

# In-Place Editing for Digital Archives

## Literature Review

Laylaa Varachia  
Department of Computer Science  
University of Cape Town  
Cape Town, South Africa  
vrclay001@myuct.ac.za

## ABSTRACT

As it stands, the administrative functionality for the Simple DL digital archiving toolkit is limited and does not allow the in-place editing of files that are part of an archive. As a result, administrators need to download the files, make the necessary changes and then reupload the file which requires unnecessary time and effort. It is hoped that by the end of this project, a feature allowing administrators to edit files in-place and online will make it easier and faster to make changes to files and thus improve the overall user experience of Simple DL.

## KEYWORDS

Digital Archive, In-place Editing, Data Management, Administrative System

## INTRODUCTION

Simple DL is a system that essentially has no software being executed at runtime, making it a ‘software-less’ solution to implement digital libraries with ease. The Simple DL toolkit provides a means for archivists to create digital libraries that do not require much computational power or maintenance, enabling them to be implemented successfully in low-bandwidth environments [9].

The goal of this project is to produce a usable administrative system for Simple DL, allowing administrators to change configurations, manage files and edit files online, otherwise referred to as in-place editing. While the focus of this research is mainly on the in-place editing of files, digital archiving in general, including existing digital archiving tools, as well as the shift from custom software solutions to toolkits will be reviewed. Figure 1 shows the current administrative interface of Simple DL.

Figure 1: Simple DL administrative interface

Development of the administrative system will be carried out with Emandulo [7] as the client. Emandulo is a platform that brings the past five hundred years of southern African history together by digitalising scattered material, contextualising it and making it available online [9]. Emandulo was initially developed using AtoM, which was later successfully replaced by Simple DL [9].

This review aims to explore the current methods used for in-place editing and how it can be applied to the administrative side of Simple DL to make the in-place editing of files possible.

## BACKGROUND INFORMATION

### Step 1: Importing

XML files are created for metadata that is read in from source XML files and spreadsheets [9]. Based on the original structure of the source directories, a directory structure is created using nesting rules specified in metadata entries by AtoM. Additionally, user profiles are generated automatically in XML format.

### Step 2: Indexing

The information retrieval engine is capable of faceted search [12], which allows users to filter search results based on different facets of the data [8]. User files and XML metadata are indexed on the information retrieval engine. Full text, which is the entire body of a digital document, is extracted from PDF files and cached when necessary [9].

### Step 3: Generating

The files containing metadata, users and website pages are converted from XML to HTML files using the XSLT stylesheet [9]. The template website is also duplicated. Thumbnails are generated for individual items and subcollections.

This sequence of events occurs each time a new metadata subcollection is added to the system. If no changes to a spreadsheet have occurred, it will not be imported again, and an HTML file will not be regenerated from its original XML source if it is up to date.

## Simple DL

Logged in: [username] | [Edit Profile](#) | [Admin](#) | [Logout](#)

[Site Home](#)

Manager Options: [Moderate](#) | [Manage](#) | [Import](#) | [metadata](#) | [force](#) | [clean](#) | [Generate](#) | [metadata](#) | [force](#) | [Index](#)

### Manage Datasets

Datasets: [spreadsheets](#) | [collection](#) | [carousel](#)  
Options: Create [F]older, [D]elete file/folder, [U]pload file or ZIP

Select a file to attach (for uploads): [Choose file](#) | No file chosen

```
[...] ROOT/ [f]
[ ] ... AMAFA/ [f] [d]
[ ] ... JAG/ [f] [d]
[ ] ... JAG_institutional_materials_v2.csv [8839 bytes, 09 May 2021 14:31:01] [d]
[ ] ... Brenthurst_v3.csv [365457 bytes, 09 May 2021 14:31:01] [d]
[ ] ... Maritz_v2.csv [526790 bytes, 09 May 2021 14:31:01] [d]
[ ] ... Jacques_v2.csv [25625 bytes, 09 May 2021 14:31:01] [d]
[ ] ... Brodie.csv [7073 bytes, 09 May 2021 14:31:01] [d]
[ ] ... Horstmann_v2.csv [271199 bytes, 09 May 2021 14:31:01] [d]
```

## 1 Digital archiving

A broad overview of digital archiving and the existing toolkits available discussed. Since this project is building on the SimpleDL software and improving it, possessing a certain degree of knowledge regarding digital archiving, and having a high-level understanding regarding SimpleDL and how it compares to other digital archiving toolkits will be useful to understand the project.

### 1.1 Web-based toolkits

DSpace [14] and EPrints [5] are Web-based, open-source toolkits for creating digital repositories. They are two of the most popular Web-based toolkits [9]. They each consist of a Web user interface and database [11,13]. DSpace makes use of a relational database implemented using PostgreSQL [14] to store digital artifacts. EPrints uses the UNIX file system to store the files and a MySQL database to store everything else such as user information and metadata [13]. Simple DL makes minimal use of Web applications, which is where it differs from these Web-based toolkits. Simple DL has no database or database management system. Additionally, spreadsheets and XML files are used to store structured data while unstructured data is stored as flat files [9].

### 1.2 Offline toolkits

The Greenstone Digital Library Software from the New Zealand Digital Library (NZDL) project [10] provided an innovative method of organising digital artifacts and making them accessible. The Greenstone software can be accessed through any browser through the World Wide Web if there is internet access. In the cases where there is no internet access, collections can be accessed on CD-ROM through a browser but doing so requires users to install software. The ability of the information to be packed on a CD-ROM makes the Greenstone software accessible offline. Simple DL and Greenstone share the key feature of being accessible to low bandwidth communities.

### 1.3 The progression of digital archives over time

In the past, digital libraries were built using custom software solutions to meet the needs of a specific project. However, these needs could be generalised and applicable to a wide variety of projects [9]. EPrints [5] was one of the first digital archiving toolkits, making it easy for users to create their own digital archives [15].

Switching from custom solutions to general toolkits saw the rise of many new digital archiving toolkits, each with its own primary goals addressing separate needs within the digital archiving space [9]. DSpace [14] and AtoM are the most commonly used digital archiving toolkits, but they do not meet the needs of all potential users wanting to create a digital archive. Neither of these solutions

work offline, are simple enough and requiring little to no maintenance. This left a gap for Simple DL to fill.

The Digital Bleek and Lloyd Collection [8] was created in 2006 as a custom software solution [15]. It was designed based on the following set of principles: network could not be assumed, software systems should not be used for mediation and that static representations should be created by pre-processing data. Simple DL was built based on these principles and the lack of digital libraries in developing countries demonstrate the need for a digital archiving toolkit with these characteristics.

## 2 Editing files

### 2.1 What is in-place editing

In-place editing is sometimes referred to as live site editing, in-context editing or in-situ editing [3]. It is commonly used for the editing of Web pages. The in-place editing model allows users to make changes to the content of Web pages. Sparrow [3] was the first in-place editing system for the Web. Other examples of in-place editing systems for Web pages are DirectEdit and ISAWiki [3]. All these systems share philosophies of a Wiki but what differentiates them is that Wikis, at the time of development, required users to have HTML knowledge where these systems did not.

### 2.2 In-place editing of Web pages

Sparrow [1] in-place editing focuses on community shared Web pages, a Web page created by a single user that other users can then modify. The changes made to the page are carried out within the browser, i.e., they are edited in-place. As a result, a user who wishes to contribute to a page or make any changes can do so independently without relying on the original author to make the changes.

Although the functionality required by Simple DL is slightly different, the functionality of in-place editing is still required. Administrators on Simple DL should be able to edit documents forming part of an archive in-place and on a browser just like Sparrow allows users to edit Web pages. Administrators should be able to edit documents without having to download them, make the necessary changes and then reupload them.

While editing using Sparrow, the remainder of the Web page is left untouched. Items that are being edited have edit regions that are close by to ensure that the item is clearly associated with its edit region [3]. The remainder of the content that is not being edited is remains as is and visible, giving users the opportunity to continue browsing the Web page, even while editing. Similarly, while editing files on Simple DL, administrators should be able to scroll and browse through the remaining content of the file that they are in the process of editing.

Another feature of Sparrow that relates to Simple DL is that users editing documents on Sparrow do not need to know or see any HTML. Users simply fill out forms that are then processed and used to apply the necessary changes to the Web page. Similarly, administrators on Simple DL should be able to edit the files in the archives without knowing or seeing any code. The information that the administrator edits on the files should be used to update the archive. Figure 3 shows an example of Sparrow being used to edit a Web page.

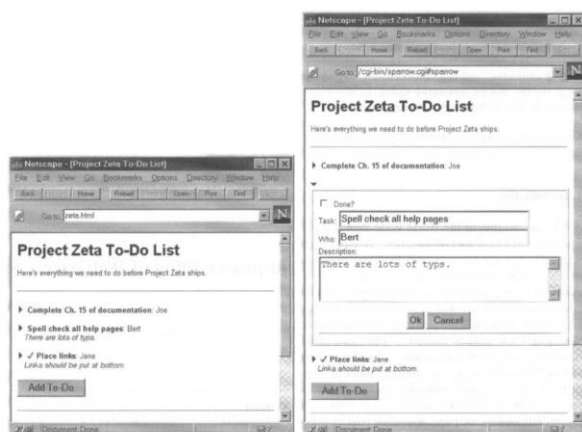


Figure 2: Editing content using Sparrow

The functionality offered by Sparrow differs from that required by Simple DL in several ways. Firstly, Sparrow focusses on community-shared documents meaning that there is a single author who is the creator of a document and then users that form part of the community can all edit the content. Simple DL will only allow administrators to make changes to any documents that are part of a specific archive. Secondly, what is meant by 'documents' differs in the context of Simple DL and Sparrow. Sparrow refers to a web page as a community shared document while 'documents' in the context of Simple DL would be files in CSV or XML format that form part of an archive.

## 2.3 Back-end editing

Back-end editing is the opposite of in-place editing [3]. While both can be carried out through a browser, back-end editing requires the user to interact with a new and separate interface make changes to the content of a Web page. The most popular example of back-end editing is a Wiki.

A Wiki is a linked set of Web page that was created collaboratively and can be improved iteratively [4]. The goal of a Wiki is to allow authors to collaborate on Web documents within a browser by creating, editing, and deleting pages using an editor

[2]. Like a Wiki, Simple DL requires to allow users to edit files within a browser.

Anyone should be able to contribute to a Wiki, making it necessary for creating and editing to be a simple task. However, it is debatable whether this was achieved since editing early Wikis required users to have some HTML knowledge [3]. Unlike like a Wiki, only an administrator should be able to edit files.

Additionally, for Simple DL, no HTML knowledge or any coding knowledge should be required to edit the files. Editing of files should be a simple and intuitive task. Figure 3 shows the user interface of an early Wiki.

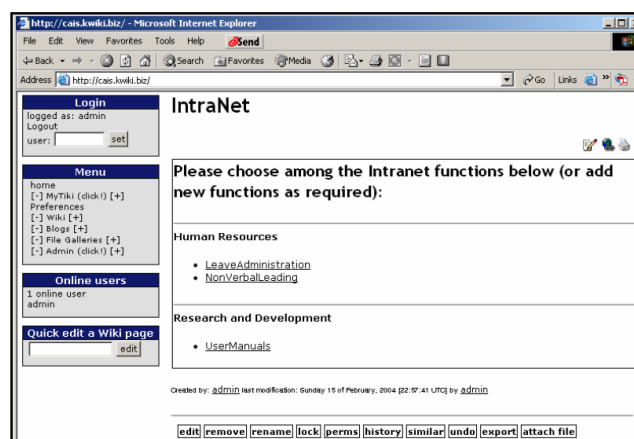
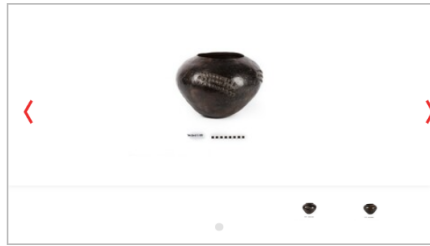


Figure 3: Wiki screen

To edit a Wiki, a user is directed to a new interface (a back-end). The way files will be edited using Simple DL will differ to the way a Wiki is edited. The user will not be redirected to a new page, rather, the edits will take place in-place. An "edit" button will be present next to the content they may want to edit, in this case a file. An example of a system that uses in-place editing similarly is Wikipedia [3]. Essentially, the administration system is a back-end, but within the administration system, the edits will take place in-place.

## 2.4 Metadata editing

Digital archives are made up of individual artifacts, metadata and thumbnails [9] that goes along with each one. The metadata included depends on the digital archive being accessed, but using Emandulo [7] as an example, the following are typical fields that could be included: title, material designation, reproduction conditions and description. Figure 4 shows the metadata listed for an item on Emandulo.



#### \*\*\* Metadata \*\*\*

Title	Ceramic vessels [ Source of title : Nessa Leibhammer using JAG materials ]
Material Designation	Object
Institutional Identifier	2012-6-137
Reproduction Conditions	Creative Commons License: CC BY-NC-ND <a href="https://creativecommons.org/licenses/by-nc-nd/4.0/">https://creativecommons.org/licenses/by-nc-nd/4.0/</a>
Descriptions and Notes	Description [Source : Maritz, N.G. Relics of War: A Collection of 19th Century Artefacts from British South Africa and Southern Rhodesia. Pretoria: Salut Africa CC, 2008, used without alteration by JAG, 2015. JAG Description: Ceramic vessels, North Nguni, 19th Century, Material: Wood]

**Figure 4: Emandulo item metadata**

The metadata of an item is subject to change, and it should be possible to update it in-place. Editing metadata involves creating forms, validating the input data, and generating the correct storage format for the data [7].

With the current functionality of Simple DL, it is already possible for administrators to edit metadata in place. This functionality needs to be extended to the editing of files.

XML has become the standard storage representation for metadata [8]. Simple DL captures metadata as XML files. These XML files are converted to HTML files and the same is done to all other files [9]. The source files containing the information are CSV files which are read in during the *Importing* phase. To edit files on digital archives built using Simple DL, editing of the CSV files is necessary.

## 2.5 Seamless editing

Seamless editing [3] is a new approach and different approach to editing the content of Web pages. The current approach forces authors to switch from “editing” to “previewing” and finally to “publishing”, making it a tightly structured and extremely moded process. Seamless editing is entirely modeless and does not require authors to switch between modes, making the process less complex and more intuitive.

Seaweed [14], short for seamless Web editing, is a framework for editing contents of Web pages directly within any Web browser

and it has been shown that Seaweed can be integrated into any content management system.

As an experiment, WordPress was used along with a plugin that was developed for the Seaweed framework to create a seamlessly editable environment [3]. Users using the plugin and carrying out tasks within the environment were observed and compared to users who performed the same tasks without the plugin. Participants of the experiment found that editing Web pages, in this case the users’ blogs, with the Seaweed software was easy and preferable.

Seamless editing allows users to edit Web pages in a matter of seconds [3]. From the experiment conducted, it is also shown that this method of editing is preferable and was found to be intuitive for the participants. When carrying out this project and implementing the functionality to allow in-place editing of files using Simple DL, some aspects from Seaweed such the lack of or avoidance of modes could be adopted to some extent. While having no modes at all may not be possible because the edits made to the files need to be validated before being published, other less essential modes such as “previewing” could be omitted.

## 2.6 Issues with in-place editing

There are three main factors that disorientate users when using in-place editing [3]. These are namely, the appearance disconnect, the interaction disconnect and the delay disconnect. These are termed as “disconnects” and interrupt users when they initiate the editing process and when they evaluate the results of the changes that they have made.

The appearance disconnect can be avoided by maintaining the same CSS styles and position of content when switching to edit mode. By reducing the actions that need to be carried out by a user from the time they decide that they wish to make an edit until the time they start editing can avoid the interaction disconnect. Lastly, keeping delays to a minimum while a user is carrying out tasks can avoid the delay disconnect.

Simple DL will be susceptible to the issues mentioned above once the feature allowing administrators to edit files in-place materialises. It is important to be aware of these issues and keep them in mind when developing this project to ensure that they are being avoided to the best extent possible.

## 2.7 Crash consistency

Crash consistency [16] requires data to be recoverable in adverse conditions when a system crashes. There can be major consequences of having no proper crash consistency in place.

In the case of Simple DL, the event of an adverse condition such as a network connection being lost or a power outage while a file

is being edited by an administrator needs to be taken care of to prevent the contents of the file being corrupted or lost.

Since Simple DL is used for digital archives, the contents of the files are important collections of information such as research and history. For example, as mentioned earlier, Emandulo [7] is a collection of pre-colonial Southern African history that was scattered across the country. All these documents had to be collected, converted to digital copies, edited, and assembled with all their metadata by staff of the Archives and Public Culture Initiative.

Assembling the Emandulo archive required a great deal of time and effort and losing any of the data while editing files due to a lack of proper crash consistency would be a major inconvenience.

## SUMMARY

The body of this paper is made up of two main sections. An overview of digital archiving and the editing of files. Before getting into the body, some background information provides a short explanation of how Simple DL works was presented to gain a better understanding of the existing system and its current functionality.

In the first section, a broad look at digital archiving toolkits, including those that work online and offline were discussed with examples and compared to Simple DL. The shift from custom solutions to toolkits was also discussed.

Secondly, matters relating to the editing of files were discussed in the third and final section. In place editing was explained and how it is used for Web pages and metadata was reviewed with examples. A brief comparison between in-place editing and back-end editing was reviewed. Seamless editing was then discussed followed by the issues that come with in-place editing.

With the knowledge and understanding gained through this review, a suitable feature to edit files in-place on Simple DL will be implemented.

## REFERENCES

- [1] Bay-Wei Chang. 1998. In-place editing of Web pages: Sparrow community-shared documents. *Computer Networks and ISDN Systems* 30, 1 (1998), 489-498. DOI:[https://doi.org/10.1016/s0169-7552\(98\)00118-4](https://doi.org/10.1016/s0169-7552(98)00118-4)
- [2] Brenda Chawner and Paul H. Lewis. 2006. WikiWikiWebs: New Ways to Communicate in a Web Environment. *Information Technology and Libraries* 25, 1 (March 2006), 33-43. DOI:<https://doi.org/10.6017/ital.v25i1.3327>
- [3] Brook Jesse Novak. 2010. Seamlessly Editing the Web (Thesis, Master of Science (MSc)). *University of Wakaito, New Zealand* (2010).
- [4] Christian Wagner. 2004. Wiki: A Technology for Conversational Knowledge Management and Group Collaboration. *Communications of the Association for Information Systems* 13, (February 2004), 265-289. DOI:<https://doi.org/10.17705/1cais.01319>
- [5] Christopher Gutteridge. 2002. GNU 2 Overview. *11th Panhellenic Academic Libraries Conference* (2002).
- [6] EMANDULO. Retrieved March 24, 2023 from <http://emandulo.apc.uct.ac.za/metadata/Presentations/7549/index.html>
- [7] Hussein Suleman. 2003. Metadata Editing by Schema. *Research and Advanced Technology for Digital Libraries* (2003), 82-87. DOI:[https://doi.org/10.1007/978-3-540-45175-4\\_9](https://doi.org/10.1007/978-3-540-45175-4_9)
- [8] Hussein Suleman. 2007. Digital Libraries Without Databases: The Bleek and Lloyd Collection. *Research and Advanced Technology for Digital Libraries* (2007), 392-403. DOI:[https://doi.org/10.1007/978-3-540-74851-9\\_33](https://doi.org/10.1007/978-3-540-74851-9_33)
- [9] Hussein Suleman. 2021. Simple DL: A Toolkit to Create Simple Digital Libraries. *Towards Open and Trustworthy Digital Societies* (November 2021), 325-333. DOI:[https://doi.org/10.1007/978-3-030-91669-5\\_25](https://doi.org/10.1007/978-3-030-91669-5_25)
- [10] Ian H. Witten, Stefan J. Boddie, David Bainbridge, and Rodger J. McNab. 2000. Greenstone. *Proceedings of the fifth ACM conference on Digital libraries* (June 2000), 113-121. DOI:<https://doi.org/10.1145/336597.336650>
- [11] Jihyun Kim. 2006. Finding documents in a digital institutional repository: DSpace and Eprints. *Proceedings of the American Society for Information Science and Technology* 42, 1 (October 2006). DOI:<https://doi.org/10.1002/meet.1450420173>
- [12] Jonathan Koren, Yi Zhang, and Xue Liu. 2008. Personalized interactive faceted search. *Proceedings of the 17th international conference on World Wide Web* (April 2008), 477-486. DOI:<https://doi.org/10.1145/1367497.1367562>
- [13] Mayukh Sarkar and Sruti Biswas. 2020. Exploring Archives Space an Open-Source Solution for Digital Archiving. *DESIDOC Journal of Library & Information Technology* 40, 5 (September 2020), 272-276. DOI:<https://doi.org/10.14429/djlit.40.05.16330>
- [14] R. Tansley, M. Bass, D. Stuve, M. Branschovsky, D. Chudnov, G. McClellan, and M. Smith. 2003. The DSpace institutional digital repository system: current functionality. *2003 Joint Conference on Digital Libraries, 2003. Proceedings.* (2003), 87-97. DOI:<https://doi.org/10.1109/jcdl.2003.1204846>
- [15] Stevan Harnad. 2001. The self-archiving initiative. *Nature* 410, (April 2001), 1024-1025. DOI:<https://doi.org/10.1038/nature28061>
- [16] Yanyan Jiang, Haicheng Chen, Feng Qin, Chang Xu, Xiaoxing Ma, and Jian Lu. 2016. Crash consistency validation made easy. *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering* (November 2016), 133-143. DOI:<https://doi.org/10.1145/2950290.2950327>
- [17] Charalampos Chelmis. 2009. Mutable Peer-to-Peer File Systems: Analysis and Evaluation. *First International Conference on Advances in P2P Systems* (October 2009). DOI:<https://doi.org/10.1109/ap2ps.2009.21>