This test consists of two parts. Part 1, Developing a Web Application and Part 2, Technical Questions. You will have 3 hours to complete both parts.

**Part 1, Developing a Web Application (120 points)**

The general purpose of the web application is to process MS Word (docx) document using Lucene.NET and generate comments with number of occurrences for every paragraph appearing more than once.

**Functional requirements:**

* As a user I would like to upload MS Word (docx) file
* As a user I would like to have my file processed using Lucene.NET in a way that every paragraph repeated several times in the document gets a comment added to it with the number of occurrences.

**Technical requirements:**

* Use ASP.NET MVC (with C#) framework to create the web application
* Use Lucene.NET for full text indexing and searching
* Use OpenXML SDK for MS Word (docx) processing

**The application will be assessed based on the following criteria:**

* Each requirement will be scored based on completeness and correctness. Each requirement is worth 50 points, for a total of 100 points
* Organization of components – how well was the software designed according to object oriented design principles(DRY, SOLID), consistent naming conventions and aesthetics of application (10 points)
* Security – no login is required, however, there should not be any potential security flaws (10 points)

**Part 2, Technical Questions (80 points)**

1. When developing a web application that will be accessed by users in different locations/time zones, what considerations would you take when planning the application design and infrastructure? (20 points)

**Answer**:

Availability should be taken into consideration, as users will access the application at different times. Even more important is establishing a consistent model to save Datetime variables, for example saving all times as Utc Datetime.

For countries using characters in the Unicode spectrum to display correctly web pages we should consider a page encoding appropriate, generally UTF-8.

For display Arabic right-to-left text probably a custom view is needed.

Chinese/Japanese can read up-to-down or left-to-right (even if they read the page last-to-first) so probably web page should stick with left-to-right.

1. Describe the most common security vulnerabilities to a web application and how you can prevent them (15 points)

**Answer**:

The most common security threats in web application as for OWASP 2017 are still: injection and cross-site scripting.

SQL injection as well as XML parser injection or script injection occur when untrusted data from the client side is processed or stored without a previous decode/parse phase. XML is relevant because there are many SOAP web services which accept XML and they can target also vulnerabilities in older XML processors to allow remote requests from the server, extraction of data et al.

To mitigate SQL injection parameterized queries should be used and generally untrusted data should be sanitized before storage or before reflecting it to the client, as is the case of cross site scripting.

XSS is the second most prevalent issue and occurs when untrusted data is reflected to the client or stored and later displayed as is. In the case of Javascript scripts this can cause a wide tange of damage as malicious script is directly executed.

Cross Site Request Forgery occurs when an attacker steals auth cookie to forge a request to be send and act impersonating another user. it can be mitigated adding a random token to the request, so as to make more difficult for the attacker to replicate a valid request as there is another piece beside the auth cookie that the server checks.

SSL should be used, also for the login page as the scenario where the login page is http and the rest of the site is in https is susceptible to man in the middle attacks.

(3) Describe the most common approaches used for MS Word (docx) files processing (15 points)

**Answer**:

OPEN XML SDK is open and more flexible.

Visual Studio Tools for Office VSTO uses COM interops and allows to create add-ins.

(4) Describe multi-tenancy and how it could be useful for a project which is offering SaaS (Software as a Service) (20 points)

**Answer**:

Multi-tenancy means using a single database for various installation. It can be useful in the sense that it speeds up the time to enable a new client to use the platform and limits the time needed instead to maintain different installations, but multi-tenancy limits also the capability of customization, especially when clients have unique requirements that need to be addressed for the on-boarding.

(5) Have you ever used an ORM? If so, which one(s). Provide characteristics of the ORMS you have used. (e.g. benefits/disadvantages) (10 points)

**Answer**:

I have used Entity Framework Code First. The advantage of using an ORM is a rapid life-cycle of adapting to changes. Also the fact that it allow to use an object model, with LINQ, to make queries to the db makes it easier. Usage of an ORM comes , at the cost of less control over the queries that the ORM produces. Nevertheless EF make it possible to execute directly raw SQL or to execute custom stored procedures. The latter may be the case when reporting needs to produce sets of data as the result of advanced SQL constructs, as when a stored procedure uses PIVOT constructs with or without the presence of dynamic columns, in which case STUFF FOR XML and sp\_executesql inside the stored procedure have no ORM counterpart.

**Delivery Protocol:**

* Zip your solution and name the file firstname\_lastname.zip. For example, if your name is Mary Smith, the file should be named mary\_smith.zip
* In the zip file include a backup of any databases which might be needed
* Include answers to the five questions in the email
* **Please be sure to email whatever you have completed as soon as the allotted time is up.** You may continue to work afterwards and provide a more complete solution later. However, the work that was delivered at the deadline will be considered first.
* Upload the zip file(s) to the location specified in the email you received. **Please do not replace files which were already uploaded, as we will check the dates to ensure that they were delivered on time.** You can of course continue to upload files with another name.