

## **Dr. Rob Donald - Curriculum Vitae**

Email: [rob@statsresearch.co.uk](mailto:rob@statsresearch.co.uk), Web: <http://www.statsresearch.co.uk>  
Mob: 07780 650 910

### **Profile – Data Scientist, Systems and Data Analyst**

In my current role I am a senior data scientist with a team adapting a self-monitoring blood glucose system to meet new regulatory demands. My responsibilities are to ensure the timely collection and statistical analysis of experimental data. A major part of my work involves data validity checks and visualisation programs for other scientists in the group. I have written a suite of tools for these tasks using MATLAB, R, and Python, drawing data from an Oracle database.

In my last role I was a senior data scientist working in healthcare informatics. My projects involved the construction and analysis of population level datasets from long-term disease cohorts. This role used all the skills of the emerging discipline of Data Science, requiring a deep and broad understanding of ETL processes accessing data from multiple sources. On a day-to-day basis I used R, SQL and Python.

In a recent role, on a neurosurgical research project, I used R and MATLAB for statistical modelling research and the analysis of clinical trial data. I was heavily involved in the design of the Phase 1 and Phase 2 clinical studies including the required statistical techniques and sample sizes.

The software systems for the project used C#, Java and C++. I designed and supported a C# web application used by research nurses for manual data entry. I was required to design and implement a Java based application producing reports for each patient on the study. For the C++ section of the project I was required to integrate new functions into an existing MFC based C++ application.

I am particularly skilled at quickly analysing the “big picture” and I have extensive experience in the creation and presentation of material describing complex problems. I am skilled at presenting this type of material to a non-specialist audience often in an international setting.

I have recently been awarded a PhD in Statistics by the School of Mathematics and Statistics, University of Glasgow, Scotland. My supervisor was Dr Ludger Evers.

### **Skill Sets**

I have experience of systems analysis, software design and development using the following technologies:

- R Statistical Programming Framework, MATLAB
- Python, C#, Java, C++ (ANSI and MFC), C
- Microsoft SQL Server, Postgresql, MySQL, Oracle
- HTML/CSS, ASP, JSP, PHP
- Unix system administration, PERL and Unix Shell Scripts
- Project documentation using MS Office, MS Visio and LaTeX text processing

## **Education**

Nov 2007 – August 2014

PhD in Statistics, University of Glasgow. My thesis title is “Predicting Hypotensive Episodes in the Traumatic Brain Injury Domain”. The thesis compares the use of Bayesian neural networks and logistic regression models, which were built to predict hypotensive (low blood pressure) events in a hospital ICU setting.

Jan 2007 – Oct 2007

Open University, 2 x 30 Credit Modules.

M249 Practical Modern Statistics, M346 Linear Statistical Modelling

Jan 2006 – Oct 2006

S205 The Molecular World, SXR 205 Exploring The Molecular World

Open University, 1 x 60 Credit Module. 1x summer school

Sep 1980 – Jun 1984

University of Strathclyde, BSc Electrical and Electronic Engineering (Hons 2:1)

## **Recent Project History**

**Sept 2012 – Present**

**Project Title:** Self-monitoring blood glucose systems

(<http://www.lifescan-scotland.co.uk>)

**Roles:** Data Scientist, Statistical Analyst using MATLAB and R, Developer using Python, SQL, Oracle

**Technologies:** MacBook Pro and Windows 7 platforms. Cornerstone SVN

Our small team is responsible for processing data collected by a custom made MATLAB based laboratory information system.

This role involves checking and cleaning data before assembling it into a standard format. This standard format data is then used for investigations and visualizations. If standard techniques suggest further analysis is required, I provide custom views of the data and assist other scientists in the group with statistical techniques.

Because of my background, I am frequently drafted on to ad hoc investigations where I provide advice regarding the design of experiments and help with the assessment of actually acquiring the data. I also help with the presentation of the results and suggestions for follow on projects. As I am used to presenting complex information to upper management, I am often asked to attend meetings as the subject matter expert on the collection and analysis of complex data.

I am also involved in some longer term projects where I support the architectural design of systems that are being considered as the business moves into an environment with a large number of complex data sources. This role has required that I meet with internal project teams and outside suppliers.

### **May 2011 – Sept 2012**

**Project Title:** Healthcare Informatics Product (<http://www.aridhia.com>)

**Roles:** Data Scientist, Statistical Analyst using R, Developer using SQL, Python, Windows 2008 Server admin

**Technologies:** MacBook Pro, Windows 7 and Windows 2008 Server platforms. Tortoise SVN

The team was responsible for the information quality assessment and analysis of a daily feed of data from health board level data collection systems in Scotland and Kuwait. These systems provided demographic and healthcare encounter data for patients along with laboratory requests and results. This work was carried out in a secure VPN based environment and full information governance protocols were in place.

I was part of a newly formed data science team as the company recognized the importance of ‘big data’ in new healthcare projects. The projects were mainly concerned with long term disease conditions especially diabetes and cancer.

Our remit was to investigate new techniques and keep up-to-date with emerging technologies but with the constraint of making projects work in a commercial environment.

As well as my background in software development and statistical modeling, I brought an element of domain knowledge to the group from my previous research project with neurosurgical teams. I introduced the concept of reproducible research and moved the group away from the use of ad hoc methods that tended to be used to provide results quickly but proved difficult to build on when further analysis was required.

### **Jan 2008 – Dec 2010**

**Project Title:** AvertIT (European medical research project, <http://www.avert-it.org>)

**Roles:** Systems programmer, Statistical Analyst using R, MATLAB, Developer using Java, C++ and C#, Windows 2008 Server admin

**Technologies:** NetBeans 6.9 with SubVersion, MS Visual Studio Tools, ASP.NET web application

Worked on EU funded medical research project as the statistical analyst and systems programmer.

My project team worked with neurosurgical ICU clinicians in six teaching hospitals around Europe and data security specialists from the National eScience Centre at the University of Glasgow.

Techniques used: neural network research using MATLAB, R, C and C++; Java programming using the NetBeans environment with SubVersion source code control;

Windows 2008 Server admin; C# Data processing using MS Visual Studio; ASP.NET Web based data collection application.

I presented progress reports to the neurosurgeons and their clinical teams in the six countries involved in the study and also presented the team's work to external research groups.

I contributed to full technical progress reports required for yearly formal EU evaluation and have attended the actual review meetings in Brussels and Glasgow. I also analysed the trial data and presented interim results at meetings held in each of the participating hospitals around Europe and at an international scientific conference (ICP 2010, Tübingen, Germany).

The project produced encouraging research results, which were then taken to clinical trial stage. A Phase 1 clinical study with 30 patients has been completed and the project has now progressed to a Phase 2 clinical study, aiming to recruit another 46 patients.

#### **Nov 2004 – Dec 2007**

**Project Title:** Data Collection Product, (C3 Global, SME providing enterprise performance solutions, <http://www.c3global.com>)

**Roles:** Product Architect, Java Developer, JSP Developer

**Technologies:** Eclipse with SubVersion, Apache Tomcat Web Server

Worked on a Linux based self-contained data acquisition system for international supermarket refrigeration supplier (Danfoss, <http://www.danfoss.com>).

Starting from a limited client market requirement, I produced the concept design and helped with the project proposal. I developed the secure communications layer used to acquire data from the embedded systems on site. I supported the system integration tests and handover to client. After a successful delivery, I carried out second line support investigations as required.

#### **Jan 2001 – Oct 2004**

**Project Title:** Support Centre Alarm Handling Application, (C3 Global, SME providing enterprise performance solutions, <http://www.c3global.com>)

**Roles:** Product Architect, Java Developer, JSP Developer

**Technologies:** Eclipse with SubVersion, Apache Tomcat Web Server

Worked on an enterprise level data acquisition and service centre application for a major supermarket refrigeration supplier (Danfoss, <http://www.danfoss.com>).

Designed the initial architecture of the system from client's limited specification and led the small team that produced the pilot system for the client. I was the technical lead in the larger team that took the project to production level across two data centres (USA/Europe) serving the global reach of the system. The system consists of a heterogeneous server environment running Linux and Microsoft Server systems.

## Full Employment Summary

Dates	Company	Roles
Sept 2012 – Present	LifeScan Scotland Ltd	Senior Data Scientist, Statistical Analysis and Software Programming, Data cleaning and validation. Data visualization.
May 2011 – Sept 2012	Aridhia Informatics Ltd	Senior Data Scientist, Statistical Analysis and Software Programming, ETL (extract, transform, load) systems programming
Jan 2011 - Present	Stats Research Ltd	Consultant Data Scientist, Statistical Analysis and Software Programming
Feb 1997 – Dec 2010	C3 Global Ltd (SME providing enterprise performance solutions)	Statistical analyst, product architect, developer (Java, C#, C++)
Aug 1989 – Jan 1997	Diagnostics and Measuring Systems, (SME providing vibration data analysis solutions)	Engineering Manager. Main company activity was the analysis of vibration signal information
Jun 1986 – Jul 1989	Honeywell Control Systems	Software Design
Jul 1984 – Apr 1986	Smiths Aerospace and Defence	Hardware Design
Sep 1980 – Jun 1984	Full Time Education	University of Strathclyde, BSc Electrical and Electronic Engineering (Hons 2:1)
Sep 1973 – Sep 1980	Denholm Ship Management/Overseas Containers Ltd	Marine Engineer