

# From real to digital communities: Building a software archive for Trinity Community Arts

Annie Berry, Edson Burton, Ryan Northey  
Trinity Community Arts, Trinity Road, Bristol, BS2 0NW  
[\[annie.edson.ryan\]@3ca.co.uk](mailto:{annie.edson.ryan}@3ca.co.uk)

**Abstract.** The Heritage Lottery funded project, *What's Your Trinity Story?* documents the history of the Trinity Centre in Bristol from the 1960s to today. Trinity embodies and reflects the social and cultural changes that have taken place around it. Since its de-consecration in 1976, Trinity has passed through the hands of a number of groups and owners, whilst developing into a major music and community arts venue for the city. One aim of the project is to establish an online archive for Trinity. A digital archive can offer a more accessible means of engaging communities than a traditional physical archive. Moreover, it creates links between different items, eras, ideas and cultures, challenging physical, temporal and cultural divides. By adopting open standards in the development of the archive, it is hoped that we will allow wider inter-operability and re-use of content, as well as allowing replication of the infrastructure by other groups. This paper discusses some of the challenges encountered in building this archive.

**Keywords:** open source software, digital archiving, Bristol history

## 1 Introduction

First opened in 1832 as Holy Trinity Church, the building was de-consecrated in 1976. From this time onwards, Trinity has been in use under a covenant for the service of 'youth, community and the arts,' with its success ebbing and flowing under each change of management. Amongst varied community and arts activities, Trinity has consistently since been used as a music venue, having played host to many famous artists, including U2, The Wailers, Crass, Echo & the Bunnymen and Prodigy. Significantly, Trinity was a key venue in the locally inspired and internationally exported 'Bristol Sound' of the 1990s, with regular appearances from the likes of Roni Size, Massive Attack and Portishead.

Following the building's most recent closure in 2001, the centre was re-opened by Trinity Community Arts (TCA) in 2004. The building has since undergone substantial refurbishment downstairs for use as a music venue and for community events and meetings; and upstairs for IT and music technology training, playing host to a number of heritage-based community projects. The HLF project, *What's Your Trinity Story?* is a one-year project to document the history of the Trinity Centre from the 1960s to today. One aim of the project is to develop a digital archive for Trinity. The archive provides Trinity staff with a facility for documenting and storing current



The content of this paper is issued under a Creative Commons Attribution-Share Alike License. See [www.creativecommons.org](http://www.creativecommons.org) for more information

events as well as historic documents. It will also enable public submission of items via the Internet, encouraging greater community involvement in the creation of the centre's history. The archive can be found at <http://archive.3ca.org.uk>.

Given its wide range of uses and users, it is challenging to define Trinity's communities: Is it the people living around it; is it the people using the building; is it the group who manages or owns the building; or is it those living in its shadow? Furthermore, Trinity lies at the juncture of a number of inner-city districts and communities. The boundaries of these, as well as the cultural backgrounds of their constituents, has developed and changed with time. With the archive growing, these communities –with their divisions and changes – have begun to emerge through the stories and artefacts that the archive holds. However, aiming for open access to and re-use of material has been problematic, particularly in the case of personal stories. The technical development of the archive has therefore needed to reflect individual requirements.

## **2 Our approach to software development**

One of our aims as an organisation is to use and promote Free Software, and we develop resources and provide training to that end. Code that we develop is committed to a public repository and licensed under the GPL. We adopt open standards wherever possible, and avoid the use of proprietary ones such as Flash. We also promote and use creative commons licensing for other types of creative content.

The software and coding used for our online archive is Python/Zope/Plone on the GNU/Linux operating system. We use this software for our existing systems and projects in order to enable ease of content and code re-use. It is highly extensible, scalable and extremely secure, integrating well with other systems and services, and easily deployed. This software is widely used in commercial, public and third sectors, particularly due to its strong emphasis on indexing and cataloguing; and its tight integration with a caching framework. It provides workflow-based permissioning to allow review of content and nuanced publishing, thus offering a system with the capacity for multiple users and complex roles within the site.

## **3 Working with open standards**

The entry schema follows the Community Archives and Heritage Group guidelines (see <http://communityarchives.org.uk>) as closely as possible. This in turn is guided by ISAD(G): General International Standard Archival Description, thus ensuring the data we create provides a useful and transferable reference source. A further design goal was to allow interoperability with other online archive specifications such as that provided by the Open Archives Initiative.

We take the long-term view that accessibility is best served by adopting open standards, which has influenced the design of the archive and our approach to working with the content. In particular, we use HTML with CSS and javascript, and avoid the use of Flash or other proprietary elements.

In practice, however, adopting ‘open’ standards is often not the same as adopting ‘common’ standards or even ‘working’ standards. Flash provides probably the easiest way to embed multimedia content within a web page, and it could be argued that this is more accessible as it is so widely implemented.

It is hoped by many that HTML5 will provide an open standard for working with multimedia, so we have adopted that for rendering multimedia elements. As this would break the compliance of the page markup, we add them dynamically using javascript.

The actual specification of HTML5, and in particular for multimedia codecs, has been disputed by the browser-makers. Some of the specification has therefore been left open, much in the way that it is for image elements in x/html currently. However, this leaves the browsers natively supporting different multimedia codecs.

While we always keep the original files as submitted, we decided to use the Ogg Vorbis and Ogg Theora codecs for online presentation as they are open standards and are therefore implemented by ‘open’ browsers. If we were to adopt a proprietary codec such as mp3 or H264, we would effectively mandate to our users that they must legally buy a license from the patent-holders of those technologies in order to use any of our archive. In practice, buying such a license is transparent to users as it is embedded in the cost of their hardware or software.

Ogg is not currently supported by all browsers, and has limited hardware support, which can be frustrating for users if they, for example, want to listen to an entry in their car. It is, however, becoming more common, and it is hoped that Google’s controversial takeover of On2, the company controlling the licensing of more recent versions of the standard, is an indication of their long-term support, meaning that it will become more widely implemented.

#### **4 Archotyping real-world artefacts**

Although we tried to keep closely to the standard schema definitions, we made some extensions in order to meet our requirements. This is to allow for digital presentation and online referencing, and to accommodate some of the significant information we held about our content.

The software we used provides basic ‘types’ such as Page, File, Image, Link, Folder, Event. We defined a number of archive ‘subtypes’ based upon these basic types, allowing us to adapt the ways in which they are viewed, indexed and managed.

A key issue we had to deal with was disambiguating namespaces. A ‘collection’ or a ‘contributor’ can mean different things to the software or the specification, and even to us. We had to settle on some names early on, and in some cases compromise between the specifications.

In developing an information architecture there is often a tension between keeping data structures simple, against providing a richer dataset or more complex functionalities that might limit accessibility, or make meanings ambiguous. We therefore tried to limit the number of subtypes that we created for the sake of simplicity. New subtypes were generally only created according to information storage requirements, rather than to reflect the intentions of the author. The software

has definitions to model most of the cultural items that Trinity stores, such as reviews, interviews, flyers, photographs, audio and video.

## **5 Data mining and making the content accessible**

We want to enable as many people as possible to locate and access our online archive and its content. The software that we use provides for relatively semantic content, therefore making it well understood by search engines. Internally, the site indexes and catalogues significant fields, and provides search functionality and a number of systems to list and facilitate finding relevant content. We also developed some software to make the content more visual by generating images from files and associating images in listings. Furthermore, we use a number of AJAX technologies to speed up the users' site experience and allow the content to be browsed more quickly.

In terms of its practical functioning, it has been very important for the archive to enable new connections to be drawn between the past and present communities and groups who have used the building. The linking of related content is crucial to this, and we have developed a system for relating items and tagging content according to subject, places and related artists. This allows us to generate tag 'clouds' or other custom listings, providing a more visual representation and making the content easier to navigate. RSS and other web services are also available, enabling automatic syndication of content.

While the archive is physically held at Trinity, we created a copy of the published content within our website, which is updated as the content in the archive changes. This allows us to store all the original content and perform the resource-intensive activities, such as converting files, on servers at Trinity. At the same time it also allows for the content to be served to the public from an internet server with higher bandwidth capacity.

## **6 Identifying and working with archive users**

We identified and implemented several roles within the archive, such as administrators, editors, reviewers and submitters. Key amongst these is the reviewer, as they are responsible for checking that content meets our guidelines, is authoritative and properly sourced, and is not legally infringing. As these users have significant responsibilities and access to confidential material, we are developing a training system for inducting new reviewers.

We also created a number of different levels of access and categories of users. In order to hold content in cases where we have been asked not to release it publically, or where we feel that we cannot publish it for some other reason, we created a category of content which is restricted to 'research' access, and have set out a policy for granting access. Research access may be given to academic or community historians, journalists or other local people on the recommendation of a higher education establishment, or at TCA's discretion and according to the policy. We also

created a 'restricted' level where no access is granted to anyone other than the archivists, for a set period of time.

## **7 Attribution, copyright and permissions**

We have had to strongly consider the licensing of content to enable its use beyond the archive. With new oral interviews conducted within the What's Your Trinity Story? project, this has meant facing the challenge of negotiating consent with interview participants. In particular, we have to be careful not to make public defamatory or incriminating content. Generally, for this type of content we hold the copyright ourselves, according to the wishes and any conditions set down by the participants. In the case of oral interviews, we hold a full, unedited original archive copy of the interview, which is only accessible to those with research access. Public users will have access to the same interview, divided into more user-friendly sections. Defamatory or incriminating data is to be silenced out of these smaller sections at our discretion and according to the wishes of interviewees.

Likewise, we want public contributors of other items such as flyers and photos, to be included in choices about whether their stories and artefacts can be shared online. As a shared representation of Trinity's history, we want to enable users to fully draw from the archive and even create new cultural artefacts from it; to reuse and recreate Trinity's past, into the future. We therefore license the content that we hold under a creative commons license, and actively encourage those using content from it to permit the same.

By default we have opted for a non-commercial, share-alike license, which means that any derivative work must carry a similar license and further permission must be sought for commercial use. We adopted this in part to protect contributors from commercial exploitation and in part to make the license seem more reasonable to contributors who are unsure about open licensing.

## **8 Summary**

There are similarities between a community archive such as ours, and sites such as YouTube. Both sites host users' content, allow feedback, and build communities with an affinity based around their use or enjoyment of particular content. The community archive is, however, a fundamentally different locus of knowledge: It is owned and constructed by its communities and it reflects a specific theme or locality through the histories of its constituents.

In order for a community archive to be successful and valued by its communities, and to host quality contributions, it needs to gain the trust of potential contributors. They must trust that their rights will be respected and the content will be managed appropriately. Software is 'disinterested' to a greater degree than many other cultural artefacts, and is therefore easier to license liberally. Licensing a photograph, for example, can have greater implications for its subject than for its author. We have

therefore tried to find a balance between our role in making the content accessible and our role as legal and moral guardians of what can be sensitive content.

By hosting well-researched and authoritative content, making the content widely accessible, and providing references to other information stores, we have tried to develop a model for peer archiving. It has already proved to be a vital resource for compiling and researching Trinity's history, giving insight to Trinity's local and national significance. We also hope that these discussions around artefacts, users, access and permissions will open discussions and possibilities for other cultural and community archive projects, long into the future.