Centre for Work Health and Safety

Project Brief – iLab 2023

Can you introduce yourself?

- The Centre for Work Health and Safety's (WHS) mission is to deliver solutions that prevent harm in the workplace. It's as simple as that – and we do this in several ways:
    - the traditional research process partnering with academics,
    - the less traditional Research & Development (R&D) pathway of partnering with developers and innovators,
    - and through data science, thanks to an in-house team of experienced data scientists, we develop data tools and predictive models for SafeWork NSW regulators in providing WHS intelligence from internal and external data sources.
- If you check our website https://www.centreforwhs.nsw.gov.au/ you will see a number of safety innovations, we have delivered in the last few years such as:
    - the world-first real time silica detector,
    - a WHS Management Tool for AI - to assess and manage risks in working with AI systems.
    - working safely with collaborative robots – a set of guidelines for Safe Collaborative Robot Design and Implementation.
- 'Safety innovations' are our end goal with everything we do. We identify a need and bring together industry, government, academic, and development partners to come up with the best way to fill the gap and ultimately prevent harm.

What is the challenge you have put forward for iLabs?

- Simply put, we provide the students with a data science problem, and we expect them to come back to us with an innovative solution.
- The challenge is part of a new project, which aims to provide data intelligence to SafeWork NSW - the data science team as well as regulators for WHS in NSW.
- The need here is for the regulator to improve its use of data intelligence to better inform decisions at a strategic and operational level.

- The focus of this iLab challenge we are proposing is,
    - to retrieve a list of businesses in NSW (in the form of ABNs) by querying specific work activities or business services of interest (e.g., bricklaying, engineered stone).
    - As a starting point, we are providing students with a set of work activities or business services relevant to the two SafeWork Regulatory Priorities.
    - Our expectation is that the iLab participants will come up with an intelligent way to lookup businesses in identifying ABNs from public data, based on keywords, that is of work activities or business services, relevant to the businesses. The lookup tool should cover public information that could be obtained from the businesses' websites, social media pages, etc. Further, to allow options to refine the search (e.g., by industry, location / postcode, etc).
- What does success look like? There are two key things when defining success here:
    - Quality = we need to be able to collect the relevant info out there, and analyse it to create meaningful intelligence.
    - Sustainability = resource friendly – here we would love to see a process of data collection and analysis that is less manual, and easier to achieve.

Why is this important?

- The student challenge is central to the Business Intelligence (BI) function in SafeWork NSW (SafeWork).
- The intelligent business lookup solution would allow the Centre to link the business to other data sources (e.g., SafeWork incident notifications, workers compensation, etc) and provide intelligence to SafeWork based on specific work activities or business services.
- The expectation is that, through this, SafeWork will be better informed on at risk businesses around regulatory priorities, for example safety around moving plant, exposure to harmful substances, falls, etc.
- If SafeWork is better informed, it can allocate resources to high-risk industries and businesses more efficiently.

- This student challenge will ultimately help prevent workplace harm for our workers in NSW.

Do you have potential directions or inspirations for student teams to take up?

- All we are providing to the students is a starting point with a list of work activities or business services relevant to the two SafeWork Regulatory Priorities, our 2023 RADAR report to provide some domain-specific knowledge, and our expectations.
- What we would love is for them to take this away, bring their innovative and challenge thinking into it, obviously their data science skills and surprise us with something 1) that is of quality = relevant = insightful and 2) sustainable.

Do you have specific resources or domain knowledge that you would like to share with students?

- Students can get a feel of this by looking into the Work Health and Safety Roadmap published by SafeWork NSW (https://www.safework.nsw.gov.au/roadmap), and the SafeWork Regulatory Priorities for 2023 (https://www.safework.nsw.gov.au/about-us/safework-nsw-regulatory-priorities-2023) - which outlines the priority areas we are after.
- You may also look into the Centre for WHS's R&D Blueprint which outlines who we are and what we do (https://www.centreforwhs.nsw.gov.au/about/research-blueprint).
- Students will have access to the National WHS Radar reports, an initiative by the Centre that regularly delivers the latest insights on work health and safety.
- Students will also know more about our data science work, for instance we developed a 5-star compliance rating for all businesses in NSW, which is estimated using predictive modelling and machine learning, and gives SafeWork a way to better allocate resources to those businesses the most at risk.
- Students can learn more about the Centre's projects on our website. At www.centreforwhs.nsw.gov.au.

Are you looking for specific skills and attributes in student teams that work on your project?

- We're eager to work with students who are open to trying new methods and challenging the status quo. They should be excited by new possibilities and not afraid to ask questions and seek answers to tricky problems.

- Most of all, we want to work with students who aren't afraid to suggest smarter approaches, beyond the limitations of existing tools and technologies.

- We pride ourselves on taking a multi-disciplinary, collaborative approach in the Centre, so this project will suit students who will thrive while working across cross-functional teams.

What skills can MDSI students develop during the project? How do you think it can help prepare students for their careers/ how does it apply to the wider data science community?

- Students will experience what data science means in the real world – the fun part is the analytics part, but as a data scientist, you are also often best placed to drive and improve data collection processes > and improve the quality and format of data you will analyse.
- Students will explore real-world data which are a lot messier than data they typically get through a curriculum.
- They'll also have the opportunity to develop a range of practical hard and soft skills that comes with the role of data scientist in any organisation, such as:
    - Data sourcing and processing
    - Data analysis, modelling, and visualisation
    - Use of data tools and programming languages
    - Git, which is a source code management tool
    - Problem solving
    - Critical thinking
    - Collaboration.

Anything else you would like to add?

- At the Centre for WHS, we really appreciate the value of the unique perspectives students bring.
- Last year, we launched our own Academic Studies Program, which brings funds, resources and opportunities to increase collaborations with students and bring their innovative and challenging thinking into WHS.
- We are working with 7 major Australian universities and so far, we have had 53 students doing their internships/university projects with the Centre.
- There will also be networking opportunities with other students, as well as research and data science professionals from across Australia.

Materials

Students are provided with the following files:

- Regulatory priorities across risk types and sectors SafeWork NSW we will focus for 2023 (SafeWork NSW Regulatory Priorities 2023.pdf).
- Example work activities / keywords from two regulatory priorities - Safety around moving plan and Exposure to harmful substances (Example Search Keywords.pdf).
- The latest report on pertinent issues, trends and insights related to the current state of play regarding WHS in and outside of Australia, as well as potential WHS issues in the future world of work (WHS Radar NATIONAL - April 2023.pdf)
- Source code in Python for text analytics using Reddit data (TextAnalytics_Reddit.zip). For examples of text analytics – data sourcing and processing, word clouds, sentiment analysis, concordance, topic modelling.
- WHS Academic Stream Confidentiality and IP Deed Poll_iLab_2023.docx – to be signed by all students.