DataBytes

COMPANY MISSION

DataBytes is a multi-faceted corporation that focuses on creating groundbreaking applications and platforms driven by data. The company's mission revolves around utilizing the latest advancements in artificial intelligence, data science, and digital finance to address global challenges. Their expertise spans numerous sectors such as environmental sustainability, personalized education, online advertising, fintech, open banking, consumer data rights, data portability, retail, e-commerce, and payments.

COMPANY DIRECTORS

- Trimester 1 – Mr Jesse McMeikan (Manager, Industry Projects)
- Trimester 2 – Dr Sergey Polyakovskiy (Senior Lecturer, Computer Science)
- Trimester 3 – Dr Elicia Lanham (Senior Lecturer in Information Technology)

COMPANY STRUCTURE

The DataBytes company is currently operating two active projects:

- Project Echo
- DolFin

These projects share a unified workspace on Trello and GitHub, bringing about centralisation of tasks and software results, which in turn cultivates a shared identity among DataBytes team members.

Team members associate with our wider organisation and contribute to any or all products within our active project communities. Each DataBytes team member has a main project affiliation, along with possible additional contributions to secondary projects. The primary project sets the focus for each member during the trimester, facilitating substantial and enduring contributions.

PROJECT OVERVIEW

Project Echo (Pause in Trimester 3 2023)

Project Echo focuses on creating advanced bioacoustics tools, specifically audio classification tools, to assist conservationists globally. The project aims to discover, monitor, and track endangered species and their predators in a non-invasive manner within their natural environment.

Initiated in Trimester 3, 2022, Project Echo aims to aid conservationists by developing AI/ML tools for bioacoustic analysis. By Trimester 2, 2023, the team successfully developed a proof-of-concept audio classification model, packaged it for Python, and created an API for audio file predictions. The project has made significant strides in audio processing and aims to further enhance its capabilities in the coming trimesters.

Product Owner: Dr Michael Hobbs (Associate Head of School Geelong)
Industry Partner: Hannah Wiggs (Project Officer – Wild Otways | Corangamite Catchment Management Authority)

Tech Stack & Key Resources:

- Tensorflow / Keras
- Tensor Serving
- Librosa
- Fast API
- MongoDB
- Bootstrap
- Docker Desktop
- Google Cloud Platform
- Node JS
- Redis
- Yamnet
- JWT

Project Echo Showcase (T2 2023)

DolFin

DolFin, a pioneering fintech solution, was conceived with the ambition to transform the traditional banking landscape. By amalgamating fragmented banking operations, it offers a unified and intuitive platform. This innovation harnesses the power of open banking data, presenting users with a holistic view of their financial standings, enriched with detailed insights and visual analytics.

Building on its foundational goal of streamlining financial management, DolFin has made significant strides since its inception. By 2023, the platform seamlessly integrated an AI-driven prediction model and introduced a chatbot, aptly named "DolFin", to dispense tailored financial advice. These enhancements underscore DolFin's commitment to continually refine the banking experience for its users.

Product Owner: Dr. Yang Li (Data Scientist at ANZ)

Tech Stack & Key Resources:

- Flask (Software)
- Jupyter Notebooks (Software)
- AWS Cognito (Software)
- AWS S3 (Software and Hardware)
- AWS Lambda (Software and Hardware)
- AWS SageMaker (Software and Hardware)

Project DolFin Showcase (T2 2023)

New Discovery Project - Privacy Technology for Financial Intelligence

1. Background

Complex financial crimes are frequently orchestrated by well organised and experienced criminal groups. These illicit activities often encompass multiple business entities and individuals, with financial activities spanning various countries, effectively concealing the illegal operations through layers of subterfuge. Money laundering is one of the most severe financial crimes, exacting an annual economic toll on Australia estimated at 10 − 15 billion dollars.

Presently, the responsibility for combating financial crime extends beyond the purview of a single government agency. Instead, it relies on a collaborative effort between various government bodies and financial institutions to systematically counter these financial offenses. This collaborative approach offers a mutual benefit, especially when it comes to exchanging vital information during the investigative phase. By doing so, the involved agencies and organisations can harness powerful computer algorithms and models for a more thorough analysis of hidden illegal financial activities within millions of daily transactions.

Nonetheless, the practice of sharing information also carries inevitable privacy and security risks. As a result, governments and financial institutions worldwide have established data privacy regulations and policies to safeguard individuals' data, restricted or prohibited certain data to be shared or matched between agencies and industries. This has led to ongoing discussions about data privacy and security throughout collaborative financial crime investigations. For instance, when sharing AUSTRAC's list of suspicious individuals with banks to access their financial records, there exists a risk that banks might inadvertently alert the individuals under investigation, which is also known as tipping off. From a customer's standpoint, even if AUSTRAC's final investigation exonerates them of any wrongdoing, a bank may decide to terminate its services to a suspicious individual, thereby implementing a risk-avoidance measure that can be construed as discriminatory.

2. Project Aim

The project aims to explore prominent privacy technologies commonly researched and implemented in the field of
financial intelligence. Some of these technologies include, but are not limited to, the following.

- Secure multiparty computation (SMC or MPC) - A protocol that allows multiple parties to work together to achieve the common goal, whilst keeping secure of the information that are not meant to be shared. For example, private set intersection (PSI) only releases the common entities between two data bases.
- Homomorphic encryption (HE) - An advanced cryptographic technique that allows certain arithmetic operations (e.g., addition and multiplication) to be performed on encrypted data (i.e., ciphertext) without needing the secret key. A common scenario of HE is to encrypt the sensitive data then outsource the encryption to a third party with computational or analytical advantages.
- Differential privacy (DP) - A statistical method to release data base query outcomes with calibrated noise to guarantee no individual information can be reverse engineered from the outcomes. For example, carefully calibrated noise is added to the average income of a group of people to stop individual income to be inferred, should another query is allowed to exclude a certain individual from the data base.

There are other popular privacy techniques that are often discussed and employed in financial intelligence, include hashing techniques, federated learning and so on.

Product Owner: Dr. Yang Li (Data Scientist at ANZ)