



## Gareth Burns

Senior Data Scientist

### Technical Skills

R 10+ yrs

Data Analysis 10+ yrs

Shiny 6 yrs

RMarkdown 8 yr

Machine Learning 3 yr

Git 7 yr

AWS 5 yr

### Other Skills

AGILE Project Management 5 yrs

Requirments Gathering 10+ yrs

Mentoring & Training 10+ yrs

Product Owner 5 yrs

Data Governance 6 yrs

Scientific & Coporate Communication 10+ yrs

## Summary

An award-winning Doctorate with over 10 years of commercial experience of successfully implementing statistical and data tools across a range of industries that deliver exceptional business value. I have extensive experience the in the full data life cycle from data governance, requirements gathering, prototyping, implementation of a range of data tools to developing communication and maintenance strategies to maximize adoption and user experience.

I have extensive experience in working and leading multi-disciplinary teams across industries and geographical boundaries using AGILE methodologies. My applied knowledge of the full data life cycle coupled with extensive Project Management experience allows me to translate the intricacies of a business problem into a scalable technical solution.

Conveying complex information to a range of stakeholders using a diverse range of methods is my forte - winning awards presenting at Scientific Conferences to making the regional finals of a Famelab, a public engagement in Scientific Communication Competition. I'm experienced in producing data visualisations from creating dashboards to track metrics to bespoke data storytelling for improved engagement on social media.

## Key Achievements

Some of my key achievements demonstrating sustained success across multiple industries:

**Developed a predictive model that is used as a national standard**

Data Scientist

I developed a Near Infrared Reflectance Spectroscopy calibration models to predict the nutritive quality of forage grasses using feature extraction, data pre-processing and advanced regression techniques. The resultant predictive models are the industry standard and used as commercial model for a national variety evaluation schemes to provide 10,000s of results per annum to farmers.

**Work Package Leader: Synergies and Future Development**

H2020 Project: InnoVar (21 partner 5 million funded international project)

I organised an annual Stakeholder Forum to gather requirements and feedback from a diverse range of stakeholders to steer the direction of national policies. I developed the roadmap for the implementation of drones for the collection of field plot data to be implemented across many geographical regions and political jurisdictions.

**Invited Speaker to a national conference in Data Science**

Jumping Rivers: Shiny in Production 2022

Presented on how to develop production grade Shiny applications for Data Science applications. This included writing modular code, how to deploy on cloud-based computing services (AWS), how to carry out automated tested and addressing edge cases.

**Workpackage Leader for 'Software and algorithm development' on an Innovate UK Project**

Expansion of a clinical trial modelling and simulation platform to optimise trial design and facilitate synthetic data generation

Developed two production-grade (Shiny) applications to improve in the extraction and transformation of real world data for use Exploristics flagship simulation software. These applications has a bespoke API that securely communicated with KerusCloud using a single login and reduced transcribing of data.

## Education

2008 - 2012

**Ph.D. Chemometrics**

Queen's University of Belfast

Variation in the nutritive quality of Lolium as assessed by Near infrared reflectance spectroscopy

2004 - 2007

**Environmental Biology B.Sc. Hons**

Queen's University of Belfast

## Contact

📍 Belfast, Northern Ireland

✉ XXXXX

🔗 GABurns

🌐 DrGarethBurns

🐦 GarethBurns4

## Experience

**Senior Data Scientist**

Product Development Team  
Exploristics

08/2021 - Present

Developer of statistical features of Exploristics flagship cloud-based (AWS) Software as a Service (SAAS) clinical trial simulation engine using production grade R code. This involves researching novel statistical design and analysis features and implementing in production grade software (R) using OOP principles. Product Owner and principle developer of the KerusCloud ecosystems - a collection of R packages and Shiny tools to assist in data wrangling and statistical functionality using simulated datasets. I composed a collection of online help pages (astro starlight, quarto, packagedown) for each data tool.

**Senior Data Scientist**

Grassland and Plant Sciences  
Agri-Food and Biosciences Institute

06/2020 - 08/2021

Managed a national statutory testing program (18 staff, c.600 k per annum) to collect and analyse complex biological timeseries and weather data (c. 4 million datapoints collected per annum). Developed a strategy to modernise dated data collection, quality control and storage systems using novel technology and cutting edge techniques (R and SQL). Technical expert on multiple national and international committees covering policy, governance, and emerging technologies.

**Team Lead**

Data Strategy  
Exploristics

04/2019 - 06/2020

Leading a team of 3 Data Scientists to develop innovative tools and solutions (R, python, Shiny) for complex simulated clinical trial data. This included developing roadmaps, gathering requirements and implementing a cloud based (AWS) production grade application (docker and Shiny) to carry out feature extraction and algorithms to estimate distribution parameters for a patient level data sets. This was supplemented with demonstrations to stakeholders and writing user documentation (R Markdown, packagedown).

**Data Scientist**

Marine Fisheries  
Agri-Food and Biosciences Institute

04/2016 - 04/2019

Responsible for the Coordination, collation (SQL) and analysis (R) of fisheries-independent and fisheries dependent data and meta-data from multiple sources (c. 10 million biological data points per annum). Delivery of specialist advice to international technical and scientific committees. Automation of c. 1,000 Data Quality Control and Analysis and reports per annum (R Markdown) ensuring reproducible research. Development of dashboards (Flexdashboard and Shiny) to track key project metrics. Development of a software application (Shiny) to carry out data wrangling, aggregation and allometric modelling of complex biological data.

**PhD Candidate**

School of Biological Sciences  
Queen's University of Belfast

04/2008 - 04/2012

Development of predictive calibration models to assess the nutritive quality of perennial ryegrass. This involved data collection of 15,000 biological samples each with 692 data points from wet laboratory analysis. Model development involved pre-processing, feature selection and validation. The calibration models were commercially used for a national forage evaluation scheme.