Library Data Quality Maturity (IIUM as a Case Study)

Hewa Majeed Zangana

(Department of Computer Science, College of Computer Science and IT/ Nawroz University, Kurdistan Region of Iraq)

Abstract: *Data and information are important assets to any organization. However it is more important for some organization than others because this data and information in its core and heart, that means the work of such organizations focused on data. In library specifically, and owing to its un-profit nature works, the information about the number of books, researches, thesis, articles, periodicals, and papers and the move of these assets between 'Borrow and Return' or the way to archive it along with the electronic services nowadays which is consider one of the essential service that any library should has. All this reasons make the database system is the critical factor to determine the level of library's performance. IIUM library is not an exception in this field. Its object's to consolidate the education process in IIUM specifically, and in the whole country generally. Recently, and for mentioned reasons, IIUM library moved from old system (Horizon 7.3) to new one (Symphony 3.2) to meet the world standards and enhanced library performance by providing students with best service. Considering the number of IIUM library assets, this was a very difficult project and it's normal to face some problems during the change operation. Problems like the unfamiliar staff with the new system, modification the old data structure to suit the new one and inconsistent data that associated with it. These problems are expected and certainly need solving, which will take time, and efforts. The change to the new system does not mean that is the final step one order to achieve library aims, but it means this is the first step in the way to get the optimization. The new system needs to clean its data and to be developed during time to optimize the services to help in supporting the education in IIUM. Many libraries including IIUM library are using the Symphony system in both of establishing a new or upgrading libraries system such as Utah State University, Wabaunsee Community College, and Endeavour College of Natural Health. And the system takes their service further toward and enhances the services they offer. The purpose of this paper is to show and assess the data quality maturity level of the Library of International Islamic University Malaysia.*

***Keywords:*** *Quality, Data, Maturity, Symphony.*

I. Introduction

Quality plays an important role as one of the competitive powers for those companies which run businesses in the Information Society (Drucker. 1985). With this trend, data quality of the information system is regarded as more important because it is the basis of the information quality. Low quality data bring several negative effects to the business users such as dropping of customer satisfaction, high running cost, inefficient decision making process, and sluggishness of the corporate strategic performance (Rye et. tal. 2005).

Many efforts have been made in order to assure the quality of the running software, although not of the data. But this situation is changing due to new status of data: at last, organizations have realized data is one of the most important assets for them (Huang et tal. 1999). Library was established in 1983, when the Library flapped opens its door to users, it was a modest floor area with a very small seating capacity. Thriving on simplicity, the working area for the few staff was located in the basement level of the three level mini libraries.

In 1986, it moved to its own four (4) storey building where it began to steadily developed its collection and expand its services and facilities. February 1997 witnessed the final move to the new library building in Gombak campus with a space area of more than 25,000 square meters with the capacity to cater for about 1 million printed materials. Today, IIUM Library is considered among the strongest academic library in Malaysia with a collection of more than two (2) million titles in print and electronic formats supported by a staff strength of more than 200 spread over seven (6) campuses.

(Figure 1) The Library of International Islamic University Malaysia

II.   Structure

Library is made up of the following departments:
1.   Acquistions Division (S48 NA).
2.   Technical Services Division (S48 ZMN).
3.   CRculation & Coll Management Division (S48 HH).
4.   Information Services Division (S48 ZAS).
5.   ISTAC (S48 NH).
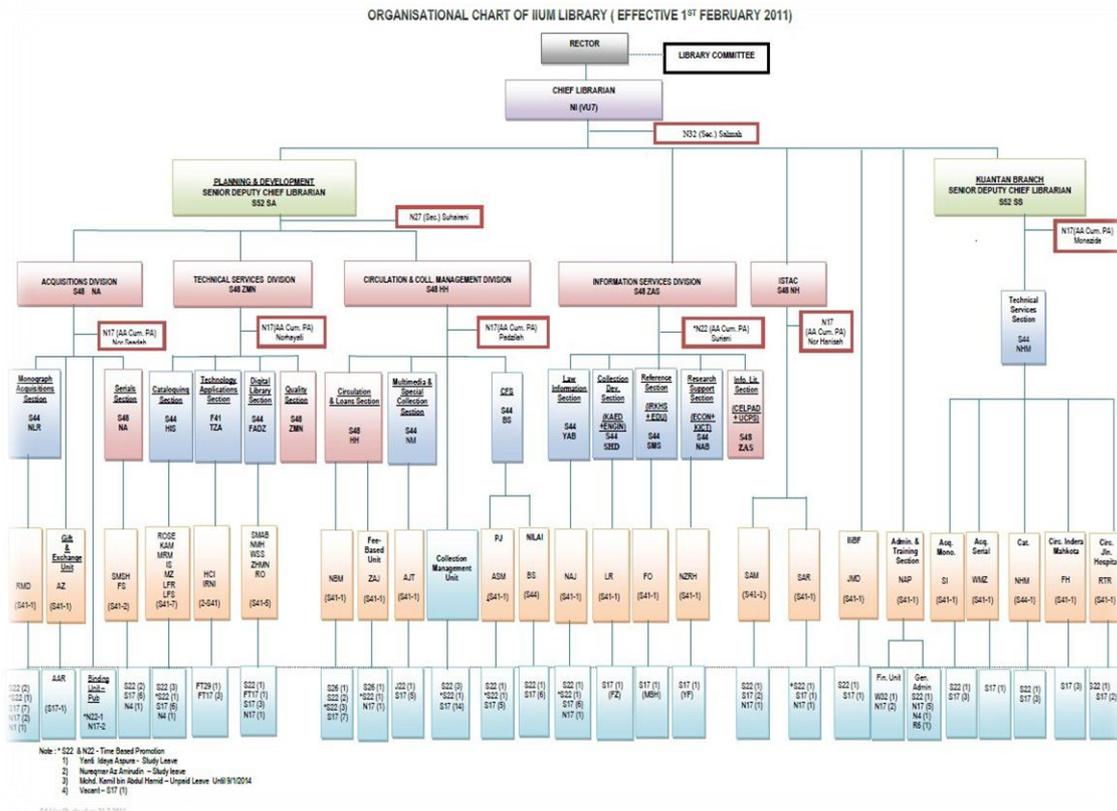
(Figure 2) Organization Chart of the Library

III.   Management Commitment

The Management always gives its commitment to the development and implementation of the quality management system and continually improving its effectiveness by:-
-   Communicating to the organization on the importance of meeting customer needs by following the stated rules and regulations.
-   Establishing and reviewing the quality policy.
-   Ensuring that the quality objectives are established and monitored.
-   Conducting management reviews.
-   Ensuring the availability of resources.

IV.   Data Quality Maturity

Quality of data or Data quality is defined as follows: data has quality if it satisfies the requirements of its intended use. Quality plays an important role as one of the powerful competition advantages for those companies that run businesses in the information society. For an easy way to determine organization's level of data quality maturity is to look at the current data quality improvement activities (fig. 1) (Adelman et.al., 2005)
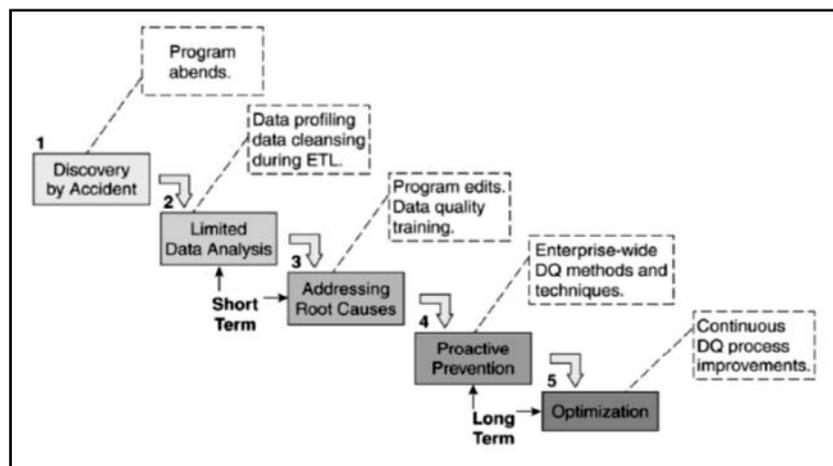
(Figure 3) Data Quality Maturity Levels

Quality level according to Adelman can be defined as the following five levels:
Uncertainty, Awakening, Enlightenment, Wisdom and Certainty.

Level 1: Uncertainty- At this level, the organization is unsteady over data defects as its programs fail or its information consumers express their dissatisfaction. There is no control over data quality improvement process in order to make it better, no data quality group, and no financial resources. The organization refuses to admit the truth about existence of any data quality problems and considers data analysis a waste of time. Or the CIO has prepared for retirement and doesn't want anything that will interrupt him. In short, the organization is asleep and doesn't want to be awakened.

Level 2: Awakening- At this level, the organization carry out some restricted data analysis and data correction activities, such as data profiling and effort in making to get rid of an unwanted data. There is still no wide project in support of data quality improvement, no data quality group, and no financial motivation. However, a few number of individuals admits about the truth of their dirty data and want to include data quality disciplines in their projects. These individuals can be the data administrators, database administrators, database developers, or people in the business circle.

Level 3: Enlightenment- At this level, the organization begins to deal with the root causes of its dirty data through program editing and data quality training. Also, at this level the organization creates a data quality group provides funds for data quality improvement project. The data quality group immediately carry out a wide project of data quality assessment on their critical files and database, giving more priority to data quality improvement activities. This group also establishes several data quality code of behavior and starts a comprehensive data quality training programs across the organization.

Level 4: Wisdom- At this level, the organization works actively towards the prevention of future data quality defects by adopting more data quality code of behavior to its quality improvement program. Individuals' managing the organization accepts personal responsibility for data quality. A chief officer (CIO, COO, and CFO) is assigned in charge of the data quality group. Metric are put in place in order to measure the quantity of data defects produced by other staffs, and the organization take these metrics into account in appraising the staff's job performance. Incentives for data quality improvement have replaced incentives for cranking out systems at the speed of light.

Level 5: Certainty- At this level, the organization enters the optimized circle by through the continuous monitoring and improvement in its data defects processes. Data quality is a necessary and important part of all business processes. Every job description needs to consider data quality, reporting of data defects, determining the main basic causes of the defects, improving the affected data quality processes to get rid of the basic causes, and monitoring the effects of the improvement. Basically, the art and intellectual achievement of the whole organization has changed.

V. Problems in the Library

Inconsistency data in IIUM Library catalog.
As we mentioned before library catalog is one of the drawback of the new system. During surfing in the library catalog you can find many inconsistency data there. We classified this inconsistency in two major groups. The first one is inconsistency in Latin-based latter (English, French, etc), and the second one is inconsistency in Arabic-based latter (Arabic, Urdu, Persian, etc). Both of them have two problems in term of spelling, wrong spilling or inaccurate format (the focus on the author name). We gave a sample about each problem to be clearer.

5.1 Latin-based latter problems.

1.1 Wrong spelling: as example for such problem when we were looking for an author under name "**Micheal**" the result of the search showed '9443' results. Whereas the results of looking under "**Michael**" back with '38' results. One of the results is "**Tuggy, Micheal**" and when we looked for this book in amazon books search we found that the real name is " **Tuggy, Michael**".

5.1.2 Inaccurate:

Some authors' names is not accurate and you cannot find it like the name "**francios beguin**", either by first or last names, because it is saved under "**Be?guin, Franc?ois**" with question marks which it did not replace any latter! And also if you typed like this way(With question marks) you cannot find it.

5.2 Arabic-based latter problems

5.2.1 Wrong spelling

This looks has more problems than the previous one, because it has two parts, one is Arabic names which translated to the English-based latter format, and the other is the Arabic names in Arabic-based latter format.

5.2.1.1 Translated Arabic-based latter names:

As an example for this one and the most common is the spelling of name "**MOHAMMAD**". When we typed the name in formula "**Muhammad**" the result backed with (29005) results. Whereas "**Mohammad**," backed with (1039) results. And "**Mohamad**" came with (646). Another example of the same problem is the name "**Ahmad**" which had (14197) results and the other spelling format "**Ahamad**" returned with (26) results. This problem also repeated with the name "**Mahamood**" and "**Mohmod**."

5.2.1.2 Arabic-based latter names:

This problem is supposed not to be existed in IIUM library. We can understand the above problem because there are many different translated from Arabic to English, but in Arabic itself that cannot be understood. Example for such problem is the name "داود" which is the wrong spelling for the name "داوود", when we were looking for the name "داود" we found (297) results and when we tried the other formal "داوود" the result were (36).

5.2.2 Inaccurate:

As an example for this kind of problems is the deferent between the name "عبدالرحمن" with no space between the two word's parts, and the same name with the space between the two word's parts. The result is different. In the first case the result was (193) while it was (2477) in the second case. Actually the two formulas refer to the same name.

Other problems that we observed in IIUM library's catalog, is that when you are looking for an author name، the result of your research comes inaccurately. As example for that some time you find the name you are looking as an author in the title and this result comes in the top of result.

The table below shows some issues that the library faced during the migration:

Issue	Horizon	Symphony Workflows
Manage changes in the new system	Functions in Symphony Workflows a bit different with Horizon. Staff may need some time to adopt the new system environment, terms used and functions. Example: Cataloguing Module Some functions in Horizon may not offer in Symphony Workflows. Symphony workflows are upgrade version, so some of the functions may no longer use in these days.	
Clean Data	Some data that are not cleaned in Horizon is inherited in Symphony Workflows. Example : Circulation Module There are students that are no longer in IIUM are migrated in Symphony Workflows.	
Data Consistency	After migration, data in Symphony Workflows are consistence with Horizon. It is very much related to field mapping, should the field mapping is consistent the migration result also will be consistent. Example : Circulation module All charged books for current and previous in customers' borrowing record are migrated in Symphony Workflows.	
Diacritic	At the beginning, some Arabic diacritics are not read by Symphony Workflows.	
Field Mapping	Field mapping from Horizon to Symphony Workflows are discussed and documented before sent to vendor for migration.	

Findings – Data Quality Maturity

In summarizing the data we have collect, we found the following points:

1) Since its inception, the Library adopts the policy of keeping abreast with Information Technology. Proposal to automate the library operations was first submitted in 1985. However, due to financial constraints the budget was only approved in 1990 and in April 1991, the library's first integrated library system mainframe based DOBIS/ LIBIS was installed. Since then, IIUM Library has migrated to the Horizon Library System with web-based capabilities and in 2010 the Library migrated to the Symphony Library System which possesses syndetic features.

Library security gates were installed to ensure security of library materials. The Digital Library services, introduced in 2004, enable users to search full text documents of various materials: examination papers, theses, Islamic Economic articles, contents pages of printed subscribed journals, IIUM publications, faculty publications, etc. Via this service, users are also able to view online digital videos on various subjects.

To facilitate circulation activities, the Library acquired the Radio Frequency Identification (RFID) technology based equipment (staff workstations, self-check machines and the smart book drop).

2) Power Supply and Backup system:
- Power supply: UPS.
- Backup system: Automatically backed up all of the relevance data into data tape (every night). The used tape being replaced with a new one by the technicians after backup, there's around 30s data tapes provided every month.

3) Inconsistency in the system:
- After the replacement the old system (SirsiDynix Horizon Library System) into new system (SirsiDynix Symphony Library System):
• Adaptability might be slow
• The system still on upgrading process
• Catalogues of the system modules are not better than the old one.
- Migration from Sybase:
• Need to do API for integration after migration
  With the SirsiDynix Symphony® API, your library gains full access to any and all information held within your system, plus the ability to extend, manipulate, share and present that information however you see fit. The Symphony API includes tools to perform batch transactions, as well as utilities for gathering, editing, loading, exporting and formatting data. Among its many benefits, the API enables a library to:
➢ Increase efficiency, by integrating the ILS with other library software or third-party systems;
➢ Minimize staff time spent on data entry, thanks to global editing capabilities;
➢ Manage new branch openings or temporary closure easily with collection mergers or removal tools;
➢ Write custom reports or automate processes for smoother departmental operations.
  The SirsiDynix Symphony® API is the ultimate in control, providing full open access for a system administrator or developer to customize, tweak, build and explore to his heart's content. Fully documented APIs give libraries the ability to modify, extract or update data to meet specific organizational needs – within the library facility itself, or for applications on the Web or otherwise.
• Buying a new license and training for staffs needs more costs.

4) Position of the database storage system:
- Library Data Centre (Level 1), inside Technology Application Section.

5) The structure of library:
- Using metric structure.

VI. Recommendations

The libraries today have a diverse, complex and ever-changing need. These needs largely driven by the community a library serves- reflect everything from the current economic climate to evolving technology to shifting patron demographics. Meanwhile, each library's services, collections, staffs and facilities are equally sundry.

Some recommendations we suggest to solve the problems that we found in the library of IIUM are:

1) Language (diacritic)
Description – Multilanguage support for symphony 2.0.7 and higher.
This symphony extension adds a Multilanguage support layer on top of symphony 2.0.7 and higher.
Installation: - create a directory called Multilanguage in your extension directory and place these files in there.
How it works-: in preference multiple language codes can be added, which will become available in the sites (like 'en, 'de,'nl, etc). When a section is created, each field will have an extra parameter called 'Multilanguage', making the field multilingual. Then create entries, you'll notice that are extra tabs on top of the page, each representing its own language. Also a new data source is available in 'Components' called 'available language', which represents different languages as supplied in preferences. The first parameter of the url will be language code (eg *http://mysite.com/en/my-page)*. This parameter is added to the parameter pool called $url-language-code. This represents the code of the current language shown. If no language code is provided in the URL, by default the first language as given in preference is chosen. According to the chosen language, the right content is automatically placed in your data course, so there is no need for filtering language in the XSL templates.

OR ADD LANGUAGE REDIERCTION TO SYMPHONY (LANGUAGE REDIRECT)

Description-: language Redirect provides an event to redirect visitors based on browser settings, values in cookies or default settings. It adds language and region parameters to the parameter pool and renders those parameters in a clean way at the beginning of the URL.

2) Learning Symphony

Before perusing for a serious project like symphony, there are some additional basics that need to be covered by the staffs

- They should make sure they get a good grasp on XSLT, so they can write clean, elegant templates.
- They should they understand the various <u>field types</u> available and how they work so they can design a good model.
- Familiarize themselves with all the options available when creating <u>data sources</u>, so they can be sure they are getting everything they need from their data.
- They should understand how parameters work so they can keep their project lean and dynamic.
- They should know what extensions are available; they can make their life much easier.

3) Removing User

Removing user records from the system is something that will be done automatically each year based on user activity. However, there may be occasion where you may need to remove a User record before those schedules.

To remove user record

- In the user group, click the Remove User wizard
- Do one of the following:
➤ In the User ID box, scan the user ID.
➤ Click the <u>User Search Helper</u>, and search for the user.
➤ Type the user ID in the User ID box, and click get user information.
➤ If the Current button is available, click the button to select the user that was displayed, modified or added.

Information about the selected user record displays in the Removing User Window

➤ Verify that the displayed user record is the record you want to remove, and then click Remove user.

A dialog box displays to verify that the user has been removed.

When you finished removing the user record, select one of the following options:

➤ Remove Another User to remove another user
➤ Close to exit wizard

4) Field Mapping

To resolve the issue of field mapping:

- Review that the mapping of the record is correct
- Check the following
a. Maintain source field- verify the structure
b. Maintain field mapping and conversion rules- verify that each field (if not defaulted with a value) is mapped into the structure defined in the previous step.
c. Verify that the file has the same fields in the structure.

5) The database system is located in the level 1, and this is not a good location to such important section in library, so we suggest moving it to level 3 or level 4.

VII. Conclusion

At the end of this research, we found that Library has made a great leap by moving to Symphony system. Although the upgrading to the new system came as a result of migration of two big companies Sirsi and Dynix, library has got many benefits since the new system came with new reporting tools, new modules, library catalogue with syndetic features. The main features for the new system are:

1- Efficiency: batch editing and on-shelf hold processing, right-click menus, and   group management tools.

2- Security: extreme control over viewing privileges and polices.

3- Flexibility: the system gives the staff a control of what goes where, when, and how. Also provides the user with fully documented APIs and open source platform.

4- More services to the users such as, books by mail, self-service holds mobile circulation support.

As we said the system in upgrading process and the problems is a time matter. A part from some problems we had mentioned before the system and library service students quite good.

In the end we would like to add a point that maybe very important to help the library to success and prosperity even though it does not relate to the data and information, but it is an obstacle in the way, which is the lack of finance.

References

[1]. Adelman, S., Moss, L., and Abai, M. (2005) "Data Strategy", Addison-Wesley.
[2]. Huang, K.T., Lee, Y., Wang, R. (1999) Quality Information and Knowledge. Prentice Hall, Upper Saddle River.
[3]. Kyung-Seok Ryu, Joo-Seok Park, & Jae-Hong Park. (2005). A Study on Data Quality Management
[4]. Maturity Model. Electronics & Telecommunications Research institute (ETRI).
[5]. lsmael Caballero and Muiioz-Reja Mario Piattini Velthuis. (2003). Data Quality Management Improvement
[6]. Peter F. Drucker. (1985). Playing in the Information-Based –Orchestra. The Wall Street Journal, June 4, 1985.
[7]. About Library. IIUM Website (http://www.iium.edu.my/lib/about-us/overview), accessed (May, 2016).