SharePoint Analyzer Tool – A Windows Desktop Application Using Data Extraction and Exportation

Sharad Narayan Sharma

B.Tech Student, Computer Science and Engineering, Vellore Institute of Technology, Katpadi Road, Vellore, Tamil Nadu-632014, India.

Senthil Kumar R

Head of Department, Computer Science and Engineering, Vellore Institute of Technology, Katpadi Road, Vellore, Tamil Nadu-632014, India.

Abstract

A majority of the Fortune 500 firms around the world use SharePoint. It is employed by them for different purposes. It is not just a particular software or program but rather provides a platform for various kinds of programs. It can be thought of as a back-end system that congregates all the employees' PCs and mobile devices in unison. This allows them to simultaneously synchronize their efforts and helps them to communicate with each other as well. The main goal of SharePoint is to help a company with hundreds and thousands of employees to work with the same level of coordination and efficiency as with ten people working from one office. Users can either comply with the SharePoint On-Premise version (server based storage) or the Online version (cloud based storage).

All this leads to large amounts of data, sometimes redundant, to accumulate up in the servers or in the cloud itself. Such data is of no use to firms since it can be neither expressed meaningfully to examine patterns to avoid hindrances nor used to plan future strategies on the basis of the data collected over years. Our primary goal is to make sense of this data present in SharePoint servers. For this, we create a desktop application which can retrieve data for the server and client sides and express this data meaningfully in Excel Sheets. The summary

of all the data will be represented graphically along with the customizations present. The tool will be created using Visual Studio. The tool will go through the farm, web application, site collection and sub-site level for SharePoint On-Premise. For the client side, it will extract data from the site collection and sub-site levels. The data collected will also help users and developers to migrate from one version of SharePoint to another. The tool will hence, provide a concise view of the SharePoint contents in a user – friendly manner.

Keywords: SharePoint On-Premise, SharePoint Online, Site Collections, Farm, Web Application, Data Extraction

I. INTRODUCTION

SharePoint is an application which was launched by Microsoft in 2001. It is web-based. Our project involves making a data mining and data comprehension tool based on information extracted from SharePoint Online and SharePoint On-Premise. We'll be working with the 2013 version of SharePoint. SharePoint Online is also known as the cloud version of SharePoint. There is a stark difference between the two.

SharePoint Online consists of data servers and data centres. These are maintained by Microsoft itself. The On-Premise version includes data on the company servers or data centres which are managed by the Information Technology professionals employed by the firm itself. The cloud also reduces the need for hardware and hence, internal resources are reduced. The On-Premise version helps access direct control over information and standards to be complied but it is more dependent on internal abilities. The cloud version has to adhere to industry standards and may be even verified by third parties. SharePoint On-Premise has data organized at three levels: farm, web application and site collections. The site collection contain further sub-sites as well. On the other hand, SharePoint Online includes only site collection and its subsequent sub-sites. Out tool will work on both client side (SharePoint Online) and server side (SharePoint On-Premise) levels.

The Server Object Model includes a set of SharePoint objects. They are executed in the server side. It is the costliest API set in SharePoint 2013. The main assembly is Microsoft.SharePoint.dll. We make use of other dynamic link libraries as well such as those for Excel Utilities, Data security, etc. Client Side Object Model is rapidly taking over and is used widely by SharePoint professionals and developers. It is developed by Microsoft. It is available in Nuget. It involves "Microsoft.SharePointOnline.CSOM". Its most important dynamic link libraries include- Microsoft.SharePoint.Client.dll,Microsoft.SharePoint.Client.Runtime.dll, Microsoft.Online.SharePoint.Client.Tenant.dll.

Knowledge of ASP.NET framework is equally important. C#.NET is used for ASP.NET framework to create the desktop application. ASP.NET is a part of the .NET Framework. Applications coded in ASP.NET can make use of the available classes in the .NET Framework. ASP.NET has three such frameworks: ASP.NET Web Forms, ASP.NET MVC, and ASP.NET Web Pages. PowerShell scripting is used for direct updating of farms and adding of content to the farm, etc. PowerShell was developed by Microsoft as a shell language for automating task and for management purposes of configuration. Powerful also depends on the .NET framework. It envisages a command-line shell as well as a scripting language.

II.   LITERATURE REVIEW

Barry Prescott et al. are concerned about the ITER Organization, which is dedicated to the transfer of data to a particular worker or a set of workers, and remote collaborators. ITER has an associate authenticated-access web site. They state that in 2007 and 2008, the number of users who registered on the website grew by a factor of 10. This amounts to over 3000 users presently, with around 900 distinctive users surfing the web site per month. They mention that ITER has made and can still produce a colossal quantity of data which will be generated from anyplace at intervals within the organisation. SharePoint features a learning phase that is sometimes steep for some authors. It often needs a particular author to try and do their jobs effectively and for this they have to induce to grips with several ideas reminiscent of Sites, Web Pages, Document Libraries, Folders, Document Libraries, Content sorts, Lists, Permissions, SharePoint teams, Active Directory security teams, Web Parts, Placeholders, and Navigation. A fusion organisation additionally should contend with gigantic volumes of scientific knowledge. They found the interface for knowledge management present provides an easy approach for individuals to extract their knowledge from documents and store it in a very relative fashion, the SharePoint engine is limited to datasets, and to activity complicated aggregations on the information. They require to overcome this learning curve which can be achieved with the help of a data generation and exportation tool which not only collects data related to the above libraries, but also subjectively organizes the data according to these library headers. This allows for faster and easier interpretation and learning. It also helps to organize the data associated with the ever-increasing number of users. [1]
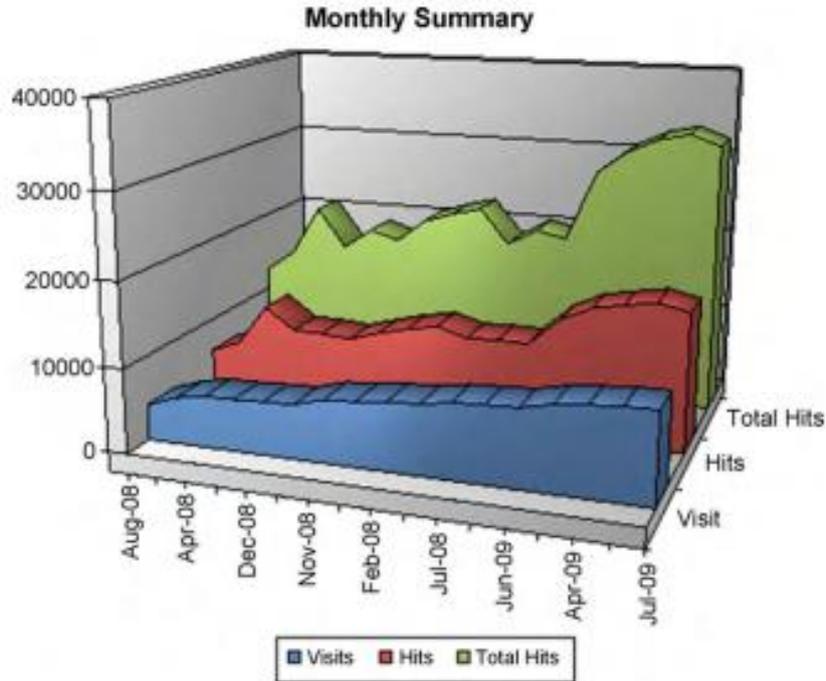
Figure 1: Daily requests on an average, month wise.

DD-ID	PBS	NewPBS	Package	Title
A.01	55._A.01	55.AA._ _	PA02	A.01: Outer-Vessel Magnetics
A.02	55._A.02	55.AB._ _	PA02	A.02: In-vessel Magnetics
A.03	55._A.03	55.AC._ _	PA02	A.03: Divertor Magnetics
A.04	55._A.04	55.AD._ _	PA02	A.04: Continuous Rogowski coil
A.05	55._A.05	55.AE._ _	PA02	A.05: Diamagnetic Loop
A.06	55._A.06	55.AF._ _	PA02	A.06: Halo Current Sensors
B.01	55._B.01	55.BA._ _	PA03	B.01: Radial Neutron Camera
B.02	55._B.02	55.BB._ _	PA11	B.02: Vertical Neutron Camera

Figure 2: Complexity and quantity of data increases as every cell leads to more information.

Mohd Amir Azmi et al. point out that the education system has undergone a shift from learning inside a classroom to learning from and in the web. Learning Management Systems (LMSs) were designed to fulfil the need for an appropriate platform to help improve student-teacher interaction in a technology-based environment. They talk about the development of a gamified LMS Associate in Nursing. They talk about how Microsoft SharePoint has been employed to demonstrate concepts of gamification within Associate in Nursing LMS. The system was built round the core components of SharePoint. It made use of lists, which acted as information providers and workflows which were used to program the system logic. A large majority of the environment was designed on SharePoint Designer 2013, with the interface choices being offered on the SharePoint web site itself. This in addition creates a draw for the students to play with the system for longer hours and provides a wholesome experience. This is where the gamified LMS differentiates itself from typical LMSs. The education system mentioned above can be strengthened if the same data can be used across different versions of SharePoint available. A tool which helps in SharePoint migration by extrapolating data can be useful for LMS. [2][3]

Figure 3: A screenshot of LMS using SharePoint

Abdul Kadhar Sherif et al. discuss the demand for migration to SharePoint. They say that it continues to grow because the platform speedily becomes the central hub for organizations searching for a centralized data repository and cooperative space. However, migrating the vast amounts of data from one SharePoint atmosphere to another isn't an easy task. Several things need to be considered before an organization gets involved in a migration. Abdul et al consider an Enterprise Content Management (ECM) heritage system migration to SharePoint, the challenges associated with it as well as the advantages and technical issues of a SharePoint migration. The success of

finishing a migration to an environment of SharePoint on-line from associated older SharePoint environment needs specialised designing and analysis. Many different types of migration can take place, each associated with different kinds of data and information. As discussed earlier, our analyser tool provides the much-needed analysis and designing of data before the migration can take place. [4]
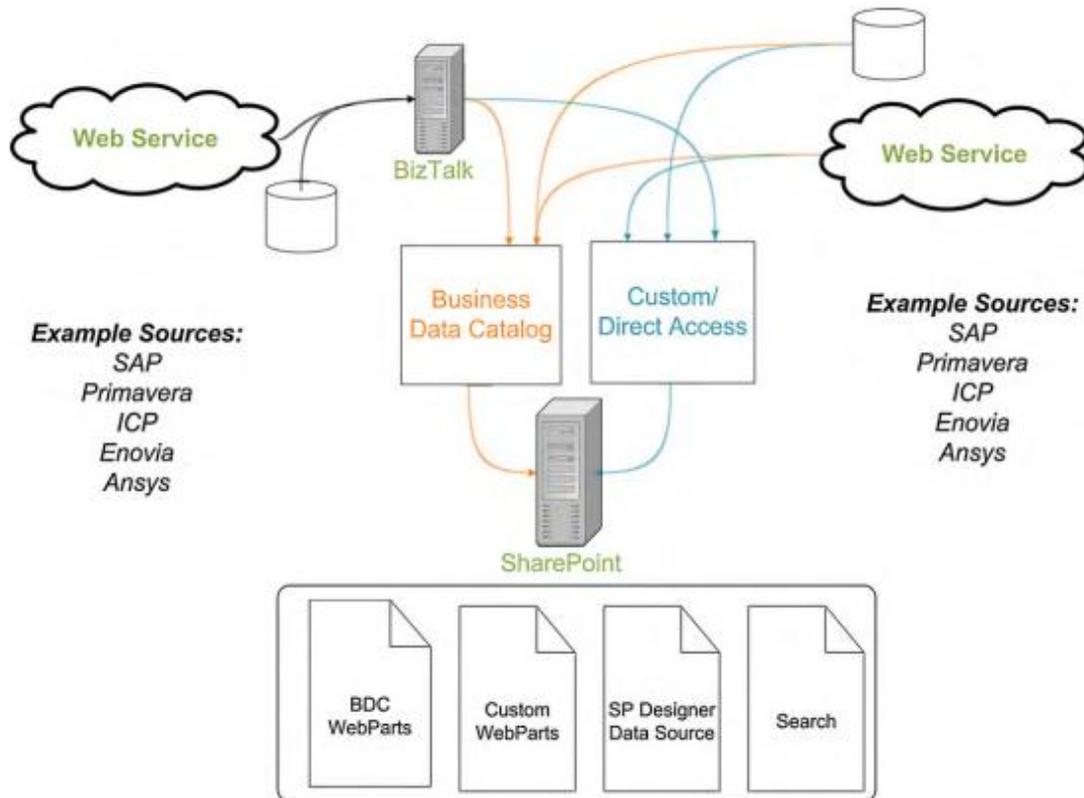

Figure 4: Architecture used for gathering data

III. OUR CONTRIBUTION

In this paper, we have discussed about an analyser tool which extracts data from both server (SharePoint On-Premise) and client side (SharePoint Online) and presents it through Excel Sheets and via a graphic output involving a pie chart and bar graphs. an overview of our process. For the server side, the user has the option to collect inventories for the whole farm or for specific web applications. One can further choose to gather and examine specific site collection inventories. These site collection inventories will also have inventories related to sub-sites along with inventories related to customized solutions, customized web-parts, customized features, content types and info-path forms. The roles, user and permissions' inventories related to the site collection chosen by the user are also collected.

These inventories will be generated in Excel Sheets. Each output will generate a single Excel workbook with multiple worksheets, each worksheet containing information about a specific inventory. Each worksheet will have headers related to properties of that inventory. For example, the custom solutions inventory will include headers such as custom solution name, its ID, the library to which it belongs, etc. Such collection and augmentation of data will help a user or an organization to examine trends or pattern within an area of SharePoint or of the farm as a whole.

Figure 5: Excel Sheets for farm inventory

The graphical content shows a pie chart comparison between out-of-box contents and customized contents. It further presents a graphical cone breakdown of the components that make up the customizations, that is, web parts, content types, etc. The importance of this tool also lies in the fact that is user friendly and knows what the user expects out of the tool. For example, when the user chooses to gather all the inventories on the farm, then automatically the tool will generate all inventories related to all the web application, all the site collections, all the sub-sites and all the customizations as well. The same hierarchy is followed when the user chooses a particular web application. The tool generates sheets having inventories related to that particular web application as well as inventories of all the site collection, sub-sites, and customizations within that particular web application. Similarly, on choosing a particular site collection, the particular site collection's inventory as well as inventories related to its sub-sites and customizations are generated.
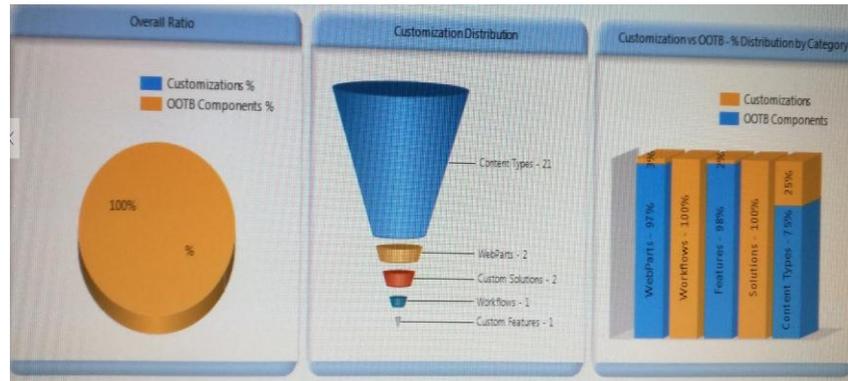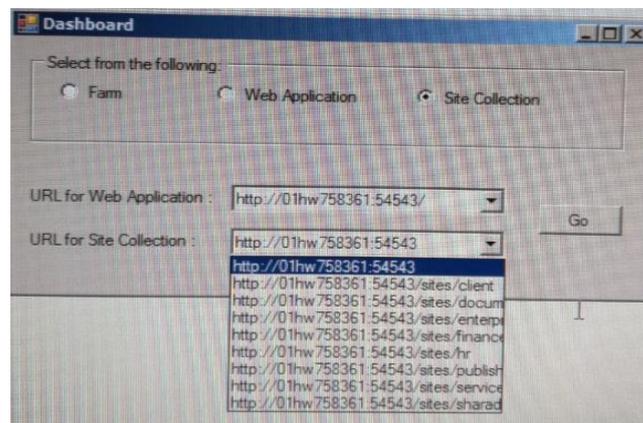
Figure 6: Graphical output

Figure 7: Dashboard for SharePoint On-Premise

Once the user has the data available, he or she can identify patterns in it with ease due to the way the data has been segregated into separate inventories in separate worksheets of an Excel workbook. This data helps organizations as well as individuals to strategize future plans, avoid obstacles encountered in the past and increase productivity, efficiency by following trends which have been shown to benefit the individual or the organization in the past. The tool also works well on security aspects. It only generates data for which a particular user has clearance level. For example, it will not generate data for a site collection for which the user is not an administrator or the owner.

Another aspect of the tool is that it works to collect data from the cloud as well (SharePoint Online). Here it allows the user to enter the tenant details from whom he or she wishes to extract the data. The user is further asked to enter the username and the password, which has been encrypted by the tool, thereby strengthening security. The cloud version of SharePoint has only data available om site collections and its sub-sites. It also has fewer customizations when compare to the On-Premise version.
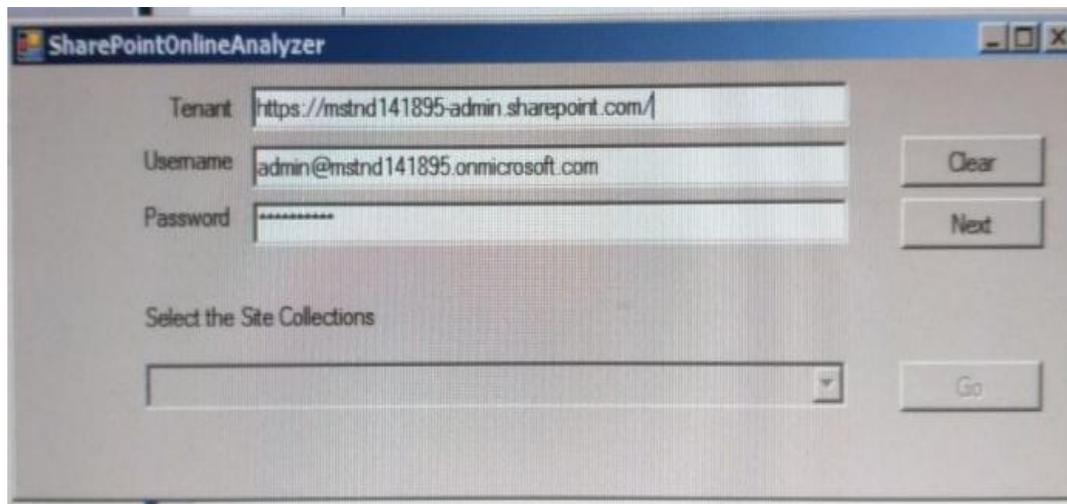
However, our tool takes care of all these complications.

Figure 8: Dashboard for SharePoint Online

On entering the user credentials, the user has the option of either continuing with the same credentials or enter a new one. Once the user is satisfied, he can click on the valid button provided by the tool. This enables a drop-down list for him or her which allows one to choose either a particular site collection by selecting that site collection's URL or all of the site collections by clicking on "All". If the user goes for a particular site collection, only the inventories related to that site collection will be generated. However, if he or she wishes to go for "All", the inventories related to all the site collections will be generated. The inventories generated for SharePoint Online are also Excel sheets. The graphical output displayed here is similar to that of the On-Premise. Another thing common to both SharePoint On-Premise and Online is that there will be a pop-up box which will keep the user informed as to which inventory is being created by the tool at that point of time. The analyser takes care of the difficulties that one might face when working with models mentioned in the above literature review. This single tool will not only solve the problems of the above mentioned three models but will also provide impetus for further research on how effectively data collected from SharePoint can be used.

IV.   SUMMARY AND CONCLUSION

The tool provides a comprehensive extraction, gathering and exportation of data. Its working can be judged on the following parameters-

- Safety – Data security while data is being retrieved. Only the user and groups

for that particular SharePoint farm will be allowed access to the data. There is encryption provided for password settings.

- Time – 1 min for summary, 5-10 min for site collection and web application inventory generation and 15-20 min for farm level inventory generation. (On-Premise). It requires 10-15 min to extract all the site collection data from the cloud.

- Resources – Includes out of box as well as customized web parts, custom solutions, custom features, info path forms.

- Scope – Scope ranges from sub- site level to the entire farm level. It includes web application and site collection inventories too.

- Quality and actions - Provides an easy to use user interface and hassle free Excel Sheet generation.

The Excel Sheets generated and visual representation of the output provide a full in-depth view and analysis of the data lying on SharePoint servers and on the cloud.

REFERENCES

[1] Barry Prescott, James Downing, Marco Di Maio, John How (2010, July).Using SharePoint to manage and disseminate fusion project information: An ITER case study In Fusion Engineering and Design, Volume 85, Issues 3-4.

[2] Mohd Amir Azmi and Dalbir Singh (2015, 25 May). Schoolcube: Gamification for Learning Management System through Microsoft .In International Journal of Computer Games Technology, Volume 2015, Article ID 589180

[3] D. McIntosh, Vendors of Learning Management and e-Learning Products, Trimeritus eLearning Solutions, 2015.

[4] Abdul Kadhar Sheriff, A.Shameem Fathima and Ashraf Mohammad Badawood (2014,November). A study on Best Practices for a Successful SharePoint Migration or Up-gradation. In UCSI International Journal of Computer Science Issues, Volume 11,Issue 6, No 2