Part IB Group Project Design Briefs 2017

Air Quality Radar
Client: Matthew Smith, Microsoft
Rain radar apps are familiar and popular - allowing people to plan their activities by approaching rain. This project will create the first 'air radar' that uses monitoring of multiple air quality parameters to predict air movement and local variations in pollution. We can provide access to recordings of air quality data streams from multiple locations around Cambridge, and also the Microsoft Cortana Intelligence Suite under Azure. You should use machine learning techniques to predict what happens next. This will involve collecting and integrating other data such as time of day, traffic movements, temperature, wind speed and other factors. Your current situation maps and future trend predictions should be useful not only to cyclists and residents, but to business analysts and city planners concerned with future environmental health.

Antarctic Chasm One
Client: Matt Polaine, British Antarctic Survey
The Brunt Ice Shelf in the Antarctic is growing a massive chasm at the rate of 1.7km a year, that may lead to an area of nearly 1,000 sq km breaking off and taking the Halley VI base with it. Your task is to make a 3D immersive visualisation of this growing chasm, allowing viewers to descend into it in the way that BBC's Peter Gibbs did on camera (goo.gl/48cVds). You will have access to aerial scans, photographs, video, multispectral satellite imagery, ground penetrating radar of the crack and other data from technical specialists at British Antarctic Survey (BAS). You will use these to extrapolate places where data is missing, and also to model the growth of the chasm for a speeded-up view of the future. For a full immersive effect, spatialized audio would add further realism.

Auto-Archive
Client: Richard Watts, Kynesim
Everybody has a different approach to organising email and documents, but nobody has time to do it properly. Your challenge is to create a new kind of mail-plus client that uses unsupervised machine learning techniques to help run a small business. It should recognise clusters of email that the user can deal with as a batch, using the IMAP protocol to create and populate archive folders on the mail server, or (if feeling brave) automatically reply to them. Since much email content simply repeats existing documents (whether as attachments or with repeated text), your system should also scan for valuable data that is *not* in the email archive - for example as disk files or on the local network. These ought to be archived too, and somehow associated with relevant email. But take care not to waste space by backing up too many similar copies - prioritise information that seems to be important and unique.

Class-y Action
Client: Yuki Cheung, Cambridge Legal Risk Analytics
Many companies cause harm and damage to customers in small ways, but it's not worth employing a lawyer unless a large number of customers aggregate their compensation claims. The ability to pursue this kind of class action was introduced to British law in 2015. There is a business opportunity in identifying which companies can be most profitably sued. Thousands of people will thank you for it, and the long term effect would be to improve products and services for everyone. You will need to use social media - Twitter, Facebook, Instagram, and anything else that comes to hand - to work out which companies are most likely to lose a class action suit in relation to their products or services. Your client can provide expertise in data mining of legal cases, in order to optimise training of your statistical models.

Drone Safety
Client: Lawrence Gripper, Altitude Angel
When drones fly beyond direct human control, they must plan routes that meet operating and safety criteria. For example, a delivery drone needs to make a number of visits within time constraints, while public safety and policing may require repeated surveys of a specific area. Different drones have different capabilities; some can hover, others cannot, while others require more regular charging. Routing drone traffic safely is also challenging: avoiding airspace restrictions, manned aviation and ground hazards, like crowded roads and pedestrian areas. Your goal is to create a cloud-based air traffic control system that registers flight plans, modifies them in response to hazards based on safety data maintained by Altitude Angel, and makes route alterations in real-time to avoid other drones, as well as gatherings of people identified automatically from real-time Cambridge ground data including WiFi network registrations and public transport locations.

Eco-Location
Client: Tim Wilkinson, UNEP World Conservation Monitoring Centre
Wildlife and critical ecosystems around the world include over 230,000 legal 'protected areas', ranging in size from Coldham's Common in Cambridge https://protectedplanet.net/555561770 to the Serengeti https://protectedplanet.net/555570276. Many people from tourists to researchers and conservationists need to know about these, but there is no quick way to find out whether you are near one. If one is nearby, what can be found there? Barren scrubland or dense forest? Could a park manager see evidence of agriculture where there shouldn't be, or a visiting scientist learn that the area has been flooded? Your task is to build a mobile application that enables the user to quickly orientate themselves with nearby protected areas, and deliver contextual data by accessing other services. You should specifically include data from the European Space Agencies Global Land Cover Layer to show the percentage of cover types within the protected area (e.g. 40% Mosaic vegetation, 60% Mosaic grassland), and you may include other information such as geolocated photos from Flickr, or species from the IUCN red list of threatened species.

Energy with Social Conscience
Client: Yuichi Abe, Informetis
Although electricity generation in the UK relies increasingly on renewables, these only provide low baseline capacity. To meet demand spikes, electricity companies must burn gas and oil. In principle, smart meters would allow a social consensus to save the planet – we just can't afford 10 million people getting out of bed to switch on their kettles at the same time! Your client has smart technology already detecting household current consumption. Your task is to create machine-learning appliance detection, real-time data infrastructure and front-end services that could minimise fossil fuel consumption across the whole country, while keeping everyone happy through consensus pricing markets and social media.

Financial Battlefield
Client: Ewan Kirk, Cantab Capital
Attempts to make financial graphs or political trends more entertaining are often disappointingly unimaginative. We have seen too many VR mountain ranges and city skylines. Your task is to communicate the excitement of financial trading data by using the latest gaming technology. We'll provide powerful GPUs for you to use with the Unreal game engine and a budget for off the shelf game art and graphics assets. You will render large scale investment data based on thousands of data streams as competing armies of simulated characters whose bodies, weapons and AI behaviour are determined by key market attributes.

History Phone
Client: Ina Pruegel, Cambridge Museums
There are historic objects around the Gates building that have stories to tell - if only they could speak! Imagine walking up to one of the display cases (like the one to the right of the front door), and using your phone to start a chatbot session as if you were speaking to the object itself. We will provide Bluetooth location beacons so that an Android phone can detect which object it is near. Your application should respond to questions, and maintain a conversation, with style and content that is customised to each object. The project team will have a chance to visit some of the famous museums elsewhere in Cambridge, to learn how curators design interactive content.

Hololens Escape Room
Client: William Bakker, IMC
Escape Room facilities are popular for student social outings and professional team building, but the props and decorations needed to create an engaging theme are difficult to move from place to place. Your task is to create an authoring tool for Hololens that can be used to turn any physical space into an escape room, using only simple portable apparatus. Using holographic projections, clues and puzzles should be provided that the player must navigate in order to unlock the exit. The Hololens can be connected to a screen to share the vision of the wearer to the other players. IMC will provide one Microsoft Hololens.

Learn to be an Alien
Client: Edward Ashton, Frontier
Procedural landscape generators are often used in videogames, for example to create new unexplored planets. What would a procedural creature generator look like? Recent experiments demonstrate that a neural network can learn how to control robot arms, if it has enough to explore with. Your task is to create a 3D editor that allows users to define a new kind of alien creature - legs, claws, sensors whatever - which then learns how to behave through experimentation with the laws of physics applying to its own body in a simple environment (furniture, food etc).

Micro-Volunteering
Client: Michael Elliott, JPMorgan
Many charities rely on volunteer networks, but small local charities sometimes don't have the staff to maintain those networks. On the flip side, many people would like to volunteer to help a local charity but aren't able to commit a set number of hours per week. Micro-volunteering allows individuals to offer up their skills on an ad-hoc basis and for charities to take advantage of that, and is extremely valuable to smaller charities or community organisations. Your task is to design a micro-volunteering exchange, that can match volunteers and their skills to opportunities and needs in their local area. Locality is critical in micro volunteering, perhaps a graph database (like neo4j) would be interesting here. Remember also that skills and needs often involve different terminology - how will your system understand that the skill of 'simple wiring jobs' should be matched to the need of 'broken bulb in shelter'?

Neural Guide
Client: Jan Kis, IMC
The NeuralTalk Model Zoo provides pre-trained deep neural net models that can be used to generate text descriptions of unseen images. In principle, such sentence generation could be used to assist blind people, allowing them to point their mobile phone at a scene, upload the camera image to a server, and receive the predicted text as synthesised speech. The results are likely to be far less reliable than for images from the NeuralTalk demo database, so you will probably have to provide audio or tactile feedback on image quality, prediction confidence, and guidance to help the user point the phone in a more productive direction.

Personal Reality
Client: Mark Ogilvie, Jagex
It is often hard to 'read' people in social gatherings. But virtual reality can help us understand the relationship between different personalities and social situations. Your task is to create a virtual reality environment, in which people can create models of their perceived self, using abstract objects of various shapes, colours and sizes to mimic how their self-perception changes under various influences. You might consider aspects of the five-factor model of personality, to assess how these are affected by the personality of other people in the same space. The result will be a new kind of social gathering, allowing people to meet and interact without the guessing games that are usually necessary to interpret body language and subtle social cues. You might add ways to model other individuals, showing how our personal observations differ from other people's models of themselves, and shape the data to deliver meaningful trends and conclusions to sociologists.

Retail Startup Automator
Client: Harry Collard, The Hut Group
Online retailers invest heavily in reformatting information from product catalogues into their own databases, and this is a barrier to entry for new businesses. But it should be possible to automate the startup process, by using artificial intelligence to extract the common data structure for standard products that get sold by dozens of different retailers. You must first create a machine learning algorithm that can automatically infer page templates by comparing product detail pages for the same product being sold on a variety of existing retail sites. Those templates can then be applied to extract the semi-structured data from retail sites for further products that have not been seen in the training phase. The resulting database will then be automatically published as a brand new shopping site, offering instant startup at the press of a button, allowing specialist entrepreneurs to cater for niche retail markets such as toothpaste-compare.com, muesli-and-sandal-world, or toysforhamsters.

Science for AD2500
Client: Chris Fell, Cambridge University Press
Everyone knows that peer-reviewed publication is the gold standard for scientific facts. But it is too slow for the millennial generation, who prefer fresh opinions and data 'verified' by Facebook likes, GitHub pull requests, up-votes, blog comments or TED talks. We need a new model for science that is agile and open, but also solid enough to last another 500 years. Your task is to prototype the next Royal Society or Cambridge University Press, providing democratic public access for millenials, rapid quality control that would satisfy the next Isaac Newton, and PDF for permanent paper archives to survive the Apocalypse of 2499. Elegance and attention to detail are critical on both paper and screen - if it looks like Buzzfeed, nobody will believe it was the Transactions of the Royal Society. Don't forget the essential academic attributes of accurate and secure certificates for time, authorship, and reference to prior work.

Simulation and Warning for Cyclists
Client: Andy Fawkes, Bohemia Interactive Simulations
Too many cyclists are injured or killed by collisions with trucks and buses, particularly at junctions. We have data that could help. All Cambridge buses have real time GPS tracking, and cyclists increasingly wear GPS-connected devices like the Pebble smartwatch. We propose a single online simulation and warning system, accessed from screens and that exploits both historic and real data, to help drivers and cyclists avoid accidents. Using historical data the simulation can be used to give drivers and cyclists an appreciation of the dangers and provide general warnings of dangerous traffic conditions for the time of day. The real time data can be fed back to cyclists out on the streets by the warning system, with coded buzzes on their wrist helping them plan routes and avoid any danger ahead.

Surgery in the Cloud
Client: Rob Hague, Cydar
New hospital treatments often involve as much computer imaging as they do scalpels. Local company Cydar manages high volumes of images and video data in the cloud using Amazon Web Services, and can even track real-time video from an operating theatre, but configuring those data flows is time-consuming and confusing. Your goal is to create a video cloud service management tool inspired by Stuart Taylor's VPlay system for live video remixing (http://vimeo.com/2738692). Users should be able to review and modify the data flow architecture on AWS virtual machines, while also getting a live view of the current streams and archive content to see where the problems might be.

Survey Swarm
Client: Ashkan Tousimojarad, ARM
Autonomous robots often have to work with an incomplete model of the world, and high-end devices such as the Dyson 360 Eye do their best to construct accurate image maps of the rooms they work in. But if there are many cheap robots, they can collaborate in a swarm to share information. Your task is to create sensing and coordination infrastructure so that a swarm of Pololu 3pi robots can use their reflection sensors to work out the overall design of a black and white floor pattern that they are moving over. To start with, they will have to use dead reckoning to collect single data points for the shared model. Eventually, they can reduce uncertainty (for example, if some robots seem less precise than others) by delegating individual robots to go and make further observations, navigating by the map of previous attempts.

The Deep Learning Society
Client: Oli Bage, Morgan Stanley
Technology companies invest billions in self-driving cars and self-playing computer games, but surprisingly little in real social problems. Your task is to use the latest deep learning technologies to create an intelligent social work assistant that can recognise and act in situations of real need. Your client will provide a GPU-accelerated system suitable for use with deep learning frameworks like Google TensorFlow. You will train it using data from online social networks such as MumsNet, to recognise and anticipate situations where people are going to use words like 'hopeless', 'depressed' or 'suicide'. With the help of deep learning, even a simple Bag of Words, together with metadata such as time of day, location and comment feedback will be sufficient to recognise trigger conditions in large data sets and mobilise assistance.

VR Algorave DJ
Client: Matthew Johnson, Frontier
Why can't a DJ be more like an orchestra conductor, remixing instrumental sections and adding new expressive content, instead of simply selecting from a library of prerecorded tracks? In the future this will be possible, using gesture controllers in virtual reality. Your task is to create a VR space in which the DJs of the future will edit and configure algorave-style music synthesis programs. You can use APIs for the Sonic Pi music language from Cambridge's Sam Aaron as a back end to produce professional-standard musical results. We can provide an Oculus Rift and Myo gesture control armband for an immersive experience (and some professional audio gear for the demonstration day).

Who's at my party?
Client: Richie Jones, Boeing
It's always disappointing when you've had a great night out, but you can't remember the details the next day. It would be creepy to photograph everyone, but simple to take a short video with a quick scan of the room. Your goal is to build a social media platform with video upload as the main user interface. Your back-end processing will apply open source machine vision libraries such as OpenCV, OpenFace and others to capture faces, match them against known friends, see who is talking to who (from head orientation) and so on. You should package the results as a fun and engaging social media site. Consider making a plugin using the WeChat Video API - or perhaps this could be the next Facebook!