# 2017/18 Vacation Scholarships

|  |  |
| --- | --- |
| **Job Title:** | CSIRO Undergraduate Vacation Scholarships – **Data61** |
| **Reference No:** | 43691 |
| **Classification:** | CSOF1.1 |
| **Stipend:** | $1462.77 per fortnight (before tax) |
| **Location:** | Please refer to the list of ***Projects*** at the end of this document |
| **Tenure:** | 8 to 12 weeks from November 2017 to February 2018 |
| **Role Purpose:** | The 2017/18 Vacation Scholarship Program is designed to provide students with the opportunity to work on real-world problems in a leading R&D organisation.  Participation in the Vacation Scholarship Program has influenced previous scholarship holders in their choice of further study and future career options. Many have gone on to pursue a PhD in CSIRO or to build a successful research career within CSIRO, a university or industry. |
| **Project Description:** | Please refer to the list of ***Projects*** at the end of this document.  *If you require more information please contact the person listed for the project.* |
| **Eligibility/**  **Pre-Requisites:** | To be eligible to apply you must be an Australian/New Zealand Citizen, Australian Permanent Resident, or an international student who has full work rights for the 8 to 12 weeks duration (does not require visa sponsorship).  Vacation scholarships are for students who:   * are currently enrolled at an Australian university; * have completed at least three years of a full-time undergraduate course (however exceptional second year students may be considered); * have a strong academic record (credit average or higher); and * intend to go on to honours and/or postgraduate study. |
| **How to Apply:** | Please apply online at [www.csiro.au/careers](http://www.csiro.au/careers). **You will be required to:**   1. select your **top 2 preferred research projects** in order of preference; 2. submit a **resume/cover letter** (as one document) which includes:  * the reasons why the research project/s you have selected are of interest to you; and how your previous skills/knowledge and experience meets the project requirements; * an outline of your longer-term career aspirations and detail how this program will help you achieve them; and * using the project numbers listed below, list in order of preference, **all of the projects** you are interested in.  1. upload your **academic results** in the ‘***Requested Information’*** field.   **Referees:** If you would like to include referees (either work or university lecturers/ tutors)in your application, please add their name and contact details into your resume**.**  If you experience difficulties applying online call 1300 984 220 and someone will be able to assist you. Outside business hours please email: [csiro-careers@csiro.au](mailto:csiro-careers@csiro.au). |

**There are 45 projects available in Data61:**

|  |  |  |
| --- | --- | --- |
| **Project No.** | **Location** | **Project Title (see the following pages for more information)** |
| [**Data61 1**](#_Project_Number) | Innovation Central Perth, WA | Innovation Central Internships |
| [**Data61 2**](#_Data61_2) | Kensington, NSW | Modelling and Verification of Wireless Mesh Network Routing Protocols |
| [**Data61 3**](#_Data61_3) | Kensington, NSW | Object-Capability Systems on the seL4 microkernel |
| [**Data61 4**](#_Data61_4) | Eveleigh and Paddington, NSW | Immersive Experiences for Scenario Planning and Decision Making |
| [**Data61 5**](#_Data61_5) | Marsfield, NSW | Secure Storage for IoT enabled Industrial Control Systems (ICS) |
| [**Data61 6**](#_Data61_6) | Kensington, NSW | Secure and Usable seL4 System Development |
| [**Data61 7**](#_Data61_7) | Eveleigh, NSW | Blockchain: Live Observations or Application Development |
| [**Data61 8**](#_Data61_8) | Eveleigh and Paddington, NSW | Immersive Data Interaction Framework |
| [**Data61 9**](#_Data61_9) | Eveleigh, NSW | Software Dependency Analysis |
| [**Data61 10**](#_Data61_10) | Docklands, VIC | Cyber security in a box |
| [**Data61 11**](#_Data61_11) | Clayton, VIC | Connecting Workspace to the Science App Store |
| [**Data61 12**](#_Data61_12) | Sandy Bay, TAS | Analytics for Automation, Trust and Workload assessment toolkit |
| [**Data61 13**](#_Data61_13) | Docklands, VIC | Channel Constraints |
| [**Data61 14**](#_Data61_14) | Sandy Bay, TAS | Collaborative Interaction Design for Virtual Reality and Augmented Reality |
| [**Data61 15**](#_Data61_15) | Eveleigh, NSW or Black Mountain, ACT | Designing behaviourally robust digital market platforms |
| [**Data61 16**](#_Data61_16) | Marsfield, NSW | Domain Adaptation in Social Media |
| [**Data61 17**](#_Data61_17) | Clayton, VIC | Hedging FX risk LSV-style! |
| [**Data61 18**](#_Data61_18) | Docklands, VIC | Reverse logistics to avoid food loss |
| [**Data61 19**](#_Data61_19) | Clayton, VIC | Simulating Heat Transfer in Granular Flows of Variable Shape On Combined CPU /GPU Architectures |
| [**Data61 20**](#_Data61_20) | Docklands, VIC | Virtual Reality Based Simulation Setup and Data Exploration |
| [**Data61 21**](#_Data61_21) | Clayton, VIC | The science of evacuation studied through computational modelling |
| [**Data61 22**](#_Data61_22) | Clayton, VIC | High resolution satellite image analytics for asset classification and disaster management applications |
| [**Data61 23**](#_Data61_23) | Canberra, ACT | Developing an assistive medical technology for better inspection of internal organs through 3D reconstruction from endoscopy videos. |
| [**Data61 24**](#_Data61_24) | Marsfield, NSW | Measurement and modelling of digital TV received signal level height variation using drone and/or balloon |
| [**Data61 25**](#_Data61_25) | Eveleigh, NSW | Evaluating methods to generate differentially private synthetic data |
| [**Data61 26**](#_Data61_26) | Sandy Bay, TAS | Cloud Mining – Easy Machine Learning tools for IoT |
| [**Data61 27**](#_Data61_27) | Sandy Bay, TAS | Bees with Backpacks |
| [**Data61 28**](#_Data61_28) | Sandy Bay, TAS | Choppy Brain Waves : Data Quality Framework for EEG data |
| [**Data61 29**](#_Data61_29) | Marsfield, NSW | Additive Manufacturing for Millimeter-Wave Electromagnetic Applications |
| [**Data61 30**](#_Data61_30) | Pullenvale, QLD | Creating bespoke robot parts |
| [**Data61 31**](#_Data61_31) | Eveleigh, NSW | Traffic Simulation for Congestion Management in Sydney |
| [**Data61 32**](#_Data61_32) | Eveleigh, NSW | Transport data analytics and visualisation |
| [**Data61 33**](#_Data61_33) | Eveleigh, NSW | Visual analytics of machine learning performance |
| [**Data61 34**](#_Data61_34) | Swinburne/Marsfield, NSW | Data Mining to Analyse Performance in Team Projects |
| [**Data61 35**](#_Data61_35) | Eveleigh, NSW | Knowing Customers’ Expectation by Modern Machine Learning Models. |
| [**Data61 36**](#_Data61_36) | Eveleigh, NSW | Machine Learning in Transportation Models |
| [**Data61 37**](#_Data61_37) | Eveleigh, NSW or Canberra, ACT | User Experience Design Research |
| [**Data61 38**](#_Data61_38) | Eveleigh, NSW | Machine Learning on Graphs |
| [**Data61 39**](#_Data61_39) | Eveleigh, NSW | Privacy-preserving machine learning via fully homomorphic encryption |
| [**Data61 40**](#_Data61_40) | Eveleigh, NSW | Smart Infrastructure Systems |
| [**Data61 41**](#_Data61_41) | Eveleigh, NSW | Add Blockchain Ledger for N1 |
| [**Data61 42**](#_Data61_42) | Eveleigh, NSW | Blockchain Meets Confidential Machine Learning |
| [**Data61 43**](#_Data61_43) | Eveleigh, NSW | Adversarial validation for machine learning in production |
| [**Data61 44**](#_Data61_44) | Eveleigh, NSW | Data Visualisation for Networks |
| [**Data61 45**](#_Data61_45) | Eveleigh, NSW | Deep learning with partially homomorphic encryption |

Select the **Project Numbers** above to take you directly to the project details, including relevant fields of study, Project Duties/Tasks and Locations for these projects (which are on the following pages).

Pease read though these and decide **which 2 projects are your preferred choices** as you will need to enter these into your application. If you require more information please contact the person listed for each project.

Note: CSIRO are advertising vacation scholarships by the different business units we have. You can apply for more than one CSIRO business unit, but your application for **Data61** should only refer to Data61 projects, such as Data61 1, Data61 2, etc.

|  |  |
| --- | --- |
| Project Number | **Vacation Scholarships Project Details** |
| **Data61 1** | **Project Title**  Innovation Central Internships  **Project Description**  Blockchain for distributed power systems  This project will assess, in the City of Fremantle, how cities of the future can use big data and blockchain technology to integrate distributed energy and water systems. Highly resilient, low carbon and low cost systems–solar assets, water storage and treatment, and battery storage–will be installed and connected via blockchain to offer the community financial and essential service sustainability while still engaging the private sector.  Atom-by-atom visualisations of geological materials in 3D  Atom probe Microscopy (APM) is a powerful analysis tool for extracting the 3D atomic-scale structure of sample materials. It can be used to visualise and analyse volumes containing up to several hundred million atoms, with each atom identified and positioned in 3D space. Rendering these massive datasets dynamically in three dimensions presents an interesting and rewarding challenge.  This project is looking for a student to develop innovative software tools for rendering data from the newly established Geoscience Atom Probe facility, situated within the John de Laeter Centre. This unique facility is the first of its kind to use APM specifically for geoscience studies, and is providing new insights across many areas of current research.  Although some limited visualisation tools are available, there is scope for significant development to provide a workflow from raw data to high-quality visualisations that can aid in the understanding, education, and promotion of this high-impact research.  SWARM - Secure Wireless Adaptive Resilient Modern motes  The intern would work with Cisco and Woodside teams to:  - Review different wireless standards  - Review different sensors types and applicability to a wide range of LNG measurements  - Using Curtin as a sandpit connect a wide range of sensors to the SWARM network  - Work with Data61 researchers to incorporate CLAP into the deployment  Power Smoothing of Energy Use  Curtin's power usage is very peaky and during summer the peak load will require the construction of a new power distribution station.  The cost of this is multiple tens of millions.  It is believed that by modelling user behaviour using CMX, modelling thermal heating of buildings and predicting future weather it may be possible to pre-cool buildings thus smoothing power use.  CSIRO Undergraduate Vacation Scholarship – Water Optimisation Monitoring  Innovation Central Perth is working with Newton Orchards, located in Manjimup WA, to develop an orchard soil moisture monitoring platform. This will enable the local management to track the application of irrigated water so that the soil moisture levels remain at a level to promote accelerated plant growth and optimal fruit development.  The project will evaluate a range of sensors, probes and instruments, to determine the most effective combination and application, to provide decision support to the orchard management on maintain effective soil moisture and to track the impact on the plan growth.  CSIRO Undergraduate Vacation Scholarship – Synergy Coal Mill Maintenance  Synergy is keen to investigate the collection of data and the creation of machine learning algorithms which can predict the degree of wear and tear on a mill without having to shut the operation down and open it up.  **Project Duties/Tasks**   * Project specific   **Relevant Fields of Study**   * Data science * Blockchain/distributed ledger * Maths and statistics * Energy * Civil and mechanical engineering * Built environment and informatics * Physics   **Location:** Innovation Central Perth, WA  **Contact:** For more details please contact **Chris Nelson** on phone on (08) 9333 6025 or email [chris.nelson@data61.csiro.au](mailto:chris.nelson@data61.csiro.au) |
| Data61 2 | **Project Title**  Modelling and Verification of Wireless Mesh Network Routing Protocols  **Project Description**  Wireless Mesh Networks (WMNs) are a promising technology that is currently being used in a wide range of application areas. Specifications of routing protocols (for WMNs) are usually written in plain English. Often this yields ambiguities, inaccuracies or even contradictions. Moreover, no formal guarantees can be given based on such a description.  The use of Formal Methods such as process algebra avoids these problems, leading not only to a precise description of protocols, but also allowing formal reasoning. The project's work will be in this area; it can include   * Modelling Routing Protocols: So far, we have modelled one of the standard protocols using process algebra, namely AODV. The project's work could include the formalisation of a second standard protocol such as OSLR or HWMP. * Verifying Routing Protocols: Based on a faithful specification that has been given, the work could include the verification of basic properties of the routing protocol: packet delivery for example guarantees that a packet, which is injected into a network, is finally delivered at the destination (if the destination can be reached). * Tool Support: The generation of tools for (semi)automatic reasoning is of high interest. For our work. We have used the interactive theorem prover Isabelle and the model checker Uppaal. So far, the use of these tools includes manual work. The project's work could also aim at more automation: this can be automatic translation software from a process algebra language to a model description for Uppaal; or the development of proof-tactics for Isabelle.   The concrete topic is set according to the applicant's interests and strengths.  **Project Duties/Tasks**   * Model or verify a routing protocol for WMNs, using formal methods * Present results to the group at the end of the project; write project report.   **Relevant Fields of Study**   * Computer Science; the candidate should have experience in formal methods, e.g. process algebra or model checking.   **Location:** Kensington (UNSW), NSW  **Contact:** For more details please contact **Peter Höfner** on phone on +61 2 9490 5861 or email [Peter.Hoefner@data61.csiro.au](mailto:Peter.Hoefner@data61.csiro.au) |
| Data61 3 | **Project Title**  Object-Capability Systems on the seL4 microkernel  **Project Description**  There are two separate but related sub-projects under this overall topic, they can be done by separate students, together they are too much for a single student.  Virtualising the seL4 ABI  Capabilities are opaque object references with implied access rights. A capability system should allow transparent interposition, e.g. for security monitoring.  In a capability-based operating system, this should allow complete virtualisation of the system-call interface. This project is to put seL4, the world's most trustworthy operating system, to the test: Is seL4 truly virtualisable? I.e. can an arbitrary program written to run natively on seL4 run correctly in a virtualised environment where all system calls are intercepted by a security monitor? If so, what is the inherent cost of virtualisation? The simplest proof of virtualisability is to construct a minimal wrapper around a user process that intercepts all capability transfers in and out of the process and replaces them by endpoint capabilities. Those endpoints are invoked by the process as if they represented kernel objects, but instead send a message to the wrapper which then invokes the actual operation on behalf of the kernel. The virtualisation is complete if the wrapper can proxy all such invocations without any knowledge of the operation the client process' intended. Validate this by running the seL4 regression test suite as the client process. This work demonstrates that seL4 indeed benefits from the nice properties of classical object capabilities. A good job can produce publishable results.  Capability-oriented programming interface for seL4  The formally verified seL4 microkernel is arguably the world's most secure operating system. Capability-based access control is a core enabler of security, as it provides fine-grained control over access rights. Existing middleware for seL4 provides a programming model similar to traditional (Posix-like) systems, effectively reverting to ambient authority and thus abandoning most of the security advantages of capabilities. The recently introduced object-capability features in JavaScript, the language widely used for web pages as well as in the embedded space, provides the opportunity for a programming interface to seL4 that avoids the above sacrifice. This project is to develop and evaluate a JavaScript embedding of the seL4 API, using the "embedded" XS JavaScript virtual machine.  **Project Duties/Tasks**   * Familiarise yourself with seL4, both in terms of concepts of the kernel as well as its development environment * Develop a user-level layer that implements the required functionality * Evaluate performance * Present results to the group at the end of the project; write project report.   **Relevant Fields of Study**   * Computer Science; the candidate should have experience in OS.   **Location:** Kensington (UNSW), NSW  **Contact:** For more details please contact **Gernot Heiser** on phone on +61 2 9490 5850 or email [Gernot.Heiser@data61.csiro.au](mailto:Gernot.Heiser@data61.csiro.au) |
| Data61 4 | **Project Title**  Immersive Experiences for Scenario Planning and Decision Making  **Project Description**  This project will explore various futures for rapidly growing digital economy, use of immersive experiences to support decision making and scenario planning. Virtual Reality will be used to showcase plausible futures, and enable wiser choice of action to take. Examples could be: How Sydney CBD will look like in 2050? How traffic will be changed with autonomous vehicles driving around in the city? How traffic could be optimized by using smart network monitoring traffic lights and movements of cars? Future of medicine?  Successful Vacation Student will have opportunity to work closely with researchers in Data61 on core science, and engineers and postdocs in the Expanded Perception and Interaction Centre (EPICentre) on high performance visualisations techniques, to create tangible immersive experiences. It is very unique opportunity for student to be across leading edge Australian R&D centres.  Background about EPICentre:  EPICentre has brought together world-leading leading researchers in visualisation, to build the world’s highest resolution omnispatial interactive and immersive 6.47m diameter 3D 340-degree visualisation environment. The Expanded Perception and Interaction Cylinder (EPICylinder), is a cluster-driven system of 120 million pixels in 3D. Another key visualisation environment in this facility is DomeLab, a 6m diameter 4K fulldome. EPICentre also aims in using mixed reality techniques that would combine existing visualisation systems with the power of augmented reality (e.g. using Microsoft HoloLens’s in the mix). All these systems offer unique opportunities for transformative data-driven research and accelerated scientific discovery.  **Project Duties/Tasks**   * Literature review, design and concept planning. * Execution, and evaluation of the project. * Write scientific report from the experiments and research findings.   **Relevant Fields of Study**   * Computer Science * Design * Engineering   **Location:** Australian Technology Park (Eveleigh) and EPICentre UNSW (Paddington), NSW  **Contact:** For more details please contact **Tomasz Bednarz** on phone on (02) 9490 5760 or email [tomasz.bednarz@data61.csiro.au](mailto:tomasz.bednarz@data61.csiro.au) |
| Data61 5 | **Project Title**  Secure Storage for IoT enabled Industrial Control Systems (ICS)  **Project Description**  This project aims to develop secure and reliable Internet of Things (IoT) data storage on multiple untrusted and unreliable cloud storages through edge computers deployed within the proximity of IoT devices in the 4th generation Industrial Control Systems (ICS).  Despite its well proven benefits in large scale data processing and storage, trust and data communication are two major drawbacks in the cloud-based computing paradigm.   1. *Trust:* Cloud storage has been cursed to be untrusted for three major reasons: who accesses it, where is it stored and how is it stored? There have been tremendous efforts to develop cryptographic methods to secure the cloud data. Even when we assume that these attempts will eventually succeed to answer the above mentioned issues, nothing can stop a ransomware locking the cloud data of an ICS for extortion. Secure and redundant data dispersal to multiple cloud storages appears to be a promising solution. Data can be encrypted using methods that eliminate the need for key management, e.g., message-locked encryption. Such a solution provides perfect secrecy in the sense that less than certain number (threshold) of cloud storages cannot have sufficient information to recover the whole data object. Redundancy ensures that even under ransomware attacks, data can be recovered. 2. *Data Communication:* Either because of limited cloud communication bandwidth or cloud availability, critical systems may experience severe latencies. The 4th generation industrial control systems (ICS) will be examples of such critical systems. Edge computers deployed between cloud and IoT devices, within the proximity of IoT devices, can orchestrate multiple-cloud storage with redundancy to ensure data availability and timely transmission. Edge computers in the IoT enabled ICS can ensure business continuity by handling (down/up)-stream data to/from IoT devices when up to a certain number (threshold) of clouds slows down or becomes unavailable.   This project assumes a network topology where large number of IoT devices accesses edge computers in their proximity. Edge computers collaborate to aggregate, secure, disperse, cache and store data to multiple cloud storage sites with redundancy. This project will built upon CSIRO’s existing solutions (e.g., TrustStore) and implement a novel data dispersal schemes using design theory. Security of the dispersed data will be ensured through our novel combinatorial message-locked encryption scheme.  **Project Duties/Tasks**   * Developing threshold-based redundant data dispersal and message-locked encryption schemes. * Developing methods to prevent malicious encryption attempts. * Developing multiple-cloud data storage, update, recovery, retrieval and disposal solutions. * Developing human-machine interface for human operators. * Developing machine-machine interface for IoT enabled ICS.   **Relevant Fields of Study**   * Information Security * Data Science   **Location:** Data61, Marsfield NSW  **Contact:** For more details please contact **Dr Seyit Camtepe** on phone on (02) 9372 4325 or email [seyit.camtepe@csiro.au](mailto:seyit.camtepe@csiro.au) |
| Data61 6 | **Project Title**  Secure and Usable seL4 System Development  **Project Description**  The formally verified seL4 kernel is the best base for building secure systems. But it could be even better.  The seL4 kernel is formally verified, so we have high confidence that it doesn't have bugs. Other parts of the seL4 platform (including libraries, drivers, etc.) have not yet been verified, so may have (possibly many) bugs in them.  A goal of this project is to use a fuzzer (or some other technique) to analyse some of this non-verified code to find bugs and potential security vulnerabilities in it. If you are so inclined you can create exploits for these bugs and/or fix them too.  In order for seL4 to be an effective base for building secure systems it also has to be usable – i.e. building systems with it shouldn’t be too difficult, or no one will do it. Data61 has developed CAmkES (a component-based platform for developing microkernel-based systems on seL4) which simplifies development of seL4 systems somewhat. However, CAmkES systems quickly become too complicated to read and manipulate in a text format when the system becomes non-trivial.  Another goal of this project is to develop a graphical editor to design such componentised systems: allowing users to draw new components and connections and manipulate existing ones, then generate the code that represents their drawn system.  Note that this project does not require all of the above to be done, it is sufficient to tackle part of the problem (e.g. just the security or just the usability).  **Project Duties/Tasks**   * Develop a graphical editor for CAmkES * Search for vulnerabilities in the non-verified part of the seL4 platform * Possibly create exploits and/or fixes for those vulnerabilities * Opportunity to explore new techniques for program generation * Present results to the group at the end of the project; write project report.   **Relevant Fields of Study**   * Computer Science; the candidate should have experience in OS.   **Location:** Kensington (UNSW), NSW  **Contact:** For more details please contact **Ihor Kuz** on phone on +61 2 9490 5882 or email [Ihor.Kuz@data61.csiro.au](mailto:Ihor.Kuz@data61.csiro.au) |
| Data61 7 | **Project Title**  Blockchain: Live Observations or Application Development  **Project Description**  Blockchain is an emerging technology for decentralised and transactional data processing across a large network of participants. It enables new forms of distributed software architectures, where agreements on shared states can be established without trusting a centralised party. Public blockchains like Bitcoin and Ethereum have seen a massive increase in transaction volumes over the last months. For many applications, however, private or consortium blockchains may be more applicable.  There are two possible directions available for this project:   1. Data61's AAP team has implemented a number of tools to speed up blockchain application development, including through model-driven development (MDD). The goal is to implement applications on the basis of these tools, e.g., to support public government registries, supply chain processes, or interaction with various devices that are deployed in the physical world. 2. Data61's AAP team has two prototypes for observing blockchains live as well as the new *Red Belly Blockchain* in collaboration with the University of Sydney. This blockchain achieves 400,000 transactions per second whereas the observation prototypes have a number of limitations, including:    * Limited scalability: recent increases in transaction volumes have highlighted bottlenecks and limitations in the used techniques for visualization    * Blockchain-specific solutions: the current solutions are specific to a given blockchain protocol   The goal is to overcome both of these limitations, while gaining deep insights into cutting-edge technology.  **Project Duties/Tasks**   * Familiarization with blockchain technology and deciding which additional blockchains or applications to target – this may include Red Belly blockchain * For direction 1. above: deciding on new use cases to implement, then architecting and implementing the use cases, utilizing the AAP team's MDD tools * For direction 2. above: re-architecting and refactoring an existing observation tool to improve scalability, and porting the observation and visualization technologies to additional blockchain protocols.   **Relevant Fields of Study**   * Computer Science * Computer Engineering * Software Engineering.   **Location:** Data61, NSW, Eveleigh (Australian Technology Park)  **Contact:** For more details please contact **Xiwei Xu** or **Vincent Gramoli** on phone 02 9490 5654 or email both [Xiwei.Xu@data61.csiro.au](mailto:Xiwei.Xu@data61.csiro.au) and [vincent.gramoli@data61.csiro.au](mailto:vincent.gramoli@data61.csiro.au) |
| Data61 8 | **Project Title**  Immersive Data Interaction Framework  **Project Description**  In order to cope with increasingly complex, multidimensional data sets, researchers have devised diverse visualisation techniques. These include various dimensionality reduction methods or creatively configured spatial arrangements, coupled with corresponding interaction mechanisms. In addition to such algorithmic and statistical methods, the Information Visualisation discipline also benefits from a variety of creative techniques used in arts and sciences.  However, within the context of immersive environments, including large interactive displays and AR/VR, such techniques are still mostly unexplored. Meanwhile, these environments hold great promise for expanding human perception and cognitive abilities. Therefore, this project aims to explore visual representations and interaction mechanisms suitable for immersive data analytics. The project will build on an existing visualisation framework, to gain access to diverse visual representations and enable rapid prototyping. Building on an existing data model and specification language will allow to focus primarily on aspects of presentation and interaction.  Therefore, the central aim of this project lies in developing a suitable vocabulary of interactions, and integrating them into the visualisation framework to allow for direct manipulation of highly dynamic visualisations. The main objective is to deliver fluent, uninterrupted experiences, which can be characterised as visual analytics at the pace of human thought. Through continuous feedback and natural interactions, the resulting visualisations aim to fully immerse the analyst while exploring complex, multidimensional data sets.  **Project Duties/Tasks**   * Investigate and compile existing best practices and guidelines for interactive visualisation, through literature and contextual reviews. * Iteratively design and implement interaction techniques for AR/VR environments, building and extending upon existing visualisation frameworks. * Evaluate the proposed interaction techniques with potential target users.   **Relevant Fields of Study**   * Computer Science * Design * Engineering.   **Location:** Australian Technology Park (Eveleigh) and EPICentre UNSW (Paddington), NSW  **Contact:** For more details please contact **Tomasz Bednarz** on phone on (02) 9490 5760 or email [tomasz.bednarz@data61.csiro.au](mailto:tomasz.bednarz@data61.csiro.au) |
| Data61 9 | **Project Title**  Software Dependency Analysis  **Project Description**  Modern software generally includes libraries and functionality of third-party modules. Bugs in these can lead to critical vulnerabilities that are hard to detect for a software development team.  In this project, we are going to integrate existing databases and tools for dependency analysis into a broader solution that analyses dependencies. Depending on the successful candidate’s interests, we can focus more on applying the analysis into a Continuous Deployment Pipeline (CDP) or on data mining of software dependencies to uncover hidden patterns.  **Project Duties/Tasks**   * Integration of existing dependency analysis tools into a production-ready broader solution that includes vulnerability tracking * Analysis of software dependencies to uncover hidden risks * Integration into a Continuous Deployment Pipeline.   **Relevant Fields of Study**   * Computer Science * Software Engineering (focus security and/or Data Science).   **Location:** Data61, Eveleigh (Australian Technology Park), NSW  **Contact:** For more details please contact **Dr. Ingo Weber** and **Dr. Ralph Holz** and email both [ingo.weber@data61.csiro.au](mailto:ingo.weber@data61.csiro.au) and [ralph.holz@sydney.edu.au](mailto:ralph.holz@sydney.edu.au) |
| Data61 10 | **Project Title**  Cyber security in a box  **Project Description**  Data61 is in the process of developing a modular and fully customisable cyber security awareness training experience to clients. The proposed portfolio has a number of different components that can all contribute to a rounded cyber security awareness experience within the client’s environment. This focus of this vacation scholarship project would be to design and develop a computer and/or mobile game that can be used to enhance cyber security awareness amongst the target audience. The game type can be envisioned by the student (i.e. role play, virtual world, question based, action based, etc.), but needs to incorporate cyber security knowledge and skills transfer, as well as graphics that are on an acceptable level for use by corporate clients.  **Project Duties/Tasks**   * Plan game plot and general concept development. * Design game storyboard. * Populate game content. * Develop the game on the identified platform. * Test. * Demonstrate.   **Relevant Fields of Study**   * Computer Science, with programming background, either for web based or mobile development.   **Location:** Docklands, Melbourne, VIC  **Contact:** For more details please contact **Dr Marthie Grobler** onphone on (03) 9518 5953 or email [marthie.grobler@data61.csiro.au](mailto:marthie.grobler@data61.csiro.au) |
| Data61 11 | **Project Title**  Connecting Workspace to the Science App Store  **Project Description**  [**Workspace**](https://research.csiro.au/workspace/) is a powerful scientific workflow and application development framework which has been developed by CSIRO Data61 for over 10 years. Originally designed to construct workflows for scientists in the computational fluid dynamics space, Workspace recognises the importance of interactivity, visualisation, scalability, and ease-of-use in a scientific workflow system.  The **Science App Store** is a web-service developed by CSIRO Data61 that allows researchers to publish science software. It allows users, e.g. using a Virtual Laboratory, to discover, browse, and execute the code.  In this project you will be adding functionality to **Workspace** to allow it to interface with the **Science App Store** to access workflows and workflow modules and create and execute new workflows from those components.  **Project Duties/Tasks**   * Become familiar with Workspace and Science App Store architectures and implementations to understand the potential of current App Store data model to the needs of Workspace * Make necessary modifications to the App Store to be able to support Workspace requirements * Modify Workspace to support browsing and retrieving modules and workflows from SSSC * Stretch Goal: Modify Workspace to be able to add workflows to the App Store.   **Relevant Fields of Study**   * Software Engineering, Computer Science. Will require knowledge of Python and C++.   **Location:** Clayton, VIC  **Contact:** For more details please contact **Stuart Woodman** onphone on (03) 9545 8259 or email [Stuart.Woodman@csiro.au](mailto:Stuart.Woodman@csiro.au) |
| Data61 12 | **Project Title**  Analytics for Automation, Trust and Workload assessment toolkit  **Project Description**  The need to understand the relationship between automation, trust and workload is a key requirement across numerous workplaces and industries.  While the physical elements of workload are generally well understood, the concept of cognitive workload is less self-evident. Implementations of real time systems for classifying workload are the next logical step, and while some work has been done in this direction, this is a field with possibilities for novel research and systems design.  This project will focus on further developing a toolkit to measure workload in real-time and testing it in an experimental user study. Tasks will include optimising signal processing and machine learning algorithms to classify psycho-physiological measurements of operator workload, including varying levels of system autonomy, with applications in Human Machine Interaction and general Human Factors.  **Project Duties/Tasks**   * Define research scope with supervisors and background research * Develop and/or adapt existing signal processing tools * Develop and/or adapt machine learning algorithms * Design and conduct small scale validation study with users.   **Relevant Fields of Study**   * Computer science, engineering, human factors or related discipline with interest and skills in one or more of the following areas: Signal processing and Machine learning (required), prototype development (using Python or similar), Human-Computer Interaction and Human Factors.   **Location:** Sandy Bay, Hobart, TAS  **Contact:** For more details please contact **Andreas Duenser** onphone on (03) 62375678 or email [andreas.duenser@csiro.au](mailto:andreas.duenser@csiro.au) |
| Data61 13 | **Project Title**  Channel Constraints  **Project Description**  Channel constraints appear as scarce resources in real-world problems such as port and railway scheduling, which are mainly connected to mining in Australia. Often in a transport network, there are passages in which overtaking or passing of two vehicles is not possible, for example, single railway tracks, or tight river channels. The throughput of such a passage highly depends on the number of times two consecutive vehicles have an opposite travel direction. Current Constraint Programming solvers do not exploit this information for reducing the search space and thus poorly estimate the minimal time required for a set of vehicle trips going one or the other direction.  **Project Duties/Tasks**   * Implementation of a channel propagator in the open-source solver Chuffed written C++ * Testing and evaluation of the channel propagator on synthetic and real world problems.   **Relevant Fields of Study**   * Computer Science.   **Location:** Docklands, VIC  **Contact:** For more details please contact **Peter Stuckey** and **Andreas Schutt** by email on[Peter.stuckey@csiro.au](mailto:Peter.stuckey@csiro.au) and [Andreas.schutt@csiro.au](mailto:Andreas.schutt@csiro.au) |
| Data61 14 | **Project Title**  Collaborative Interaction Design for Virtual Reality and Augmented Reality  **Project Description**  Virtual reality (VR) and augmented reality (AR) systems are increasingly considered as effective interfaces for visual analytics of complex data sources. Collaborative frameworks aid in sharing an environment between multiple users at disparate locations to effectively work together. On account of different interaction approaches in VR and AR interfaces, there is an urgent need to investigate data interaction and manipulation techniques that facilitate a seamless workflow between users and improve their immersive experience for an effective collaboration. In this project, the student will develop and evaluate interaction techniques to support a collaborative visual analytics environment between a VR interface and an AR interface. A user centred design approach will be taken and a formal user study will be performed to evaluate user experience and task performance for the system.  **Project Duties/Tasks**   * Develop interaction techniques for data exploration in a collaborative environment between a VR interface and an AR interface. * Perform a user study to evaluate user experience and task performance within the collaborative system. * Disseminate the outcomes of the development and study in a conference paper, journal article, and/or CSIRO report.   **Relevant Fields of Study**   * Computer Science * Human Computer Interaction * Immersive Systems (AR/VR).   **Location:** Sandy Bay, TAS  **Contact:** For more details please contact **Huyen Nguyen** and **Ulrich Engelke** by phone on 03 6237 5691 or email [huyen.nguyen@csiro.au](mailto:huyen.nguyen@csiro.au) |
| Data61 15 | **Project Title**  Designing behaviourally robust digital market platforms  **Project Description**  Through the Digiscape Future Science Platform CSIRO is developing a digital market platform to help Australian farmers to trade carbon and other environmental products and services. As a designed market, correctly defining the functional details will be critical to its success, both in terms of how efficiently carbon can be traded and how equitably the benefits will be shared. This project provides an opportunity for a computer scientist to help create an experimental market platform which can be used to test how human participants respond to alternative market designs. The student scientist will use the Python-based oTree ([www.oTree.org](http://www.oTree.org)) platform to set up experiments, and work with behavioural economists to apply them.  **Project Duties/Tasks**   * Coding market experiments in oTree (Python) * Assisting with behavioural experiments.   **Relevant Fields of Study**   * Computer Science or similar.   **Location:** Eveleigh, Sydney or Black Mountain, Canberra  **Contact:** For more details please contact **Dr Todd Sanderson** by phone on (02) 9490 5805 or email [todd.sanderson@data61.csiro.au](mailto:todd.sanderson@data61.csiro.au) |
| Data61 16 | **Project Title**  Domain Adaptation in Social Media  **Project Description**  Machine learning techniques have been widely used to tackle various natural language processing or information retrieval problems. Their superior performance is achieved using human-annotated data in a supervised way. However, it is time-consuming and labour-intensive for human-beings to manually annotate data. Furthermore, the annotated data in one domain (source domain) is not usually useful for a different domain (target domain) to train machine learning models.  For this reason, we could ask the question; do we need to annotate data for every single domain? Domain adaptation methods have been proposed to alleviate this problem, adapting a machine learning model built from source to target data. In particular, there are various multiple domains in social media such as politics or sport. While it is fast and less expensive to acquire a large amount of unlabelled data, it is a slow and costly process to obtain human-annotated data for each domain in social media.  The goal of this project is to build a machine learning model for domain adaptation for social media. The outcome of this project will allow us to deploy a single model on multiple social media domains without loss of performance.  Requirements:   * Excellent communication skills * Good programming skills * Experience with Java and Python * Familiarity with natural language processing techniques * Familiarity with machine learning concepts.   **Project Duties/Tasks**   * Software engineering * Natural Language Processing * Machine Learning Research tasks.   **Relevant Fields of Study**   * Computer Science * Natural Language Processing * Machine Learning.   **Location:** Marsfield, NSW  **Contact:** For more details please contact **Mac Kim** by phone on (02) 9372 4253 or email [mac.kim@csiro.au](mailto:mac.kim@csiro.au) |
| Data61 17 | **Project Title**  Hedging FX risk LSV-style!  **Project Description**  In the foreign exchange market, accurate pricing models are essential to the creation and operation of hedging portfolios. This project will provide the first analysis of the vega hedging performances of a new pricing model termed the lognormal local-stochastic volatility model.  **Project Duties/Tasks**   * Task 1: Setup vega hedging portfolios using existing testing interface * Task 2: Compute hedging performance over a range of market inputs and compare with previously computed delta-hedged and delta-gamma hedged LSV results * Task 3: Document hedging performance for possible publication.   **Relevant Fields of Study**   * applied mathematics * probability theory * computer engineering.   **Location:** Clayton, VIC  **Contact:** For more details please contact **Geoff Lee** by phone on (03) 9545 8040 or email [Geoffrey.Lee@data61.csiro.au](mailto:Geoffrey.Lee@data61.csiro.au) |
| Data61 18 | **Project Title**  Reverse logistics to avoid food loss  **Project Description**  As the world population continues to grow, the need for nutritious food produced sustainably can only increase. Today, the only conceivable way of increasing food production is through improvements in efficiency and reductions of food loss, as there is no more land to clear left in the world. In this project, the student will assist in the development of a mathematical programming model to optimise the supply chain of a perishable product in order to reduce food loss. The model is expected to use actual production, transportation and loss data to better plan the operations of the involved companies.  **Project Duties/Tasks**   * Participate in the development of the mathematical formulation, to reflect the actual optimisation needs of the stakeholders; * Code the model into software, to test the validity of the assumptions and verify the model.   **Relevant Fields of Study**   * Mathematics * Engineering * Computer Science * Physics * Economics.   A strong background in linear algebra is required, as well as interest in modelling industrial operations. Experience with any programming language is desirable, as well as basic knowledge of Unix/Linux operating system.  **Location:** Docklands, VIC  **Contact:** For more details please contact **Rodolfo García-Flores** by phone on (03) 9545 8059 or email Rodolfo.Garcia-Flores[@csiro.au](mailto:xx.xxx@csiro.au) |
| Data61 19 | **Project Title**  Simulating Heat Transfer in Granular Flows of Variable Shape On Combined CPU /GPU Architectures  **Project Description**  Granular flows are prevalent throughout many industrial, agriculture, defence and environmental applications. Numerical simulation has become a valuable tool in analysing granular phenomena with the Discrete Element Method (DEM) being one of the most successful techniques. CSIRO has a world-leading DEM solver that models the mechanical behaviour of granular materials with non-spherical shape. Yet the ability to model both thermal and mechanical behaviour in variably shaped granular materials is still in its infancy – despite its high relevance in defence, industrial and geological applications. We have recently developed a new capability to model heat transfer through variably shaped granular materials by combining our F90 DEM solver (which models the collisions between particles) with an OpenCL solver (which models the heat transfer within particles) in order to utilize both CPUs and GPUs in the simulations. This brand new capability needs significant testing and profiling in order to advance it to real-world industrial applications. Results from the project will be valuable in tuning/optimizing this capability to run on CSIRO’s combined CPU/GPU clusters.  **Project Duties/Tasks**   * Adding valuable profiling/timing metrics in the OpenCL and DEM solvers * Trialling different methods of transferring particle data/memory between the CPU and GPU * Extensive testing to determine the optimal combination of simulation size, particle resolution and CPU/GPU number for two benchmark industrial simulations on our CSIRO GPU clusters.   **Relevant Fields of Study**   * Computer Science * Software Engineering * Applied Mathematics.   **Location:** Clayton (Victoria)  **Contact:** For more details please contact **Dr Sharen Cummins** by phone on (03) 9545 8061 or email sharen.cummins[@csiro.au](mailto:xx.xxx@csiro.au) |
| Data61 20 | **Project Title**  Virtual Reality Based Simulation Setup and Data Exploration  **Project Description**  Virtual Reality environments offer the potential to more deeply investigate scientific data and visualise simulation output in ways that were previously impossible. This project will build on existing data analytics and visualisation capabilities to create an integrated tool for simulation setup and visualisation and analysis of simulation data in an interactive Virtual Reality Environment. A shared multi-user VR environment will be developed that will allow for collaborative setup of simulations at real world scale and immersive cooperative analysis of simulation output. The project has places for two students and will focus on developing VR tools for two main use case scenarios:   * Setup and visualising detailed biomechanical models of elite springboard divers and understanding how variations in technique (both between repetitions for an individual athlete and between different athletes) impact on the assessed quality of the dive.   Setup and visualising of simulated industrial devices including metal 3D printers.  **Project Duties/Tasks**   * Develop enhancements to existing data processing pipelines in C++ and C# from CSIRO’s [Workspace](http://research.csiro.au/workspace/) platform to the Unity engine for setup and analysis of simulation data * Develop extension to VR user interface for specification of simulation inputs, transformation of object geometry and real world scale visualisation of a full simulation scene * Report on the effectiveness of the developed capabilities in enhancing the user’s speed and accuracy in specifying simulation inputs and the enhancement to shared understanding in cooperative analysis of complex simulation output.   **Relevant Fields of Study**   * Mechanical/Mechatronics/Electronics Engineering * Computer Science / Computational Science / Physical Science * Biomechanics * Software Engineering * Information Technology (Games and Graphics Programming).   **Location:** Clayton (Victoria)  **Contact:** For more details please contact **Dr Gary Delaney** by phone on (03) 9545 8006 or email [gary.delaney@csiro.au](mailto:gary.delaney@csiro.au) |
| Data61 21 | **Project Title**  The science of evacuation studied through computational modelling  **Project Description**  This project will help unravel some fundamental aspects of crowd evacuation in an emergency. This is of extreme importance with numerous real-world applications ranging from mass evacuation due to natural or man-made disasters to building evacuation in bomb threat scenarios. Numerical modelling has been increasingly used in the last two decades for evacuation planning purposes. Even so, many factors influencing the outcome of an evacuation remain poorly understood, an example being the effect of individuals panicking in a stressful situation. In this project the student will first systematically investigate some of these factors on a simplified model of a room evacuation. Several extensions to the initial model will then be considered in order to contribute to the team objective of building a flexible simulation platform for evacuation modelling.  **Project Duties/Tasks**   * Conduct literature review as required during the project; * Setup numerical models. This will require limited programming using Python. Prior knowledge is desirable but not essential, however willingness to learn is; * Run simulations and study the effect of selected parameters in a systematic manner. A basic understanding of agent based and/or physics based modelling is desirable; * Contribute to the analysis of the results including by creating meaningful and re-usable outputs; * Provide regular progress updates to the supervisors. Good verbal and written communication skills are expected; * Write a technical report at the end of the project.   **Relevant Fields of Study**   * Computer Science * Physics * Applied Mathematics.   **Location:** Clayton, VIC  **Contact:** For more details please contact **Vincent Lemiale** by phone on (03) 9545 2980 or email [vincent.lemiale@csiro.au](mailto:vincent.lemiale@csiro.au) |
| Data61 22 | **Project Title**  High resolution satellite image analytics for asset classification and disaster management applications  **Project Description**  Satellite data imagery is now reaching quality levels that are high enough to be able to discern infrastructure types in good detail. In the area of disaster management, infrastructure assessment and granular economic development assessment, it has now become a real possibility to use satellite data to assess and classify different types of infrastructure assets. The outputs from these datasets can then be used in predictive analytics models for disasters such as floods, fires and cyclones. Once developed, the capability has the potential to be used in a global context due to the pervasiveness such datasets.  **Project Duties/Tasks**   * Evaluate a range of high resolution satellite datasets available to Data61 for a case study region in Melbourne and perhaps another one in a developing country. * Employ a python based program to assess suitability of this code to classify assets including roads, homes, critical infrastructure such as schools, hospitals and airports. * Perform calibration and validation by surveying sections of the data at ground level. * If time permits create a database of the asset classes that can be used as a general template for future expansion.   **Relevant Fields of Study**   * Engineering with strong applied mathematics component and/or Applied Mathematics. * Knowledge and use of Geographic Information Systems is a plus. * Knowledge and use of programming languages especially Python and/or C++ will be viewed very positively.   **Location:** Clayton, VIC  **Contact:** For more details please contact **Dr Mahesh Prakash** by phone on (03) 9545 8010 or email [mahesh.prakash@data61.csiro.au](mailto:mahesh.prakash@data61.csiro.au) |
| Data61 23 | **Project Title**  Developing an assistive medical technology for better inspection of internal organs through 3D reconstruction from endoscopy videos.  **Project Description**  One of the main challenges in diagnosing bowel or lung cancer is to determine whether clinicians cover all areas for cancerous lesions. Furthermore, localising lesion is important when a surgery is required. In this project we aim at generating 3D structure of colon/lung using colonoscopy/bronchoscopy video frames. It is believed that providing a 3D structure of colon/lung during the operation can improve the insight of clinicians and increase their awareness of uninspected areas.  **Project Duties/Tasks**   * Task 1 * Task 2 * Task 3   **Relevant Fields of Study**   * Computer vision or medical imaging   **Location:** Canberra, ACT  **Contact:** For more details please contact **Ali Armin** by phone on phone on (061) 262183703 or email [Ali.Armin@data61.csiro.au](mailto:Ali.Armin@data61.csiro.au) |
| Data61 24 | **Project Title**  Measurement and modelling of digital TV received signal level height variation using drone and/or balloon  **Project Description**  This project aims to characterise the variation of received signal level from terrestrial digital TV transmitters as a function of height using a drone and/or balloon. The project involves both the measurement using a software defined radio (SDR) based digital TV signal receiver and the modelling using electromagnetics theory. Due to the complex interaction of multiple signal paths (e.g. direct line-of-sight path and ground-reflection path) and the variation of radio refractivity in the atmosphere, the received signal level is expected to vary significantly as a function of receiving antenna height. The recent development in the SDR and drone technologies means that such a measurement can be performed more cost effectively to support the development of radio propagation prediction methods with a higher prediction accuracy. This project aims to contribute to the development of radio propagation prediction methods recommended by the International Telecommunication Unions (ITU).  **Project Duties/Tasks**   * Literature survey on terrestrial radio propagation modelling in VHF/UHF, signal format of digital TV, and SDR, and its report. * Investigation into use of drone and/or balloon for digital TV received signal quality measurement, and its report. * Measurement and modelling of digital TV received signal quality, and its report.   **Relevant Fields of Study**   * Electrical Engineering * Information Technology * Data Analytics   **Location:** Marsfield, NSW  **Contact:** For more details please contact **Hajime Suzuki** by phone on phone on (02) 9372 4121 or email [hajime.suzuki@csiro.au](mailto:xx.xxx@csiro.au) |
| Data61 25 | **Project Title**  Evaluating methods to generate differentially private synthetic data  **Project Description**  The Networks Group at Data61 has been developing several methods to generate synthetic data which are differentially private. These methods covers different type of data (e.g. unit record, location-based) and different types of utility (e.g. preserve one-way counts, preserve pair-wise correlations). The team has been evaluating these methods in terms of utility (of the generated data) and computational complexity. These evaluations are in different stages for different methods.  With the help of the NG team, the student working on this internship project will implement some of these methods (in R or Python), perform a series of evaluation experiments using these implementation, and discuss the results with the team.  **Project Duties/Tasks**   * Implement some of the generation methods designed by the team * Perform evaluation experiments of these methods on different data sets, and generate associated graphs and tables * Discuss the results with the team * Propose potential implementation or method improvements.   **Relevant Fields of Study**   * Computer Science * Data Analysis * Data Privacy * Cybersecurity   **Location:** Data 61, Australian Technology Park (ATP), Eveleigh, Sydney  **Contact:** For more details please contact **Thierry Rakotoarivelo** by email [Thierry.Rakotoarivelo@data61.csiro.au](mailto:Thierry.Rakotoarivelo@data61.csiro.au) |
| Data61 26 | **Project Title**  Cloud Mining – Easy Machine Learning tools for IoT  **Project Description**  Machine learning is a powerful technique for gaining insight into patterns and relationships in data, however it typically requires specialist knowledge to be utilised effectively.  This project aims to make the process of applying machine learning techniques to sensor data more user-friendly, by developing one or more adaptable machine learning workflows which can be applied across many datasets with minimal changes in configuration, and without requiring the user to have had any experience in machine learning techniques.  This aim is to lower the barriers to advanced data science methods for developers and system integrators using IoT technologies.  This project will make use of CSIRO's Senaps cloud platform technology to provide support for manipulating sensor data, and for hosting the developed machine learning workflows.  **Project Duties/Tasks**   * Develop one or more machine learning workflows * Upload model(s) into Senaps platform * Develop any required pre- or post-processing functionality * Update Senaps “Data Mining” web interface to support developed model(s).   **Relevant Fields of Study**   * Computer Science * Machine Learning * Environmental Sciences   **Location:** Sandy Bay, TAS  **Contact:** For more details please contact **Mac Coombe** by phone on (03) 6237 5688 or email [mac.coombe@csiro.au](mailto:mac.coombe@csiro.au) |
| Data61 27 | **Project Title**  Bees with Backpacks  **Project Description**  CSIRO has development a tracking system to monitor bee activity and their responses to different stressors. This technology is being applied in different countries and the data is collected and organised by Data61 | CSIRO.  To facilitate data management and information exchange between scientists working across the world a data management interface is required to integrate both engineering and scientific data, associated metadata and assets deployed in the field.  The data management tool should have a neat front-end, including dashboards displaying analytical results. This tool should allow management of experiments, assets, bee activity, metadata for experiments conducted worldwide.  **Project Duties/Tasks**   * Familiarise with existing tools developed by CSIRO Team * Design and validate the proposed data management tool, including analytics and visualisation * Implement the data management tool in Python.   **Relevant Fields of Study**   * Computer Science with interest in data analytics and visualisation   **Location:** Sandy Bay, Hobart, TAS  **Contact:** For more details please contact **Peter Marendy** by phone on (03) 6237 5670 or email [peter.marendy@csiro.au](mailto:peter.marendy@csiro.au) |
| Data61 28 | **Project Title**  Choppy Brain Waves : Data Quality Framework for EEG data  **Project Description**  Electroencephalogram (EEG) data is collected from rodents by scientists studing post-traumatic injury (PTI). This data is recorded over with sample rates from 256 and 512Hz, with 12-24 bit resolution. The data is recorded over a period of several days. As time evolves, EEG data becomes noisy as a result of biofouling, instrument drifting and other reasons. Also as time pass, an injured brain could result is pattern changes in EEG recordings that are important to be detected. The aim of this project is to develop a framework to classify the data (QA/QC: quality assurance and quality control) to determine how appropriate the data is for a given purpose.  **Project Duties/Tasks**   * Read the literature about QA/QC, including standards and different approaches used to process sensing data in real-time * Design and implement a QA/QC framework * Present results internally and with collaborators * Publish a paper.   **Relevant Fields of Study**   * Computer Science with interest in data analytics and machine learning * Maths, Statistics and Physics   **Location:** Sandy Bay, Hobart, TAS  **Contact:** For more details please contact **Prof Paulo de Souza** by phone on (03) 6237 5644 or email [paulo.desouza@csiro.au](mailto:paulo.desouza@csiro.au) |
| Data61 29 | **Project Title**  Additive Manufacturing for Millimeter-Wave Electromagnetic Applications  **Project Description**  To evaluate the performance of additively manufactured components, using multiple metal and plastic processes, over a range of frequencies up to millimetre-wave. There will be opportunity to undertake both measurement and electromagnetic simulation to evaluate losses due to conductivity and surface roughness over a range of frequencies. This will enable the student to gain experience in both metal and plastic additive manufacturing processes, three dimensional electromagnetic modelling, and multi-port Scattering parameter measurements.  **Project Duties/Tasks**   * Literature study on estimating conductivity and surface roughness losses from measurements. * Scanning Electron Microscope measurements of the surface finish of selected samples. * Three dimensional electromagnetic modelling of waveguide cavity resonator prototypes. * Measure dimensional errors of several additive manufacturing processes to determine typical tolerances. * Scattering parameter measurements on a Precision Network Analyser. * Calculation of material conductivity and attenuation verses frequency. * Report on work undertaken and results.   **Relevant Fields of Study**   * Bachelor of Engineering in Electrical or Electronic Engineering with specialty in telecommunications.   **Location:** Marsfield, NSW  **Contact:** For more details please contact **Stephanie Smith** by phone on (02) 9372 4131 or email [stephanie.smith@csiro.au](mailto:stephanie.smith@csiro.au) |
| Data61 30 | **Project Title**  Creating bespoke robot parts  **Project Description**  We are developing technologies that allow us to design bespoke robot components, e.g., the shape of a gripper or the geometry of a chassis. To push this capability, we are seeking a motivated student to create realistic simulation environments, look at ways to match the simulations more closely to reality, and to investigate how the way we represent these components affects their optimality.  **Project Duties/Tasks**   * Create code that lets us represent different types of robot components, and investigate how different representations effects robot performance * Design a number of different, physically-plausible tests for the components to undergo * Present your work to site visitors.   **Relevant Fields of Study**   * Computing / Computer Science * Software Engineering * AI * Mechatronics * Robotics   **Location:** QCAT, Pullenvale  **Contact:** For more details please contact **Dr David Howard** by phone on (07) 3327 4714 or email [david.howard@csiro.au](mailto:david.howard@csiro.au) |
| Data61 31 | **Project Title**  Traffic Simulation for Congestion Management in Sydney  **Project Description**  Addressing congestion in large cities is a major concern for traffic management centres, especially when unpredictable incidents can severely affect important corridors in the city. Traffic Simulation can be used to test various scenarios of how to deal with accidents, road closures, or deviated traffic in the city. But building such simulation models can be a true challenge especially due to large amounts of data that need to be fed into the model such as: traffic demand, public transport lines that need to be configured in terms of routes and schedules timetables, as well as traffic signal control, incident data, etc.  This project aims at using an existent traffic model from the city of Sydney, augmented by real data obtained from Transport for New South Wales. The student will work on calibrating the traffic simulation model, as well as applying various scenarios for static or dynamic traffic assignment. The first phase of the internship will require preparation for working in a very popular traffic simulation tool: Aimsun. This phase assumes some pre-existing knowledge of Python programming and traffic simulation. The student will get hands-on and traffic modelling experience from this internship which can be considered as high assets in industrial applications in transportation.  **Project Duties/Tasks**   * Gain first-hand experience in Aimsun including scripting in Python * Calibrate and validate traffic simulation model for a popular corridor in the Sydney metropolitan area * Provide a detailed technical report that summarises the methodology and outcomes of the simulation model.   **Relevant Fields of Study**   * Civil Engineering * Computer Science and Engineering   **Location:** Sydney, ATP, 13 Garden Str, Eveleigh, NSW, 2015  **Contact:** For more details please contact **Simona Mihaita** by phone on + (+61) 94 90 56 15 or email [simona.mihaita@data61.csiro.au](mailto:simona.mihaita@data61.csiro.au) or **Chen Cai** by phone on +610294905531 or email [chen.cai@data61.csiro.au](mailto:chen.cai@data61.csiro.au) |
| Data61 32 | **Project Title**  Transport data analytics and visualisation  **Project Description**  We are looking for an aspiring software engineer or data scientist, who is interested in learning from and working with a successful transport data analytics team. We have access to a lot of interesting real-world data on Sydney’s roads and public transport networks, along with talented researchers in Transportation, Machine Learning, Behavioural modelling and Simulation.  **Project Duties/Tasks**   * Complex geospatial data fusion and visualisation * Big data analytics and visualisation * Using AI or Machine Learning to monitor real-time sensor data or operational logs * Building an API and interactive visualisation for parts of our cuttingedge research.   **Relevant Fields of Study**   * Computer Science   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Ben Itzstein** by phone on (02) 94905718 or email [ben.itzstein@data61.csiro.au](mailto:ben.itzstein@data61.csiro.au) |
| Data61 33 | **Project Title**  Visual analytics of machine learning performance  **Project Description**  As machine learning (ML) is becoming a more widely used tool in various fields, adequate methods have to be provided to allow users to employ ML techniques effectively. However, various ML methods may create different results for the same data set, which makes users hard to justify ML performance from different ML methods and different feature combinations of the data set, especially if large number of features are included in the data set.  This research project focuses on the visual analytics and interpretation of ML performance. The objective of this project is to develop visual analytics approaches of ML performance and compare different visual analytics approaches quantitatively and qualitatively with different data sets. The ultimate goal is to find effective approaches for facilitating understanding of ML performance so that users make trustworthy decisions based on ML results. Such investigation is also beneficial to ML researchers in factor analysis and contributes to our transparent machine learning project.  Students involved in this project will gain comprehensive knowledge on machine learning, visualization, computing skills as well as various interesting HCI techniques.  **Project Duties/Tasks**   * Design approaches to visualize ML performance under different conditions such as different models and feature combinations. * Develop interpretation approaches for visual analytics of ML performance. * Implement the proposed approaches. * Carry out scale user studies with different ML approaches and different data sets. * Write a report to summarize approaches and results.   **Relevant Fields of Study**   * Machine Learning, Visualization, HCI   **Location:** Level 5, 13 Garden Street, Eveleigh, NSW 2015  **Contact:** For more details please contact **Jianlong Zhou** by email [Jianlong.zhou@data61.csiro.au](mailto:Jianlong.zhou@data61.csiro.au) |
| Data61 34 | **Project Title**  Data Mining to Analyse Performance in Team Projects  **Project Description**  The aim of this project is to assess the contribution of members in student software team projects. The project will involve analysing weekly contribution logs and code committed to the software repository to estimate the level of contribution and performance by each member of the team, machine learning techniques will then be used to generate appropriate feedback to team members as well as grading. A visualisation system will also be designed and implemented to present the results to the teachers and students.  On top of the challenge of natural language processing (e.g. text similarity), this project presents the challenge of possible conflict of information, for example, the difference between log presented by students versus work plan, as well as their real contribution to the team. The other challenge of the project will be to take into account each student’s work from the previous week(s) in providing the feedback.  The work from this project can have applications in similar team based work in academia and industry.  **Project Duties/Tasks**   * Text analysis of work log and code * Use machine learning to assess work submitted against plan * Visualisation of assessment for easy presentation   **Relevant Fields of Study**   * Software Engineering * Computer Science * Data Science   **Location:** Marsfield/Swinburne  **Contact:** For more details please contact **Rosalind Wang** by phone on (02) 9372 4709 or email [Rosalind.Wang@csiro.au](mailto:Rosalind.Wang@csiro.au) or **Caslon Chua (Swinburne University of Technology)** by phone on (03) 9214 5397 or email [cchua@swin.edu.au](mailto:cchua@swin.edu.au) |
| Data61 35 | **Project Title**  Knowing Customers’ Expectation by Modern Machine Learning Models.  **Project Description**  In data-driven prediction projects, customers may found that our prediction results are quite different from what they are expecting based on their past experience, then they could feel unconfident when considering to use the prediction. This may happen even when the model prediction is more accurate. Here we will design a solution to partially address this issue by discovering their possible expectations with modern data mining and machine learning techniques based on the actual data.  **Project Duties/Tasks**   * Task1: design the framework of solution; * Task2: collect and pre-processing the data; * Task3: investigate the similar problems to choose the suitable model; * Task4: implement the chosen model and apply the data;   **Relevant Fields of Study**   * Data Mining * Machine Learning   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Zhidong Li** by phone on (02) 9490 5853 or email [Zhidong.Li@data61.csiro.au](mailto:Zhidong.Li@data61.csiro.au) |
| Data61 36 | **Project Title**  Machine Learning in Transportation Models  **Project Description**  We seek the study of economic and probabilistic models representing how individuals make choices in transportation space.  These mathematical models are used to analyse modal choice to predict likely behaviour in an ever changing environment. For example, how individuals would react to the establishment of new modes of transport and other exogeneous events.  Drawing and utilizing insights from the emerging data in transportation science, however, remains a challenge due to fast moving developments, its unprecedented degrees of heterogeneity and lack of theoretical underpinning. Furthermore, with the increase of data type and depth involve, the data management and the security and privacy issues of the users are becoming a serious concern of various different authorities.  The project aims at using cutting edge machine learning techniques to complement and extend well known inference mechanisms for choice modelling in transportation..  **Project Duties/Tasks**   * Understand recent advances in transportation models and its connections to other mathematical machine learning * Provide a detailed technical report that summarises the methodology and outcomes of the findings.   **Relevant Fields of Study**   * Mathematics and Statistics * Computer Science and Engineering   **Location:** Sydney, ATP, 13 Garden Str, Eveleigh, NSW, 2015  **Contact:** For more details please contact **Le Minh Kieu (Main Supervisor)** by email [leminh.kieu@data61.csiro.au](mailto:leminh.kieu@data61.csiro.au) or **Chen Cai (Joint Supervisor)** by email [chen.cai@data61.csiro.au](mailto:chen.cai@data61.csiro.au) or **Young Lee (Joint Supervisor)** by email [young.lee@data61.csiro.au](mailto:young.lee@data61.csiro.au) |
| Data61 37 | **Project Title**  User Experience Design Research  **Project Description**  Identify and profile the types of users that Data61 creates data enabled technologies for. This could be a broad activity or for a specific project (to be determined on the timing and the applicant).  The deliverable would be a comprehensive report challenging/validating the assumptions held to date by the UX team and other project team members about data users as referenceable body of knowledge to help guide better user research and development discovery.  **Project Duties/Tasks**   * Draw together various user experience research done to date to find themes of common motivations and contexts. * Create a set of discovery questions to work with. * Capture gaps or more specific questions for the Data61 UX team to include in user interviews in their user research work on existing and future projects. * Conduct desk research and/or interviews with internal staff for further insights. * Create communication artefacts (e.g. personas, user stories)   There will be no field work (user interviews) expected.  **Relevant Fields of Study**  Sociology, anthropology or human factors with keen interests in practical application in the creation of digital technologies.  **Location:** Eveleigh or Canberra  **Contact:** For more details please contact **Hilary Cinis** by phone on +61-2-9490-5970 or email [hilary.cinis@data61.csiro.au](mailto:hilary.cinis@data61.csiro.au) |
| Data61 38 | **Project Title**  Machine Learning on Graphs  **Project Description**  Data61 is currently building a graph analytics and visualisation platform for use across Australian government and business. The platform will contain a suite of machine learning and data visualisation tools for prediction, classification and to better understand datasets. This project involves building and evaluating a machine learning algorithm for graphs (networks). The project will use Python or R to analyse a real dataset and to predict features using local network properties.  **Project Duties/Tasks**   * Perform data analytics in Python or R to understand a dataset * Use machine learning algorithms to predict features on a graph * Evaluate and summarise results   **Relevant Fields of Study**   * Computer Science * Engineering * Mathematics * Statistics   **Location:** Eveleigh ATP  **Contact:** For more details please contact **Alex Collins** by phone on (02) 9490 5963 or email [alex.collins@data61.csiro.au](mailto:alex.collins@data61.csiro.au) |
| Data61 39 | **Project Title**  Privacy-preserving machine learning via fully homomorphic encryption  **Project Description**  Machine learning algorithms a proving remarkably successful in creating systems that make intelligent predictions. These prediction systems are inveigling their way into all aspects of modern life, from personal recommendation systems to medical diagnoses to creditworthiness predictions. Even when these prediction systems are created with laudable goals in mind, there remains the question of who can see the sensitive raw data that is being used to construct the system? One very promising approach to resolving this issue is to adapt machine learning algorithms to use fully homomorphic encryption, which allows certain arithmetic operations to be performed on encrypted numbers. When a machine learning algorithm can be adapted to operate on an encrypted form of the raw data, then we can have our cake-the sensitive data remains behind the veil of encryption-and eat it too-we still obtain a useful prediction system based on the raw data.  The goal of this project is to implement and benchmark a machine learning algorithm widely used for classification problems-logistic regression-on encrypted data. The fully homomorphic encryption layer will be provided by an existing library such as SEAL or HElib.  Prerequisites: Abstract algebra, elementary number theory, calculus, algorithms, C++ programming.  Primary references:  Dowlin et al. “Manual for using Homomorphic Encryption for Bioinformatics” in Proceedings of the IEEE, Vol 105, No. 3, March 2017. <http://ieeexplore.ieee.org/document/7843616L>  **1**  Laine, et al. Secure Encrypted Arithmetic Library (SEAL) Manual, v2.2. <https://www.microsoft.com/en-us/research/publication/single-encrypted-arithmetic-library-seal-v2-2/>  Aono, et al. "Scalable and Secure Logistic Regression via Homomorphic Encryption" IEICE Transactions 99-D(8): 2079-2089 (2016) <https://eprint.iacr.org/2016/111>  **Relevant Fields of Study**   * Mathematics & Computer Science   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Hamish Ivey-Law** by email [Hamish.ivey-law@data61.csiro.au](mailto:Hamish.ivey-law@data61.csiro.au) |
| Data61 40 | **Project Title**  Smart Infrastructure Systems  **Project Description**  The smart Infrastructure Systems engineering team builds cloud software for the “internet of things” (IoT) – applying remote sensing, machine learning and analytics to a range of applications, from monitoring of civil infrastructure to precision agriculture.  You’ll get to apply and develop your software engineering skills to deliver cloud software in this cutting-edge field of IoT.  **Project Duties/Tasks**   * Cloud software development for AWS (Amazon Web Services) * Python and functional Scala programming * Software engineering including Agile development with JIRA, continuous integtation, Git, code reviews, team presentations   **Relevant Fields of Study**   * Computer Systems Engineering * Computer Science   **Location:** ATP (Eveleigh, Sydney)  **Contact:** For more details please contact **Tim Pitman** by phone on (02) 9490 5633 or email [tim.pitman@data61.csiro.au](mailto:tim.pitman@data61.csiro.au) |
| Data61 41 | **Project Title**  Add Blockchain Ledger for N1  **Project Description**  The confidential compute platform currently uses traditional logging mechanisms, this project is to add a distribution blockchain ledger to record exactly who did what in a private blockchain between N1 engines.  **Relevant Fields of Study**   * Computer Science * Engineering * Mathematics   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Stephen Hardy** by email [Stephen.Hardy@data61.csiro.au](mailto:Stephen.Hardy@data61.csiro.au) |
| Data61 42 | **Project Title**  Blockchain Meets Confidential Machine Learning  **Project Description**  We are building an analytics market place where organisations can offer their data for others to create better prediction systems and models. What differentiates our market from others is that the offered data is encrypted homomorphicly, allowing the data provider to keep their data confidential, while allowing the data “buyer” to build better ML models without needing to see the data they are using.  A core component of any market is the incentive mechanism. How can we price the value the “prediction model builders” gain from having access to the additional data provided by the market? How should we distribute the income of the market to its data providers.  In this project we want to concentrate on the incentives for the data providers. Intuitively, we expect that the different datasets used for learning a specific prediction models will contribute differently to the model’s “uplift”. We also expect that the reward for having access to more data is tied to the same uplift.  We would like to explore a model where potential consumer data will agree to a reward structure tied to some evaluation criteria of the learned model. We also assume that there is a mechanism that not only can measure that criteria but can also determine the relative contribution of each data provider to the outcome.  As we would like to minimize the need for a central, trusted authority, we would like to explore in this project the possibility of using a blockchain and smart contracts to realise the above described vision.  Prerequisites: Good understanding of standard machine learning algorithms as well as blockchain technologies with an emphasis on smart contract logic. Basic understanding of zero-knowledge proofs will most likely be beneficial.  **Relevant Fields of Study**   * Computer Science &/or Engineering   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Max Ott** by email [max.ott@data61.csiro.au](mailto:max.ott@data61.csiro.au) |
| Data61 43 | **Project Title**  Adversarial validation for machine learning in production  **Project Description**  A fundamental assumption in supervised machine learning is that an algorithm is trained on data with similar statistics as the data used to query the algorithm for predictions. However, in real-life applications this is very often not the case as the query data distribution is unknown and cannot be guaranteed a-priori. This problem is known as *covariate shift* in the machine learning literature, and using a machine learning algorithm in this situation can result in spurious and often over-confident predictions.  In this project we propose to add to the existing suite of machine learning model selection and validation tools (such as cross-validation), a suite of out-of-domain and adversarial validation techniques. These techniques will be specifically created to measure an algorithm's performance when it experiences covariate shift, and will reward algorithms that *extrapolate* or model *predictive uncertainty* well in these situations. They will also alert a user when an algorithm is experiencing covariate shift at prediction time. Depending on time, there may also be scope for looking at deep learning *domain adaption* techniques (such as generative adversarial networks and implicit models) that learn and correct for the shift between the training and testing datasets.  Ideally the candidate will be familiar with tools such as:   * A variety of machine learning algorithms * Cross validation * Python * Numpy/Scipy/Scikit Learn/Bokeh * Bonus points for being familiar with Bayesian machine learning, e.g. Bayesian linear regression and Gaussian Processes   **Project Duties/Tasks**   * A framework for adversarial validation/cross validation like Scikit Learn's model\_selection module. * A framework for prediction-time diagnostics that can warn a user when a machine learning model is predicting out of domain. * Look into the state-of-the-art domain adaption techniques to see if covariate shift can be rectified in certain scenarios.   **Relevant Fields of Study**   * Computer Science * Physics * Software Engineering * Applied Mathematics   **Location:** Eveleigh, Sydney  **Contact:** For more details please contact **Dr Daniel Steinberg** by phone on (02) 9490 5520 or email [daniel.steinberg@data61.csiro.au](mailto:daniel.steinberg@data61.csiro.au) |
| Data61 44 | **Project Title**  Data Visualisation for Networks  **Project Description**  Data61 is currently building a graph analytics and visualisation platform for use across Australian government and businesses. The platform will contain a suite of machine learning and data visualisation tools for prediction, classification and to better understand datasets. This project involves building data visualisation software for understanding large graph datasets. This project will involve developing software with WebGL or other browser technologies.  **Project Duties/Tasks**   * Building data visualisation software using GPU technologies * Collaborating with user interface and user experience teams * Testing, evaluating and presenting results   **Relevant Fields of Study**   * Computer Science * Engineering * Mathematics * Statistics * Computer Graphic Design   **Location:** Eveleigh ATP  **Contact:** For more details please contact **Alex Collins** by phone on (02) 9490 5963 or email [alex.collins@data61.csiro.au](mailto:alex.collins@data61.csiro.au) |
| Data61 45 | **Project Title**  Deep learning with partially homomorphic encryption  **Project Description**  Can we use the predictive power of deep neural networks without compromising private information contained in the training data? Can we leverage encryption mechanisms to address this challenging problem? Scientists are investigating these questions, experimenting with polynomial approximation of neuron's activations and homomorphic encryption -- with mixed results. The goal of this project is to design and implement a working prototype of a training algorithm for a neural network, based on data that has been encrypted (e.g. with Paillier partially homomorphic encryption schema) and therefore is not visible in the clear.  Required: linear algebra, calculus, Python/Java  Preferred: basics of machine learning and cryptography  1: [https://www.microsoft.com/en-us/research/publication/cryptonets-applying- neural-networks-to-encrypted-data-with-high-throughput-and-accuracy/](https://www.microsoft.com/en-us/research/publication/cryptonets-applying-%20neural-networks-to-encrypted-data-with-high-throughput-and-accuracy/)  **Relevant Fields of Study**   * Computer Science * Mathematics * Engineering   **Location:** Eveleigh, NSW  **Contact:** For more details please contact **Wilko Henecka** by email [wilko.henecka@data61.csiro.au](mailto:wilko.henecka@data61.csiro.au) |