



DECSAI

Departamento de Ciencias de la Computación e I.A.

Universidad de Granada

Ingeniería de Sistemas de Información

Curso 2020/2021

Convocatoria ordinaria, junio de 2021

Al realizar estos ejercicios, de forma explícita está suscribiendo un código de honor según el cual se compromete a presentar su trabajo personal y renuncia a compartir información sobre la realización de los ejercicios de esta prueba hasta que haya pasado su fecha límite de entrega.

En el desarrollo de las pruebas de evaluación no presencial se respetará la ética e integridad académica aplicable en el ámbito universitario y, en especial, los principios de mérito individual, autenticidad y originalidad de las pruebas. Se promoverá un uso responsable y adecuado de las herramientas de evaluación no presencial, y se evitarán las conductas fraudulentas que, en su caso, tendrán las consecuencias académicas y disciplinarias previstas en la normativa vigente.

De acuerdo con la normativa de evaluación y calificación de la Universidad de Granada, en lo que respecta a la originalidad de los trabajos y pruebas (artículo 15), los trabajos y materiales entregados por parte de los estudiantes tendrán que ir firmados con una declaración explícita en la que se asume la originalidad del trabajo. El plagio, entendido como la presentación de un trabajo u obra hecho por otra persona como propio o la copia de textos sin citar su procedencia y dándolos como de elaboración propia, conllevará automáticamente la calificación numérica de cero en la asignatura en la que se hubiera detectado, independientemente del resto de las calificaciones que el estudiante hubiera obtenido.

NOMBRE Y APELLIDOS

DNI

FIRMA

Diseño arquitectónico

1. **Arquitectura del sistema:** Se nos ha encargado el diseño de un sistema de integración de datos que dará soporte a un metabuscador de libros de segunda mano (esto es, un comparador de precios de libros). Entre los requisitos de nuestro sistema se encuentran los siguientes:

- a. El usuario accederá al sistema a través de dos interfaces, una aplicación web y una app móvil.
- b. El sistema debe estar diseñado para facilitar la incorporación de nuevas fuentes de datos (por ejemplo, usando web scraping o algún API ofrecido por la fuente).
- c. El sistema almacenará localmente los datos con los que trabaje con el objetivo de ir construyendo una base de datos que permita la incorporación de nuevas funcionalidades (p.ej. al estilo de camelcamelcamel.com).

Diseñe la arquitectura de un sistema que satisfaga los requisitos anteriores utilizando los patrones de diseño que considere necesarios. Proporcione un diagrama en el que se reflejen los requisitos anteriores (vista lógica de acuerdo al modelo 4+1 de Kruchten o punto de vista funcional en el modelo de Rozinsky y Woods).

Sistemas distribuidos

2. **Teorema CAP:** De acuerdo al teorema CAP, indique a qué categoría pertenece cada uno de los sistemas siguientes y explique el porqué:

- El sistema de control de inventario de una empresa de venta por correo (postal o electrónico).
- El sistema de control de inventario de una empresa de comercio electrónico (a través de una página web).
- Una aplicación web implementada en HTML básico.
- Una aplicación web implementada en HTML5 con almacenamiento local de datos.

Integración de datos

Nuestro sistema de integración de datos combinará información de distintos vendedores de libros usados en un esquema integrado simplificado como el siguiente:

autor (nombre, fechaNacimiento, biografía)
obra (autor, título, añoPublicación, género, descripción) edición
(autor, título, isbn, editor, año, encuadernación, páginas, dimensiones, peso, portada)
oferta (isbn, vendedor, precio, fecha, urlOferta)
donde isbn es un código único que permite identificar una edición concreta de un libro (siempre que éste no fuese editado antes de los años 70, que fue cuando se estandarizó el

ISBN [International Standard Book Number]).

En función del último dígito de su DNI o pasaporte, utilice el vendedor de libros indicado para realizar los siguientes ejercicios:

0. <https://www.libros-antiguos-alcana.com/>
1. <https://librosmelior.org/>
2. <https://libreriallerapacios.com/>
3. <https://www.librosambigu.com/>
4. <https://www.uniliber.com/>
5. <https://www.tikbooks.com/>
6. <https://www.librousado.es/>
7. <https://www.librovicios.com/>
8. <https://www.bolsabooks.com/>
9. <https://relibrea.com/>

3. **Integración de esquemas:** Modele el esquema que represente todos los datos ofrecidos por la fuente de datos correspondiente al último dígito de su DNI o pasaporte. A continuación, modele la correspondencia entre el esquema de su fuente de datos y el esquema integrado del sistema utilizando el enfoque GLAV.

NOTA: Puede suponer que dispone de funciones como `date.now()` para obtener la fecha actual (con la que registrar el momento en el que tiene validez el precio al que su fuente de datos ofrece un libro) o `id()` para crear un identificador único con el que registrar un libro que no tenga ISBN.

4. **Extracción de datos:** Implemente un web scraper en Python o Java que permita extraer datos de su fuente de datos. Proporcione un ejemplo de URL sobre la que su web scraper funciona y la salida que obtiene ejecutando se extractor de datos. No olvide incluir el código de su implementación y presente un fichero adjunto en formato ASCII al entregar esta relación de ejercicios (con extensión `.java` o `.py` dependiendo del lenguaje que haya utilizado).

Determinadas fuentes de datos pueden resultar problemáticas al hacer web scraping. En ocasiones, por el aviso de cookies al que les obliga la legislación europea, que a menudo puede sortearse cambiando la cabecera `User-Agent` de la solicitud HTTP. En otras ocasiones, puede que la aplicación web del sitio al que accedemos haga un uso extensivo de JavaScript o incluya algún mecanismo anti-CSRF, lo que dificulta nuestro trabajo. Aunque siempre existe una forma de sortear dichos problemas (al fin y al cabo, el navegador web lo hace), si encuentra problemas al implementar su extractor de datos, descargue la página web desde su navegador y trabaje directamente con el fichero HTML local que haga descargado.

Middleware

5. **Diseño de brokers:** Imagine que le encargan el diseño e implementación un bróker. Describa cómo diseñaría, en un bróker, la gestión de eventos asociados al envío y recepción de mensajes para maximizar la escalabilidad del sistema, poder ofrecer garantías de calidad de servicio (QoS) y, de paso, intentar que sea robusto frente a ataques por denegación de servicio.

Seguridad

6. **Inyección SQL:** Acceda al siguiente formulario (vía VPN desde fuera de la red de la UGR)
<https://elvex.ugr.es:8443/sa2020/login.html?url=login4.jsp>

Diseñe ataques por inyección de código que permitan hacer lo siguiente:

- Acceder al sistema sin conocer la contraseña de acceso.
- Volcar los identificadores de usuario, contraseñas y direcciones de correo electrónico de los usuarios registrados en la base de datos de la aplicación.

Para cada uno de sus ataques, proporcione la URL completa que permite ejecutar el ataque con éxito.

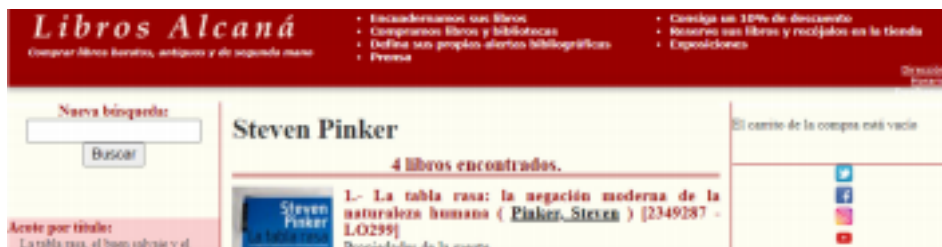
PISTA: Puede ayudarse del siguiente formulario para ejecutar consultas SQL sobre la base de datos SQLite que da soporte a la aplicación web:

<https://elvex.ugr.es:8443/sa2020/query.jsp>

Apéndice: Fuentes de datos

Libros antiguos Alcaná

<https://www.libros-antiguos-alcana.com/>



Intensa en la máquina
La table rose

Aste por editar:
Prados Iberica
Prados.

Encuadernación:
Tapa blanda
Tapa dura

Últimas novedades

Ofertas:

Pañós. Barcelona. 2012. 24 cm. 794 p. d., gráf.
Encuadernación en tapa blanda de editorial ilustrada.
Colección 'Pañós transiciones'. Parker, Steven 1954.
Traducción de Roc Filalla Escoté. Bibliografía: p. [633]-684.
Índice. Título original: The blank slate. Herencia y medio.
Continúa ferso adhesivo en la cubierta y sello en libro del anterior propietario en
al interior. ISBN: 9788449322990 Si desea recoger personalmente
este libro en la librería, solicítelo, por favor, con 6 horas de
antelación.

Recomendaciones:

Precio: 39,22 €

 **carrito**

<https://www.libros-antiguos-alcana.com/resultados.do?autor=steven%20pinker>

Libros Melior

<https://librosmelior.org/>

 **GRATIS 35.** NOVEDADES  OFERTAS  MENOS DE 1€  [Inicio sesión / Crear cuenta](#)

 [Mi lista de deseos](#)

LIBROS SOLIDARIOS
POR UNA EDUCACIÓN MEJOR

BUSCAR "SAGAN"
5 resultados encontrados.

CATEGORÍAS

- Artes
- Biblioteca amiga
- Biografías - Memorias
- Ciencias
- Ciencias sociales
- Clásicos
- Cocina y gastronomía
- Deportes
- DVD y CD

CONTACTO. CARL SAGAN. REF.184979
1,00 € **MÁS**

LA CONEXIÓN CÓSMICA. CARL SAGAN. REF.191036
2,00 € **MÁS**

EL CEREBRO DE BROCA. CARL SAGAN. REF.190563
2,00 € **AÑADIR AL CARRITO**

https://librosmelior.org/busqueda?controller=search&orderby=position&orderway=desc&search_query_cat=0&search_query=sagan&submit_search=

Librería Llera Palacios

<https://libreriallerapacios.com/>

 **Librería Llera Pacios**  [ACCEDER / REGISTRARSE](#) **CARRITO: 0,00€** 

NOVEDADES **LIBROS BARATOS** **SEGUNDA MANO** **ANTIGUOS** **COMPRA-VENTA**

INICIO / PRODUCTOS / RESULTADOS DE BÚSQUEDA PARA "BECQUER" **Mostrando 1-24 de 46 resultados** **Relevancia**



https://libreriaallerapacios.com/?s=b%C3%A9cquer&post_type=product

Libros Ambigú

<https://www.librosambigu.com/>



https://www.librosambigu.com/libros/?titulo=&autor=b%C3%A9cquer&materia=&editorial=&palabra_clave=&action=buscar

uniliber

<https://www.uniliber.com/>



https://www.uniliber.com/buscar/libros_y_coleccionismo?query=steven%20pinker

tikbooks

<https://www.tikbooks.com/>



https://www.tikbooks.com/libros/?texto=conan+doyle&action=busqueda_rapida&buscar=buscar

<https://www.librousado.es/>

[Aviso legal](#) / [Política de Privacidad](#) / [Condiciones de compra y envío](#) / [Formas de Pago](#) [Entrar en su cuenta de usuario](#) / [su cuenta](#)



buscar por título, autor, ISBN, palabra clave... 
» [BUSQUEDA AVANZADA](#)

0 ARTÍCULOS
(0,00€) 

[LIBROS](#) [NOVEDADES](#) [OFERTAS](#) [NUESTRAS LIBRERÍAS](#) [NOTICIAS](#) [QUIÉNES SOMOS](#)



[Inicio](#)

- » ANIMALES
- » ANFOLOGÍA
- » ANTHROPOLOGÍA
- » ARQUITECTURA
- » ARTE
- » ARTICULOS
- » ASTROLOGÍA
- » AUTOAYUDA

En nuestro **fondo de Librería** encontrarás todas las temáticas.

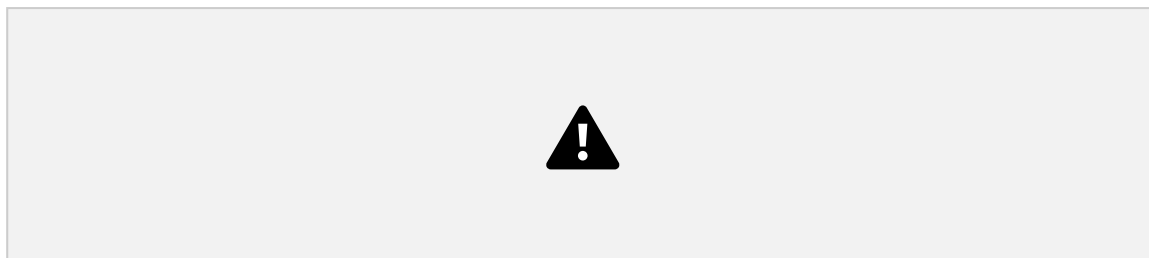
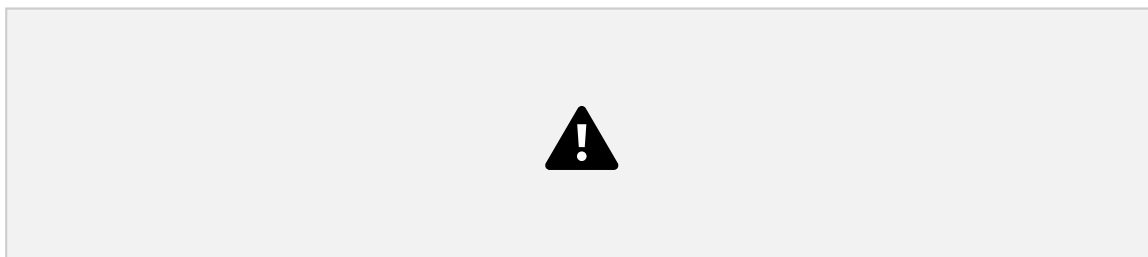


buscar 

https://www.librouzado.es/libros/?texto=becquer&action=busqueda_rapida&buscar=buscar

librovicios.com

<https://www.librovicios.com/>



<https://www.librovicios.com/jolisearch?s=conan+doyle>

bolsabooks

<https://www.bolsabooks.com/>



https://www.bolsabooks.com/es/products?search_api_views_fulltext=pinker

relibrea

<https://relibrea.com/>



<https://relibrea.com/libros-de-segunda-mano?kws=steven-pinker>

English translation (for Erasmus students)

Architectural Design

1. **System architecture:** We have been assigned the task of designing the data integration

system that will support a used book metasearch engine (i.e., a book price comparator). Our system has the following requirements:

- a. The user will Access the system through two different interfaces, a web application and a mobile app.
- b. The system must be designed to ease the addition of new data sources (e.g., using web scraping or the API provided by a data source).
- c. The system will store data locally, so that a database will be created to support the addition of new features (e.g., something like camelcamelcamel.com).

Design the software architecture of a system that satisfies the previous requirements using the design patterns that you deem suitable for the task at hand. Include a diagram showing how the previous requirements are addressed (the logical view according to Kruchten's 4+1 model or the functional viewpoint in Rozinsky and Woods' model).

Distributed Systems

2. **CAP Theorem:** According to the CAP theorem, determine the category of each of the following systems and explain why:

- The stock control system of a mail order company (postal mail or electronic e-mail).
- The stock control system of an e-commerce business (through a web app).
- A web application created using basic HTML.
- A web application created using HTML5 with local data storage.

Data Integration

Our data integration system will combine data from different used book vendors in a simplified integrated schema such as the following one::

```
author (name, birthDate, biography)
work (author, title, publicationYear, genre, description)  edition
(author, title, isbn, publisher, year, binding, pages, dimensions,
weight, cover)
offer (isbn, seller, price, date, offerURL)
```

where isbn is a unique identification number for a particular book edition (provided that the book was not edited before the 1970's, when the ISBN [International Standard Book Number] was standardized).

According to the last digit in your ID card (Spanish DNI or Passport), use the corresponding book vendor to complete the following exercises:

0. <https://www.libros-antiguos-alcana.com/>
1. <https://librosmelior.org/>

2. <https://libreriallerapacios.com/>
3. <https://www.librosambigu.com/>
4. <https://www.uniliber.com/>
5. <https://www.tikbooks.com/>
6. <https://www.librousado.es/>
7. <https://www.librovicios.com/>
8. <https://www.bolsabooks.com/>
9. <https://relibrea.com/>

3. **Schema integration:** Model the schema representing all the data offered by the data provider corresponding to the last digit in your ID card. Next, model the correspondence between your data source schema and the integrated schema using the GLAV approach.

NOTE: You can assume that you can resort to available functions such as `date.now()` to obtain the current date (to record when the book price offered by your data source was valid) or `id()` to create a unique identifier for books without an ISBN.

4. **Data extraction:** Implement a Java or Python web scraper to extract data from your data source. Include an example URL where your web scraper works and the output obtained from that URL by running your web scraper. Do not forget to include the source code of your implementation and attach an ASCII file when you submit your solutions to this problem set (with `.java` or `.py` extension depending on the programming language of your choice).

Some data sources might be problematic for web scraping. Sometimes, due to the cookie warning required by the European legislation, which can often be sidestepped by changing the `User-Agent` header in your HTTP request. On other occasions, however, the web app from your data source might make an extensive use of JavaScript or even include some anti-CSRF mechanism, which makes web scraping harder. Even though there is always a way to circumvent those obstacles (after all, your web browser does it), if you find problems running your web scraper on your web data source, you can download the desired web page using your web browser and work directly on the downloaded HTML file.

Middleware

5. **Broker design:** Imagine that you have been requested to design and implement a broker. Describe how you would design, within a broker, its event management subsystem (i.e., handling the events related to sending and receiving messages from multiple sources) in order to maximize your system scalability, provide quality of service (QoS) guarantees and, as a side effect, make it more robust against denial-of-service attacks.

Security

6. SQL injection: Access to the following web form (through a VPN outside the UGR network)
<https://elvex.ugr.es:8443/sa2020/login.html?url=login4.jsp>

Design SQL code injection attacks to attain the following goals:

- a. Access to the system without knowing the user's password.
- b. Display user IDs, passwords, and e-mail addresses of all the users who are registered in the web app database.

For each of your attacks, specify the complete URL that is successful when performing the attack.

HINT: You can use the following web form to run SQL queries over the SQLite database that supports the web app under attack: <https://elvex.ugr.es:8443/sa2020/query.jsp>