

**BCS THE CHARTERED INSTITUTE FOR IT**

BCS HIGHER EDUCATION QUALIFICATIONS  
BCS Level 6 Professional Graduate Diploma in IT

**WEB ENGINEERING**

Monday 21<sup>st</sup> March 2016 - Morning

Answer **any** THREE questions out of FIVE. All questions carry equal marks.

Time: THREE hours.

**Answer any Section A questions you attempt in Answer Book A  
Answer any Section B questions you attempt in Answer Book B**

*The marks given in brackets are **indicative** of the weight given to each part of the question.*

Calculators are <b>NOT</b> allowed in this examination.
---

**The figures referenced in this question paper are provided in a separate booklet.**

**Section A**

Answer Section A questions in Answer Book A

**General Comments**

There is a large disparity between well-prepared candidates and those who were not ready to sit this examination. One comment from previous examiners' reports warrants repeating, since the Examiners continue to see answer pointers from old papers quoted verbatim in a small number of answer scripts:

*"It is important for candidates to know that whilst on occasion questions may look similar to those in past papers, the context and approach is often significantly different, which means that previous answers cannot simply be restated, thus it is not appropriate to memorise and restate past paper answers. Additionally, the answer pointers provided here give guidance and are only a guideline and should not be merely quoted by candidates but applied to the topic of the question."*

Moreover, when developing an argument (e.g. discussing the benefits or drawbacks of a certain approach), candidates should ensure they fully understand the meaning of the technical words they are using. In particular, the term "cost-effective" repeatedly appeared in a number of answers this year but used in a manner which suggested the candidate really meant "costly".

Candidates should also be aware that questions require application of knowledge rather than just restating knowledge.

## Section A

**A1.** Web developers face a number of technical choices when developing a website.

- a)** A solution stack is a complete set of software components required to run an application.
- i)** LAMP is a well-known solution stack for dynamic web sites. Give the name of every component in this stack and explain their role within the stack.  
**(4 marks)**
  - ii)** Give an example of an alternative to LAMP  
**(1 mark)**
- b)** Server-side scripting can be performed using a variety of programming languages.
- i)** Give the names of FOUR different server-side scripting languages.  
**(2 marks)**
  - ii)** Give THREE distinct factors which may lead a web developer to choose (or avoid) a particular server-side scripting language when embarking on a brand new project. Clearly explain the impact (risk/cost or benefit) of each factor on the project.  
**(3 marks)**
  - iii)** Are there any *different* factors that must be considered when choosing a client-side scripting language? Justify your answer with a clear explanation.  
**(2 marks)**
- c)** When developing complex web applications, it is often advised to reuse existing code (and supporting tools). These additional software components are typically distributed as libraries or structured as frameworks. An alternative to using these well-established solutions would be writing all the code yourself.
- i)** Give the name of a well-known web application framework (indicate the language it is written in)  
**(1 mark)**
  - ii)** Explain, with appropriate justification, when it is preferable to reuse external code, and when it is preferable to write your own. Give at least TWO arguments in favour of each approach.  
**(4 marks)**

- d)** Your company intends to launch a database driven website. A senior manager has suggested the company should buy its own server and host the website itself, using its existing broadband connection. Write a short presentation (no more than 10 slides) analysing this proposal. The presentation should:
- consider one or more alternatives
  - compare and contrast available options, considering a range of business and technical factors
  - conclude with a recommendation

**(8 marks)**

### **Answer pointers (A1)**

**a) i)**

LAMP originally means *Linux Apache MySQL PHP*:

- Linux is an operating system
- Apache is a web server
- MySQL is a database
- PHP (can also be Perl or Python) is a server-side scripting language

**a) ii)**

Any relevant stack, e.g. WISA (Windows Server, ISS, Microsoft SQL Server, ASP.NET) but also WAMP or MAMP (equivalent to LAMP on Windows or MacOS X)

**b) i)**

Could be any of PHP, Perl, Python, Ruby, ASP.NET, JSP, etc.

**b) ii)**

This is an open-ended question. Any objective, appropriately justified factor is acceptable.

For instance:

- how easy will it be to find a web host supporting the language? A well-supported language means a wider range of hosts to choose from, which can mean better value for money (more competition to offer lower prices or additional services)
- is the developer already familiar with the language (and if not, is it hard to learn?). Lack of familiarity and a difficult learning curve can mean a longer development time, and a greater risk of bugs
- how well documented is the language, and how easy is it to get support help (from the vendor - if applicable - or from the community). Again, poor support can lead to longer development time and the development of suboptimal solutions
- does the language offer strong security features, or is it open to vulnerabilities. Poor security could lead to the web site being compromised. This is particularly critical for certain types of applications (e.g. e-commerce)
- are there many libraries available? The existence of relevant libraries can greatly decrease development time.

**b) iii)**

Yes. In particular, we must consider whether the language is likely to be widely supported by the client (different web browsers on different types of devices, some of which may have limited computing capabilities), and how easy it would be for users to upgrade their client (e.g. install a plugin?) so that it becomes compatible.

**c) i)**

For instance Flask or Django (Python), Ruby on Rails (Ruby), Zend Framework or CodeIgniter (PHP)

**c) ii)**

The question is relatively open-ended.

Possible arguments in favour of reusing external code:

- it's often quicker (provided the tool is easy to learn, and does what you want out of the box, or is easy to configure to meet your needs. Good documentation and/or support is key)
- it can offer advanced features which the developer doesn't have the expertise to implement on their own
- it's relatively secure (a lot of time will have been dedicated, by developers and users alike, towards implementing reliable solutions to well-known and not so well-known security issues)
- it's easier for another developer to join the project (if they are already familiar with the reusable part of the code, or if the learning curve is not too steep)
- it can be seen as a badge of quality (reassuring, in terms of image)
- etc.

Possible arguments in favour of writing your own code:

- better understanding of the code, and greater freedom to customise it to meet your needs (more efficient)
- opportunity to become a more proficient programmer (learning experience)
- sometimes more secure (if a vulnerability in the reusable component becomes well-known and doesn't get fixed rapidly)
- greater control on the overall development environment, and less compatibility issues (relying on several reusable components can lead to compatibility issues when certain components are upgraded but others are lagging behind. E.g. if a library is only compatible with an older version of the scripting language or database)
- no dependence on a third party (e.g. the licensing terms of external component may change and make it unusable or too expensive)
- etc.

**d)**

This is an open-ended question.

The alternative to in-house hosting is outsourcing the web hosting to a third party company (different types of web hosting services may be considered, e.g. shared hosting, dedicated hosting, cloud hosting, etc.)

Possible factors to take into account:

- financial
  - o buying (up front cost) VS renting (long term cost)
  - o different pricing factors for third party hosting (data storage, bandwidth)
  - o impact on in-house infrastructure (electricity and broadband connection, physical space) + staff dedicated to server maintenance
- availability (mirrors/redundancy, reliability of the line)
- protection against data loss (back up strategy)
- flexibility (total freedom with own server, different degrees of freedom with third party providers - often correlated with price)
- etc.

The final recommendation should be consistent with the factors listed.

### Examiners' Guidance Notes (A1)

There was a wide disparity in the quality of answers for the entire question, with some candidates obtaining almost full marks, other candidates obtaining almost no marks, and most candidates falling somewhere in between.

The same kind of disparity was observed at the sub-question level: every sub-question was answered well by at least one candidate, but no single sub-question received perfect marks from all candidates.

The evidence shows that a few candidates performed equally well (or equally poorly) in all sub-questions. However, a larger proportion of candidates performed fairly well on some of the sub-questions, but answered poorly (or didn't answer at all) the rest of the question.

Such wide variations (especially in term of individual performance across different sub-questions) could be due to gaps in the knowledge of most candidates.

a) i) Only answered by about half of the candidates. When answered, the meaning of the acronym was usually correct (aside from occasional typos, in particular for "Apache"), but the role of each component was often missing or incorrect.

a) ii) Only answered by about half of the candidates. Acronyms such as "WAMP" had to be explained (e.g. "W = Windows") in order to attract full marks.

b) i) With a few exceptions (e.g. candidates listing client side scripting languages, mark-up languages, or similarly off-topic technologies), the evidence shows that most candidates could give the names of at least one or two languages, but often struggled to name more.

Note: languages which are not specific to server side scripting (e.g. C#, or Java) only received full marks if they were described in a server-side scripting context (e.g. C#.net, or JSP)

Frameworks such as .NET received no marks (a specific *language* should be mentioned, e.g. VB.NET)

b) ii) With a few exceptions, the question was generally poorly answered, and frequently misunderstood. In particular:

- answers confusing "language" and "source code" (e.g. written as if listing factors to choose or avoid a particular library/CMS/framework)

- answers listing the respective benefits and drawbacks of server-side scripting and client-side scripting

Some of these off-topic answers appeared to be quoted verbatim from previous reports.

Even when the question was correctly understood, there is evidence that candidates often failed to clearly explain (in practical terms: e.g. time gain/loss, risk of errors, etc.) the impact of each factor on the project.

b) iii) This question was often skipped (especially by candidates who misunderstood the previous question). When it was answered, factors such as client compatibility were often alluded to, but usually in vague terms. A clear explanation (again, mentioning practical impact, e.g. “the script won't execute correctly, possibly making the page unusable, and resulting in user dissatisfaction”) was required in order to attract full marks.

c) i) About half of the candidates could not name a framework: either the question wasn't answered at all, or a different type of web/Internet technology was mentioned (e.g. data format such as JSON, or web page editor such as Dreamweaver)

Among candidates who could name a framework, only about half could correctly indicate the language it was written in.

A number of answers attracted only partial marks, as they didn't exactly fit the usually accepted definition of “web application framework”:

- Wordpress is a CMS (but could, arguably, be used as a framework)
- the .NET framework is not specific to web applications (ASP.NET, however, is specific to the web and attracted full marks)

c) ii) This question received mixed answers. It was skipped by a number of candidates, and the evidence shows it was misunderstood by a small proportion of candidates, who instead discussed low level mechanisms such as external stylesheets, or the syntax to define reusable functions and variables.

When the question was correctly understood, a significant proportion of candidates failed to give a balanced account of benefits and drawbacks for both approaches.

As in other questions, the practical impact of a particular factor (e.g. time gain/loss, risk of errors, etc.) was often insufficiently explained.

d) This question was sometimes skipped (or barely answered). When attempted, the quality of answers varied widely, but the majority of candidates achieved an average mark.

Relatively frequent mistakes involved:

- failing to consider an alternative to the initial suggestion
- a narrow focus on a limited set of factor (e.g. only considering cost)
- failing to conclude with a clear recommendation

A small number of answers focussed on the development of the website instead of its hosting infrastructure.

**A2.** Since August 2014, several United States-based Healthcare providers and Health Insurance companies have had serious data breaches, which may affect in total almost 100 million users. Please read the following Website article excerpt on the topic and then answer the questions below:

### **Health Care Info Used to Hack Phone NFC**

A survey done by tax advisory firm KPMG reports that 81 percent of health care executives say their organization has been compromised by a data breach in the last two years. And the effects of those breaches can be seen reaching into pockets, wallets and phones — all at the same time.

After polling 223 chief information officers, chief technology officers, chief security officers and chief compliance officers at health care providers and health plans, KPMG "found the number of attacks increasing, with 13 percent saying they are targeted by external hack attempts about once a day and another 12 percent seeing about two or more attacks per week. More concerning, 16 percent of healthcare organizations said they cannot detect in real-time if their systems are compromised," as stated in a press release.

Consumer Affairs reports just a few of the recent attacks on the vulnerability of health care organizations. In August of 2014, Community Health Systems admitted that Chinese hackers had breached their network and made off with data on more than 4.5 million patients. The following February, the Anthem health-insurance network admitted that hackers had stolen up to 80 million medical records dating back to 2004.

In March, Premera Blue Cross admitted to a breach compromising 11 million medical and financial records dating back to 2002. CareFirst Blue Cross/Blue Shield admitted to a hacking in May. In mid-July, the UCLA Health System admitted that 4.5 million patient records were at risk from a hacking UCLA had discovered two months earlier, Consumer Affairs reports.

Malware, software designed to disrupt or gain access to private computer systems, is the most frequently reported line of attack on health care organizations during the past 12 to 24 months, according to 65 percent respondents in KPMG's survey. Botnet attacks, where computers are hijacked to issue spam or attack other systems, and "internal" attack vectors, such as employees compromising security, were cited by 26 percent of respondents.

According to the survey, the areas with the greatest vulnerabilities within an organization include external attackers (65 percent), sharing data with third parties (48 percent), employee breaches (35 percent), wireless computing (35 percent) and inadequate firewalls (27 percent).

Source: Excerpt from Forensicmag.com, "Health Care Info Used to Hack Phone NFC", <http://www.forensicmag.com/articles/2015/09/health-care-info-used-hack-phone-nfc> on 9<sup>th</sup> September, 2015

**a)**

**i)** For each of these three attack vectors on the business listed in the article, *outline* a suitable countermeasure that could be used to minimise risk or impact of:

- *Malware within the corporate network;*
- *Denial of service attack by botnets; and*
- 

*attacks from employees.*

*Internal*

**(3 marks)**

**ii)** *State and explain* THREE wired-network security risks that could result in disclosure of user data. (N.B.: when answering this question, you should not list risks that are solely client- or server-based).

**(6 marks)**

iii) Wireless networking (e.g. 802.11n) is often used for convenience. *State and explain* ONE additional security risk introduced by the use of wireless networks, and provide a suitable countermeasure.

**(2 marks)**

iv) Customers may wish to access services using mobile devices (e.g. mobile phones, tablets), where there is a real risk of losing the device. Provide TWO countermeasures that can be used to could be used to minimise risk or impact of a lost device.

**(2 marks)**

b)

i) One approach to reducing risk after a serious compromise is to disable all user accounts. With a suitable example, *describe* ONE positive consequence of this action and, ONE negative consequence of this action. (N.B.: this may be considered from either the corporate or user perspective).

**(2 marks)**

ii) After disabling all user accounts, the company might ask the user to re-activate them after proving their identity.

*Explain* why each of the following three normal methods of account verification might not be safe:

a) Sending an *activation code* to the users email address.

b) Writing a letter to the user's home address asking them to call a *telephone number to get a new password*.

c) Asking the user *questions about their family* (e.g. ages of siblings) over instant messaging.

**(3 marks)**

iii) Taking into account the weaknesses of the methods discussed in part c) ii), *propose* (with appropriate justification) a robust method of verifying user's identities that offers a good balance between *user account safety* and cost effectiveness.

**(3 marks)**

c) The article reports that one in six providers cannot detect in real time if their systems are compromised.

In no more than 300 words, *explain* what benefits and drawbacks real-time monitoring may offer to the management of security risks and – with appropriate justification – recommend whether it is worthwhile for Healthcare Providers to use real-time monitoring.

**(4 marks)**

## Answer pointers (A2)

a) i) One mark for each appropriate countermeasure provided, but answers must be sufficiently detailed and properly put into context of the question brief.

- *Malware* – prevent this by using up to date virus checkers, user education and perimeter controls to block incoming connections.
- *Denial of service attacks by botnets* – Provisioning extra capacity and using load balancing services to distribute the load, firewalls can be configured to block simple attacks e.g. SYN flooding, intrusion detection systems could be installed to try and identify DDoS attacks, compartmentalisation of the network can help mitigate the effects by limiting the amount of services that are affected by a single attack.
- *Internal attacks by employees* – proper separation of privileges, appropriate user access controls, appropriate security policies, and user education.

a) ii) Sample risks (with possible mitigation in brackets) include:

- Eavesdropping/Packet sniffing to reveal plaintext passwords (encryption).
- Man-in-the-middle attack (SSL with a mutually trusted Certificate Authority)
- Identity spoofing/IP Address Spoofing)
- Replay attack (time-stamping)

No marks will be given for non-network based attacks (e.g. social engineering, installation of key-loggers) or repetition of attacks from part i), and the answers must once again be sufficiently detailed.

Attacks which are on the client- or server-side but which are facilitated via the Network (e.g. SQL injection, session hijack) will get *partial* credit, depending on quality of explanation but must be **explicitly** related to the network.

[Up to 2 marks for each fully explained *distinct* risk to the network, simply listing a risk without explanation will gain only a half mark]

a) iii) Sample risks (with possible mitigation in brackets) include:

- over-the-air eavesdropping (use wireless encryption)
- connection by unauthorised individuals to unsecured networks (allowing misuse of the network) (use wireless encryption, MAC address white listing, not broadcasting the Access Point name)
- compromise of security keys for e.g. WEP (use a more secure wireless encryption protocol such as WPA)

- Setting up of spoof network access points as a vector for a man-in-the-middle attack (use of a Wireless Intrusion Prevention System, or use of EAP protocols)

[One mark for an appropriate risk fully explained, one mark for an appropriate countermeasure appropriately defined or explained. Risks that are not wholly related to wireless networking (in the context of 802.11) will not get credit, and merely providing a name will get minimal credit. Risks that are extremely close to ones provided in part ii) will obtain only a half mark for each repeated risk/countermeasure unless sufficiently explained].

**a) iv)** Countermeasures might include:

- Full device encryption
- Secure password policies and device wiping after 5 incorrect attempts
- Remote blocking/wiping of lost devices over the air
- Informing a web service of the current geographical location of the device

[One mark for each appropriate countermeasure provided; weak countermeasures like setting a standard password/PIN will get only partial credit].

**b) i)** Candidates should list one positive [1 mark] and one negative [1 mark] consequence with a good example, consequences listed must be relevant to the question. Some sample consequences might include:

Positive	Negative
Since all accounts have been disabled, Brute forcing the encrypted passwords will not give them access to the service.	Inconvenience to users by denying them access to the service until they can unlock the account.
Disabling the accounts will ensure that users are aware of the compromise next time they attempt to access the service.	The request to change passwords does not mitigate against other risks e.g. possible identity theft, or that the password might be used on other systems (e.g. the email provider).

**b) ii)** Answers will be marked on merit, with 1 mark for each section given for a well-reasoned justification, and a viable risk to safety (improbable answers would get minimal or no marks). An example of possible answers:

- Sending an activation code to the users email address.*

Many users have the same password or multiple systems, so there are likely to be users who re-use the password from their email account on eBay. In that instance, if the attackers can break the eBay password (e.g. by using rainbow tables) they would be able to gain access to the user's email account as well, and so could intercept the activation code.

*b. Writing a letter to the user's home address asking them to call a telephone number to get a new password.*

The address details were also compromised, so attackers may well send their own letters to the users masquerading as the service provider (perhaps to e.g. conduct additionally phishing for other personal details), or other people at the same address might open the letter and misuse the contents.

*c. Asking the user questions about their family (e.g. ages of siblings) over instant messaging.*

It is possible that the personal records contain details of siblings (indeed, this method of verification would not work unless they do have these details, and the lack of these details might be another potential problem), and moreover those siblings might also have accounts with the same provider (which will have also been compromised). Furthermore, providers of genealogical data – or public records – may provide a route for third parties (i.e. those previously not involved in the data breach) to obtain the information and hijack the users account.

**b) iii)** An open question, 1 mark for the outline of the solution, 2 marks for justification and fitness for purpose of the solution. Candidates must propose a method that has not been mentioned elsewhere in the question, and that either does *not* suffer from the weaknesses in part c) ii), or is acknowledged and with a justification as to why this is not important.

Answers that focus too much on either the side of system security (with impacts on either cost or usability), or aiming for a cheap solution (with impacts on security) will be capped at a maximum of 1 mark, unless they provide a very good justification as to why the decision should be taken, and include some acknowledgement of the weaknesses.

**c)** This is a discussion question. Marks will be allocated for:

i) The depth of discussion on the benefits and drawbacks of real-time monitoring [1 mark for each aspect, 1 discretionary mark for the quality of the overall argument], with full marks granted for a balanced consideration (including financial or opportunity costs of real-time monitoring).

ii) Stating a definite answer to the question [1 mark]; this is marked on the quality of this conclusion and degree to which is supported by the main argument.

## Examiners' Guidance Notes (A2)

The candidates provided answers that did not match the question. There is evidence that the content of these answers were from previous answer pointers and therefore achieved little or no marks.

a) i) Some candidates answered this extremely well, but others were unaware of the nature of these attacks, and many described the attacks rather than outlining a suitable countermeasure.

a) ii) Generally very poorly answered and omitted by many candidates. There is evidence that many candidates seemed unable to provide network based attacks, despite the fact that sniffing, spoofing and MITM are standard attacks in Security. Often candidates did not appreciate that the question specifically asked for *wired network* risks, and many candidates even repeated the attacks provided in part i) – which did not attract any marks.

a) iii) Very poorly answered. Candidates typically provided generic security risks (e.g. sniffing) without relating it to wireless networking.

a) iv) Generally well answered, though some candidates proposed the weaker countermeasure of password protection without expanding this into device wiping or full disk encryption.

b) i) Generally very poorly answered. Candidates did not explicitly point out the risks of user access using compromised data, but instead focussed on the data breach itself. The drawbacks were very rarely framed in terms of business impact (arguably the greatest impact), but instead user dissatisfaction due to being hacked, which would be present regardless of the method of protecting the service post-hacking.

b) ii) Typically well answered, although candidates often focused on weaknesses that were not related to safety (e.g. a user no longer having access to their email address) – these did not attract any marks.

b) iii) Reasonably answered, but very few candidates achieved full marks. Typical answers either focused too much on having a secure system (with no consideration of the drawbacks), or proposed security measures that would be ineffective (e.g. using CAPTCHA to avoid brute forcing – this was not a credible risk present in any of the previous solutions).

c) Generally omitted entirely, and of those that attempted this part it was poorly answered. Candidates were typically able to identify the cost or resource element of real time monitoring but were unable to articulate the benefits of implementing such a system, nor able to draw a well-supported conclusion.

## Section B

Answer Section B questions in Answer Book B

**B3. a)** Data can be stored in child elements or in attributes as demonstrated by the example in Figure 3.1. Briefly state 5 factors to be considered when using attributes instead of child elements. **(5 marks)**

**b)** The XML document in Figure 3.2 contains a number of errors when validated against the DTD (document data type) in Figure 3.3. Identify all the errors, and provide a solution for each. **(5 marks)**

**c)** An XML schema can be written as a DTD or as XSD (xml schema document). State the benefits and drawbacks of each approach. Convert the DTD in Figure 3.3 to an XML schema. **(5 marks)**

**d)** Referring to XML document in Figure 3.4, write a DTD which enforces the following constraints on `cbcatalogue`:

- Sequence of elements is as shown in the XML document (i.e. `cbcatalogue` is a container of `subject` elements).
- `subject` element must be present zero or more times.
- Each `subject` element contains a `subject_title` and zero or more `books`.
- The two attributes in `book` are mandatory.
- A `cover image` for each `book` that captures attributes of height, width and location of image
- `book_url` is optional.
- `book_url` must have a `page` attribute.
- For each `book`, exactly one `book_title` and one `Publisher` element must be present
- For each `book`, there must be one or more `authors`.
- An `author` has a `name` and optionally a `short biography` and an `email`.
- A `name` consists of both a `first_name` and a `last_name`.
- For each `book`, there must be one or more `review`.

**(10 marks)**

### Answer pointers (B3)

- a) Some of the problems with attributes are:
- attributes cannot contain multiple values (child elements can)
  - attributes are not easily expandable (for future changes)
  - attributes cannot describe structures (child elements can)
  - attributes are more difficult to manipulate by program code
  - attribute values are not easy to test against a DTD

- using attributes as containers for data, leads to documents that are difficult to read and maintain.

Use elements to describe data. Use attributes only to provide information that is not relevant to the data.

b) Errors of syntax, well-formed and valid xml.

c) DTD vs XML Schema:

DTD's are not namespace aware.

DTD's have #define, #include, and #ifdef – or, less C-oriented, the ability to define shorthand abbreviations, external content, and some conditional parsing.

A DTD describes the entire XML document (even if it leaves "holes"); a schema can define portions.

XSD has a type system.

XSD has a much richer language for describing what element or attribute content "looks like." This is related to the type system.

You can put a DTD inline into an XML document; you cannot do this with XSD.

You have far more control over what is considered a valid XML document using a schema.

You can even extend your types from other types you've created, require uniqueness within scope, and provide lookups.

```
<?xml version="1.0" encoding="UTF-8"?>
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema"
elementFormDefault="qualified">
  <xs:element name="cv">
    <xs:complexType>
      <xs:sequence>
        <xs:element ref="preface"/>
        <xs:choice maxOccurs="unbounded">
          <xs:element ref="qualification"/>
          <xs:element ref="experience"/>
        </xs:choice>
        <xs:element minOccurs="0" ref="hobbies"/>
        <xs:element minOccurs="0" maxOccurs="unbounded" ref="referee"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:element name="preface" type="xs:string"/>
  <xs:element name="qualification" type="xs:string"/>
  <xs:element name="experience" type="xs:string"/>
  <xs:element name="hobbies" type="xs:string"/>
  <xs:element name="referee" type="xs:string"/>
  <xs:element name="companyAddress" type="xs:string"/>
</xs:schema>
```

d)

```
<?xml version = "1.0" encoding="ISO-8859-1"?>
<!ELEMENT cbcatalogue (subject*)>
<!ELEMENT subject (subject_title,book*)>
<!ELEMENT subject_title (#PCDATA)>
<!ELEMENT book (book_title,image*, author+,book_url?,publisher, review+)>
<!ATTLIST book isbn CDATA #REQUIRED edition CDATA #REQUIRED>
<!ELEMENT book_title (#PCDATA)>
<!ELEMENT image EMPTY>
<!ATTLIST image height CDATA #REQUIRED>
<!ATTLIST image width CDATA #REQUIRED>
<!ATTLIST image
src CDATA #REQUIRED >
<!ELEMENT author (name, email, bio)>
<!ELEMENT name (Last_name,First_name)>
<!ELEMENT Last_name (#PCDATA)>
<!ELEMENT First_name (#PCDATA)>
<!ELEMENT email (#PCDATA)>
<!ELEMENT bio (#PCDATA)>
<!ELEMENT book_url EMPTY>
<!ATTLIST book_url page CDATA #REQUIRED>
<!ELEMENT publisher (#PCDATA) >
<!ELEMENT review (#PCDATA)>
```

### **Examiners' Guidance Notes (B3)**

Most candidates were able to answer this question and competently demonstrated the knowledge regarding the topic. There was a minority who did not know the difference between XML schema and DTD.

Most candidates had a good attempt at writing the DTD, however for some, nesting of elements and use of optional and required attributes caused problems.

**B4.**

- a) i) Briefly explain how XPath expressions can manipulate an XML document.  
**(2marks)**
- ii) With reference to the XML file shown in Figure 4.1, write an XPath expression to show all details of all products supplied by Coste.  
**(3 marks)**
- b) The XML document in Figure 4.2 (list of journals and articles) needs to be displayed as a web page as shown in Figure 4.3. Using the HTML template provided in Figure 4.4, in your answer book provide the missing code. Use the section marked **<!--TO BE COMPLETED -->** to accomplish this. You may assume that the `journal.css` file exists and no style code needs to be written.  
**(5 marks)**
- c) Modify the HTML template further so that the XML file is rendered as shown in Figure 4.5. The journal title and the publisher information are to be active links to the respective websites. In your answer book write the code to do this.  
**(5 marks)**
- d) Modify the HTML template further so that the XML file is rendered as shown in Figure 4.6. The editor email address is an active email link. In your answer book write the code to do this.  
**(5 marks)**
- e) The last stage in rendering the XML file is to display the details of the articles within each journal as shown in Figure 4.7. In your answer book write the code to do this.  
**(5 marks)**

## Answer pointers (B4)

XPath expression specifies a pattern that selects a set of XML nodes such as Root, Element, Text, etc. XPath expressions are used in XSLT to transform XML document to HTML pages.

"/ReadyMeals/\*/\*[@supplier='Coste']".

b-e).

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:template match="/">
    <html>
      <head>
        <title>Catalogue of Journals</title>
      </head>
      <body>
        <h1>Catalogue of Journals - updated January 2016</h1>
        <table border="1">
          <tr bgcolor="#9acd32">
            <th align="left">Title</th>
            <th align="left">Journal Details</th>
            <th align="left">Editor</th>
            <th align="left">Publisher</th>
          </tr>
          <xsl:for-each select="JournalCatalogue/Journal">
            <tr>
              <td align="left">
                <a href="{Journal_url/@page}"
                target="_blank">
                  <xsl:value-of
                    select="Journal_title"/></a></td>
                <td align="left">
                  <xsl:value-of
                    select="@ISBN"/><br/>
                    <xsl:value-of
                    select="@Issue_no"/><br/>
                    <xsl:value-of
                    select="Editor"/><br/>
                    email:<a href="mailto:{email}">
                      <xsl:value-of
                        select="Editor_email"/>
                    </a> </td>
                <td align="left">
                  <a href="{Publisher_url/@page}"
                  target="_blank">
                    <xsl:value-of
                      select="Publisher"/></a></td>
              </tr>
            <tr>
              <td colspan="4" align="center">
                <table border="1"
                bgcolor="#e7e767" width="90%">
```

```

Title</th><th>Abstract</th><th>Author(s)</th></tr>
select="Article">
select="Article_title"/></td>
select="Abstract"/></td>
select="Author">
select="Name/First_name"/>&#xa0;<xsl:value-of select="Name/Last_name"/>
<br/><em><xsl:value-of select="Occupation"/></em>
</li>
</xsl:for-each>
</ol>
</td>
</tr>
</xsl:for-each>
</table>
</td>
</tr>
</xsl:for-each>
</table>
</body>
</html>
</xsl:template>
</xsl

```

## Examiners' Guidance Notes (B4)

Some candidates produced excellent solutions to this question demonstrating their competence in manipulating XPATH expressions.

Each part of the question was built on the previous part to better render the XML data as a web page using the images to extract the relevant information.

The use of nested *for loops* was at the heart of the problem with the knowledge of how to render *email* addresses and *urls*.

Some candidates rewrote the entire stylesheet for each and every part, when it would have been sufficient to write the new lines code for each part from the previous part.

There were a minority of candidates who wrote HTML scripts for each part using the images as a guide and did not get any marks for it – the question clearly stated this was about XSLT.

**B5. a)** Define the term API and state the purpose it serves. **(4 marks)**

**b)** Companies such as eBay™ and Google™ provide public APIs.  
State FOUR benefits of APIs. **(3 marks)**

**c)** Identify the stages and issues when creating an API from a producer's viewpoint. **(4 marks)**

**d)** Data can be returned as XML or JSON by an API call. As a developer, give two distinct examples where each of these is most suitable, based on the type of device in use during the API call. **(4 marks)**

**e)** As a Web Engineer, identify at least FIVE major challenges in the design and development of an image-hosting site, where users can upload their images to a central server and the images can be retrieved via a web link or an API.

Initially you are to focus on two aspects only:

- the ability to upload (write) an image to the server
- the ability to query for an image.

For each of the challenges identified, provide an outline solution including the technology to be used. Your answer should be in the form of a slide presentation to the client (5 slides with bullet points). **(10 marks)**

**NOTE:** Your answers should focus on the various technologies and challenges in using those technologies. No credit will be given for “text book” answers that follow the Software Development Life Cycle Model (SDLC).

### **Answer pointers (B5)**

- a) Application Programming Interface. Controlled access to an organisation's data. Allows a third party developer to utilise the data to provide enhanced service, etc.
- b) Revenue, Standards, etc.
- c) Planning,
  - Enabling authentication and encryption
  - Deciding between REST and SOAP
  - Performance considerations
  - Error handling
- d) XML easy to read, verbose, more processing. JSON light, efficient ideal for mobile devices e.g. weather on a mobile phone.
- e) Availability; Performance; Reliability; Scalability; Manageability; Cost

### **Examiners' Guidance Notes (B5)**

This was the least popular question. Most candidates had a good knowledge of APIs but did not fully articulate the benefits to the producers and instead focused on benefits for users.

The stages and issues in designing an API were not fully explained in answers and the quality of responses varied immensely.

The XML vs JSON question was reasonably well answered.

The challenges and issues for image-hosting site were better articulated and much more specific. Some candidates still approached this from a Software Development Life Cycle Model despite the advice in the question and scored no marks.

**BCS THE CHARTERED INSTITUTE FOR IT**  
**BCS HIGHER EDUCATION QUALIFICATIONS**  
**BCS Level 6 Professional Graduate Diploma in IT**

**FIGURES TO ACCOMPANY 2016 EXAMINATION PAPER IN  
WEB ENGINEERING**

**option 1:**

```
<booking reference="20120521" date="12/05/2012">  
  <agent>M.Miles</agent>  
  <client_name>Jane Doe</client_name>  
  <client_email>jdoe@jmail.com</client_email>  
  <advice>No changes allowed</advice>  
</booking>
```

**option 2:**

```
<booking reference="20120521">  
  <date>12/05/2012</date>  
  <agent>M.Miles</agent>  
  <client_name>Jane Doe</client_name>  
  <client_email>jdoe@jmail.com</client_email>  
  <advice>No changes allowed</advice>  
</booking>
```

**option3:**

```
<booking reference="20120521">  
  <date>  
    <day>12</day>  
    <month>05</month>  
    <year>2012</year>
```

```

    </date>

    <agent>M.Miles</agent>

    <client_name>
        <first_name>Jane</first_name>
        <last_name>Doe</last_name>
    </client_name>

    <client_email>jdoe@jmail.com</client_email>

    <advice>No changes allowed</advice>
</booking>

```

**Figure 3.1**

**Figure 3.2**

```

1 <?xml version="1.0"?>
2 <!DOCTYPE cv SYSTEM "cv.dtd">
3 <cv>
4   <preface> Personal Data</preface>
5   <preface> Name: Jane Doe</preface>
6   <Qualification>BSc</qualification>
7   <experience>Charity work</experience>
8   <qualification>MSc</qualification>
9   <hobbies>running</hobbies>
10  <experience>Charity work</experience>
11  <qualification>MSc</qualification>
12  <referees>Joe Blogs, BSc</referees>
13  <!--<hobbies> kite flying </hobbies> -->
14 </cv>

```

**Figure 3.3**

```

1 <!ELEMENT cv (preface, (qualification | experience)+, hobbies?,
referee*)>
2 <!ELEMENT preface (#PCDATA)>
3 <!ELEMENT qualification (#PCDATA)>
4 <!ELEMENT experience (#PCDATA)>
5 <!ELEMENT hobbies (#PCDATA)>
6 <!ELEMENT referee (#PCDATA)>

```

## Figure 3.4 Figure 4.1

```
<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<!DOCTYPE cbcatalogue SYSTEM "cbcatalogue.dtd">
<?xml-stylesheet type="text/xsl" href="cbbook05.xsl"?>
<cbcatalogue>
  <subject>
    <subject_title>Web Technology </subject_title>
    <book isbn="0137001312" edition="1st">
      <book_title>JavaScript for Programmers</book_title>
      <image src="javascript.jpg" height="100" width="100" />
      <author>
        <name>
          <Last_name>Deitel</Last_name>
          <First_name>Paul</First_name>
        </name>
        <email>deitel@deitel.com</email>
        <bio>Paul J. Deitel, Chief Technical Officer of Deitel
and Associates.</bio>
      </author>
      <author>
        <name>
          <Last_name>Deitel</Last_name>
          <First_name>Harvey</First_name>
        </name>
        <email>deitel@deitel.com</email>
        <bio>Chairman and Chief Strategy Officer of Deitel and
Associates.</bio>
      </author>
      <book_url page="http://www.deitel.com/books"/>
      <publisher>Prentice Hall</publisher>
      <review>Demonstrating use of commonly available tools to
create dynamic webpages and
server side technologies.</review>
    </book>
    <book isbn="067232797" edition="3rd">
      <book_title>Sams Teach Yourself XML in 24 Hours</book_title>
      <image src="xml.jpg" height="100" width="100" />
      <author>
        <name>
          <Last_name>Morrison</Last_name>
          <First_name>Michael</First_name>
        </name>
        <email>morrisonm@michaelmorrison.com</email>
        <bio>Michael Morrison is a professional Java
programmer.</bio>
      </author>
    </book>
  </subject>
</cbcatalogue>
```

```

        </author>
        <publisher>SAMS</publisher>
        <review>As an introductory text on XML, it delivers.</review>
    </book>
</subject>
<subject>
    <subject_title>Databases</subject_title>
    <book isbn="9781408007686" edition="2nd">
        <book_title>Database Management Systems</book_title>
        <image src="dbms.jpeg" height="100" width="100" />
        <author>
            <name>
                <Last_name>Ward</Last_name>
                <First_name>Patricia</First_name>
            </name>
            <email>emea.info@cengage.com</email>
            <bio>Missing</bio>
        </author>
        <book_url page="http://www.cengage.co.uk/fasttrack"/>
        <publisher>Course Technology</publisher>
        <review>Exactly what is required in an introductory database
course.</review>
    </book>
</subject>
</cbcatalogue>

```

```

<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<ReadyMeals>
    <Food>
        <lunch supplier="Coste" id="1">
            <price>3.50</price>
            <quantity>20</quantity>
            <calories>1500</calories>
        </lunch>
        <dinner supplier="Daas" id="2">
            <price>4.50</price>
            <quantity>50</quantity>
            <calories>2000</calories>
        </dinner>
    </Food>
    <Drink>
        <juice supplier="Coste" id="3">

```

```

        <price>1.50</price>
        <quantity>60</quantity>
        <calories>180</calories>
    </juice>
</Drink>
</ReadyMeals>

```

## Figure 4.2

```

<?xml version="1.0" encoding="ISO-8859-1" standalone="no"?>
<?xml-stylesheet type="text/xsl" href="Journal.xsl"?>
<JournalCatalogue>
    <Journal ISBN="10104556" Issue_no="24">
        <Journal_title> Nature</Journal_title>
        <Journal_url page="http://www.nature.com"/>
        <Publisher>Wiley</Publisher>
        <Publisher_url page="http://www.wiley.com"/>
        <Editor>Bob Jones</Editor>
        <Editor_email>bjones@bcs.ac.uk</Editor_email>
        <Article>
            <Article_title>Go green</Article_title>
            <Abstract>Global warming and eco systems</Abstract>
            <Author>
                <Name>
                    <Last_name>Simpson</Last_name>
                    <First_name>Bart</First_name>
                </Name>
                <Author_email>simpsb@eco.com</Author_email>
                <Author_url page="www.eco.com"/>
                <Occupation>Academic</Occupation>
            </Author>
            <Author>
                <Name>
                    <Last_name>Simpson</Last_name>
                    <First_name>Lisa</First_name>
                </Name>
                <Author_email>simpl@eco.com</Author_email>
                <Author_url page="www.eco.com"/>
                <Occupation>Scientist</Occupation>
            </Author>
        </Article>
        <Article>
            <Article_title>Forestation</Article_title>
            <Abstract>Flood Management</Abstract>
            <Author>
                <Name>
                    <Last_name>Hobs</Last_name>

```

```

        <First_name>Harold</First_name>
        </Name>
        <Author_email>Hhobs@eco.com</Author_email>
        <Author_url page="www.eco.com"/>
        <Occupation>Engineer</Occupation>
    </Author>
</Article>
</Journal>
<Journal ISBN="10104557" Issue_no="7">
<Journal_title> Wired</Journal_title>
<Journal_url page="http://www.wired.com"/>
<Publisher>New Riders</Publisher>
<Publisher_url page="http://www.newriders.com"/>
<Editor>Bob Jones</Editor>
<Editor_email>bjones@bcs.ac.uk</Editor_email>
<Article>
    <Article_title>Convergence</Article_title>
    <Abstract>One device as a container for all other devices</Abstract>
    <Author>
        <Name>
            <Last_name>Bates</Last_name>
            <First_name>Bill</First_name>
        </Name>
        <Author_email>bates@wires.com</Author_email>
        <Occupation>Academic</Occupation>
    </Author>
</Article>
<Article>
    <Article_title>USB3</Article_title>
    <Abstract>New standards</Abstract>
    <Author>
        <Name>
            <Last_name>Jones</Last_name>
            <First_name>Joe</First_name>
        </Name>
        <Author_email>jonoes@wires.com</Author_email>
        <Occupation>Engineer</Occupation>
    </Author>
</Article>
</Journal>
</JournalCatalogue>

```

---

# Catalogue of Journals - updated January 2016

Title	Journal Details	Editor	Publisher
Nature	ISBN:10104556 Issue No:24 URL: <a href="http://www.nature.com">http://www.nature.com</a>	Bob Jones bjones@bcs.ac.uk	Wiley
Wired	ISBN:10104557 Issue No:7 URL: <a href="http://www.wired.com">http://www.wired.com</a>	Bob Jones bjones@bcs.ac.uk	New Riders

Figure 4.3

Figure 4.4

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<xsl:stylesheet xmlns:xsl="http://www.w3.org/1999/XSL/Transform" version="1.0">
  <xsl:template match="/">
    <html>
      <head>
        <title>Catalogue of Journals</title>
      </head>
      <body>
        <h1>Catalogue of Journals - updated January 2016</h1>
        <table border="1">
          <tr bgcolor="#9acd32">
            <th align="left">Title</th>
            <th align="left">Journal Details</th>
            <th align="left">Editor</th>
            <th align="left">Publisher</th>
          </tr>
          ----
          ----
          <!-- TO BE COMPLETED -->
          ----
          ----
        </table>
```

```

        </body>
    </html>
</xsl:template>
</xsl:stylesheet>

```

## Catalogue of Journals - updated January 2016

Title	Journal Details	Editor	Publisher
<a href="#">Nature</a>	ISBN:10104556 Issue No:24	Bob Jones bjones@bcs.ac.uk	<a href="#">Wiley</a>
<a href="#">Wired</a>	ISBN:10104557 Issue No:7	Bob Jones bjones@bcs.ac.uk	<a href="#">New Riders</a>

Figure 4.5

## Catalogue of Journals - updated January 2016

Title	Journal Details	Editor	Publisher
<a href="#">Nature</a>	ISBN:10104556 Issue No:24	Bob Jones email: <a href="mailto:bjones@bcs.ac.uk">bjones@bcs.ac.uk</a>	<a href="#">Wiley</a>
<a href="#">Wired</a>	ISBN:10104557 Issue No:7	Bob Jones email: <a href="mailto:bjones@bcs.ac.uk">bjones@bcs.ac.uk</a>	<a href="#">New Riders</a>

Figure 4.6

# Catalogue of Journals - updated January 2016

Title	Journal Details	Editor	Publisher
<a href="#">Nature</a>	ISBN:10104556 Issue No:24	Bob Jones email: <a href="mailto:bjones@bcs.ac.uk">bjones@bcs.ac.uk</a>	<a href="#">Wiley</a>
<b>Article Title</b>	<b>Abstract</b>	<b>Author(s)</b>	
Go green	Global warming and eco systems	1. Bart Simpson <i>Academic</i> 2. Lisa Simpson <i>Scientist</i>	
Forestation	Flood Management	1. Harold Hobs <i>Engineer</i>	
<a href="#">Wired</a>	ISBN:10104557 Issue No:7	Bob Jones email: <a href="mailto:bjones@bcs.ac.uk">bjones@bcs.ac.uk</a>	<a href="#">New Riders</a>
<b>Article Title</b>	<b>Abstract</b>	<b>Author(s)</b>	
Convergence	One device as a container for all other devices	1. Bill Bates <i>Academic</i>	
USB3	New standards	1. Joe Jones <i>Engineer</i>	

Figure 4.7