

Open Principles, Open Data: The Design Principles and Architecture of the UC CEISMIC Canterbury Earthquakes Digital Archive

James Smithies, Paul Millar, Chris Thomson

Digital Humanists are developing a tradition of disaster archiving. The trend began with the Centre for History and New Media's (CHNM) 9-11 archive,¹ started after the attacks on the World Trade Centre in New York in 2001, which eventually crowd-sourced 150,000 items and has been itself archived by the Library of Congress. CHNM, in partnership with the University of New Orleans, later developed the Hurricane Memory Bank (2005), in response to hurricanes Katrina and Rita.² Digital humanists have also been involved in responses to the devastating Tōhoku earthquake and tsunami of 2011.³ This article outlines the approach taken by digital humanists at the University of Canterbury, New Zealand, to the series of earthquakes that seriously damaged the city of Christchurch from 2010 – 2011. In many ways, these initiatives represent a continuation in the digital sphere of activities many centuries old. Disasters are, after all, primarily human events: nature may cause them, but it is the human impact that demands a response. Many of our cultural memories of great disasters were created by humanists: oral stories and subsequent myths about the destruction of Atlantis in 1620 BC, extensive literary and historical references to the Lisbon earthquake of 1755 and the eruption of Krakatoa in 1883, artistic impressions of the destruction of the pink and white terraces in the North Island of New Zealand in 1886. Although any disaster response must be regulated by the need to prioritize life and property, history shows us that humanists will normally be somewhere in the mix, collecting, preserving and commenting on the event for present and future generations. The appearance of digital tools has provided us with new avenues for our activities in disaster and post-disaster contexts, but has done little to alter the innate drive to collect, analyze,

create, and explain. This article outlines the design principles and architecture of the UC CEISMIC Canterbury Earthquakes Digital Archive in an attempt to record an approach to DH disaster response that might benefit future efforts.

1. Background

At 12:51 on Tuesday 22 February 2011 a magnitude M6.3 earthquake struck the city of Christchurch, on the east coast of New Zealand's South Island. Local geography and soil structure, combined with a series of faults under the Canterbury Plains, produced significant amounts of damage. Unlike the 7.1 earthquake experienced on September 4th of the previous year, this event produced remarkably high rates of ground movement and resulted in the loss of 185 lives.⁴ At the time of writing over 1300⁵ of the 3000⁶ buildings in the Central Business District (CBD) have been demolished, in a process that will continue well past the two year anniversary of the event. The government has noted that "much of the historic fabric [of Christchurch] has been lost, as have key facilities such as the Convention Centre, a significant proportion of the hotel capacity, and sports and recreation facilities".⁷ There was severe damage to lifelines infrastructure across the city of 450,000 people, "...including road, the water and wastewater networks and the electric systems".⁸ Eventually, entire suburbs had to be abandoned due to liquefaction, subsidence, rock fall, and a host of other geo-structural issues. The city rebuild is expected to cost NZ\$30 billion and take 10 – 15 years to complete. By August 2012 over 143,000 insurance claims have been made with the New Zealand's Earthquake Commission.⁹ As of March 2012, the disaster ranked as the third costliest insurance event in history.¹⁰ Four earthquakes of magnitude 6 or greater and over 11,000 aftershocks have been recorded in the area since September 4th 2010.¹¹ Residents continue to struggle with significant disruption to their daily lives.

2. Disaster Management and Digital Archives

Post-disaster recovery plans focus – as they must – on saving lives, identifying victims, reconstituting essential services, and providing information to residents and business owners. They are characterized by overtly vertical command and control management structures, to assist with the coordination of emergency teams, government agencies and the army. Normally only essential recovery activities are resourced in an attempt to maximize the response and minimize confusion.¹² That said, disaster management methodology highlights four phases that guide the response of governments, local body authorities and Non-governmental Organizations (NGOs) to disaster situations: mitigation, preparedness, response, and recovery. This holistic approach, that includes preparation beforehand and recovery afterwards, is particularly well suited to a response by humanists, because it acknowledges the need for long-term participation and accepts the need for cultural and educational input in order to lessen both the immediate and long-term impacts of disasters.¹³ Although every country differs in its disaster response capabilities and policies (and all will prioritize the preservation of life and property in the immediate aftermath of an event), most government agencies in the OECD are aware of the need for broad-based information gathering, education, and the development of social and cultural capital:

Memory, experience, and knowledge are critical to the development of effective response mechanisms. Knowledge of past events can condition how contemporary society not only conceptualizes the risk connected with particular events but also anticipates the impacts of future catastrophes.¹⁴

This principle is enshrined in UNESCO's *Text of the Convention for the Safeguarding of Intangible Cultural Heritage* (2003)¹⁵ and the United Nations' *Hyogo Framework for Action* (2005 – 2015), which includes a priority action to “[u]se knowledge, innovation and

education to build a culture of safety and resilience at all levels”.¹⁶ Similarly, the 2013 Vancouver Declaration explicitly “urged the [UNESCO] secretariat” to

create an emergency programme aiming at preservation of documentary material endangered by natural disasters or armed conflicts, as well as a programme for the recovery of analogue and digital heritage that is under threat of becoming, or is already inaccessible because of obsolete hardware or software.¹⁷

University of Canterbury¹⁸ researchers were heavily involved in the early disaster management response to the February earthquake, coordinating information about the status of essential infrastructure and services, providing high performance computing storage and services, and educating the public through the media. Few roles, though, were available to humanities academics in the weeks after the February 2011 earthquake. The University of Canterbury closed for several weeks, to allow staff and students to contact loved ones and attend to damaged properties. Thousands of people left the city to avoid aftershocks. A state of emergency stayed in place until April 30th,¹⁹ making it impossible to attend to business as usual. Damage to buildings and general infrastructure, and severe dislocation of regional infrastructure, created significant operational problems for many months. The loss of life and presence of international search and rescue, the military, and increased numbers of police contributed to a general sense of emergency and disorder. When the university did open, during March, staff conducted some classes in tents as contractors began building prefabricated classrooms on vacant ground; the situation was anything but conducive to the development of new digital resources. It wasn't until May, 2 months after the first major event, that Paul Millar from the Department of English was able to start thinking about possible responses. In an indication of the problems university staff were facing at the time, Millar discussed possible options with the author of this paper while he was temporarily located in Wellington, because his house was unfit for habitation.

Despite the difficulties, and as Millar was aware, a robust response to disaster situations by humanists soon after the event itself is important, not only for disaster management decision support²⁰ but also in broader cultural terms, to avoid the “digital death”²¹ of crucial artifacts produced as a direct result of events. Evidence suggests that the formation of autobiographical and collective memory in the aftermath of significant trauma is a complex affair. Psychologists²² and sociologists²³ alike point to the role emotion plays in the construction of individual and collective memories of events, so it is important that large bodies of primary material are identified and safely stored for later analysis (both by professional researchers and individuals) when the immediate trauma has passed. By doing so, individual and collective cultural memory of the event can be continually revisited and refined and a longitudinal understanding developed. Although cultural resources tend to be “ignored and neglected”²⁴ in the immediate post-disaster phase, the experience of the September 11, Hurricane Memory Bank, UC CEISMIC, and Tōhoku archives suggest there are compelling arguments for the speedy deployment of cultural heritage-related assets. Studies suggest a focus on ‘social capital’ (which includes a cultural component) can speed recovery, enhance cohesiveness, and contribute to post-disaster resilience.²⁵ Patrick Meier has pointed out the specific role collections of big data can have in these processes.²⁶ It is important to recognize the “multidimensionality”²⁷ of disasters, and the impact a loss of cultural heritage can have on communities. In some cases, “rescuing culture is essential for the mental survival of people in emergency situations, and can contribute to their overall resilience and empowerment when overcoming catastrophe”.²⁸

The problem is, of course, that easily deployable digital archives suitable for the complex task of post-disaster collection do not exist. While applications like Omeka, Islandora, Ushahidi, and Fedora Commons provide excellent starting points, developing a ‘shrink-wrapped’ solution tailored specifically to post-disaster situations would be a non-trivial task. The ideal situation would be one in which humanists, or perhaps government employees

within a central cultural heritage agency, could visit an online service provider (or download an easily deployable virtual machine to deploy on their own infrastructure) after a disaster and provision a robust, preservation-quality, archive system capable of ingesting large quantities of digital content (either crowd-sourced or through an administrative interface) according to configurable standards-based ontologies. The infrastructure would be on-demand Infrastructure as a Service (IAAS) and fully scalable. The software would be provided as a service as well (SAAS), perhaps with a modular architecture to allow administrators to deploy services as required (including integration with and archiving of social media services, and provision of downstream big data analysis). It would be capable of metadata aggregation to allow it to act as the central node in a heterogeneous federated archive, and have an easy to use interface and 'baked in' Terms and Conditions and copyright tools, to ensure broad-based usage and legal probity. Large-scale data export functionality would allow for migration of content to long-term preservation systems and dark archives. In lieu of such a service systems need to be put together very quickly by humanities or cultural-heritage teams, in the difficult circumstances of post-disaster management and with minimal funding. Sometimes, as was the case with the September 11 archive, the process works very well and great benefits accrue from small outlay; at other times, as with the Hurricane Memory Bank, a similar team can go through the same process but with significantly reduced results.²⁹ There is currently no example of a broad archive aggregating post-disaster cultural heritage, scientific, geographic and social data in a manner conducive to long-term preservation, general research by the public and academics, and computationally intensive research.

3. Governance

One issue with the development of a generic cultural-heritage disaster archive system would be how to deploy it in a variety of operational, cultural, and legal contexts. While New

Zealand's culture and local and central government structure bears many similarities to other OECD nations there are significant differences too, especially in its lack of state or federal structures, which enabled UC CEISMIC to develop a broad-ranging Consortium of both local and central government agencies. Differences with developing countries would presumably be even more profound. Detailed discussion of these issues is outside the scope of this article, however. The important thing to note is merely that the UC CEISMIC archive was designed and deployed with a particular operational context in mind. New Zealand's small cultural network meant that it soon became clear that a variety of groups were considering or actively engaged in developing archive systems to collect quake-related content. A meeting was held at Lincoln University amongst interested parties and a decision was made to form a Consortium, which eventually comprised Archives New Zealand, Christchurch City Libraries, the Canterbury Museum, the Canterbury Earthquake Recovery Authority, the Ministry for Culture and Heritage, the National Library, New Zealand Film Archive, NZ on Screen, the Ngai Tahu Research Centre and Te Papa Tongarewa: The Museum of New Zealand. The new University of Canterbury Digital Humanities Programme led the Consortium, which came to be known as the UC CEISMIC Consortium.³⁰ Responsibility for the development of the Consortium and leadership of the project as a whole rested with Paul Millar as Director, and responsibility for project management and technical development of the federated archive (and additional University of Canterbury research repository) rested with the author. The University of Canterbury (UC) provided funds for the first two years of development and operations, and Consortium members offered both technical and archival resources, and input from Chief Executives, senior managers, and general staff. The project benefitted from a remarkable degree of inter-agency cooperation and goodwill that might not have been possible in other countries. Although there turned out to be little practical need for it, the project was underpinned by a broad adherence to a concept of mutual aid. It was agreed that Consortium members would help each other

where possible, even if that meant supporting or improving a ‘competing’ archive. Similarly, if smaller nodes began failing in future years, the broader Consortium would try to step in and help, or migrate their content to more robust infrastructures. The goal was to create a radical model, where the whole was always held up as greater than its parts. Paperwork and official documentation was kept to a minimum, and any documents that were produced adhered to a ‘less is more’ principle. The founding document of the Consortium was a three page Memorandum of Understanding, which allowed for any member to leave the Consortium with two weeks notice, signed by senior representatives of all organizations. Policies and processes for ingestion into the archive are determined by the policies and processes of contributing members, with the Consortium providing advice to the UC team in establishing their new processes. In general, aside from community archives where a lower standard is accepted, the contributing archives can be said to adhere to best practice.³¹ In many ways the UC CEISMIC federation presents a classic example of the use of information federalism to manage “information and [establish] standards for cultures that celebrate empowerment and widespread participation”.³²

4. Design Principles

The first act in the technical development of the archive was the organization of an information architecture workshop involving technical personnel from a variety of Consortium partners. The workshop was held at NV Interactive, a web development company already contracted by the Ministry for Culture and Heritage to build their ‘QuakeStories’ archive,³³ and later contracted by UC to build the main web portal for the federated archive: <http://www.ceismic.org.nz>. At this workshop, general principles were agreed about how the Consortium archives would work together and it was agreed that, rather than pooling existing resources to create a single archive, the team would work to a “distributed custody model”,³⁴ storing content in a variety of existing and planned

repositories, and contributing content to a federated archive via metadata aggregation.

Although some principles were pinned down later, and no formal list was ever produced and agreed to, the following design principles were discussed:

1. **Open Access:** The concept of a federated archive would not have gained approval without this. No Consortium partners were willing to contribute their existing content to a gated archive, and it was felt that individuals and organizations would be unlikely to contribute additional content in those circumstances either.
2. **Open Source:** Some of the government agencies were already using proprietary software and would continue to do so, but the workshop evinced broad agreement that open source components should be used wherever possible, in order to foster sharing and reuse.
3. **Multi-channel:** ceismic.org.nz would be the 'front door', but it was agreed that the federation would aim for a radically multi-channel approach. The metaphor of an 'ecosystem' was used to describe a belief that all nodes in the federation were to be of equal importance. Small community archives, after all, could well contain more valuable content than large national ones. The key was to facilitate and foster a broad, healthy federation, which was capable of supporting large and small partners.
4. **Asymmetry:** Because of the support behind the UC team it was understood that the proposed QuakeStudies repository was likely to become the largest node in the federation. Other contributing organizations were constrained by their normal business as usual operations, and although the New Zealand government had directed them to prioritize quake-related activities as part of their strategic plans (and the Christchurch City Libraries team in particular had made valiant efforts to get an archive up and running very quickly after the earthquakes), it was clear the

UC team were the only ones in a position to focus their efforts on earthquake-related content ingestion for years at a time.

5. Heterogeneity: Archival 'nodes' mushroom in post-disaster situations, due to the ubiquity of web technologies and the ease with which simple archives can be established with products like Wordpress. Rather than impose uniform standards and technologies that would have stifled the development of new archives and left many small community archives outside the federation, it was necessary to embrace heterogeneity and design a solution that could cope with a broad range of technologies.
6. Extensibility: The development of an open dataset allowed for the possibility of myriad new sites and applications as the archival ecosystem developed. This was embraced and undertakings were made to encourage the development of widgets, mobile applications and satellite sites in order to broaden the reach of the Consortium as far as possible.
7. Leveraging existing assets: It was made clear that UC CEISMIC was a national project rather than a regional or University one. The post-disaster situation, involving significant loss of life and a devastated city, meant that there was no room for partisan politics. From the outset it was understood that the programme would leverage existing national digital infrastructure as much as possible. There was no reason to spend money duplicating existing solutions or services when the earthquake had apparently already put a \$10-15 billion hole in the economy.
8. Data consistency: The development of federated archives requires attention to data consistency, to aid metadata aggregation and facilitate longitudinal and computationally intensive research. Metadata consistency was relatively low across the existing archives, despite basic adherence to Dublin Core essentials. The UC

team undertook to work with member archives to improve their metadata if necessary.

9. Data Openness: For Phase 1 of the UC project (the initial build and deployment of ceismic.org.nz and quakestudies.canterbury.ac.nz) online services like Facebook and Twitter wouldn't be archived. These services were very useful in the post-disaster context, but pose difficulties for long-term preservation. It was felt better to deal with the basics first and consider social media later.
10. Geo-referencing: Damage caused by earthquakes tends to be associated mainly with built environments such as buildings and houses. Indeed, a significant proportion of the damage to Christchurch was in the Central Business District (CBD) and suburbs hit by severe liquefaction. Although time-consuming, it was agreed that efforts should be made to geo-reference as much content as possible to enable implementation of map-based discovery tools.
11. Linked Open Data: Wherever possible design efforts would enable participation in the world of Linked Open Data (LOD). Because many of the 'nodes' in the federation were already established, and not capable of LOD, most of the efforts in this regard were directed towards the University of Canterbury's new QuakeStudies repository.

None of these principles were particularly challenging to workshop participants. Conversely, they represented a set of principles – or a common language - held in common across the IT, cultural heritage, and digital humanities worlds that bonded the group. The biggest concern of many participants before the workshop was the possibility that one or more participants would not be aware of these common expectations and would demand, for instance, a gated archive using proprietary technologies that would undermine the smaller, more vulnerable archives in the Consortium. Some participants worried that the University, in particular, would take a closed approach to data acquisition and sharing; the communication of common standards, and digital humanities principles of openness and sharing went a long

way to allaying fears and allowing development to proceed. As Linda Barwick has noted, distributed systems like UC CEISMIC “can only work into the longer term if they are built on shared standards, formats, and procedures, designed for long-term viability”.³⁵

12. Architecture

UC CEISMIC relies on two main assets: the UC CEISMIC Federation, which provides a website, metadata aggregation services, and federated archive comprised of over 10 ‘nodes’, and UC QuakeStudies, a bespoke repository tailored to the collection of cultural heritage content and research data. Either of these assets could be deployed individually, but together they offer a wide-ranging solution to myriad issues. The products were built concurrently over a period of 18 months by NV Interactive and CWA Media / Learning Media Limited respectively. The University of Canterbury Digital Humanities Programme led the design and build of both systems, from procurement to requirements definition, solution design, development, testing, and deployment. Overall project management lay with UC, who also held the responsibility for final decisions on technical matters. That said, it should be made clear that external vendors and professional software developers built both assets. QuakeStudies in particular was treated as a major enterprise build, involving a vendor-side project manager, solution architect, front and backend developers, testers, graphic designers and system administrators.

i. UC CEISMIC Federation

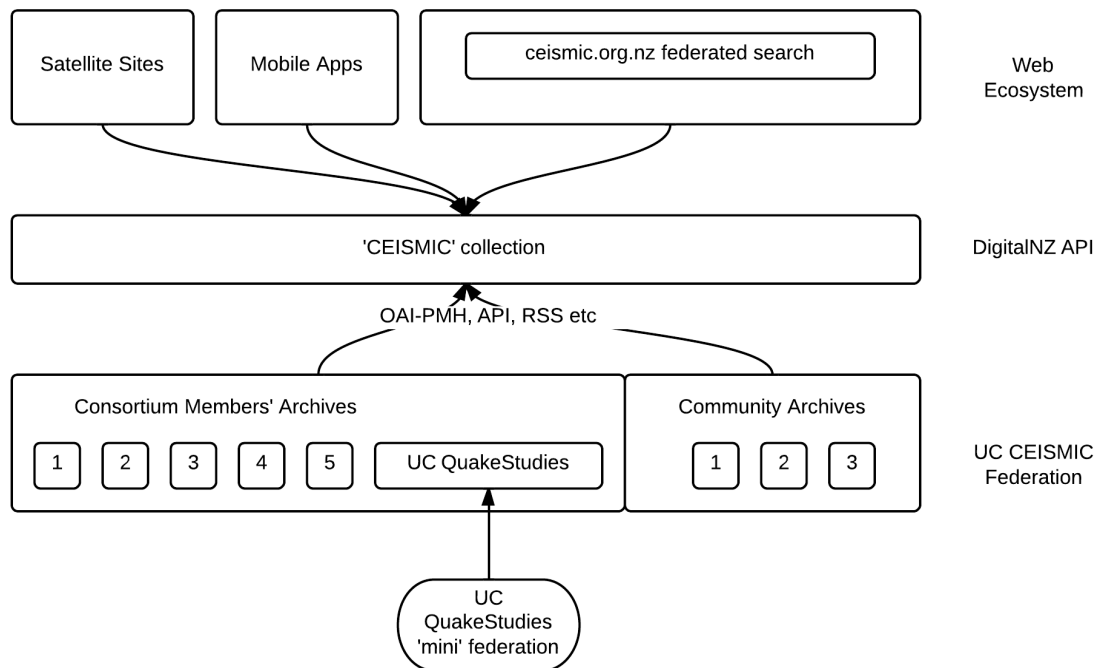


Figure 1: UC CEISMIC Federation Architecture

The UC CEISMIC federation is comprised of three separate layers. At base, a Memorandum of Understanding binds it, with signatories agreeing to “ensure that digital content deemed appropriate by them and the [Programme] Board is made available to users of [UC CEISMIC]”.³⁶ Effectively, this means that Consortium members undertake to make all their digital holdings associated with the Canterbury earthquakes available to the broader federation via seismic.org.nz or any other Consortium-related sites that might appear. For most of the members this wasn’t an issue: they are mandated by government to collect digital material related to significant New Zealand events, and make them publically available. Some organizations had less developed digital infrastructure, or no digital infrastructure at all, and would either share other members’ infrastructure or develop their own (some agencies had plans for implementing digital archives but had not yet gone live with them). At go-live 3 providers were able to deliver content, with others coming online progressively as resources allowed. The great benefit of this approach, of course, is that

nothing else needs to be done after the archives are 'plugged into' the broader federation: business as usual practices mean that contributions to UC CEISMIC will grow automatically as federation harvesting proceeds. The same is the case for any community sites added to the federation. At the time of writing, with the system operationally complete and an archive 'seed' completed, the archive includes over 60,000 items contributed by over 12 providers. Projections indicate the archive will hold upwards of 100,000 items by the end of 2013, contributed by many different content providers. The expectation is that year on year increases, over the course of the 10 – 15 year rebuild process, will result in a significant cultural asset.

The mechanism for metadata aggregation across the federation was initially a technical concern, and would be a significant challenge for any group intending to implement the UC CEISMIC model outside New Zealand.³⁷ As indicated below, a decision had been made relatively early to use Fedora Commons as the backend repository for the UC QuakeStudies repository, in part because it has native ability to act as a metadata aggregation point: QuakeStudies would aggregate content from the federation members, and seismic.org.nz search would be powered by API queries from its backend. This sounds straightforward enough, but it would have involved significant overhead. Federation members have a broad variety of metadata standards (ranging from national library and museum quality to the barest Dublin Core fields on a Wordpress instance), which would need to be massaged and mapped to a common standard. Legal and policy issues would also have been a significant issue, as the small UC CEISMIC programme team would have needed to administer content agreements and assure the University legal team that the Terms and Conditions were both robust and enforceable. As it happens, QuakeStudies is only likely to become a metadata aggregation point for archives within the University of Canterbury, or perhaps for other New Zealand universities who would like their content surfaced via UC QuakeStudies as well as seismic.org.nz. Such a 'mini-federation' would offer administrative

advantages, in that research-heavy data could be corralled into a single repository that has easy access to high performance computing services (see below).

As it happens, New Zealand has an existing metadata aggregation service based in the National Library, known as DigitalNZ. Although outside the earliest technical discussions, this team were sent the Detailed Requirements for the QuakeStudies archive early in the design process, which included an ontology based on their own metadata schema (itself based on Dublin Core), to enhance interoperability. It quickly became apparent that not only was the UC CEISMIC project well aligned to their strategic direction, they were already aggregating content from several of the federation members and, as a government agency, had robust policies and procedures that would enable the UC team to completely outsource metadata aggregation to them. In some senses this is the great ‘cheat’ of the UC CEISMIC programme: a core architectural component, crucial to the integrity of the entire system and difficult to implement, was outsourced to a government unit. And yet, it was only the underlying governing principles and clear articulation of the proposed technical architecture that made this possible. Open architectures, collaboration, leveraging existing assets, working for the public good – all these attitudes combined to point to DigitalNZ. Although it might seem an obvious choice, it would not have been possible if the project had been constituted in a different way, or been based on different attitudes: constant recourse to principles of openness and sharing offered significant benefits, but was in the final analysis a choice that could have been made differently.

DigitalNZ is a beguilingly simple service. Established in 2008, and now a business unit of the National Library of New Zealand, the service uses SOLR to aggregate metadata from hundreds of content providers across New Zealand and expose the metadata through a Ruby on Rails CMS, and simple API. The platform is capable of harvesting from a variety of formats (OAI-PMH, API, RSS), which are then mapped to the DigitalNZ metadata schema. At

the time of writing the service aggregates 25 million records from more than 120 providers.³⁸ It has recently been used as the search engine for the upgraded National Library website.³⁹ The key point in terms of UC CEISMIC, of course, is that *all* content in the federation is aggregated upstream in DigitalNZ,⁴⁰ and available at both digitalnz.org and via the DigitalNZ API, providing a radically open data model. In some ways, although it makes sense to offer the public a ‘front window’ to the UC CEISMIC collection at seismic.org.nz for ease of use and to ensure the programme has a solid web presence, there is no technical reason to do this – the content is available at DigitalNZ anyway. It should also be noted that, depending on the individual licensing restrictions placed on content, users are free to mashup, remix, and reuse UC CEISMIC content using the DigitalNZ API as well.

This open architecture offers significant advantages in the context of Web 2.0 and the movement towards a mobile web. In some senses, the UC CEISMIC archive is conceived as post-website service. The dataset itself, and the API that exposes it, are the essential components; multiple access points, or channels, can be developed as need or interest arises. seismic.org.nz, although highly functional, is merely a Microsoft Umbraco CMS with a search function that queries the DigitalNZ API; a light-weight front-end to a highly distributed data architecture. In some ways, the success of the programme will be determined by the number of channels that are built to expose the archival content: the more that exist, the more uses the content is being put to. The first mobile application – a Windows 8 ‘Metro’ application that was a winner in the Microsoft New Zealand Humanising Data competition⁴¹ - has been released and more are expected to follow. To a similar end Lincoln University Applied Computing have produced an HTML web framework that includes basic search and authentication to the UC CEISMIC collection in the DigitalNZ API. The framework is designed to facilitate the creation of ‘satellite sites’ that can be set up to showcase particular collections in the archive. Some major content providers, for instance, might want to showcase their earthquake-related content and will be able to do so using the

satellite site framework. It is one more way to open up multiple channels to the content, making it more accessible and, hopefully, more widely used. More importantly, it provides a way to encourage the development of “survivor rites”,⁴² those various cultural expressions that bond post-disaster communities and aid in the recovery, rebuild and memorialization process. If there is one ‘design principle’ that disaster-related DH projects should consider aiming to support, it is surely this.

ii. UC QuakeStudies

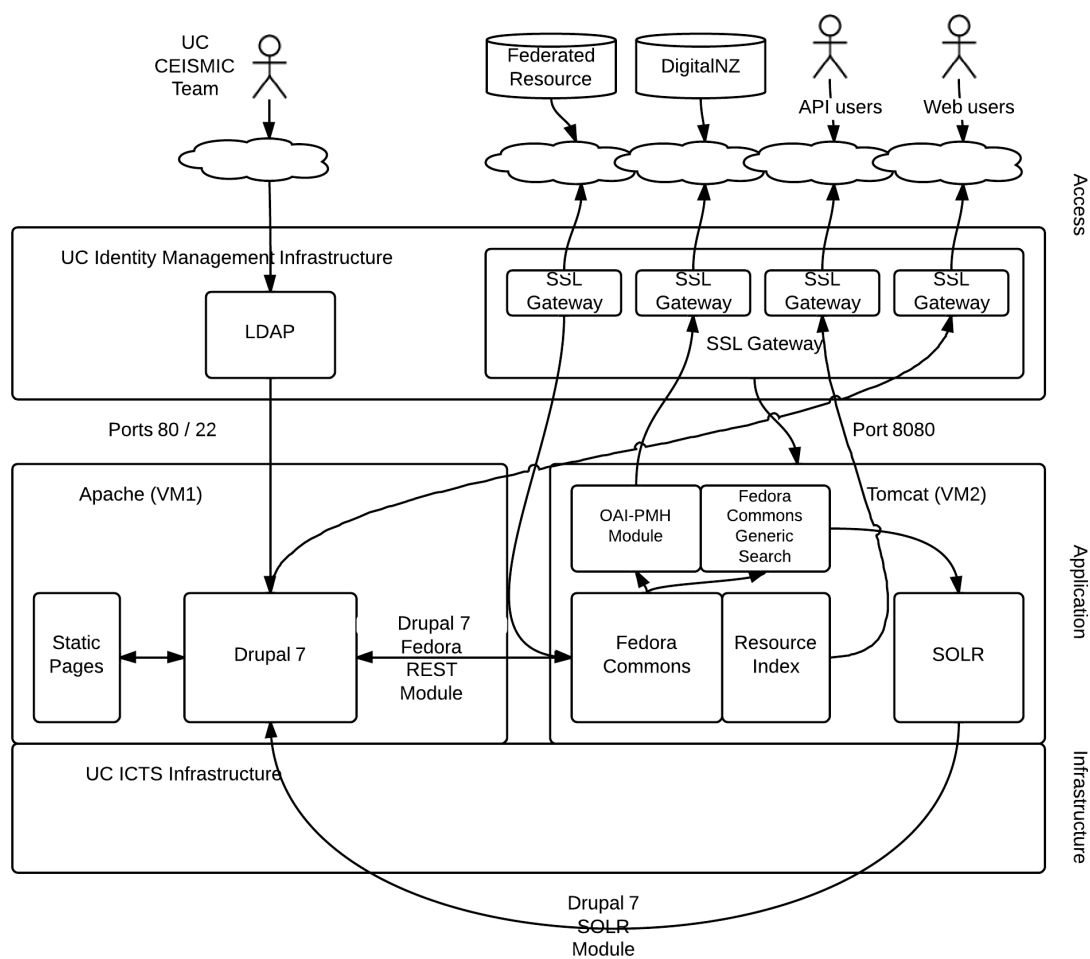


Figure 2: UC QuakeStudies Architecture

The UC QuakeStudies repository can also be described in terms of three layers, although as a single system, rather than a broad federation of systems. Following the production of high level and detailed requirements, a solution options process chose Drupal and Fedora

Commons as the preferred components. Omeka was considered as both a standalone solution and as a front-end to Fedora Commons, but Drupal and Fedora Commons were chosen as more ubiquitous technologies, with larger open source communities and therefore more opportunity for technical support. Fedora Commons fitted particularly well with requirements for metadata aggregation, native RDF indexing and API, and the ability to use no-SQL databases. Islandora was initially used, based on the assumption that it would provide a good base for further development and a desire to contribute to an excellent open source project, but it was abandoned relatively early on in favor of a 'clean' development platform of Drupal 7 and Fedora Commons, connected via a heavily customized Drupal 7 REST API module.⁴³ This created significant development overhead and introduced a higher degree of risk but, in probably the biggest decision made in the build as a whole, it was decided that the benefits of being able to use Drupal 7 and code its connection to Fedora Commons afresh out-weighed those risks. The end result is a system that could be open-sourced and added as an option alongside products like Omeka, Ushahidi, and Islandora.

UC QuakeStudies is a fairly straight-forward web application and archiving system, but its architecture needs to be described 'in the round'. The programme has three environments in total: development, testing and production. Production sits on two virtualized RHEL servers with a total of 10 CPUs and 16gb of RAM, themselves located on the university's main SAN.⁴⁴ Disk size is in the terabyte range but can (and will need to) scale to whatever is required. Most of the compute power is deployed on the Tomcat server running Fedora Commons, and is required to support Java processes triggered by requests from the Drupal frontend. Drupal application caching reduces overhead significantly, but the combination of Drupal and Fedora Commons is a heavyweight solution; cosmetic improvements are planned for the UI, but it is unlikely to ever provide a highly responsive user experience. This was a considered decision in light of predicted capacity requirements: the system would need an architectural upgrade to achieve it, but could theoretically scale

to hold 10 million objects. It should also be noted that the 'post-website' approach to the project as a whole meant that the Fedora Commons Resource Index / API and OAI-PMH feed are the most important components in the application stack: as long as these are live and able to feed DigitalNZ, and from there ceismic.org.nz, mobile apps and satellite sites, the broader 'ecosystem' will be healthy. The QuakeStudies front-end could in fact be removed entirely, with little impact on broader operations beyond a loss of the full-text and advanced search functions provided by the QuakeStudies SOLR component. Although not entirely desirable, this provides an excellent low-cost option should maintenance costs for the Drupal component become unsupportable. It may be that in twenty years time, when content ingestion is no longer a priority, all UC CEISMIC content is surfaced via ceismic.org.nz, with QuakeStudies being scaled back to the bare Fedora Commons data store.

Cloud options were considered for infrastructure, but using normal operational infrastructure offers significant benefits. Firstly, it means the UC CEISMIC Programme Office (the team that runs the archive) is supported by normal UC IT support networks. System maintenance and upgrades are all submitted to the central University Change Advisory Board (CAB), and the development expertise paid for through a Service Level Agreement with the vendor is augmented by system administration and network expertise on campus. The system also uses the standard University disaster recovery processes, and backups to tape occur as a part of business as usual. Because the infrastructure is virtualized, additional storage can be provisioned through a simple university helpdesk request and then deployed using minimal vendor support. University infrastructure also offers easy access to high performance computing services at the UC Blue Fern super-computer⁴⁵ and New Zealand's national grid computing network (NeSI),⁴⁶ via the KAREN high-speed research network.⁴⁷ It is worth noting that this choice of infrastructure was not a given, but the result of a carefully considered solution options assessment specifically related to infrastructure. Although

development of the system itself proceeded using an Agile methodology, major decisions like this involved more formal ‘waterfall-like’ methods. It is only now that the system is live that it has become fully clear just how seriously poor decision-making at crucial moments would have negatively impacted the project.

In terms of functionality, the Drupal front-end is fairly basic. Its main purpose, besides offering users browse and search access to the collections held in Fedora Commons, is as a tool to allow UC CEISMIC administrators to archive items. The ingestion process is governed by detailed policies and processes developed in consultation with the UC Human Ethics Committee and the UC CEISMIC Research Committee (composed of senior academics from all areas of the university, along with additional members from the Otago Medical School and Massey University).⁴⁸ Administrators can upload items individually, or using a bulk ingest facility capable of ingesting up to 300 items an hour with individualized metadata (for each item if necessary) included on a .csv file. During early requirements definition a lot of emphasis was placed on the inclusion of high-quality metadata (it remains one of the key performance indicators for the service), which requires the Programme Office team to spend considerable time scoping, defining and improving metadata before ingestion. The goal was never, after all, to try to collect everything produced as a result of the earthquakes but to curate a large, high-quality archive that would be useful for both the general public, but also researchers. Done properly, the hope was that the QuakeStudies repository would generate a series of use cases requiring high performance computing time. Several such use cases have already been identified.

Metadata requirements were a key focus of the early design efforts, led by the lead architect, Jason Darwin. A range of options were considered, including Dublin Core, ICOM-CIDOC, MARC, METS and the emerging “international standard for digital archiving”,⁴⁹ OAIS, in an attempt to find a commonly accepted standard that could both provide the necessary

descriptive elements, and facilitate data sharing. While many projects have difficulty finding a standard suitable for their specific heritage purposes,⁵⁰ the situation was complicated by the post-disaster context, which required event-related information rarely required in business as usual heritage contexts and therefore not included in any of the various cultural heritage standards. FOAF⁵¹ was seriously considered as well, but abandoned due to concerns about ethics and privacy; it was felt that it would be more sensible to develop specific social media projects that could use FOAF on a case-by-case basis than implementing it as a system-wide feature. The Fedora Commons Digital Object Model was extremely useful in this regard, because it offers the ability to connect multiple ontologies (or 'datastreams') to a single digital object.⁵² This allowed the team to implement a single 'base' ontology, but offer content providers the opportunity to use a reference standard of their own choice to describe their collection. With this realization, it became apparent that there was no pressing need to choose only one of the metadata standards listed above: the key was to develop a base ontology that would satisfy the immediate requirements of post-earthquake Christchurch. From there it was a relatively simple step to implement a combination of Dublin Core (satisfying the basic requirements for international data transfer) and DigitalNZ's bespoke standard (itself based largely on Dublin Core), making use of their aggregation service considerably easier. Additional standards would be attached as subsidiary datastreams when required. It was this decision, when communicated to DigitalNZ, which led to the use of DigitalNZ as the aggregation point for the entire UC CEISMIC federation: a case of internal project decisions aligning well to external service options.

The design focus placed on metadata was related to a significant long-term goal to create a dataset conducive to analysis by high performance computers (HPCs). While in some senses a *lack* of structured data would have offered more use cases for HPCs, which would have been required to derive structure and meaning programmatically, it was felt

that effort should be directed towards providing a base layer of human-entered metadata to not only facilitate basic content curation, search, and usability but to act as a control against future implementations of crowd-sourced and machine-derived metadata: users would be able to toggle crowd-sourced and machine-derived metadata on or off, giving them the ability to interpret the content through three different 'lenses' and three different levels of reliability. Another requirement called for the implementation of the Resource Description Framework (RDF) to provide researchers with semantic meaning; this was also implemented using native Fedora Commons functionality. As can be seen in Figure 1, Fedora Commons includes a REST API that can be used to access content in the backend directly, by-passing the user-interface. A web service known as Resource Index Search "exposes the relationships described in the Quake Studies Repository data model, allowing relationships between the QSR classes to be queried and navigated"; it can be programmatically accessed via HTTP GET or POST.⁵³ Tuples and Triples can be returned using a variety of query languages, offering the opportunity to engage in rich data analysis and create a broad range of data visualizations. Although it will be desirable to use computers to programmatically derive additional meaning (especially given the very large amount of content expected to be stored and the corresponding difficulty of finding the resources to describe it manually), the basis for solid semantic analysis is native to the system.

UC QuakeStudies went live as a Beta service on 26 September 2012. Many requirements (including the provision of a user-friendly API) were either scaled back or abandoned due to pressure of time and funding. Because of this, at go-live the list of potential development jobs was long, including improving the user interface, adding crowd-sourcing functions, using high-performance computers to analyse and improve collections and metadata, and improving access to the API. Drupal plugins are planned to allow more sophisticated browsing and searching of archival content based on maps and timelines, the administrative workflow could be improved, and functionality to allow the addition of

crowd-sourced and computer-generated metadata will be implemented if funding allows. Improvement continues at an infrastructure level, too, with changes being made to the way the virtual servers connect to the university SAN to improve performance as the archive scales. UC QuakeStudies is a significant asset to maintain, and the project team's espousal of continuous improvement over several decades brings with it considerable overhead, but this is simply another aspect of the model offered to the digital humanities community. Although a lot of work, it is possible for a small DH team to design, build, maintain, and improve a significant enterprise asset (in this case marked as such at a University-wide level).

13. UC CEISMIC Digital Archive: Current State

Despite its success, the UC CEISMIC Digital Archive is best considered a working 'proof of concept'. The key contribution of the project to the digital humanities community is not in the specific tools used (Drupal, Fedora Commons etc), but the commitment to national federation as a core principle, and the combination of that with long-term commitment, community-focused attitudes, and open access principles. Although the technical solution is elegant enough, and covers a very broad variety of use cases, further development is required to refine it and (as with most IT projects) the solution would benefit greatly from additional funding. Much of the value of UC CEISMIC lies in its status as a useful model for others to consider, and the attitudes, governance mechanisms, processes and policies that underpin the Consortium and control content curation, ingestion and sharing. Perhaps the most significant element in the project is the sheer scale envisaged by Project Director Paul Millar from inception. Unlike other digital humanities disaster archives, that have been envisaged as scalable but in important ways limited undertakings, UC CEISMIC was conceived from the outset as a vast all-encompassing archive of national and even international scope that is intended to keep collecting for as long as funding allows. In some ways, the federated architecture reflects a need to accommodate this vision; the suitability

of this approach for digital preservation in a post-disaster content is merely one happy result of it.

Lucky coincidences aside, it should be noted that if the goal was only to collect content in the immediate aftermath of the earthquakes the 'UC CEISMIC solution' would have been unwieldy and too slow to implement. The design approach was based on an assumption that the recovery of the Canterbury region will take decades rather than months or even years, and that the UC CEISMIC team will need to continue collecting for decades to come.⁵⁴ While the broad goal of creating a federated resource would suit most situations, the development of a relatively large bespoke repository to augment such a federation is a significant undertaking that will produce most rewards over the long term. The expectation is that hundreds of thousands of items will be ingested into the archive year on year, creating a dataset capable of providing valuable insights into the nature of disaster risk, resilience, and renewal. In this model, which could perhaps be more effectively implemented by governments or communities as part of preparedness programmes before disasters occur, the goal is to establish a robust and very wide-ranging 'net' capable of catching and preserving as much digital content as possible over as long a time period as possible.

One of the great barriers to success is the identification and storage of what UNESCO began referring to in 1952 as "intangible heritage"⁵⁵, largely due to difficulties associated with archiving social media services. Although the term initially referred to folklore, some commentators have pointed out that a focus on 'intangible heritage' opens up interesting opportunities in the digital era. Silberman and Purser suggest that in the future

[t]he task of heritage professionals will be rather to enable contemporary communities to digitally (re)produce historical environments, collective narratives and geographical

visualizations that cluster individual perspectives into shared forms and processes of remembering. These interactions are reminiscent of the conversations that once occurred much more frequently at corner bars, in town squares and by evening campfires (cf. Putnam 2001) as a vital part of the exercise of cultural diversity that is now seen as a central component of world heritage (UNESCO 2005).⁵⁶

Capturing this kind of “performative memory”⁵⁷ is made very difficult by the widespread use of social media services like Facebook and Twitter. These services are easy to use and Canterbury people flocked to them after the earthquakes, posting extensive comments on Facebook and using the #eqnz hashtag to comment on and organize themselves. The University of Canterbury Student Volunteer Army was almost solely organized around Facebook,⁵⁸ generating over 27,000 ‘Likes’, and BeckerFraserPhotos (the photographer of record) used Facebook to publish and develop a community around their photos, generating over 14,000 ‘Likes’. Ideally, content from services like Facebook and Twitter would be integrated into the UC CEISMIC archive, but it is very difficult to navigate licensing issues and organize local storage of the content outside those eco-systems. The current approach is threefold: rely on New Zealand’s National Digital Heritage Archive (NDHA) domain harvesting process (which takes a complete harvest of the .co.nz web domain and targets selected important sites for special treatment) and hope resources appear that will allow integration of that content into the UC CEISMIC archive in years to come; identify organizations and teams with significant local datasets of content that might be able to be made available via UC CEISMIC if progress can be made with the relevant companies; hope that the companies themselves are developing long-term archiving solutions, as is the case with Twitter and the Library of Congress; and make every effort to contact the relevant social media companies to request partnering arrangements. Taming these data ‘decahoses’,⁵⁹ and integrating them into a federated archive to facilitate research, analysis and public memory, is a difficult task. Cornelius Puschmann and Jean Burgess are correct to

note that, despite what many people might think, the “owners” of social media data are the platform providers, not the users. “[W]hile the data in social media platforms is sought after by companies, governments and scientists, the users who produce it have the least degree of control over “their” data”.⁶⁰

While this might not cause many issues for day-to-day use, it has significant implications for cultural heritage and research purposes because it makes it impossible for social media users to effectively gift their content to an archive. When people do make an attempt it soon becomes clear that the only option is to manually download and re-describe images, videos, comments and suchlike, with an attendant loss of essential context. Although the services are fundamentally useful in post-disaster contexts, the “*rhetoric of democratization*”⁶¹ that drives them often falls flat when it comes to preserving their content for posterity and research because the services are oriented towards “findability rather than preservation”.⁶² This issue is of particular concern if one accepts Scott Lash’s contention that our lives have come to be not only mediated by information and technology, but also constituted by them. If this is the case archiving the digital outputs that resulted from the earthquakes takes on a significant moral imperative.⁶³ Despite succeeding in developing a flexible national system capable of archiving a significant snapshot of the digital content associated with the Canterbury earthquakes, a vast amount of social media content is currently not accounted for. It sits in trust with commercial entities whose business drivers are by no means certain to ensure it will be preserved for the long-term. The problem isn’t solely related to the commercial nature of many web services, of course. Rather, it is part of the broader problem of digital preservation. Instant Messages (IMs), emails, Internet Relay Chats (IRC) and other detritus of the so-called “deep web”⁶⁴ are also likely to be available to UC CEISMIC for archiving, but archiving it, displaying it to users, and making it findable poses a different set of problems again. Novel approaches are required to allow archives to store and describe a wide variety of digital formats, that – in the context of the Canterbury

earthquakes – appears to include 3D panoramas, Flash, VRML and QTVR objects, CAD files, GIS mapping data, and LIDAR imagery. As with social media, “[t]he potential of virtual heritage has been hindered as much by people as by diverse technologies, poor provenance, and changing systems.”⁶⁵

14. Conclusion

Although outlier content such as social media and less ubiquitous file types have yet to be archived, the UC CEISMIC system holds great promise as a model for post-disaster digital archiving. The combination of national federated archive and a bespoke archive designed for the ingestion of research-oriented content has proved to be a powerful combination, and cemented a broad community of content providers reaching from local community sites to the largest national archives. The success of the broader programme has resulted directly from its insistence on the use of open source tools and the implementation of open access policies.

Funding

This phase of the UC CEISMIC project was supported by funding from the University of Canterbury, SysDoc, the Canterbury Community Trust, the Christchurch Press, the (London) Christchurch Earthquake Appeal, InternetNZ, and the Malaysian Chamber of Commerce.

References

- ¹ <http://911digitalarchive.org/>.
- ² <http://hurricanearchive.org/>.
- ³ <http://www.jdarchive.org/>; "Digital Humanities Efforts Range from Database Design to New Creations". *Harvard Magazine*. May-Jun 2012. Accessed January 14, 2013.
- ⁴ Eileen McSaveney. 'Historic earthquakes - The 2011 Christchurch earthquake', *Te Ara - the Encyclopedia of New Zealand*, updated 13-Jul-12.
< <http://www.TeAra.govt.nz/en/historic-earthquakes/page-13>>.
- ⁵ Charlie Gates. "Our Disappearing City Centre." *Stuff.co.nz*, September 29, 2012, sec. Christchurch Earthquake 2011. <http://www.stuff.co.nz/the-press/news/christchurch-earthquake-2011/7744317/Our-disappearing-city-centre>.
- ⁶ Weng Y. Kam, Stefano Pampanin, and Ken Elwood. "Seismic Performance of Reinforced Concrete Buildings in the 22 February Christchurch (Lyttelton) Earthquake." *Bulletin of the New Zealand Society for Earthquake Engineering* 44, no. 2 (December 2011): 243.
- ⁷ NZ Government. "Minute of Decision: Christchurch CBD Recovery." NZ Government, April 12, 2012: 4.
- ⁸ Sonia Giovinazzi and Thomas Wilson. "'Recovery of Lifelines' Following the 22nd February 2011 Christchurch Earthquake: Successes and Issues." *2012 NZSEE Conference Proceedings* (2012).
- ⁹ Building Research Advisory New Zealand. "BRANZ Bulletin 551." BRANZ, August 2012.
- ¹⁰ Jarrod Greig. "Christchurch Quake Third Most Expensive Disaster Ever - Insurer." *New Zealand Herald*, March 29, 2012, sec. National.
http://www.nzherald.co.nz/nz/news/article.cfm?c_id=1&objectid=10795342.
- ¹¹ Anna Turner. "11,000 quakes since Sept 2010." *Stuff.co.nz*, January 17, 2013.
<http://www.stuff.co.nz/the-press/news/christchurch-earthquake-2011/8190585/11-000-quakes-since-Sept-2010>.
- ¹² Jack Pinkowski. *Disaster Management Handbook*. Hoboken: CRC, 2008.
- ¹³ Coppola, Damon P. *Introduction to International Disaster Management*. 2nd ed. Burlington: Butterworth-Heinemann, 2010: 9.
- ¹⁴ Georgina H. Endfield, Sarah J. Davies, Isabel Fernández Tejedó, Sarah E. Metcalfe, and Sarah L. O'Hara. "Documenting Disaster: Archival Investigations of Climate, Crisis, and Catastrophe in Colonial Mexico." In *Natural Disasters, Cultural Responses : Case Studies Toward a Global Environmental History*, 305 – 325. Lanham: Lexington Books, 2009: 305.
- ¹⁵ UNESCO. *Text of the Convention for the Safeguarding of Intangible Cultural Heritage*. Paris: UNESCO, October 17, 2003.
- ¹⁶ United Nations International Strategy for Disaster Reduction. "Hyogo Framework for Action (HFA): Building the Resilience of Nations and Communities to Disasters." United Nations, 2005.
- ¹⁷ UNESCO/UBC. *Vancouver Declaration. The Memory of the World in the Digital Age: Digitization and Preservation*. Vancouver: UNESCO/UBC, 2013.
- ¹⁸ The University of Canterbury is the main university serving the broader Christchurch region. Other significant tertiary institutions in the area include Lincoln University and the Christchurch Polytechnic Institute of Technology (CPIT).
- ¹⁹ State of emergency lifted in Christchurch". *3 News*. 1 May 2011. Retrieved 6 May 2011.
<http://www.3news.co.nz/State-of-emergency-lifted-in-Christchurch/tabid/423/articleID/209247/Default.aspx>.

- ²⁰ Rolland, Erik, Raymond A. Patterson, Keith Ward, and Bajis Dodin. "Decision support for disaster management." *Operations Management Research* 3, no. 1–2 (March 2010): 68–79.
- ²¹ Stacey Pitsillides, Janis Jeffries and Martin Conreen, 'Museum of the self and digital death: an emerging curatorial dilemma for digital heritage'. *Heritage and Social Media: Understanding Heritage in a Participatory Culture*. 1st ed. New York, NY: Routledge, 2012: 58.
- ²² Brown, Norman R., Peter J. Lee, Mirna Krslak, Frederick G. Conrad, Tia G. B. Hansen, Jelena Havelka, and John R. Reddon. "Living in History: How War, Terrorism, and Natural Disaster Affect the Organization of Autobiographical Memory." *Psychological Science (Wiley-Blackwell)* 20, no. 4 (April 2009): 399–405.
- ²³ Eviatar Zerubavel. "Social Memories: Steps to a Sociology of the Past." *Qualitative Sociology* 19, no. 3 (September 1, 1996): 283–299.
- ²⁴ Dirk H.R. Spennemann, "Cultural Heritage Conservation During Emergency Management: Luxury or Necessity?" *International Journal of Public Administration* 22, no. 5 (January 1, 1999): 746.
- ²⁵ Daniel P. Aldrich. *Building Resilience : Social Capital in Post-disaster Recovery*. Chicago, IL: University of Chicago Press, 2012. Aldrich defines resilience as (7) "a neighborhood's capacity to weather crises such as disasters and engage in effective and efficient recovery through coordinated efforts and cooperative activities".
- ²⁶ Patrick Meier. "How to Create Resilience Through Big Data." *iRevolution*, January 11, 2013.
- ²⁷ Anthony Oliver-Smith, 'The Centrality of Culture in Post-Disaster Reconstruction'. Goldewijk, Berma Klein, Georg Frerks, and Els van der Plas, eds. *Cultural Emergency in Conflict and Disaster*. Rotterdam: NAI Publishers, 2011: 224.
- ²⁸ Iwana Chronis, Louk de la Rive Box, Eleonore de Merode, 'Cultural Emergency Response: At the Crossroads of Heritage and Humanitarianism'. Goldewijk, Berma Klein, Georg Frerks, and Els van der Plas, eds. *Cultural Emergency in Conflict and Disaster*. Rotterdam: NAI Publishers, 2011: 348. See also Coppola: 407 – 408.
- ²⁹ Sheila A. Brennan, and T. Mills Kelly. "Why Collecting History Online Is Web 1.5." *Roy Rosenzweig Center for History and New Media*. ND. Accessed January 14, 2013.
- ³⁰ 'CEISMIC' initially stood for 'Canterbury Earthquakes Images, Stories and Media Integrated Collection', but the full title was soon dropped in favour of the more easily digestible 'UC CEISMIC Canterbury Earthquakes Digital Archive'.
- ³¹ Michael Forstrom, Nancy Kuhl, Susan Thomas, Jeremy Leighton John, Megan Barnard, Gabriela Redwine, Kate Donovan, Erika Farr, Will Hansen, and Seth Shaw. *Born Digital: Guidance for Donors, Dealers, and Archival Repositories*. Media Commons Press, 2013.
- ³² Ingrid Mason, 'Cultural Information Standards - Political Territory and Rich Rewards', Ed. Fiona Cameron, Sarah Kenderdine. *Theorizing Digital Cultural Heritage: a Critical Discourse*. Media in Transition. Cambridge, Mass: MIT Press, 2007: 231.
- ³³ <http://quakestories.govt.nz>. QuakeStories, along with the Christchurch City Libraries Kete archive (http://ketechristchurch.peoplesnetworknz.info/canterbury_earthquakes_2010_2011), was one of the first post-quake archiving initiatives.
- ³⁴ Gillian Oliver, Brenda Chawner, and Hai Ping Liu. "Implementing Digital Archives: Issues of Trust." *Archival Science* 11, no. 3–4 (November 1, 2011): 313.
- ³⁵ Barwick, Linda. "Turning It All Upside Down . . . Imagining a Distributed Digital Audiovisual Archive." *Literary and Linguistic Computing* 19, no. 3 (September 1, 2004): 254.
- ³⁶ UC CEISMIC Programme Board. *Memorandum of Understanding*. Christchurch, 09 September 2011: 1.
- ³⁷ At the time of writing it seems likely that DigitalNZ will open source their metadata aggregation toolset, making it possible for the entire UC CEISMIC system to be replicated.

-
- ³⁸ <http://www.digitalnz.org/about>. Accessed 14 February 2013.
- ³⁹ <http://natlib.govt.nz/>.
- ⁴⁰ <http://www.digitalnz.org/records?text=%22CEISMIC%22>.
- ⁴¹ <http://www.nvinteractive.co.nz/news/nv-announced-as-winner-in-microsofts-humanising-data-competition>. Accessed 14 February 2013.
- ⁴² Susanna M. Hoffman. 'The Worst of Times, the Best of Times: Toward a Model of Cultural Response to Disaster'. *The Angry Earth: Disaster in Anthropological Perspective*. New York: Routledge, 1999: 143.
- ⁴³ fedora_rest by Don Gourley (https://github.com/dongourley/fedora_rest). The intention is to open source the modifications when time and resource allows.
- ⁴⁴ Consideration was given to using open source CENTOS servers, but as the University uses RHEL it was felt sensible to stay with the supported product. Because of the close relationship between RHEL and CENTOS the feeling is that the current application stack could be open sourced using CENTOS relatively easily.
- ⁴⁵ <http://www.bluefern.canterbury.ac.nz/>. At the time of writing Blue Fern[®] features IBM Blue Gene/L, and IBM Blue Gene/P and an IBM POWER7 cluster.
- ⁴⁶ <http://www.nesi.org.nz/>.
- ⁴⁷ <http://www.karen.net.nz/>.
- ⁴⁸ At the time of writing the UC CEISMIC Research Committee was chaired by Prof. and Dean of Postgraduate Studies Lucy Johnston, and included representatives from the University of Canterbury (including the Human Ethics Committee), Massey University, and Otago University.
- ⁴⁹ Mason, 'Cultural Information Standards': 226.
- ⁵⁰ Alonzo C. Addison, 'The vanishing virtual: Safeguarding heritage's endangered digital record'. Kalay, Yehuda. *New Heritage New Media and Cultural Heritage*. Hoboken: Taylor & Francis, 2007: 35.
- ⁵¹ <http://www.foaf-project.org/>.
- ⁵² Fedora Commons. *The Fedora Digital Object Model*. <http://fedora-commons.org/documentation/3.0b1/userdocs/digitalobjects/objectModel.html>.
- ⁵³ Jason Darwin, 'QuakeStudies Repository: API Documentation'.
- ⁵⁴ This assumption is backed up by research. See Hoffman, 'The Worst of Times, the Best of Times': 149.
- ⁵⁵ Barbara Kirshenblatt-Gimblett. "Intangible Heritage as Metacultural Production." *Museum International* 56, no. 1/2 (May 2004): 54.
- ⁵⁶ Neil Silberman and Margaret Purser, 'Collective Memory as Affirmation: People-centered cultural Heritage in a digital age'. *Heritage and Social Media: Understanding Heritage in a Participatory Culture*. 1st ed. New York, NY: Routledge, 2012: 14.
- ⁵⁷ Ibid.
- ⁵⁸ <https://www.facebook.com/StudentVolunteerArmy>.
- ⁵⁹ Patrick Meier. "Social Media: Pulse of the Planet?" Blog. *iRevolution*, February 2, 2013.
- ⁶⁰ Cornelius Puschmann and Jean Burgess. "The Politics of Twitter Data (January 23, 2013). HIIG Discussion Paper Series No. 2013-01.
- ⁶¹ Beer, David. "Power Through the Algorithm? Participatory Web Cultures and the Technological Unconscious." *New Media & Society* 11, no. 6 (September 1, 2009): 986.
- ⁶² Sheenagh Pietrobruno. "YouTube and the Social Archiving of Intangible Heritage." *New Media & Society* (January 13, 2013): 6.
- ⁶³ Cited in David Beer. "Power Through the Algorithm? Participatory Web Cultures and the Technological Unconscious." *New Media & Society* 11, no. 6 (September 1, 2009): 987.
- ⁶⁴ Michael K. Bergman, "White Paper: The Deep Web: Surfacing Hidden Value." *The Journal of Electronic Publishing* 7, no. 1 (August 2001).

⁶⁵ Alonzo C. Addison, 'The vanishing virtual: Safeguarding heritage's endangered digital record'. Kalay, Yehuda. *New Heritage New Media and Cultural Heritage*. Hoboken: Taylor & Francis, 2007: 37.