



Tecnológico de Monterrey

Instituto Tecnológico y de Estudios Superiores de Monterrey

Momento de Retroalimentación: Reto Datos

TC3007C.501 Inteligencia Artificial Avanzada para la Ciencia de Datos II

Profesores:

Iván Mauricio Amaya Contreras

Blanca Rosa Ruiz Hernández

Félix Ricardo Botello Urrutia

Edgar Covantes Osuna

Felipe Castillo Rendón

Hugo Terashima Marín

Equipo 2

Integrantes:

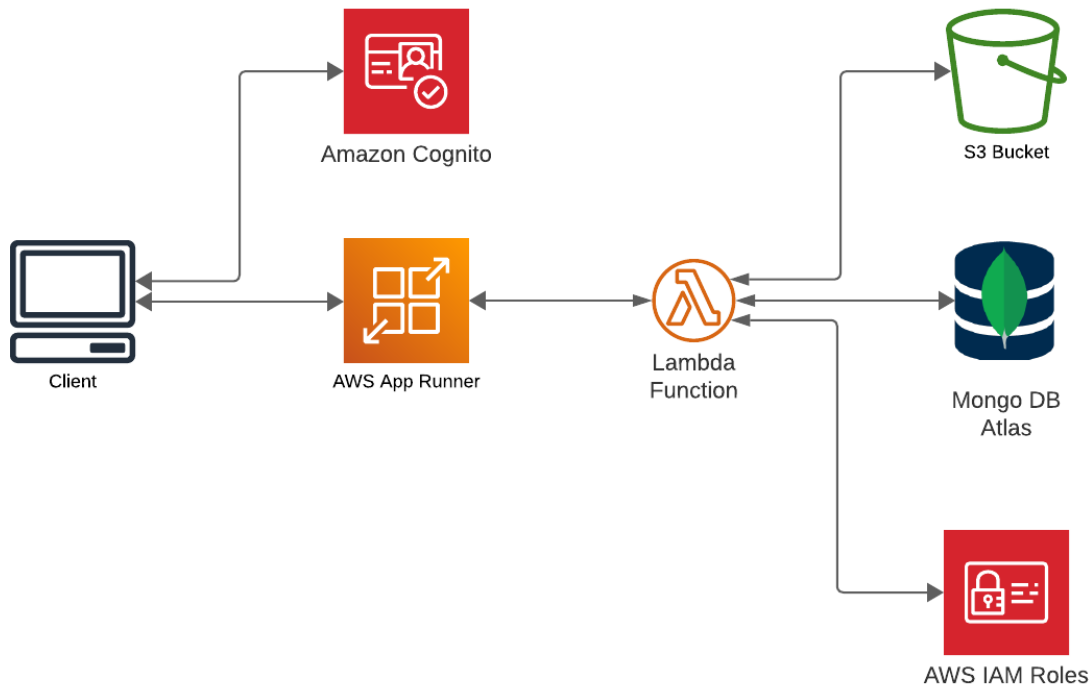
Luis Ángel Guzmán Iribe – A01741757

Julian Lawrence Gil Soares – A00832272

Alberto H Orozco Ramos – A00831719

11 de Octubre de 2023

Diagrama de arquitectura de software



Nuestro sistema consta de 3 componentes principales:

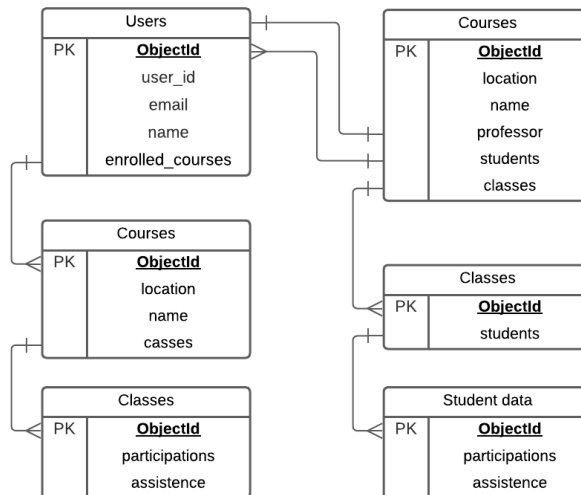
1. **SPA:** sirve como interfaz de usuario y le permite a este interactuar con el sistema. Esta SPA emplea AWS Cognito para manejar la autenticación de usuario y es desplegada utilizando el servicio de AWS App Runner.
2. **API Rest:** desplegada con AWS Lambda se conecta a los servicios de AWS IAM Roles para el manejo de credenciales de AWS y una Bucket de S3 para almacenamiento de archivos, principalmente relacionados a reconocimiento facial. También actúa como controlador de una base de datos de documentos en MongoDB.
3. **Document DB:** Base de datos de documentos que utiliza el servicio de MongoDB Atlas. Funciona de manera independiente al resto de servicios.

Todos estos servicios requieren de credenciales de acceso privadas para acceder al resto de la estructura del sistema. Garantizando seguridad para evitar el acceso de actores maliciosos a la información almacenada.

Los servicios empleados (AWS y MongoDB) emplean protocolos seguros, encriptación y restricción de IPs para el acceso a sus servicios, con la finalidad de garantizar que solamente los actores necesarios tengan acceso a la información trabajada.

Almacenamiento de datos

Empleamos una base de datos de documentos con formato JSON. Esta base de datos se divide en colecciones, y sigue la siguiente estructura.



La misma estructura representada como un archivo JSON aparecería de este modo:

```

{
  "users": [
    {
      "id": "",
      "name": "",
      "email": "",
      "user_id": "",
      "enrolled_courses": {
        "course_id": {
          "classes": [
            {
              "assisted": true,
              "participations": 0
            }
          ]
        }
      }
    }
  ],

```

```

    "courses": [
      {
        "name": "",
        "location": "",
        "professor": {
          "id": "",
          "user_id": "",
          "name": "",
          "email": ""
        },
        "students": [
          {
            "id": "",
            "name": "",
            "email": "",
            "user_id": ""
          }
        ],
        "classes": [
          {
            "student_id": {
              "assisted": true,
              "participations": 3
            }
          }
        ]
      }
    ],
  ]
}

```

Enfoque de Big Data

No consideramos que este proyecto requiera de un enfoque de big data al momento de manejar los datos, por las siguientes razones:

1. **Volumen de datos manejados:** Un enfoque de Big Data es útil cuando se manejan Terabytes o Petabytes de datos, en nuestro caso, no esperamos superar tan siquiera un Gigabyte de almacenamiento entre datos y archivos. Además de que por ser una etapa de prueba y desarrollo, no esperamos contar con más de media docena de usuarios.
2. **Limitante de costos:** Trabajar con Big Data frecuentemente implica considerar soluciones que ofrezcan grandes volúmenes de almacenamiento de datos, así como baja latencia para el acceso a los mismos. Dado que contamos con recursos limitados para la implementación del proyecto, estas soluciones pueden exceder nuestro presupuesto actual.
3. **Alcance del proyecto:** Esta fase del proyecto sirve como una prueba de concepto para probar la factibilidad de un proyecto que sirva como una solución de registro de asistencia y participación en el aula, por lo que el análisis de datos a gran escala queda como segunda prioridad luego de probar la factibilidad del producto.