

Prueba 2: Opción C: Ciencia de la Web**Temas troncales del NM y el NS**

C1 Una compañía editorial, ABC Publications, que tiene su sede en Londres, cuenta con un gran departamento de TI. Este departamento es el responsable de:

- ofrecer servicios TI a la compañía
- mantener el sitio web de la empresa
- crear y mantener recursos de aprendizaje basados en la web que posteriormente se venden a colegios e institutos.

La compañía está teniendo dificultades para contratar y retener una cantidad suficiente de personal con alta cualificación en TI para seguir realizando esas funciones a un nivel óptimo. Está considerando transferir al menos una parte de sus operaciones de TI a una solución basada en informática en la nube. Por ahora no se ha decidido qué parte de ellas debería implementarse en una nube privada y qué parte en una nube pública.

a) Defina el término *nube privada*.

[1 punto]

Los servicios de informática en la nube se ofrecen para un grupo concreto con un número limitado de usuarios.

b) Distinga entre un modelo de informática en la nube y un modelo convencional cliente-servidor, en cuanto a la provisión de servicios informáticos.

[4 puntos]

En un modelo cliente-servidor, el computador del cliente interactúa con los servidores pertenecientes a la propia organización. Estos servidores podrían estar ubicados en la LAN de las oficinas de la organización y los podría gestionar el equipo de TI. Con la informática en la nube, las organizaciones externalizan tareas como almacenamiento, mantenimiento y procesamiento de datos, usando Internet para conectar el cliente al servidor y permitiendo el acceso desde cualquier ubicación que tenga conexión a Internet.

c) Explique por qué ABC Publications podría beneficiarse de cambiar parte de sus servicios de TI a un modelo basado en la nube.

[4 puntos]

La compañía ya no tendría que seleccionar y mantener a personal altamente cualificado en TI ya que los servicios los podría gestionar la empresa externa. De esta forma se podría prescindir de los costos asociados, así como otros puestos locales necesarios para los servidores (incluyendo mantenimiento, depreciación y actualización) y, en su lugar, la compañía podría pagar un precio fijo en función de su uso. Además, el almacenamiento de la información en una ubicación externa (o incluso en varias

ubicaciones externas) reduce la gravedad ocasionada por posibles daños (fuegos, etc.) en la sede local de la compañía. ABC también se podría beneficiar de la experiencia de las compañías externas, ya que podría tener más empleados dedicados a resolver problemas de red que si estuvieran únicamente en la oficina de ABC.

d) Comente los problemas de privacidad y seguridad relacionados con el uso de la informática en la nube por parte de ABC. *[4 puntos]*

Privacidad: El uso de una compañía externa implica que cualquier información sensible esté expuesta de facto a gente externa a la empresa ABC. Esta empresa privada podría incluso externalizar el servicio a otra compañía, aumentando así la exposición de sus datos.

Seguridad: Si la “compañía de informática en la nube” se usa para almacenar datos, habrá preocupaciones inevitables sobre la fiabilidad de este tipo de almacenamiento que ahora está fuera del control de ABC. La informática en la nube es un concepto relativamente nuevo, pero ya se han hundido algunas empresas que se dedicaban a almacenar cantidades masivas de datos (de terceros), llevándose con ellos los datos.

C2 La taquilla de un teatro cuenta con un sitio web que puede mostrar las próximas producciones que se exhibirán y los asientos disponibles para cada una de ellas en particular. En el sitio web es posible comprar las entradas. La interfaz está formada por páginas web dinámicas, como la que se muestra a continuación, en que el código HTML subyacente interactúa con guiones del cliente y del servidor.

TAQUILLA

Búsqueda rápida

Desde Hasta Tipo de evento Buscar

En el calendario solo aparecerán las actuaciones y los eventos que se puedan reservar ahora.

Visite nuestra página [Programación](#) para ver listas completas de todas las producciones.

Buscar Julio de 2015 Tipo de evento

Domingo	Lunes	Martes	Miércoles	Jueves	Viernes	Sábado
11	12 10:00 - 12:00 Taller de ballet Comprar ahora	13 10:00 - 12:00 Taller de ballet Comprar ahora	14 10:00 - 12:00 Taller de ballet Comprar ahora	15 10:00 - 12:00 Taller de ballet Comprar ahora	16 10:00 - 12:00 Taller de ballet Comprar ahora	17
19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora	19:30 - 22:00 El lago de los cisnes 3 Counties Ballet Comprar ahora

Parte del código fuente para esta página es:

```
<script type="text/javascript" src="http://assets.ophse.org/ inc/popdt/init live.js"></script>
<script type="text/javascript" src="http://assets.ophse.org/ inc/popdt/src/flash.js"></script>
```

a) Identifique **una** característica del lenguaje de marcas de hipertexto (HTML). [1 punto]

Todos los navegadores de Internet pueden mostrarlo.

b) Identifique los pasos que debe realizar el servidor para que la información del calendario de eventos se pueda mostrar en el computador del cliente. *[4 puntos]*

- El servidor captura en un servidor de base de datos los eventos de los días que se quieren mostrar
- Obtiene resultados y genera HTML para mostrarlos en una tabla
- Incrusta HTML en la página
- Envía la página al navegador

c) (i) Identifique **dos** formas en que un guión del cliente puede hacerse accesible para un navegador web. *[2 puntos]*

Con un guión del cliente para que el navegador use etiquetas que están incluidas en el código HTML. Usando un archivo externo.

(ii) Describa **una** razón por la cual se preferiría usar un guión del cliente en lugar de un guión del servidor. *[2 puntos]*

Un guión del cliente no necesita acceder a un servidor remoto, así que cualquier procesamiento se realiza más rápidamente y consume menos ancho de banda. Esto reducirá la carga en el servidor.

d) Los organizadores del teatro quieren asegurarse de que sus páginas Web aparecen en los primeros lugares del ranking de los motores de búsqueda.

Sugiera si el uso de metaetiquetas puede ayudar a conseguir este objetivo. *[4 puntos]*

Las metaetiquetas son un medio de promocionar los componentes clave en un sitio web. Estas etiquetas permiten que el robot de un motor de búsqueda obtenga fácilmente el contenido de un sitio web y puede hacer que esa página aparezca antes en el ranking de búsqueda. Como los motores de búsqueda trabajan de distintas formas, es difícil cuantificar la importancia de las metaetiquetas, y algunos sitios web valoran más el número de enlaces hacia una página.

e) Es normal que las páginas web dinámicas usen guiones del cliente y del servidor. Explique cómo la interacción de HTML y los scripts de cliente y servidor permiten crear una página web como la que se muestra en esta sección. [6 puntos]

Para un sitio web dinámico como este, antes de que se genere el HTML, deben ejecutarse guiones en el servidor. Esto puede requerir el uso de un lenguaje como PHP para gestionar los datos introducidos por el usuario, recuperar la información de las bases de datos almacenadas en los servidores y realizar cálculos. Como el cliente no puede ver las acciones de los cálculos del servidor, puede gestionarse información privada o confidencial sin preocuparse sobre la seguridad de los datos. En el servidor se cotejará esta información y se enviará al navegador. La información enviada podría estar en formato HTML, pero también en otro. Una página web podría, por ejemplo, mostrar HTML con JavaScript en línea, lo que se podría interpretar posteriormente en el navegador del cliente. La naturaleza pública del cliente implica que está más adaptado a realizar tareas simples, como validación o verificación de datos, pero puede funcionar conjuntamente con guiones del servidor (p. ej. a través de Ajax) para entregar información dirigida por el servidor en tiempo real.

C3 A medida que se ha desarrollado la web, el formato de los datos, los protocolos de comunicación y los estándares como XML o SQL han demostrado ser fundamentales para su avance. Dos preocupaciones fundamentales que han sido centrales en este desarrollo son los problemas de interoperatividad y los estándares abiertos.

a) (i) Identifique **una** característica de XML. [1 punto]

No contiene un conjunto fijo de etiquetas y, por tanto, se pueden añadir otras nuevas.

(ii) Defina el término *protocolo*. [1 punto]

Conjunto de reglas y procedimientos al que deben adherirse el emisor y el receptor para permitir una transferencia de datos coherente.

b) Describa, usando ejemplos, cómo el uso de los estándares abiertos permite la interoperatividad. [3 puntos]

El uso de estándares abiertos permite que cualquiera pueda usarlos e introducirlos y consensuarlos para asegurar que se mantiene la calidad. Los estándares abiertos requieren que la interoperatividad se acuerde desde el primer momento (ab-initio) y se use a lo largo de todo el proceso. Por ejemplo, la columna vertebral de Internet se basa en el protocolo IP, un estándar consensuado, que permite que se produzca la transferencia de información.

La música se distribuye por la web de distintas formas, por ejemplo a través de redes de igual a igual (P2P: *peer-2-peer*).

c) Discuta **dos** factores que podrían afectar a la decisión de usar ya sea compresión sin pérdida o con pérdida en la transferencia de archivos a través de Internet. *[6 puntos]*

La compresión sin pérdida de datos se usa cuando no se desea perder datos durante la transmisión de archivos como los de audio. No obstante, hay ocasiones en las que los datos que se podrían perder no afectarían significativamente a la versión final del archivo al descomprimirlo, de manera que en estos casos no sería necesario usar la compresión sin pérdida. La decisión de usar la compresión con o sin pérdida podría verse determinada por las necesidades del usuario final que desee utilizar un formato como el MP3, que permite que se almacene más música en un dispositivo.

Además, el tiempo empleado en transmitir el archivo comprimido sin compresión, que podría ser tres o cuatro veces el tamaño de la compresión con pérdida, podría ser un factor importante para elegir el tipo de compresión. Por ejemplo, si hay una conexión rápida a Internet o no hay necesidad de acceder inmediatamente al archivo se podría usar la compresión sin pérdida. No obstante, a medida que aumenta el tráfico de Internet, la compresión sin pérdida podría ser demasiado lenta.

d) Explique **una** ventaja del uso de una red igual a igual (P2P) para obtener y descargar archivos de música y películas. *[2 puntos]*

Una ventaja de usar una red P2P en lugar de una red cliente-servidor para obtener y descargar ficheros es que resulta más fácil de configurar y requiere menos tiempo para configurar la red.

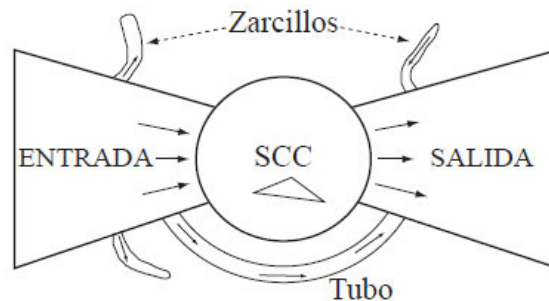
Ampliación del NS

C4 Internet (la Web) se puede representar como un grafo orientado. Esto permite que los motores de búsqueda usen algoritmos basados en la teoría de grafos.

a) Identifique cómo se puede representar la Web usando un grafo orientado. [1 punto]

La Web se representa como un grafo dirigido, en el que las páginas web se ven como vértices y los enlaces entre ellas como aristas.

La Web se puede representar como una estructura del tipo de una corbata pajarita, como se indica en el diagrama siguiente.



[Fuente: <http://nlp.stanford.edu/IR-book/html/htmledition/the-web-graph-1.html> (consultada el 13 de abril de 2011)]

b) (i) Defina el término *componente fuertemente conexo* (Strongly Connected Core, SCC) [1 punto]

El componente fuertemente conexo es la parte de la Web en que un usuario puede navegar desde cualquier página concreta.

(ii) Esboce las características de los sitios web que se encuentran en la sección ENTRADA del diagrama. [2 puntos]

Estos sitios web tienden a contar un número de enlaces que llevan a otros sitios web, como los de la zona SCC u OUT. Estas páginas pueden verse como destinos no revelados que no han encontrado las páginas del SCC o podrían ser una lista de recursos que ofrece una pasarela a las páginas del SCC.

Con el crecimiento de la Web, los desarrolladores han tomado conciencia de las preocupaciones relacionadas con la capacidad de enlazar con todas las páginas web y de cómo asegurar que una página obtiene una buena posición en los resultados de los motores de búsqueda.

c) Describa cómo las leyes de potencia sugieren que será posible enlazar desde una página web a cualquier otra a pesar de que la Web crezca tan rápidamente. *[2 puntos]*

Las leyes de potencia afirman que el diámetro de la Web está relacionado con su tamaño de forma exponencial. Esto implica que si el número de páginas web del SCC aumenta 10 veces, el diámetro del SCC podría aumentar solo en una razón de tres.

d) Explique cómo se puede determinar la importancia relativa de una página web en los motores de búsqueda. *[4 puntos]*

La monitorización del número de enlaces a un sitio web es muy importante para varios motores de búsqueda, como Google, que suele asignar al más popular un mejor puesto en el ranking y lo muestra en la zona alta de la página de resultados.

El ranking lo determina el número de votos que recibe. Esto se basa en el número de enlaces 'in', así como la importancia de la página que la vota. Para situar al sitio web en una zona alta del ranking, los programadores web intentarán maximizar el número de enlaces 'in'. No obstante, algunos algoritmos de clasificación pueden pasar por alto algunos de los enlaces, ya que podrían estar generados por granjas de enlaces y no se consideran auténticos. Además, las páginas antiguas podrían verse beneficiadas ya que habrán tenido tiempo de recibir enlaces 'in'.

El desarrollo de la Web ha modificado la forma en que interactúan los usuarios entre ellos y con la Web.

e) Mencionando ejemplos específicos, distinga entre una ontología y una folcsonomía. *[2 puntos]*

Las ontologías son un marco formal para el almacenamiento del conocimiento y la relación entre las entidades en un determinado campo del conocimiento, por ejemplo, en relación con una marca concreta de automóviles. Una folcsonomía es un marco de trabajo informal desarrollado a partir de la creación y gestión de etiquetas para clasificar información como, por ejemplo, en sitios como Flickr.

El crecimiento de la Web ha hecho posible el desarrollo de nuevas formas de resolver problemas. Uno de estos enfoques es la inteligencia colectiva.

f) Identifique **dos** características de la inteligencia colectiva.

[2 puntos]

- Puede hacer referencia a la forma de pensar del grupo
- Implica grupos de personas trabajando en todo el mundo usando datos abiertos en un entorno público

g) Con el aumento de la población mundial y los efectos de la globalización, la población mundial se enfrenta a nuevos y complejos problemas, como la reciente crisis financiera. ¿En qué medida podría ayudar la inteligencia colectiva a tratar estos problemas?

[6 puntos]

La inteligencia colectiva tiene el potencial de tratar problemas como la crisis bancaria, en los que la ignorancia, que ha ocasionado graves decisiones incorrectas que ha tomado una pequeña minoría, ha perjudicado a la mayoría de la gente. Con la inteligencia colectiva, cualquiera puede contribuir al proceso de producción, y con tanta gente capaz de ver y criticar el trabajo es más probable descubrir y resolver rápidamente problemas que no se podrían abordar usando técnicas tradicionales, como las reuniones presenciales. El elevado número de personas implicadas en los proyectos de inteligencia colectiva probablemente incluirá a expertos en distintos campos que puedan supervisar áreas concretas del diseño, lo que no sería posible o asequible en un proyecto privado. En la realidad, no obstante, la inteligente colectiva podría tener problemas por falta de motivación (ya sea financiera o de otro tipo) para que la gente participe, o que haya demasiada información y no sea fácil gestionarla o usar sistemas como foros abiertos para tomar decisiones. En conclusión, los posibles problemas generados por el exceso de información, de colaboradores y su falta de motivación pueden verse superados por la capacidad de obtener y procesar grandes cantidades de información rápidamente para resolver problemas complejos.