

Escenario: Duolingo

[Duolingo](#) es una plataforma líder en el aprendizaje de idiomas que ha transformado la forma en que millones de personas estudian. Con un enfoque gamificado, la plataforma ofrece cursos en más de 40 idiomas y ha expandido recientemente su oferta hacia otras áreas como matemáticas y música. Su modelo freemium, accesible desde dispositivos móviles y web genera una enorme cantidad de datos provenientes del comportamiento de los usuarios, interacciones con el contenido, ejercicios resueltos y mucho más.

La arquitectura de la plataforma se basa en microservicios, donde cada subsistema maneja su propia base de datos, permitiendo elegir la tecnología que mejor se ajuste a los requerimientos funcionales y no funcionales.

A continuación se presentan los subsistemas principales sobre los que trabajará el obligatorio:

- Gestión de usuarios y perfiles
- Sistema de lecciones y ejercicios
- Motor de recomendaciones
- Foros y comunidad
- Analítica y seguimiento de actividad
- Sistema de notificaciones
- Tienda virtual y suscripciones
- Privacidad y seguridad

Estos subsistemas son independientes unos de los otros por lo que la comunicación entre ellos se realiza mediante servicios y cada subsistema almacena los datos en la tecnología y base de datos que mejor se adapte a sus necesidades. A continuación se describe cada subsistema.

Gestión de usuarios y perfiles

El subsistema de manejo de cuentas y configuración de privacidad en Duolingo es una herramienta esencial para los usuarios, ya que no solo les permite gestionar su acceso a la plataforma (registro con correo electrónico y contraseña, activación de 2FA, ajustes de privacidad, etc.), sino también personalizar su experiencia de aprendizaje. Dentro del perfil de usuario, se almacenan datos clave como el nombre de usuario, la foto de perfil, el progreso en

los cursos, la racha de aprendizaje, los logros obtenidos y la lista de amigos con los que compiten o colaboran.

Cada usuario tiene un perfil principal donde se reflejan estos datos. Si además cuenta con Duolingo Plus (la versión premium de la aplicación), podrá acceder a beneficios adicionales, como la eliminación de anuncios, la posibilidad de practicar sin conexión y el seguimiento detallado de su progreso.

Desde su perfil, el usuario también podrá administrar la configuración de privacidad y seguridad. Estos ajustes son fundamentales para definir con quién puede interactuar dentro de la plataforma, qué tipo de información es visible para otros usuarios y cómo se manejan las notificaciones y recordatorios.

Es crucial que la plataforma permita la incorporación de nuevas funcionalidades en el perfil de usuario de manera sencilla, sin comprometer la seguridad y privacidad de los datos. Dado el gran volumen de usuarios en Duolingo, es imprescindible que la implementación de estas mejoras no afecte la escalabilidad del sistema ni la fluidez de la experiencia de aprendizaje.

Sistema de lecciones y ejercicios

El sistema de lecciones y ejercicios sigue una jerarquía bien definida que organiza el aprendizaje de manera progresiva y efectiva. Cada idioma está compuesto por múltiples cursos, los cuales se dividen en unidades temáticas. Dentro de cada unidad, los usuarios avanzan a través de lecciones específicas que contienen una serie de ejercicios diseñados para reforzar el conocimiento de manera interactiva.

Este modelo jerárquico permite una navegación clara y estructurada, asegurando que los estudiantes progresen de manera lógica a medida que adquieren nuevas habilidades lingüísticas. Además, la plataforma ofrece acceso secuencial, lo que significa que los usuarios deben completar ciertos requisitos antes de desbloquear nuevos niveles de aprendizaje.

Para mantener el contenido actualizado y adaptado a las necesidades de los estudiantes, Duolingo permite la edición frecuente y el versionado de sus cursos. Esto garantiza que los materiales de aprendizaje puedan ser ajustados y mejorados sin interrumpir la experiencia del usuario. Gracias a este enfoque, se pueden incorporar nuevas metodologías pedagógicas, corregir errores y actualizar ejercicios según la evolución de cada idioma.

Dado el alto volumen de usuarios y la diversidad de cursos disponibles, es fundamental que la gestión del contenido se realice de manera eficiente, asegurando que los cambios no afecten la continuidad del aprendizaje ni la estabilidad de la plataforma.

Motor de recomendaciones

Duolingo recomienda contenido educativo a los usuarios basándose en diversos criterios para optimizar su aprendizaje. Este sistema de recomendaciones no solo ayuda a personalizar la experiencia de estudio, sino que también permite detectar áreas de mejora y reforzar conocimientos clave a través de ejercicios específicos.

El motor de sugerencias de Duolingo analiza el perfil del usuario, su progreso en los cursos, los errores más frecuentes que ha cometido y sus preferencias de aprendizaje. A partir de estos datos, se pueden generar ejercicios personalizados, reforzar habilidades que requieren mayor práctica y sugerir nuevas unidades o desafíos adicionales.

Para lograr esto, Duolingo utiliza distintas técnicas avanzadas que permiten detectar similitudes entre los patrones de aprendizaje de distintos usuarios. Si un estudiante tiene dificultades con una estructura gramatical específica, el sistema puede recomendar ejercicios que han sido útiles para otros con problemas similares. De la misma manera, si un usuario ha mostrado interés en cierto tipo de contenido, se le pueden sugerir actividades relacionadas para mantener su motivación.

Este enfoque adaptativo permite mejorar la experiencia del usuario y aumentar la eficiencia del aprendizaje. Además, gracias al análisis de datos a gran escala, es posible optimizar continuamente el sistema, asegurando que las recomendaciones sean cada vez más precisas y efectivas.

Foros y comunidad

Duolingo proporciona espacios de discusión para fomentar la interacción entre estudiantes, permitiendo el intercambio de conocimientos y el refuerzo del aprendizaje a través de la práctica activa del idioma. Estos espacios están diseñados para incentivar la lectura intensiva y la escritura frecuente, brindando a los usuarios la oportunidad de aplicar lo aprendido en un contexto real.

Las discusiones se organizan en hilos con mensajes ordenados cronológicamente, lo que facilita el seguimiento de las conversaciones y permite a los usuarios participar de manera estructurada. En estos foros, los estudiantes pueden hacer preguntas sobre gramática, vocabulario o pronunciación, recibir explicaciones de otros usuarios y compartir estrategias de estudio.

Además, estos espacios funcionan como una comunidad de apoyo donde los aprendices pueden motivarse entre sí, compartir experiencias y resolver dudas en conjunto. La escritura constante en un entorno colaborativo ayuda a reforzar la retención del idioma y a mejorar la expresión escrita de los usuarios.

Para garantizar una experiencia enriquecedora, el sistema puede incorporar moderación automática y filtros de contenido, asegurando un ambiente respetuoso y adecuado para el aprendizaje. A medida que la comunidad crece, es fundamental mantener la accesibilidad y escalabilidad del sistema para que las discusiones sean eficientes y beneficiosas para todos los usuarios.

Analítica y seguimiento de actividad

El subsistema registra de manera continua diversos eventos durante la interacción del usuario con la plataforma, lo que permite mejorar la calidad del contenido y comprender mejor el comportamiento de los estudiantes. Entre los eventos capturados se incluyen los intentos de ejercicios, el tiempo de uso de la aplicación y los patrones de error más comunes.

El análisis de estos datos permite ajustar la dificultad de los ejercicios, optimizar la progresión de los cursos y ofrecer sugerencias personalizadas basadas en el rendimiento de cada usuario. Por ejemplo, si un estudiante comete errores recurrentes en una estructura gramatical específica, el sistema puede reforzar ese contenido mediante ejercicios adicionales o explicaciones más detalladas.

Además, estos registros son clave para evaluar la efectividad de nuevas metodologías de enseñanza y mejorar la experiencia general en la plataforma. A través del análisis de tendencias y comportamientos de los usuarios, Duolingo puede identificar qué tipos de ejercicios son más efectivos, en qué momentos los estudiantes tienden a abandonar una lección y cómo mantener su motivación a lo largo del curso.

Gracias a este enfoque basado en datos, la plataforma evoluciona constantemente, asegurando que cada usuario reciba una experiencia de aprendizaje adaptada a sus necesidades y maximizando la eficiencia del proceso educativo.

Sistema de notificaciones

Duolingo mantiene el compromiso y la motivación de sus usuarios mediante el envío de recordatorios, notificaciones de logros y actualizaciones relevantes. Estas notificaciones tienen un tiempo de respuesta optimizado para garantizar que los mensajes lleguen en el momento adecuado, maximizando su efectividad y evitando interrupciones innecesarias.

El sistema de recordatorios juega un papel clave en la retención de los estudiantes, enviando alertas sobre la racha de aprendizaje, lecciones pendientes o la necesidad de reforzar ciertos temas. Además, las notificaciones de logros refuerzan la motivación al informar a los usuarios cuando alcanzan nuevos hitos, como completar una unidad, desbloquear un nivel o superar su mejor desempeño.

Para asegurar que la información sea relevante y oportuna, el sistema maneja un tiempo de vida (TTL) adecuado para cada tipo de notificación. Esto significa que los recordatorios y actualizaciones se entregan en el momento óptimo y se descartan cuando dejan de ser útiles, evitando saturar al usuario con información desactualizada.

Gracias a este enfoque, Duolingo mantiene a los estudiantes comprometidos con su aprendizaje, reforzando hábitos de estudio y asegurando una experiencia fluida y personalizada dentro de la plataforma.

Tienda virtual y suscripciones

Duolingo ofrece una gestión eficiente de cuentas premium para aquellos usuarios que optan por Duolingo Plus, asegurando un proceso fluido en la administración de suscripciones, transacciones y el historial de compras.

El sistema permite a los usuarios suscribirse, renovar o cancelar su membresía en cualquier momento, manteniendo un registro detallado de todas las transacciones realizadas. Además, el historial de compras proporciona información sobre fechas de pago, duración de la suscripción y métodos de pago utilizados, permitiendo un fácil acceso a los datos en caso de consultas o reembolsos.

Para garantizar una experiencia segura y confiable, el sistema de pagos emplea protocolos de seguridad avanzados, protegiendo la información financiera de los usuarios. Asimismo, las renovaciones automáticas y las notificaciones sobre vencimientos ayudan a evitar interrupciones en el acceso a las funciones premium, asegurando que los usuarios disfruten de

beneficios como la eliminación de anuncios, prácticas sin conexión y seguimiento detallado de su progreso sin inconvenientes.

La correcta gestión de las cuentas premium y sus transacciones es esencial para mantener la satisfacción del usuario y la estabilidad del modelo de negocio de Duolingo, optimizando continuamente la experiencia de quienes eligen mejorar su aprendizaje con la versión Plus.

Privacidad y seguridad

Duolingo implementa un sólido sistema de control de acceso a datos para garantizar la privacidad y seguridad de la información de sus usuarios, cumpliendo con normativas internacionales como el **GDPR** (Reglamento General de Protección de Datos) y **COPPA** (Ley de Protección de la Privacidad Infantil en Línea).

El acceso a los datos está estrictamente regulado, asegurando que solo los usuarios autorizados y los sistemas internos puedan interactuar con información sensible. Los usuarios pueden gestionar sus configuraciones de privacidad, definir qué datos desean compartir y solicitar la eliminación de su información cuando lo deseen, en cumplimiento con el **derecho al olvido** establecido por GDPR.

Además, Duolingo realiza evaluaciones en tiempo real para garantizar que el tratamiento de los datos personales cumpla con las regulaciones vigentes. Esto implica auditorías continuas, monitoreo de acceso y la aplicación de medidas de seguridad como el cifrado y la anonimización de datos sensibles.

El cumplimiento de estas normativas no solo protege a los usuarios, sino que también refuerza la confianza en la plataforma, asegurando que el aprendizaje se desarrolle en un entorno seguro y transparente para estudiantes de todas las edades.

Requerimientos

En base al escenario planteado se solicita:

1. **Recomendar una o la cantidad de bases de datos que considere necesarias para cada uno de los subsistemas mencionados.** Se deberá indicar que motor se usará, los motivos para haberlo seleccionado y que ventajas y desventajas presenta su uso.

2. **¿Qué tipo de base de datos es Weaviate? ¿Podría aplicarse en algunos de los escenarios planteados?** Deberá describir el modelo de datos que Weaviate soporta, ejemplos concretos de ese modelo y analizar si es aplicable o no para los escenarios que se plantean en este obligatorio.
3. **Modelar la o las entidades usadas para el motor de recomendaciones.** Se deberá modelar en función de la base de datos seleccionada para su almacenamiento y de las necesidades planteadas en la descripción del subsistema.
4. **Implementar el almacenamiento para alguno de los subsistemas.** A modo de ejemplo, la implementación puede ser un dashboard de Kibana o un notebook Jupyter. **NO** podrá realizarse una implementación usando un RDBMS tradicional orientado a filas.

En caso de no conocer Duolingo es altamente recomendable que cree una cuenta y que se familiarice con la plataforma.

Entregables

Los requerimientos 1, 2 y 3 serán entregadas en un informe escrito.

El requerimiento 4 podrá ser implementado en el lenguaje de programación de su preferencia. Deberá entregar todos los elementos necesarios para poder ejecutar los programas (soluciones, código fuente, drivers, manual de instalación etc). El código podrá estar hosteado en un repositorio online como GitHub o Bitbucket y la aplicación, en caso de ser web, podrá ser hosteada en cualquier proveedor de servicios cloud.

El obligatorio consta de dos entregas:

1. 21 de mayo
 - a. Requerimiento 1: Tienda virtual y suscripciones, Analítica y seguimiento de actividad y Privacidad y seguridad
 - b. Requerimiento 2
2. 09 de julio
 - a. Totalidad del requerimiento 1
 - b. Requerimientos 3 y 4
 - c. ¿Realizaría algún cambio de recomendación en alguno de los subsistemas que presentó en la primera entrega? ¿Por qué?

Las entregas se realizan a través de Gestión y se permiten grupos de hasta 3 estudiantes.

Defensa

La defensa consistirá en una presentación al grupo de no más de 10 minutos. En la misma se explicará cuáles fueron los motores seleccionados en general y en particular para la parte 4 se efectuará una demo de lo implementado. Esta instancia tiene un **puntaje extra al obligatorio** y será de 10 puntos. **Fecha de realización:** 16 de julio en horario de clase via Zoom.

Escuela de Ingeniería	
Obligatorio: Modelos avanzados de bases de datos	Código de materia: 7239
Fecha: 02/04/2025	Hoja 9 de 10

Rúbrica

Área	Evaluación	Puntaje Obligatorio 1	Puntaje Obligatorio 2
Presentación del trabajo	Se espera que se utilice como base el formato 302 de la Facultad, no es necesario el uso completo, como mínimo contener una carátula, índice, introducción, desarrollo y bibliografía.	2	2
Definición de BD a utilizar	Se espera un análisis del problema, indicando las principales necesidades a cumplir, ventajas y desventajas del uso de la BD elegida así como una explicación de porque cumple el caso, y como las desventajas son subsanadas.	20	20
Investigación de BD.	Se espera una introducción explicando qué tipo de BD es y sus casos de uso. Las ventajas y desventajas que tiene. El modelo de datos así como un ejemplo del mismo. Comparación con los subsistemas y si aplica a alguno y por qué.	8	No aplica
Modelado del subsistema	Se espera una justificación de las decisiones de diseño realizadas. La presentación del modelo utilizando el diagrama esperado para modelar, según el tipo de BD seleccionado. El diagrama completo utilizando la correcta nomenclatura del mismo.	No aplica	8
Implementación	Se espera una aplicación que se conecte con la BD y permita la inserción de datos o un dashboard de Kibana que muestre la realidad descrita. Un acceso al código generado. Una descripción de lo hecho y las decisiones tomadas. Se debe incluir un diagrama y explicación del modelo generado.	No aplica	15

Escuela de Ingeniería	
Obligatorio: Modelos avanzados de bases de datos	Código de materia: 7239
Fecha: 02/04/2025	Hoja 10 de 10

	Total	30	45
--	-------	----	----

Demo (Defensa)	<p>Cada equipo deberá realizar una demostración de su trabajo. Dentro de los aspectos que se esperan se encuentran:</p> <ul style="list-style-type: none"> ● Estar listos para comenzar cuando se indique. ● Los datos y la funcionalidad a mostrar deben estar preparados de antemano. ● Conocer las capacidades de cada uno de los integrantes del equipo, pudiendo preguntar a cualquier integrante sobre la solución, su diseño, el código y sobre cómo fue construida. ● Todos los integrantes deben conocer toda la solución. Esta demostración al cliente hará a la vez de defensa del trabajo. ● La defensa tendrá una duración máxima de 10 minutos. Luego de transcurridos los mismos se podrá restar puntos al trabajo. ● Si bien pueden tener datos generados previo a la defensa se espera que se carguen o se generen datos en tiempo real. <p>NOTA: El incorrecto funcionamiento de la instalación puede significar la no corrección de la funcionalidad.</p>	10
	Total	10