Capstone projects COMP5703 – Semester 2 2018

Project Number	Project Title	
CP 1	Clustering Analysis of Time Series Data	Individual supervised project
CP 2	Non-parametric Bayesian Model based Point Process Analysis	Individual supervised project
CP 3	Multimedia Data Analysis and Understanding	Individual supervised project
CP 4	Human Motion Analysis, Modelling, Animation, and Synthesis	Individual supervised project
CP 5	Multimedia Forensic and Security	Individual supervised project
CP 6	AR Techniques and Applications	Individual supervised project
CP 7	Understanding IoT Security Threats through Network Forensics	Group projects
CP 8	Phishing Website Detection Through Image Hashing	Group projects
CP 9	Deep Neural Network (DNN) Surgery in Edge Computing	Individual supervised project
CP 10	Scalable visual analytics of big complex data	Individual supervised project
CP 11	Enabling Federated Learning in Mobile Edge Computing	Group projects
CP 12	Analysis on digital wellbeing	Group projects
CP 13	How late is too late for a Bus?	Group projects
CP 14	Machine Intelligent Financial Trading	Group projects
CP 15	SkoolBot: Visual Programming Environment for 3d- Printed Robots	Group projects
CP 16	The Teleporter 1: Scene stitching	Group projects
CP 17	The Teleporter 1: Face stitching	Group projects
CP 18	The Teleporter 3: Eye Tracking Optimizer	Group projects
CP 19	The Teleporter 4: Network streaming	Group projects
CP 20	The Teleporter 5: User Interface	Group projects
CP 21	Notion AI in Strategy	Group / individual
CP 22	An Online Web-Streaming Service for Bitcoin-Exchanges	Group projects
CP 23	Visualisation and prediction for Bitcoin-Exchanges	Group projects
CP 24	Education Analytic Platform	Group projects
CP 25	Connect	Group projects
CP 26	Data Extraction from DBLP: Computer Science Bibliography to IRMA	Group projects

CP 1 - Clustering Analysis of Time Series

Description: Time series is a common type of dynamic data that naturally arises in many different scenarios, such as stock data, medical data, and machine monitoring. Cluster analysis is a task which concerns itself with the creation of groups of objects, where each group is called a cluster. Clustering algorithms can be organized differently depending on how they handle the data and how the groups are created. For the time-series clustering problem, the major issue is the covariance between observations in different time. There are several clustering algorithms in time-series clustering including hierarchical clustering, partitional clustering and fuzzy clustering. We need to test the performance of these algorithms on time series data and compare the results with an improved clustering algorithm.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 2 - Non-parametric Bayesian Model based Point Process Analysis

Description: Point processes are a standard model when the objects of study are the number and repartition of otherwise identical points on a domain, usually time or space, with widely used applications such as astronomy, forestry and finance. The Poisson process and Hawkes process are two commonly used point processes. Hawkes process can be interpreted as the superposition of Poisson processes. However, when the knowledge is unsubstantial for domain experts to determine the latent model, an exact non-parametric Bayesian model which enables inference on Poisson process is needed. In this project, we plan to apply the model to Hawkes process. The inference method should scale linearly with the number of data points meaning it must be fast.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 3 – Multimedia Data Analysis and Understanding

Description: Multimedia data has been increasingly acquired in almost every domain, ranging from our daily experiences through smart phones, lecturing videos, footage of security surveillance, music performances, and sports videos, to surgical videos and brain imaging. This has imposed great opportunities for knowledge discovery from massive multimedia data and new interactions with media such as through VR and AR. The challenging tasks towards this goal include object detection, tracking, and recognition, event detection, video summarization, multi-modal retrieval, video captioning, and affective analysis. This project is to investigate novel techniques to advance media driven data analytics and knowledge discovery.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 4 – Human Motion Analysis, Modelling, Animation, and Synthesis

Description: People are the focus in most activities; hence investigating human motion has been driven by a wide range of applications such as visual surveillance, 3D reconstruction and animation, novel human computer interaction, sports, and medical diagnosis and treatment. This project is to address a number of challenge issues of this area in realistic scenarios, including human tracking, motion detection, recognition, modelling, animation, and synthesis to advance human behaviour analysis. Students will gain comprehensive knowledge in computer vision (e.g. object segmentation and tracking, and action/event detection and recognition), 3D modelling, computer graphics, and machine learning.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 5 – Multimedia Forensic and Security

Description: Multimedia data has been widely used to store information in almost every domain, from photos shared on social media platforms and transaction receipts to electronic health records. Meanwhile, advances in digital media processing have produced a large variety of intelligent tools for manipulating media content, such as enhancing visual quality and removing or adding an object from an image. However, the processed media content could be used to convey false, misleading, or hidden information, which has increasingly challenged the saying "seeing is believing". This project aims to develop advanced multimedia computing and machine learning techniques to identify the forensic and security trails and improve security of multimedia data.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 6 – AR Techniques and Applications

Description: AR (Augmented Reality) has demonstrated great potential in many domains such as education, entertainment, and manufacturing. This project aims to explore this emerging field by developing novel computer vision techniques for better AR experiences, and building exciting AR applications with mobile phones or other AR devices such as Microsoft HoloLens.

Technique skills: strong programming skills and math

Supervisor: Dr Zhiyong Wang (Zhiyong.wang@sydney.edu.au)

CP 7 – Understanding IoT Security Threats through Network Forensics

Description: With the advent of the Internet of Things, a myriad of devices is increasingly being installed in our physical environment, and connected to cloud-based services via the Internet. This rollout has happened with little concern for security. Governments, industry, and consumers face a grand challenge: with IoT technology proposed or already in use in Industry 4.0, smart cities, smart homes, and many other fields, the very same devices may well be vulnerable to attacks. The potential for serious damage, nationally and globally, is enormous - yet too little is known about the actual attack surface and the real threat landscape. While the Mirai botnet has demonstrated that IoT devices without the most basic protection are already being turned into massive botnets, it is unknown (beyond such widely publicized events) which other attack vectors exist, which forms of attacks or attack preparations are being tried, and which device classes they target.

The goal of this project is to address these uncertainties by carrying out network data forensics on passively collected wireless IoT network traffic. The group is expected to capture unique traffic signatures of IoT devices

such as IP addresses, protocols, DTLS/TLS headers, packet sizes, flow intervals, etc. for various IoT devices using AirPcaP NX wireless packet sniffer. Access to packet sniffers and variety of IoT devices will be provided. These traffic signatures will then enable us to assess what attack vectors are possible, already being tried, and which risks the devices' connection to the cloud (and their manufacturer's servers) is exposed to.

CP 8 – Phishing Website Detection Through image hashing

Phishing websites (websites that are visually similar to popular websites) are an ongoing problem in the internet and are being increasingly used to steal financial information, personal information, corporate espionage, spread malware, and in some cases as means to penetrate corporate networks. These websites may be able to mislead even the most tech-savvy users if they did not focus on the subtle changes in the user interface or the connected URLs when prompted through social networks or email. For example, in Australia alone, in 2017 there were over 25,000 phishing incidents reported causing damages of over 1 million AUD.

Existing commercial solutions available in anti-virus suites and browser plugins for phishing website detection majorly focus on the similarity of the URLS (e.g. www.google.com vs. www.goole.com). However, more sophisticated phishing attempts can prompt the users with highly visually similar websites to popular websites or embed forged visual content such as logos to present misleading information.

This project aims to develop a phishing website detection system based on the visual similarity of websites at Internet-scale by leveraging image hashing methods that capture the perceptual similarity of web pages. The group is expected to develop a web page crawler that will crawl arbitrary web pages in the internet using headless browsers and save screenshots of those pages. Then various image hashing methods (E.g. average hashing and perceptual hashing) will be evaluated with respect to how good they are in identifying visually similar web pages.

CP 9 - Deep Neural Network (DNN) Surgery in Edge Computing

Description: Recent advances in deep neural networks (DNNs) have substantially improve the accuracy and speed of video analytics. The maturity of cloud computing, equipped with powerful hardware like GPU, becomes a typical choice for such kind of computation intensive DNN tasks. One obstacle, however, is the large amount of data volume of video streams. For example, a self-driving car can generate up to 750 megabytes of sensed data per second, but the average uplink rate of 4G, fastest existing solution, is only 5.85 Mbps. In order to avoid the effects of network delay and put the computing at the proximity of data sources, edge computing emerges. As a network-free approach, it provides anywhere and anytime available computing resources. Nevertheless, edge computer itself is limited by its computing capacity and energy constraints, which cannot fully replace cloud computing.

To combine the benefits of both edge computing and cloud computing, DNN surgery is a very recent solution. For a typical DNN, the data volume of some intermediate results (the output of intermediate layers) are significantly smaller than that of raw input data. DNN surgery allows DNN partition, so that one portion of DNN inference is processed at the edge and the rest is processed at the cloud. The place of partition is carefully determined so that the communication data volume between the edge and the cloud is also small.

Supervisor: Dr Wei Bao (wei.bao@sydney.edu.au)

CP 10 – Scalable visual analytics of big complex data

Technological advances such as sensors have increased data volumes in the last few years, and now we are experiencing a "data deluge" in which data is produced much faster than it can be used by humans.

Further, these huge and complex data sets have grown in importance due to factors such as international terrorism, the success of genomics, increasingly complex software systems, and widespread fraud on stock markets.

Skills: programming (Python, Java, C++)

Supervisor: Prof Seokhee Hong (Seokhee.hong@sydney.edu.au)

CP 11 – Enabling Federated Learning in Mobile Edge Computing

Federated Learning aims to train the global (centralized) model utilizing participation of a large number of nodes with unevenly distributed training data. Under this setting, the mobile devices are considered as a local host for performing computation on their local data to facilitate the required updates to improve a global model. In such distributed setting, while it remains with the utmost importance to have a fair number of participating nodes with enough local data - preferably balanced in nature, a coexisting pragmatic challenge would be to handle the communication efficiency while assimilating updates for improving the global model. This project will study an incentive mechanism assisting Federated learning to involve a number of participating mobile devices with local training data for improving the global model considering communication and computation efficiency of the Mobile Edge Computing system.

Skills required (if applicable): Python (or Julia) coding skills for data analytics.

CP 12 Analysis on digital wellbeing

People are replying on mobile phone heavily now, and some of them have become excessive dependent on the phone. Both Google and Apple have updated their mobile OS recently to give people a better understanding on their phone usage. The project requires you to leverage the available system data and explore what are the major factors to influence people's digital wellbeing.

This project requires the students to conduct a survey with phone users (the number of respondents is proportional to your team size), and compare the system log with survey responses. The students are encouraged to use theories to explain the discrepancy or consistency between the two sets of data.

CP 13 – How late is too late for a Bus?

Description: The public transport provides APIs to give real-time data about public transport. This interface can be used to gather data and compare against given schedules.

The idea of this project is to gather real-time data from the public transport servers, build appropriate datamodel for the data in form of relational databases,

and compute statistics that are visualized on geographic maps in real-time. An initial prototype for this type of project has been build: https://bus-shaming.com

Skills: Data Analytic Skills, databases, and programming skills

CP 14 – Machine Intelligent Financial Trading

Description: Financial markets including stocks and forex constantly fluctuate. Apart from individual company performances, various external factors such as human psychology of the masses, local and international news contribute to the dynamics of the markets. Forecasting the level of the impact of these factors on financial markets is a dream of many decision makers. Development of an artificial intelligent (AI) agent based on pattern recognition can make this dream a reality. Elliott Waves is one of the pattern which describes the repetitive patterns and impact of human psychology on financial markets. However, the manual analysis of these patterns is very laborious and subjective.

Al has been shown to successfully unveil the hidden patterns in data. Al implementation to financial data can help to identify social mood of the masses and thus the movement in financial markets, which in turn can lead to stabilising of the socionomics of the society.

Therefore, in this project, the students will study the socionomics patterns by digging deep into the historical data and market reaction to predict future direction of the markets. They will explore the use of machine/deep learning to correctly identify the price-action trend. The students will also investigate the correlation and the level of the impact of news, human sentiments or company performances on a sector.

Requirements (if applicable): Machine Learning / Data Mining

CP 15 SkoolBot: Visual Programming Environment for 3d-Printed Robots

With the increasing focus on Science, Technology, Engineering and Maths (STEM) in schools across Australia, there is a need to provide primary and high schools with educational STEM platforms for their students. It has been shown that robots are the most suitable learning aids conveying STEM skills to primary and high-school children. However, existing platforms are expensive and/or limit the children's creativity to design, build, and program robots. With the presence of cheap consumer 3d printers (cf. Aldi's Cocoon Create for \$500), and cheap capable microcontroller such as Raspberry Pi Zero W

(cf. https://raspberry.piaustralia.com.au for \$15)/BBCbit, we would like to create a new affordable open-source STEM platform for programmable 3d-printed robots at The University of Sydney.

A new open-source collaborative software suite will give students the capabilities to interactively design, build, and program 3d-printed robots in a step-by-step fashion. With the availability of cheap 3d-printer filament, sensors, actuators, and cheap microcontrollers, the material costs per 3d-printed robot will be less than \$50 AUD.

Description: The aim of this project is to develop a collaborative visual design and programming environment for 3d printed robots. The visual programming environment is geared towards children in the age between 8 and 16. With the visual programming environment, the programming of self-printed 3d robots should become fun, educational, and easy-to-learn. Children will design their robots in an easy CAD, and program their robots by dragging blocks from a toolbox and attaching them to other blocks like a mosaic. The method of programming is usually referred to as visual programming / drag-and-

drop programming exposing the syntax rules visually rather than on a textual level. The student's programs can be either kept privately or shared publicly such that other children can modify/remix the code. Google has introduced an open-source initiative called Blockly (cf. https://developers.google.com/blockly/), which is an interactive visual programming language similar to MIT's Scratch. Blockly is a library and, hence, can be extended with ease.

Requirements: Javascript, Python, Linux

CP 16 – The Teleporter 1: Scene stitching

The Teleporter is a new project to create the most accurate version of being in another place.

We combine live 3D VR cameras with holographic displays to create the feeling of actually being somewhere with someone, when you are far away.

This a solution for families separated by distance to see their children. It is a solution for scientists working on remote field ops on the planet or in space. It is a solution for field engineers working in dangerous and remote locations to access support from their whole team.

While working in the Augmented Reality industry, this was the number one requested use case: Remote Help.

Description: Use off the shelf VR cameras, place them on a helmet, and use C/C++/Computer vision techniques to stitch together the camera feeds. The goal is high resolution at the front, balanced with a full 3D picture.

Meta Augmented Reality Glasses

HTC Vive.

Building the Interface in Unity 3D, will require a clever design to combine the 3D stitching of the face, with the environment in a way that feels natural and normal.

The future challenge will be to then share the other user back into the environment, so both parties see each other in a hologram style, in their environment. A good user interface design will include how you see each other, when you see each other, what interactions between people are possible.

The other challenge connected with this is loading all the streaming data into unity, so the team should include at least one person who feels they can work with a large amount of mesh data arriving into Unity.

CP 17 – The Teleporter 2: Face stitching

The Teleporter is a new project to create the most accurate version of being in another place.

We combine live 3D VR cameras with holographic displays to create the feeling of actually being somewhere with someone, when you are far away.

This a solution for families separated by distance to see their children. It is a solution for scientists working on remote field ops on the planet or in space. It is a solution for field engineers working in dangerous and remote locations to access support from their whole team.

While working in the Augmented Reality industry, this was the number one requested use case: Remote Help.

Description: Ever noticed how all video chats have the same problem - you never make eye contact? We will use 4 cameras positioned around the face on a helmet, to stitch together a 3D face and then place a virtual camera directly in front of the eyes to make real eye contact with a remote video. This is the most mathematically challenging project as it involves complex 3D geometry in real time to create a real 3D face mesh. The applications are to fix every video conferencing solution.

Building the Interface in Unity 3D, will require a clever design to combine the 3D stitching of the face, with the environment in a way that feels natural and normal.

The future challenge will be to then share the other user back into the environment, so both parties see each other in a hologram style, in their environment. A good user interface design will include how you see each other, when you see each other, what interactions between people are possible.

The other challenge connected with this is loading all the streaming data into unity, so the team should include at least one person who feels they can work with a large amount of mesh data arriving into Unity.

CP 18 – The Teleporter 3: Eye Tracking Optimizer

The Teleporter is a new project to create the most accurate version of being in another place.

We combine live 3D VR cameras with holographic displays to create the feeling of actually being somewhere with someone, when you are far away.

This a solution for families separated by distance to see their children. It is a solution for scientists working on remote field ops on the planet or in space. It is a solution for field engineers working in dangerous and remote locations to access support from their whole team.

While working in the Augmented Reality industry, this was the number one requested use case: Remote Help.

Description: To prioritize video and share experience we need to know where people are looking. The current systems may solve the problem but many are weak and bulky. You will investigate use of current eye tracking systems and feed in their data to Project Team 1, to provide a real-time resolution optimization and Project Team 4, to prioritize what we send over the network.

CP 19 – The Teleporter 4: Network Streaming

The Teleporter is a new project to create the most accurate version of being in another place.

We combine live 3D VR cameras with holographic displays to create the feeling of actually being somewhere with someone, when you are far away.

This a solution for families separated by distance to see their children. It is a solution for scientists working on remote field ops on the planet or in space. It is a solution for field engineers working in dangerous and remote locations to access support from their whole team.

While working in the Augmented Reality industry, this was the number one requested use case: Remote Help.

Description: Use a combination of 4G networks to stream 4K+ worth of data, to give real time video. The project will require bridging multiple network connections, working with experimental 4GX, possibly 5G and other network solutions, bandwidth aggregators like Speedify and more to create. reliable mobile ultra HD (~4K+) worth of data streaming system, on Linux.

CP 20 – The Teleporter 5: User Interface

The Teleporter is a new project to create the most accurate version of being in another place.

We combine live 3D VR cameras with holographic displays to create the feeling of actually being somewhere with someone, when you are far away.

This a solution for families separated by distance to see their children. It is a solution for scientists working on remote field ops on the planet or in space. It is a solution for field engineers working in dangerous and remote locations to access support from their whole team.

While working in the Augmented Reality industry, this was the number one requested use case: Remote Help.

Description: In our office or home to connect with the remote person, we need a high-quality setup. Options include:

Passive Glasses Free Hologram Displays

Meta Augmented Reality Glasses

HTC Vive.

CP 21 - Notion | AI in Strategy

Description: This project aims to develop tools for automating the production of business intelligence from large quantities of text. This project is an industrial project and the business client is Notion. Notion is an enterprise intelligence start up with large financial services clients on board. Notion's project is at the forefront of natural language application and an Al-based approach to enterprise intelligence. This project is a pilot project for the deployment of machine learning and artificial intelligence.

In this project, you will have the opportunity to collaborate with and learn from a core management team of experts from business strategy, technology and customer experience design.

Some of the key areas of investigation are outlined below:

Topic clustering

- Use text-based articles and NLU entity and concept extraction ("nodes") to identify clusters of nodes using machine learning algorithms (for example k-means clustering).
- Calculate the relatedness between, or proximity of, the resulting clusters.
- Identify articles and nodes that are central to each cluster.

Conversational interface

Take natural language input (text or voice to text), analyse for intent and extract concepts and entities. Classify the intent and disambiguate the concepts and entities. Clarify with the user using questions and answers until a confidence level is passed. Confirm with a summary of the requested task.

Automated summaries

Using a group of entities and concepts, a small number of key sentences, and the source articles from where they were derived, generate a short topic summary.

Relevance scoring

Taking clusters of articles, entities and concepts, a set of topic maps, and users ratings and behaviour, train a model to predict the relevance of unseen articles to each cluster. Train another model to predict the relevance of unseen entities and concepts to each cluster.

Key skills

- Firm understanding of machine learning techniques including regression and classification, supervised and unsupervised learning, for model training, clustering and NLU.
- Python preferred.
- Exposure to ML as a service offering such as AWS ML, IBM Cloud, or Google Cloud.
- Interest or experience in deep learning, RNN and other techniques.
- Interest or experience in TensorFlow, Caffe or similar.
- Other useful skills include NodeJS and MVC front-end frameworks.
- Clear communicator and self-motivated researcher able to evaluate emerging techniques and present options to technical management.

CP 22 - An Online Web-Streaming Service for Bitcoin-Exchanges

In recent years, there has been a surge of interest in cryptocurrencies such as Bitcoin from private investors, major financial institutions and governments. Investors require a mechanism to convert between fiat currency and cryptocurrency (and between cryptocurrencies) so several public exchanges (e.g. GDAX) have been created to enable cross-currency trading. Like any exchange, cryptocurrency exchanges have attracted day traders employing a variety of computational finance techniques.

In this project, students will retrieve, store, and visualize data from various cryptocurrency exchanges. Students will also employ data modelling techniques to find suitable data-structures to represent order books of exchanges, will store them in a relational database, and visualize them via a web-interface using java-script.

Required: python, some web-server technologies (preferable the ones based on python) Helpful: git, familiarity with bitcoin, familiarity with stock/currency exchanges. Student cohort: This project is suitable for MIT/MITM students

CP 23 - Visualisation and Prediction for Bitcoin-Exchanges

In recent years, there has been a surge of interest in cryptocurrencies such as Bitcoin from private investors, major financial institutions and governments. Investors require a mechanism to convert between fiat currency and cryptocurrency (and between cryptocurrencies) so several public exchanges (e.g. GDAX) have been created to enable cross-currency trading. Like any exchange, cryptocurrency exchanges have attracted day traders employing a variety of computational finance techniques.

In this project, students will retrieve, store, and analyse data from one or more cryptocurrency exchanges using data visualisation and analysis techniques. Students will also employ various machine-learning techniques to make predictions about price movements using tensor flow, and test their predictions against live exchange data.

No actual trading will take place during this project; although several exchanges offer a "sandbox" environment, which could be incorporated into a project.

Required: python, some knowledge of popular machine learning techniques. Helpful: git, familiarity with bitcoin, familiarity with stock/currency exchanges Student cohort: This project is suitable for MDS students

CP 24 – Education Analytic Platform

Description: To develop a data warehouse to store extract information from an institute 's central database. To develop a tool to automate the preparation of standard reports, as well as an easy way to carry out other adhoc comparative analysis as well as data visualisation.

Preferred skills: good programming skill, database and D3; ideal to know some data mining techniques.

CP 25 - Connect

This project is part of a Multi-Phase rollout and builds upon an existing capstone and working prototype. This project is a web (and potentially mobile) application.

Description

Information submitted by the project proposer/supervisor will go up on web site; (Please keep as closely to this model as possible as all project proposals need to be uniform to go onto the web site):

Connect is a mobile-and-tablet app, with elements of a blended virtual learning environment. This app will change the way we communicate and engage with prospective future undergraduate students in a modern channel they use and understand. It will complement, streamline and consolidate our outreach and marketing efforts, and enable us to achieve broader reach than physical on campus activities. It will provide an ongoing communication and engagement platform allowing further contact following an interaction, either directly or by referral (to NCSS, GPN, First Robotics, etc). It will contain content to help students understand the possibilities our engineering and IT degrees afford. It will alert them to events they may be interested in and enable event

registration and attendance tracking. It will track and reward user interactions.

We will gather user contact and behavioural data (learning analytics) which we will mine.

From a business perspective, Connect could be viewed as a CRM and loyalty program.

Multi-year strategy:

Engage Attract Students and help them understand what engineering and it and University of Sydney is all about in a game. years 8-10 Nurture Maintain and build the relationship over time through news, insights, events and programmes. years 9-12 Prepare Help students get ready so that they can start at the University and hit the ground running from day 1. year 12

Years 8-10

Attract and engage through a mobile game

Gather user and behavioural data

Identify potential students based on game play and other criteria

Years 9-12

Introduce and facilitate other engagement opportunities, open days, events, programs

Maintain communication through news and updates

Include gamification elements to link back to the game and maintain engagement

Years 11-12

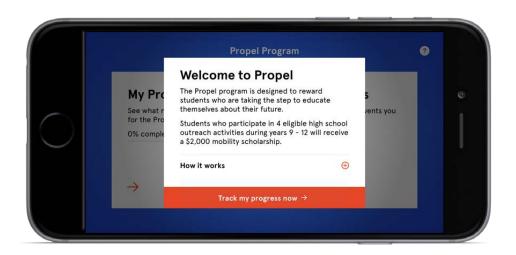
Encourage visits to the University

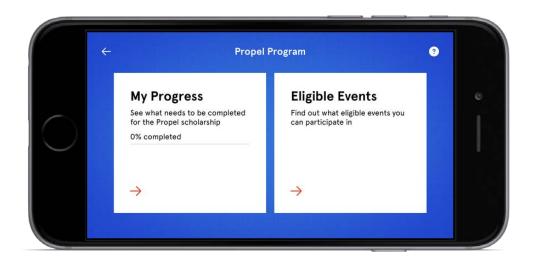
App transforms into utility and sales tool

Maintain direct communication

Continue to include gamification elements, keeping it fun

Connect incorporates 'Propel' - https://sydney.edu.au/engineering/propel











Propel is a sub-app and has its own specific requirements.

How it works:

Students who participate in 4 eligible high school outreach activities during years 9-12 will receive the Propel Mobility Scholarship.

Student attends
min 4 eligible
outreach activities
between year 9-12

Student completes
HSC & receives
offer from Faculty
of Engineering & IT

Student enrols &
commences UG
program

Student is eligible
for \$2000 mobility
scholarship

Progress to date:

We have already built a working prototype of Propel within the Salesforce CRM platform, with integrations to other 3rd party web API's including MailChimp, EventBrite, Formstack, Google Drive and others. Our custom CRM captures the following:

Min Specs.

Name

High School

Year at School

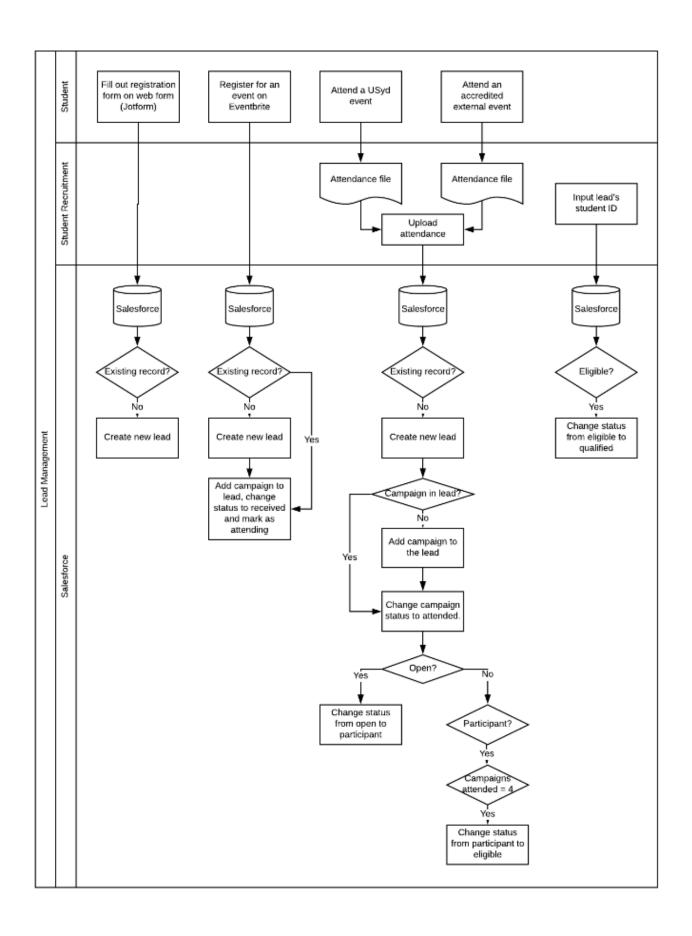
Email address

Activities completed (and points gained, akin to a simple loyalty program)

Calendar of upcoming activities with links to registration

SYDNEY Connect Project - MVP Rollout Plan





What we need for this capstone:

This project is part of a Multi-Phase rollout. With CRM implemented, we need a web application to interface with the CRM via API and build a simple portal for Propel members to join and/or login to track their progress and view their membership details / status.

The purpose of this project is to prepare a detailed functional specification for the web application client-server platform (client-server-server, database, and potentially with advanced analytics) and working prototype that underpins the game design brief and interfaces with the Salesforce CRM.

We require a technical design for the platform, preferably including:

Client environment built upon Unity3D, deployable to iOS and Android for mobile and PC and Apple desktop/laptop

User database with integration to Salesforce CRM via API

Back end architecture on AWS ec2 and S3 (or suitable alternative)

Build scripts for (AWS) architecture

Investigation of suitability of AWS GameLift - https://aws.amazon.com/gamelift/

Github repository to support dev/staging/prod environments

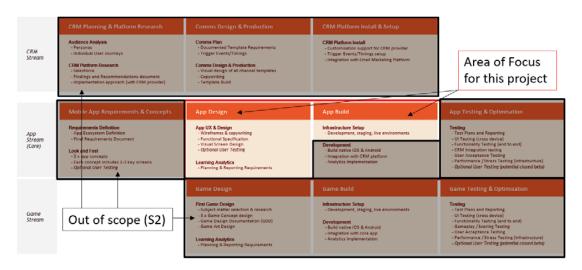
Web Services (REST?) + mobile game

Integration to Facebook gameroom

Learning analytics – snowplow analytics https://snowplowanalytics.com/ or suitable alternative

SYDNEY

Connect Project - MVP Rollout Plan



Deliverables:

Data model documented. Clickable prototype or better developed.

Specific requirements/skills (if applicable):

The target CRM is Salesforce. The target mobile development environment is Unity 3D. A role based permissions system will operate to control access to the relevant functions.

Able and willing to meet with partner is necessary for satisfactory completion of the project; English language competency; Experience in modern programming languages, namely C/C++, JavaScript, SQL; Experience in iOS and Android mobile development environments; Experience working with UNIX Operating System, Windows Operating Systems. Experience establishing and maintaining code repositories.

Reference Material/Links (if applicable):

https://sydney.edu.au/engineering/propel

Detailed wireframes have already been created and will be supplied to this project for reference, extension and improvement.

A prototype Unity 3D game has also been developed that will be available to the project team.

CP 26 - DATA EXTRACTION FROM DBLP: COMPUTER SCIENCE BIBLIOGRAPHY TO IRMA

1. BACKGROUND

dblp: computer science bibliography is an on-line service that provides open bibliographic information on major computer science journals and proceedings. dblp is a joint service of the University of Trier and Schloss Dagstuhl. It has evolved from an early small experimental web server to a popular open-data service for the computer science community. dblp's mission is to support computer science researchers in their daily efforts by providing free access to high-quality bibliographic meta-data and links to the electronic editions of publications.

Researchers largely within the School of Information Technologies (SIT), anecdotally report a very high number of their outputs via dblp, as do a growing number of researchers in the School of Electrical and Information Engineering (EIE).

2. CONTEXT

The University of Sydney collects and uses data on publications for a number of purposes: populating academic profiles (which impacts on the University's research profile and academic progression), internal benchmarking of School and Centre research activity, as well as to highlight research successes, and for inclusion in the Excellence in Research for Australia (ERA) return and new Engagement indicators. It is envisaged that the expanded range of outputs collected will reflect a greater range of and diversity in the research undertaken by University of Sydney academics.

The Research Publication Collection has undergone changes following the demise of the publications component in the Higher Education Research Data Collection (HERDC), now known as the Sydney Research Collection (SRC). In the case of SIT specifically, researchers frequently direct the Research Support Officer (RSO) to their dblp page in order to access and extract their publication outputs, for entry to the Integrated Research Management Application (IRMA), essentially copying the information held in one database into another. This process of extraction is currently a very manual process, and one that could be done more effectively automated.

3. OBJECTIVE

To automate dblp data extraction to create IRMA records for inclusion in the SRC. This will have the three-fold effect of (1) ensuring that all eligible SIT and EIE records are captured, (2) increase the overall size of the research collection for the University, (3) minimise manual intervention of effort.

4. SCOPE

Work with select SIT and Research, Reporting, Analysis, Data and Systems (RRADS) staff to develop ongoing procedures by which relevant bibliographic information can be extracted from dblp, and imported to IRMA by macro.