A Proposed Methodology for Identifying and Evaluating an E-Forms Application

Prof. Dr. Imad H. AL-Hussaini¹, Mustafa Ismail Ibrahim²

¹Iraqi Commission for Computers and Informatics / Informatics Institute for Postgraduate Studies, Baghdad, Iraq
²Computer Center / Federal Board of Supreme Audit, Baghdad, Iraq

Abstract: In order to develop a specific methodology that can be followed by any organization or individual who needs to choose the appropriate E-Forms application that could meet the required needs, this research proposes a methodology which consists of three phases: Identification phase, Evaluation phase and Testing phase. In the first phase, a general survey on E-Forms types and technologies is introduced with focus on the recent trends in order to identify which type mostly meets the user's general requirements. In the second phase, an evaluation to make an accurate selection is made to some of the top competing applications supporting the type identified in first phase by using a specific evaluation methodology, in this research three of the widely used evaluation methodologies are discussed, in addition to a simple evaluation methodology which has been proposed and implemented using "Microsoft Excel Sheets" application. Finally in the third phase, a real test is made using the winner application resulting from the second phase, where a four-stage testing plan is proposed in order to achieve this purpose. As an example, the E-Forms technology which is based on "Extensible Markup Language (XML)" is identified due to its features and capabilities. Then the evaluation is made to the top applications supporting this technology, where "Adobe's XML Forms Architecture (XFA)" technology is selected as a winner. This winner application then is tested, where samples of forms were designed to show the extraordinary features this technology can offer including digital signature. The test includes the use of a proposed offline forms' data gathering technique and uses "Oracle 10g" as the back-end database. It also includes a website to publish the designed forms in addition to an application developed using "Oracle Developer 10g" to process the forms' data for one of the designed samples in order to prove these forms' accuracy.

Keywords: E-Forms, Methodology, Identification, Evaluation, Testing, XML

1. Introduction

Collecting information is considered a basic and important step in daily activities of mostly every public and private business. Most of the interactions of citizens with public authorities are performed through forms, which may range from simple documents with less than ten fields, such as a statement for address change, to highly complex document sets, such as tax return forms or social claims forms. Claims are everywhere and according to Gartner, 85% of business processes rely on forms [1]. With the growing rise in online business and electronic commerce (E-Commerce), organizations need a simple, cost effective way to capture and track transactions to successful conclusion. Studies carried out by Gartner and Microsoft clearly show significant and tangible savings where electronic forms (E-Forms) are deployed [2]. Electronic forms are a key component in the information system for today's enterprises and are central to the development of electronic government (E-Government), being as a basic mean for implementing most of the public services considered as required for local and central public administration authorities [3].

1.1 E-Forms Solution

According to Gartner, electronic forms will always save money, always save paper and the environment, always create more productive, happier users and always make data more accurate. Data accuracy alone can save millions of dollars in most business processes. The world has been waiting for this technology [4]. E-Forms can be as simple as a digital version of a paper form or as complex as a web application. The simple paper replacement effort is where
many organizations begin and often end because they achieve results quickly, the benefits of a digital form over a paper form are clear and unambiguous [5]. Here are some of the benefits which could be obtained from applying E-Forms technology:
- Reduce costly paper handling and manual routing
- Tracking, auditing, and process awareness
- Reduce errors and improve accuracy
- Expedite collection of quality information
- Accelerate the delivery of form-based Information

1.2 Ideal E-Forms Application
The end user experience of the form filling process is a key to success or failure. Users must be intuitively led through the process of filling a form from beginning to end [2]. As an ideal E-Forms application, users expect the following offers:
1. Provides forms which can be accessed via the web.
2. Provides both online and offline forms filling.
3. Forms must be smart, interactive, dynamic and guide the user.
4. Forms that intelligently change according to each individual user session. Depending on one answer, different choices are given for the next question.
5. Users can be led through the form filling process from beginning to end with a point and click interface. As a user answers questions, a form is built electronically.
6. Preview, print, and email facilities to “see how the form will look” including the ability to save and retrieve.
7. Supports intelligent population of form fields and also supports calculations and file attachment.
8. Supports data validation and error alerting.
9. Clear indicators on unanswered questions (i.e. indicator for required fields).
10. Forms link intelligently to any supporting documents.
11. Form design should not require programming, although it is understood that more complex forms will require some IT input, thus users who understand the business process are ideally placed to design forms.
12. The E-Forms application is independent of any back-end business system, although it can be integrated; consequently, it can be used to collect data for any back-end system, including centralized information repository systems like Electronic Document Management System (EDMS), Enterprise Content Management System (ECMS), Web-Portal and other systems.
13. Allows digitally signing the completed form and archiving it.
14. Allows users to access and search for the form based on data populated in the form, directly from the information repository.
15. Receipt of the completed forms may trigger back end workflow applications.

1.3 E-Forms Maturity Levels
Moving from a paper-based business process to an automated one requires an understanding of an organization’s current state of readiness. The current state includes both an inventory of current capabilities as well as an assessment of near term plans for new systems, applications and programs. The assessment phase also includes an examination of forms which are no longer in use, or have been superseded by other methods [5]. In order to assess the current state of any organization, a level of readiness is defined. This is often referred to as a “maturity” level and is described using a simple ordinal scale (typically 0 to 5, with “0” representing the least mature and “5” representing the most mature level).

Gartner, a leading IT research firm, defines one such scale for E-Forms (as shown in Fig.1). An assessment of the E-Forms maturity levels across organization's departments is critical in determining the direction and investment strategy for next steps. The further along an agency is on the maturity curve, the greater the returned value. However, the value proposition must be assessed relative to the required cost.

![Figure1: E-Forms Maturity Level [5]](image)

1.4 Need for a Methodology
Choosing the appropriate electronic application or technology that could meets the current and may be future needs of any organization or individual, was and will always considered as a business problem. Where the inaccurate planning leads to a bad decision making which in turn may leads to a wrong choice that could not achieve the requirements and may cost extra money. For this reason, some guide or methodology is needed when deciding to purchase or to select an appropriate electronic application or technology like E-Forms application that could meet the organization or individual current and may be future needs.

A good E-Forms strategy should consider not just conversion of information from paper to electronic format, but should also identify opportunities to eliminate paper as the vehicle for collecting information. A mature E-Forms solution will manage the entire life cycle of a form, including all processes involved, and ensure that information is collected in a secure manner [5]. Methodology is generally a guideline system for solving a problem, with specific components such as phases, tasks, methods, techniques and tools [6]. In another definition, a methodology is a documented, respectable and tested set of methods, practices and rules used to collect and analyze information [7]. A proposed methodology for identifying, evaluating and testing E-Forms applications will be discussed in section 3.
2. Related Work

In this research three of the prominent evaluation methodologies which had been used in evaluating E-Forms applications are reviewed:

1. Forrester Research Group (March 2006) [8] using their Wave™ methodology has evaluated the strengths and weaknesses of top E-Forms software vendors. Forrester after examining past research, user need assessments, vendor and expert interviews, developed a comprehensive set of approximately 100 evaluation criteria, which then have been grouped into three high-level buckets: current offering, strategy, and market overview. Forrester evaluation included five Extensible Markup Language based (XML-based) forms product vendors in this assessment: Adobe Systems, Cardiff Software, File Net, IBM, and Microsoft. Forrester has used a combination of data sources to assess the strengths and weaknesses of each solution:
   - Hands-on lab evaluations.
   - Strategy and product road map briefings.
   - Customer reference calls.
Forrester according to their methodology's lab based evaluation results has found that Adobe's LiveCycle Forms and IBM's Workplace Forms are the clear market leaders.

2. Gartner Inc. (2010) [9] has applied their famous "Magic Quadrant" methodology to evaluating the top Business Process Management Suites (BPMS), where a XML-based E-Forms is one of the technologies included implicitly within BPM Suites. After considering more than 60 providers worldwide, Gartner found that 25 of them met the inclusion criteria which had been defined for this evaluation. As a two-dimensional graphical framework, Magic Quadrant places vendors of a specific industry sector into a strategic matrix. Multiple objective and subjective criteria are used to evaluate individual vendors, presented on two axes: Ability to Execute and Completeness of Vision. Where the "Ability to Execute" considers that the vendors must deliver strong functionality in all areas of capability to achieve long-term leadership in the market, while the "Completeness of Vision" in the market considers the vendor's vision and plans for addressing buyer needs in the future. From this evaluation, Gartner has found that IBM Lotus Forms and Adobe LiveCycle ES2 are highlighted as leaders in this evaluation.

3. Info-Tech Research (2011) [10] by using their famous Vendor Landscape™ (VL) Methodology, has evaluated a number of the top E-Forms applications from vendors like Adobe, IBM, Autonomy, Microsoft, Perfect Forms, Formatta and Link Ware. Each evaluated product has a different set of features designed to cater to different needs based on process criticality and use cases. Info-Tech in this evaluation case study has defined a Basic Table Stakes which includes the basic features that represent the most required needs by many of today's businesses. At the beginning and after establishing a vendor/product short list, Info-Tech uses general evaluation criteria and weighting factors table to describe E-Forms product weighting, where it is classified into two high level categories, one according to Product and the other according to Vendor, and each one consists of a specific number of criteria.

Then each vendor and product across these categories and criteria is given a score (called absolute or non-normalized) on a scale of (0-10) points.

These scores are then weighted according to weighting factors. The weighted scores are then averaged for each of the two high level categories: Vendor score and Product score.

A plot called Vendor Landscape Diagram of these two resulting scores is generated to place vendors in one of four categories: Champion, Innovator, Market Pillar, and Emerging Player. From this evaluation, Info-Tech found that Adobe and IBM are the champions in this evaluation with their products Adobe LiveCycle Forms and IBM Forms.

3. The Proposed Methodology

Electronic forms have been around for years, but the term refers to a wide variety of technologies, from scanned image applications to Hyper Text Markup Language (HTML) and new XML-based forms, that are not at all similar and far from equal in their ability to accelerate and smooth business processes [11]. According to this variety of E-Forms types and technologies and the increased number of applications which have been designed to support this rapidly growth technology, a specific methodology is needed to be followed when any organization or individual needs to select a specific electronic forms application that could meet this customer's currently and may be future needs.

For this purpose, a methodology that consists of three phases: Identification phase, Evaluation phase and Testing phase is proposed, where the first phase is already has been discussed in the previous chapter, while the other two phases will be discussed in detail over the current and the next chapter.

Each phase has a specific role in the whole process, where in the first phase and in order to identify the appropriate E-Forms type, a general survey about E-Forms types and technologies is discussed with focus on the recent technology trends through studying its features, architecture and security issues.

The second phase has a significant role among the other two phases, where an evaluation using one of the famous approved and reliable evaluation methodologies or using a proposed one is applied on the top competing E-Forms applications that support the identified type in order to make an accurate selection of the most appropriate application according to detailed features and evaluation categories.

The third phase is considered as an additional phase, where its purpose is to give the customer a high degree of assurance in order to make this customer or end user feel confident about the selected or winner application resulted from the evaluation phase, through applying a real test using this application.

The proposed methodology's schema is illustrated in (Fig.2).
3.1 Identifying E-Forms Technology

Electronic forms are generally classified into two main classes (As shown in Fig.3):

1. HTML-based Forms: Also called (Web based forms) this kind of forms can be classified into other two classes:
   - Pure HTML-based Forms: This kind is widely used and has the following specifications, advantages, disadvantages and uses:
     - Specifications
       - Accessed, filled and submitted online through web browsers.
       - Designed using HTML and CSS.
       - Validations and simple calculations can be added using embedded scripting language like JavaScript code.
       - Can be either static, or dynamic if designed using server-side language like ASP, ASP.net, PHP or JSP.
       - HTTPS feed method can be used for secure data submitting.
       - Must use a server-side CGI application or Servlet to connect with the back-end database.

     - Advantages:
       - Need no client filling application and can be accessed/filled using web browser only (i.e. thin client).
       - Easy to be filled and submitted directly to the associated back-end system.
       - Can provide data validation against a database in real time.

     - Disadvantages:
       - More suitable for light weight data.
       - May suffer difficulties in access and submission when there is an internet access problem or web server overload (i.e. no. of sessions exceeded the max. allowed limit).
       - Client-side validation and calculations script may be by default blocked by some browsers.
       - HTML pages often do not print perfectly or get cut off due to page size.
       - When session ended due to network failure or session time expire before submitting the data, then the form must be reloaded because this kind of forms is accessed, filled and submitted only online.
       - Cannot be digitally signed.
       - Cannot be integrated with enterprise systems and applications, except in case of XHTML.

     - Uses:
       - Usually used in on online registration and status checking like in airline booking, flight status check, bank account access, online shopping and other similar cases.

   - XML Embedded HTML-based Forms: This kind has the same specifications, advantages, disadvantages and uses of the pure HTML but can also ensure good data exchange (i.e. integration) with enterprise systems and other applications, since it is designed using the World Wide Web Consortium (W3C) standard XHTML based on XML language.

2. Non-HTML Forms: It's the kind that is usually named (E-Forms) by most vendors and references because it's compact, portable and usually retains the same look and feel of the corresponding paper forms. It can be also classified into two types:

   - Non-XML Forms: Includes forms like Microsoft Word, Excel, Adobe FDF (called AcroForm) and similar office forms which have the following advantages, disadvantages and uses:

     - Advantages
       - Easy to fill and user friendly and familiar.
       - Portable (i.e. can be saved on portable storage media).
       - Look exactly on the screen and on paper.
       - Can contain interactive elements (e.g. AcroForm).
       - Can be digitally signed (e.g. AcroForm).

     - Disadvantages
       - But cannot contain dynamic contents.
o Data cannot be submitted online since it is offline fillable forms.
o Could be connected to local database only but with additional requirement and limited capabilities.

**Uses**
o Usually used in daily activities.
- **XML-based Forms**: Represent the recent forms technology which includes a mixed specifications of non-XML and HTML-based forms, here are some of its advantages, disadvantages and uses:

**Advantages**
o Pixel perfect quality that delivers the exact page, fidelity.
o Portable and can contain interactive elements.
o Can be static like non-XML forms or dynamic and intelligent.
o Can include advanced calculations through programming.
o Can be filled offline and submitted later when convenient.
o Also can be filled online and act as a web based forms, but additional software or programming is needed.
o Can be digitally signed.
o Form's data can be captured using a 2D barcode (i.e. 2 dimensional graphical data representation).
o Can be archived (i.e. returned forms can be stored as files)
o Can be easily integrated with enterprise systems and applications depending on web services protocols.

**Disadvantages**
o The only disadvantage, is the need for a client form filling application to be available at the client side.

**Uses**
o Usually used in collecting large amount of data like in purchase orders, invoices, ship notices and even checks.

![Figure 3: Tree of Main Electronic Forms Classes](image)

**3.2 E-Forms Application Evaluation**

After identifying the appropriate technology, and in order to select the most suitable E-Forms application based on the identified technology, and depending on the ideas and methods used by some of the famous research groups, a simple methodology can be proposed for evaluating EForms applications.

This methodology borrowed its idea from the same method which is being used for calculating the final average for under and postgraduate study, where each subject is given a specific constant weight or a unit depending on this subject importance, and then each subject given also a variable degree or mark in a specific range (usually 0-100) depending on the mark obtained by the student. And to find the total average, all that is needed is to calculate the sum of the products of each subject by its weight and divide the result by the sum of all weights.

Similar steps can be suggested to be used as an evaluation methodology for E-Forms products and even any other products. These steps can be summarized as follows:

1. Specifying the main evaluation criteria category that includes the most important characteristics which must be available to describe the product's general capabilities.
2. Specifying the most important factors and features categories which consist of the current trends of evaluation standards that must be included within each criterion.
3. Giving a specific weight to each criterion within a scale of (1-10) according to its importance.
4. Identifying the selection mechanism of factors and features within each criterion.
5. Determining a constant maximum value (i.e. max points) on a scale of (1-100) to each factor or feature within the same criterion according to its importance and returned benefit.
6. Ensuring an accurate mapping of maximum values to factors and features, such that the total value of the selected features are governed by the selection mechanism within the same criterion not exceeding the criterion max points limit (i.e.100 point).
7. The evaluation **final result** is calculated through: first finding product of each criterion total value obtained after selecting the supported features by the corresponding weights, and then the final result is obtained through dividing the sum of all products by the sum of all weights.

As an example of applying this methodology to evaluate E-Forms applications, seven criteria may be suggested, each criterion consists of some factors and features which represent the current trends and techniques in E-Forms market. In this evaluation example, three of the top XML-based EForms products from leading vendors have been chosen to compete. These products are: LiveCycle Designer Enterprise Suite from Adobe, Lotus Forms from IBM and InfoPath Forms from Microsoft.

An Excel based worksheet will be used as a simple and powerful product evaluation tool, where the seven steps mentioned above must be followed in designing this evaluation worksheet. The following Excel based tool have been designed to assist in applying this methodology and show the evaluation results for this example:

1. An Excel sheet is designed to be the main evaluation tool, it includes the needed mathematical computations suggested by the methodology and designed to provide a customized way for selecting supported features easily and can be applied for evaluating any EFoms product within this capabilities limit.
2. This tool must be used for evaluating each competing product separately in order to calculate its final evaluation result.

Now to perform the evaluation process to the three competing E-Forms products: first the designed evaluation
sheet is applied to evaluate each product using a separate copy of this sheet. According to the competing three products, the absolute final results (i.e., depending on the predetermined features' values and no modification made) were as follow: Adobe LiveCycle ES got the highest result (88 point) like results in previously discussed methodologies, IBM Lotus Forms got (73.4 point) and MS InfoPath got (69.6 point) depending on the selected features that are evaluate Adobe Live Cycle Design ES, where in supported by each product. (Fig.4) shows a screenshot of the evaluation sheet when it is used to be supported by each product. (Fig.4) shows the same manner the other two product had been screenshot of the evaluation sheet when it is used to be evaluated.

4. E-Forms Application Testing

Although the selected application was chosen as a result of a deep survey that included an identification and evaluation process which took into consideration all aspects and details including even the cost issues, customer opinion and vendor reputation which seems enough to make the final decision on the adoption of the selected application, but an additional phase that could be named the testing phase is recommended or needed, because the selected application must be passed through a number of organized testing steps somewhat similar to applying the verification and validation (V&V) principle.

This additional phase may be needed for reasons such as:

1. To make the forms designer really tests the level of improvements and facilities carried by this form application.
2. To make the programmer tests the level of ease and the flexibility of the programming language which this application depends on, and also to test the level of usefulness, obtained from the additional tools which might be included to reduce the amount of the needed lines of code.
3. To give the database, web server and network and enterprise system's (if any like EDMS, ECMS....) administrators an idea on how easy it is to interface and integrate this type of forms with the existing system and what additional software or hardware must be provided.
4. Designing a number of forms samples that reflect some of the daily used forms by the customer and presents it to this latter, because that can give more confidence to the customer in the chosen application, since the end user represents the true beneficiary.

4.1 A Proposed Testing Plan

A four-stage testing plan (As shown in Fig.6) is proposed to achieve this purpose. It represents a general plan that can be used for testing any type of E-Forms application.

![Figure 4: Screenshot of Evaluation Sheet used to Evaluate Adobe LiveCycle Forms ES](image)

Second a bar chart is designed to be plotted automatically depending on evaluation sheets' results for the three evaluated products. (Fig.5) shows a screenshot of this chart.

![Figure 5: Screenshot of Results Bar Chart](image)

In this plan, each stage has a specific role as follows:

1. Forms Design/Deploy Stage:
   - In this stage, a number of forms samples that reflect some of customer's needed forms are designed using the selected application.
   - The designed forms must include different types of control elements and actions besides the needed programming codes, in order to show the application abilities.

2. Forms Filling/Submitting Stage:
   - In this stage, the designed forms samples must be filled by actual customer's data in order to make the customer sure that all form's elements, data validation, calculations, data lists and actions are working correctly and as needed.
Then the completed form's data or completed forms as a whole (depending on the used form's filling/submitting online or offline scenario) must be submitted to the back-end server.

3. Data Collecting Stage:
☐ In this stage, the returned form's data is collected depending on some code instruction lines or by using a specific application to import the data from the forms and storing it in the database table(s).
☐ In addition to the data collecting operation, this stage ensures that there is no error or corruption with the collected data, because the import and store operations detect and show actually errors like data type errors, data format errors, data replication errors and missing data errors.

4. Data Testing Stage:
☐ This stage may be considered an optional or additional stage, where its aim is to make an additional check of the quality and accuracy of the collected data through processing these data.
☐ An application is needed to be developed using a suitable programming language in order to process these data and outputs some useful reports of the results.

5. Conclusions

According to the proposed methodology discussed above, the following conclusions can be drawn:

1. The proposed methodology arranged in a specific manner that ensures an easy implementation through organizing the methodology's overall schema into three dependent phases by distributing the subtasks in such a way that provide better understanding.
2. With these three phases, the whole decision making process is facilitated, since each phase according to its role, gives a particular decision to be depended in the next phase.
3. The proposed sequence for these phases ensures an accurate application selection, since the testing is applied only to the selected application.

References