, statisticians and engineers from INRAE, the French national research institute for agriculture, food and environment, have teamed up to create a new solution for managing and hosting Shiny applications.

Shiny has become very popular in our institute, widely used for sharing, showcasing, and democratizing scientific work. However, the enduring challenge of establishing scalable, secure, and sustainable hosting for these apps had yet to be addressed.

So, after realizing that different research labs had each implemented their own local and makeshift solutions, we put on our thinking caps and decided to craft an open-source institutional solution. Our mission? Break down silos, unite the R community at INRAE, and make hosting applications easy for Shiny developers with no IT backgrounds.

The SK8 infrastructure allows to host Shiny code on a GitLab instance opened to all INRAE staff. We've got pipelines (GitLab CI/CD), stability ({{renv}}), containerization with Docker, scalability and seamless deployment in a Kubernetes cluster. All of this is developed, managed, and maintained by the SK8 team using open-source solutions.

Using SK8 is a piece of cake – just toss your application code into a dedicated GitLab project and hit the “play” button.

In this talk, we will be speaking about the project itself, the ecosystem that's making it all happen and how you could replicate this in your own company.



## Jeremy Horne

Founder at Datacove, Brighton, United Kingdom.



### Shiny Policies: Dashboards to Aid British Government Decisions

In collaboration with Natural England, Datacove developed a bespoke Shiny dashboard for informed government decision-making, covering Health and Wellbeing, Nature, and Sustainability (HWNS). This presentation will outline three major topics: project and data management, our approach to customization, and the route taken to enhance usability.

The first phase involved project and data management to establish clear expectations. By engaging with Natural England stakeholders, we ensured that the envisioned product met their specific needs and provided a tangible preview of the dashboard's functionality and design. We connected to government APIs and used R to extract, process, and transform multiple sources of HWNS data, bringing this information into one place for localised decision-making.

In the second phase, we focused on customisation to ensure seamless integration with Natural England's existing webpage. Using the brand guidelines and custom CSS/JavaScript, we ensured that the dashboard had the same look and feel as other products built outside of Shiny. This step was crucial in maintaining a cohesive user experience by complementing their established digital ecosystem. Thus, making it easy to access and increasing the likelihood of use.

In the third phase, we emphasized making the dashboard accessible to all, regardless of data literacy. We implemented user-friendly design principles, pre-calculated dynamic stats, and intuitive navigation. For example, we built interactive charts using libraries such as Leaflet and Highcharts, this ensured that comparisons were clear and easy to dynamically explore. We will demonstrate our tips for easy interactive visualisations.

Throughout the project, we adopted best practices in data interpretation and are looking forward to sharing our insights at Shiny in Production.



## Osheen Macoscar

Data Scientist at Jumping Rivers, Newcastle upon Tyne, United Kingdom.



### Using Google Lighthouse to Analyse Shiny Applications

Your shiny app is slow. What should you do? ....

Google Lighthouse is a tool that assesses the time taken to start any web app (among other things). It provides a report that can be analysed in R. Here, Lighthouse is used to assess a range of different Shiny apps to see if there are any common issues that slow down these apps.

The start-up performance for over one hundred applications from the Shiny Contest 2021 were analysed using Lighthouse. For those apps that were still working, Lighthouse report-generation was automated using R.

Unsurprisingly, simpler apps tended to load quicker than more complicated apps, although several complex apps were found with very good Lighthouse scores. Consequently, I sought to quantify the impact on start-up times of making a shiny app more complex. This was done by creating a simple app and iteratively adding more features (using {plotly} and {DataTable}). Lighthouse was used to reveal the effects of making an app incrementally more complex, with the additional widgets increasing start-up times.

This will demonstrate how to use Google Lighthouse on your own Shiny apps. I will cover the analysis performed on the data obtained from the Lighthouse reports. Finally, I will cover the results from my experiments on the impact of different features on Shiny apps and discuss the areas where Lighthouse may not give a complete picture of the start-up issues in your application.



## Alasdair Morgan

*Public Health Scotland, Edinburgh, United Kingdom.*

### **Monitoring and Improving Posit Workbench Usage Behaviour at Public Health Scotland**

This talk covers user behavior in Posit Workbench at Public Health Scotland, focusing on data collection from daily RStudio Server sessions across the organization.

This data is used to generate a monthly “usage behavior” report aimed at promoting more efficient resource use.



## Astrid Radermacher

*Data Scientist at Jumping Rivers, Newcastle upon Tyne, United Kingdom.*



### **PEAR: Jumping Rivers' package quality evaluation platform for regulated industries**

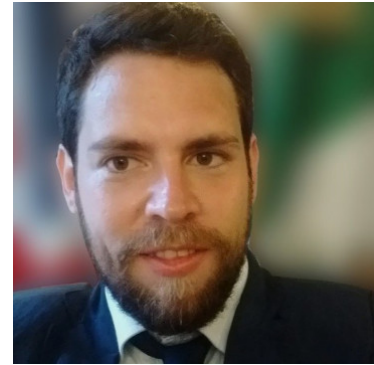
In regulated industries, leveraging new software—especially R-based tools—requires a risk-based approach to ensure compliance with strict regulatory standards. At Jumping Rivers, we are developing an automated pipeline to assess the risks of R packages, providing a safer and more efficient environment for our clients to innovate within these regulated spaces. This lightning talk will serve as a companion to Pedro Silva's presentation, offering further insights into the pipeline we are building to help clients implement risk-aware practices. I will discuss the technical steps we are taking to automate package validation, covering key elements such as risk scoring, automated testing, and report generation. Attendees will leave with a clear understanding of how this solution enhances both safety and productivity in compliance-critical workflows.





## Juan Ramon Vallarta Robledo

Data Scientist at FIND, Geneva, Switzerland.



### Chagas diagnostic algorithms: an online application to estimate cost and effectiveness of diagnostic algorithms for Chagas disease

Chagas disease, caused by the *Trypanosoma cruzi* parasite, is a significant public health concern in Latin America, with an estimated 6-7 million people affected and increasing incidence rates worldwide. Examining the available diagnostic tests and their cost-effectiveness is essential for improving early diagnosis, which is crucial in managing the disease and preventing severe chronic conditions. To address this, FIND, a non-profit organization dedicated to facilitating equitable access to reliable diagnosis, developed Chagaspathways (<https://github.com/finddx/chagaspathway>) to provide guidance for Chagas disease testing. The application is entirely built using Shiny and it incorporates a separate R library (patientpathways - <https://github.com/finddx/patientpathways>), developed by FIND that contains all the analysis algorithms. It is designed to let users select different scenarios and specify parameters about the target population they are analyzing, like prevalence, testing costs, and the type of test used. The results show the recommended testing approach, the expected number of diagnosed cases, the cost per diagnosed case, along with the positive and negative predictive values. A comprehensive outcomes table is included in the results section and users have the option to download the results as an html report, to help them with further dissemination. The Chagaspathways application is designed to be a user-friendly tool for public health professionals, recommending the most economical testing approaches to maximize resources and achieve the best results for patients and healthcare infrastructures. The application is intended to expand its scope to cover additional diseases, aiming to become an essential asset in global health initiatives for disease diagnostic modeling.



## Sponsors



Jumping Rivers has delivered quality data insights from day one. Based in Newcastle and founded in 2016, the company is bringing a fresh approach to the world of data analytics. Our trainers and consultants come with over 100 years combined experience in R, Python, Stan, Scala and other programming languages. We have worked with some of the largest (and smallest) companies in the world.



The National Innovation Centre for Data (NICD) was created in 2019 with £30 million of funding from the government and Newcastle University. Based in the state-of-the-art Helix science district in Newcastle, our mission is to transfer data skills to the UK workforce. Our team of PhD-level data scientists work to ensure that organisations across the country are equipped to reap the benefits of the global data-driven revolution.



The central mission of the R Consortium is to work with and provide support to the R Foundation and to the key organizations developing, maintaining, distributing and using R software through the identification, development and implementation of infrastructure projects.





# ROYAL STATISTICAL SOCIETY

DATA | EVIDENCE | DECISIONS

Founded in 1834, we are one of the world's leading organisations advocating for the importance of statistics and data. We're a professional body for all statisticians and data analysts – wherever they may live.

We have more than 10,000 members in the UK and across the world. As a charity, we advocate for the key role of statistics and data in society, and work to ensure that policy formulation and decision making are informed by evidence for the public good.

