# Kete and Koha: Integration Built on Open Standards

by Walter McGinnis with Joann Ransom

# Citation

Walter McGinnis, Joann Ransom, (2010) "Kete and Koha: integration built on open standards", OCLC Systems & Services, Vol. 26 lss: 2, pp.114 - 122

http://www.emeraldinsight.com/journals.htm?articleid=1864123&show=abstract

# **Abstract**

### **Purpose**

This study explores the benefits of using Open Standards to support the gathering of relevant data from a variety of sources from across an organization's software applications and also from completely external information providers.

## Design/methodology/approach

Looks at the history of Koha and Kete projects and their interactions. Features particular concentration on Kete's development.

## **Findings**

Identifies that a focus on supporting Open Standards and general interoperability not only makes it simple to integrate instances of the two software projects, but also creates larger possibilities for data sharing with a number of web sites and services.

## Originality/value

Examines how a piece of software can act as a productive part of a larger information ecosystem.

# Introduction

Kete is the second Open Source Software project started as a collaboration between Katipo Communications and Horowhenua Library Trust. It represents a new way for libraries, and other organizations, to engage their patrons where material contributed by

community members, i.e. user generated content, takes prominence. Kete is a logical companion to Koha, Katipo and HLT's first Open Source collaboration.

One of the guiding principles for the development of Kete was that it should easily share data with other computer programs including other instances of itself. A complimentary design goal was that Kete should be able to reuse data from external software to provide a richer experience for its visitors.

These ideas were largely inspired by HLT's desire to have a site where resources from its Koha catalog were accessible alongside Kete's user generated content; conversely users searching in Koha would be made aware of relevant Kete items. Therefore a main objective for HLT was the free flow of metadata about its collections between the various software that they used to support its community.

Open Standards have played key roles in achieving this objective. Since Koha and Kete use Open Standards fundamentally, this goal was accomplished between the systems with minimal coordination.

# **Context**

#### Koha

Here's how the Koha project describes itself on its About page:

"Koha is the first open-source Integrated Library System (ILS). In use worldwide, its development is steered by a growing community of libraries collaborating to achieve their technology goals. Koha's impressive feature set continues to evolve and expand to meet the needs of its user base."

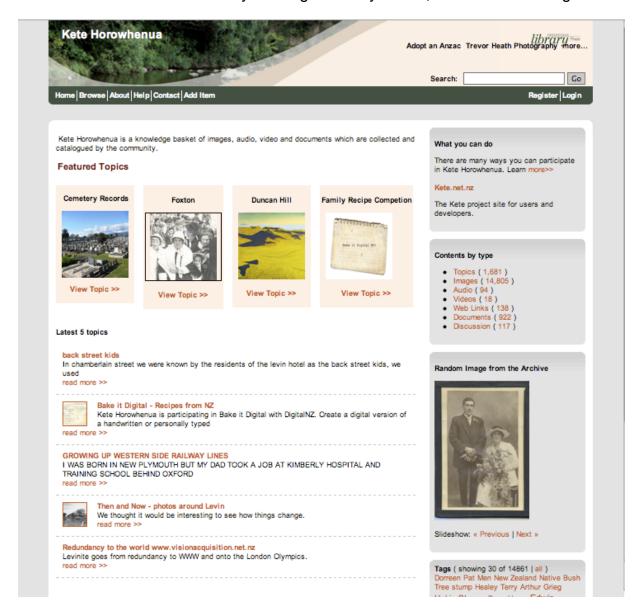
"In use worldwide in libraries of all sizes, Koha is a true enterprise-class ILS with comprehensive functionality including basic or advanced options. Koha includes modules for circulation, cataloging, acquisitions, serials, reserves, patron management, branch relationships, and more. For a comprehensive overview of features visit the Koha feature map."

The Koha project is now ten years old. Most readers will have some familiarity with its history even if they haven't used it themselves. Therefore this article will talk primarily about Kete, the lesser known and younger project.

#### Kete

Kete software project development was started in September 2006 in order to build Kete Horowhenua. Kete is an open source wiki-inspired multimedia web application. Users write topics and upload images, audio, video, and documents. All items can be licensed with the Creative Commons or other copyright licensing options appropriate to a given Kete site. These items then can be discussed and linked together to enrich the site's knowledge and to aid in information discovery.

### Kete Horowhenua is a community built digital library of arts, cultural and heritage



resources. Its mission is to get the private collections, memories and knowledge of the Horowhenua community to sit alongside the HLT's public collections. Many thousands of hours of labour have been contributed to the project by the community resulting in *fully keyword searchable* digital images, audio and video clips, documents, comments and web links. In late 2007 Kete Horowhenua won the *3M Award for Innovation in New Zealand Libraries*, and a *Special Mention for North America and Oceania in the category e-inclusion* at the World Summit Awards in Venice.

Figure 1: The Kete Horowhenua homepage in September, 2009

After Kete Horowhenua launched, Katipo and HLT generalized the software so that it could be used by others and released it as open source under the Gnu Public License

(GPL). It has since been used by a number of community sites and has been chosen by the Aotearoa Peoples Network Kaharoa to support its affiliated library districts throughout New Zealand. The collaborative model has gained traction for arts, culture, and heritage sites as well as broader applications. The Kete community continues to expand.

## **Horowhenua Library Trust**

The Horowhenua District is a one hour drive north of Wellington, the capital city of New Zealand. The district has a medium-sized town, three small towns and a collection of villages and beach settlements. There are libraries in three towns - the district library in Levin, and branch libraries in Foxton and Shannon. Tokomaru has a tiny volunteer library. The libraries are heavily used and well loved by their community, although fewer than half of the regular visitors to Levin library actually borrow material. The libraries function very much as community centers.

While the Trust is a Council Controlled Organization in terms of reporting, it is largely self-determining within the generous confines of its Trust Deed. Being outside the multi-layered government bureaucracy, the Trust is agile in its decision making.

## **Katipo Communications**

A close relationship has developed between Katipo Communications and HLT over several years of IT consultancy. Katipo is an established independent web development firm, based in Wellington, New Zealand. Rachel Hamilton-Williams started the company in September 1996. The company has specialised in open source software solutions since its inception and has enjoyed a string of innovation as a result.

Katipo has continued to grow both in numbers and skills, servicing clients throughout New Zealand and the world. Katipo has an open, friendly, and professional working style. They are always keen to share their knowledge. They tend to become a part of the client's team which enables everyone to get the most out of the process. HLT still works very closely with Katipo, a business relationship spanning over 12 years.

# **History**

Koha was originally developed by Katipo to fulfill HLT's need for a replacement ILS at the end of 1999. Koha has gone on to be used by libraries on all continents and has an international developer community.

By the time work on Kete began the Koha developers use of library oriented open standards in the foundations of Koha was mature and ensured interoperability.

HLT and Katipo began the Kete project under similar a circumstance as Koha. HLT had a need to "scratch its own itch".

#### The Problem:

In 2004 HLT carried out an audit of Arts, Culture, and Heritage resources for Horowhenua District Council, to assess the extent of the resources currently held in the District and the long-term 'safety' of these resources for future generations. The findings were not surprising:

- There is a large amount of material in private hands.
- About half may be given to public collections but half never will.
- · Most of it is available for loan or copying.
- Lots of information is in people's heads.
- Everyone knew someone else with more material and knowledge.
- People really do care about arts, cultural, and heritage resources.
- Physical space is a real issue.

A major concern was the lack of access to arts, cultural, and heritage information; Levin has no museum, archive, art gallery, or public exhibition space.

HLT sat down and talked with a number of focus groups to clarify and confirm the problems and to envisage ideal solutions for each sector: historians, genealogists, artists, students, researchers, librarians, and council staff. HLT needed to work out which problems to address and to come up with an achievable solution. HLT defined the achievable:

- To get public collections accessible by getting them online.
- To get private collections online too.
- To get the stories out of people's heads.
- To include both historical and contemporary material.
- To create a 'virtual' exhibition space for artists and craftspeople.
- To inspire a workforce of volunteers.

#### The Solution:

The solution was to create a community built digital library of arts, culture, and heritage resources: images, video, audio, documents, web-links, encyclopedia-like articles, and discussion threads, with related material clustered together. It would contain both contemporary and historical content. It had to behave very cleverly and yet look simple and intuitive. HLT wanted it to be self-managing and monitoring as much as possible, with no layer of library expertise needed. 'By the people for the people' was the mantra. The community would decide what content they wanted to include and they would be able to upload material in any common file format and be able to describe it with common language. It had to facilitate the generating and strengthening of relationships, not just between items in Kete, but between people as well.

## **Implementation Begins**

HLT managed to source a significant grant and this, combined with donations of cash and kind from the community, enabled the project to get started. Katipo was consulted

early on in the brainstorming process and was again brought in after HLT received funding for the project.

Katipo investigated whether there was existing open source software that could achieve what HLT was after. After determining that no existing repository system had user generated content at its core while still being inclusive of multimedia, HLT and Katipo agreed to create a new open source project.

Katipo then researched what tools should be used to create the project. Ruby on Rails was determined to be the best of breed web application framework in no small part because of its excellent tools for presenting data in multiple ways for different uses.

Joann Ransom, HLT's Kete Horowhenua original Project Manager, described this period of discovery aptly during her VALA 2008 presentation:

"In thinking big first, and dreaming of how Kete could ultimately look and work, we were able to ensure that the Kete core would contain all the necessary scaffolding for future enhancements."

Once the grant got approved, the team reevaluated the proposed solutions for the project, increasing their flexibility to enable them to be easily built upon in the future.

Kete and Koha: Brief Coordination at the Beginning of the Kete Project
A core tenant of Kete from the beginning has been that data should reside with its
logical stewards -- the local community -- at the same time, Kete encourages the reuse
of the information to enrich discussions based on Kete Horowhenua content across
websites. Kete has always been seen as part of a larger data ecosystem. At the very
least a Kete site should be able to share its information with another Kete site.

HLT wanted to make the most of its existing resources, such as its library catalog and other online resources that it has built in the past and requested that data integration be a goal. At this stage an evaluation of relevant standards for data interchange was undertaken by Katipo.

Katipo discussed lessons learned from the Koha project internally and with others in the Koha community. The open source program Zebra was agreed upon as the search engine and Z39.50 service because it's "a high-performance, general-purpose structured text indexing and retrieval engine." It was theorized that querying more than one Zebra database was a way to integrate Kete and Koha thanks to the Z39.50 protocol -- an open standard.

The Kete team selected Z39.50 and Zebra for its search because of Koha's experience with them. Kete, like Koha, would also use a dual database design where an index database was kept synchronized with the web application's RDBMS. However, while Koha had eventually supported MARC as its first class indexing record type, Kete opted for Dublin Core wrapped in OAI-PMH. OAI-PMH DC's purposefully simple and

concentrated scope and its broader application beyond bibliographic records made it appealing. OAI-PMH and Dublin Core as open standards had also grown up to have a number of tools that allowed interchange with MARC and other standards. This easy exchange between data formats would eventually prove useful for Kete and Koha integration.

It should be noted that this was the last time the Kete team and Koha developers directly discussed the implementation details of the Kete software. From here onwards the Kete team concentrated on returning the best software it could for HLT and other clients' needs.

At this point Katipo began looking at OpenURL and at whether Kete should support the standard. It held promise for the ideal of distributed search by defining query parameters. However, back in 2006, there were very few existing software libraries to help implement OpenURL and none in the Ruby language. OpenURL was also a complex standard and it would have required significant effort to conform to its specifications. Additionally there seemed to be only a small number of academic users of OpenURL. For these reasons Katipo ultimately decided that OpenURL would not help Kete achieve its goals of being easily integrated with other systems. OpenURL was a specialist standard, but the Kete team was after more general standards to accomplish its objectives.

At the time of the first release of Kete other feature priorities and HLT's needing need to upgrade to Koha 3 to enable Kete to be searchable from Koha resulted in Kete/Koha integration being put on the back burner rather than being delivered. However, the building blocks based on open standards were in place and integration features were steadily added to the Kete software in subsequent releases.

Early on, HLT also identified national collections that they wanted to draw information from. This was designated as a long term goal; thus accessing other databases was put aside for the initial version of Kete, but always kept in mind when considering the future of the software.

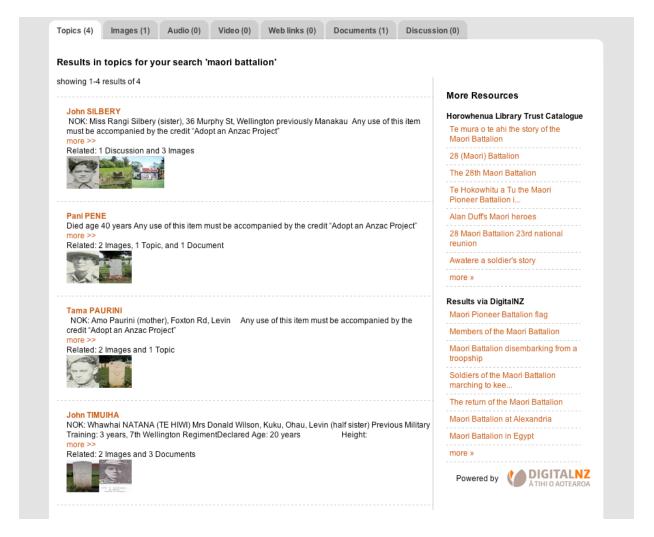
#### **RSS/Atom Feeds and OAI-PMH**

To provide updates to users, the Kete team settled on using RSS rather than conventional, but not standard, email notifications. RSS would turn out to be an excellent way to allow for Kete data reuse with minimal effort. The fact that its core elements dovetailed nicely with Dublin Core was an indication that it was going to be easy to use to redistribute Kete data with. The Ruby on Rails framework also supplied easy-to-develop ways of quickly putting together RSS feeds that matched any number of criteria a user might want.

Even before the official launch of Kete Horowhenua, the project was well known in the culture and heritage sector in New Zealand. Joann Ransom spoke about the development of the project at the 2006 National Digital Forum. As a result of this

presentation and through HLT and Katipo's contacts, a number of like-minded projects approached HLT and Katipo to see if Kete could be used for their purposes.

With other instances of Kete sites likely, interest in sharing information between the various Kete sites again came to the fore. National Library New Zealand expressed a desire to use Kete for communities across the country through the Aotearoa Peoples Network Kaharoa. One requirement was that all data could be harvested from each Kete for aggregation in the National Library's Digital New Zealand service.



As a result, Katipo, using the Ruby language's open source OAI software library, put the finishing touches on Kete's use of OAI-PMH so that each Kete site could have an OAI-PMH repository built-in. The Digital New Zealand service then could regularly and efficiently harvest each participating Kete site, thus building up the metadata records that users could search via the DNZ search site or through external services using the DNZ API.

Katipo noted that the DNZ service provided a simple RSS version of search results from the API in addition to results in an API specific XML schema output.

Figure 2: a search for "maori battalion" on Kete Horowhenua showing results from Kete Horowhenua along side results from the library's Koha catalog and the Digital New Zealand service under "More Resources"

During this period, the Chinese Association of New Zealand Auckland Branch and Auckland City Libraries approached Katipo about using Kete for a site to be called Chinese Digital Community. Auckland City Libraries and the team working with them at McGovern Online identified the need for "Contextual Search" as a new feature.

Contextual Search was to pull in relevant content from external sites and electronic services that fit the particular subject the user was viewing. Thus if someone viewed a page in the CDC titled "From Immigrant to New Zealander" the viewer would also get information from the Auckland City Libraries' Heritage Collection, or any number of sources that a site administrator could set up, that matched.

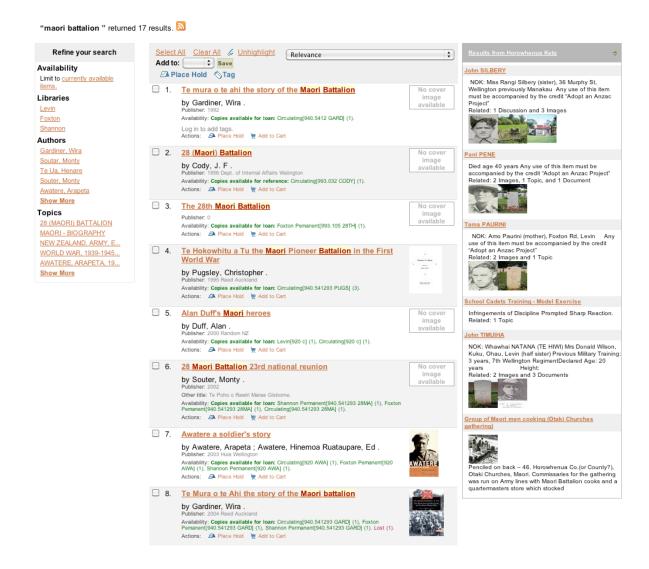
The concept for Contextual Search was similar to HLT's original idea for integrating catalog results from Koha into a Kete search result page, but with the search terms predetermined for the user by being on a particular page rather than being entered in the search input. It would turn out that both features would be implemented and they would make use of the same underlying standards -- RSS and Atom -- to retrieve their results.

While these ways of pulling in external content into Kete were being explored, it was discovered that Kete's fine-grained RSS support could be used as a simple and flexible way to showcase a Kete site's items in popular social networks like Facebook by using common RSS widgets.

It was clear that distributing Kete content via RSS was a powerful tool. The Kete team also observed the wide adoption of RSS across content sources, such as online institutional collections, and realized that being able to pull in metadata in RSS format would provide the most rapid way to "grab" information from a diverse and abundant set of sites.

A large percentage of these web site search interfaces followed two patterns; they provided a dynamically generated RSS feed of results and they had URLs that could be easily utilized as the basis of dynamic searches by appending the search terms at the end of the URL; thus, they had predictable input and predictable output that Kete could exploit programmatically.

**Note:** there is an open standard called OpenSearch that describes this pattern. Many sites adhere to OpenSearch's principles, and are largely interoperable with OpenSearch sites, without strictly following the full standard. Such was the case with Kete. However, Kete will be adding formal OpenSearch support in its 1.3 release. Kete will thus gain some of the more advanced features that OpenSearch allows for.



It became clear that Kete, by having functionality implemented around the RSS/Atom open standards, could expand the early model of a federation of Kete sites to include non-Kete sites, too. It turned out that Koha followed this pattern as well. Integrating with a Koha site was as easy as tapping into its search interface's RSS feed.

In a parallel fashion, Koha has recently implemented a feature similar Kete's external search sources and its catalog search can now be configured to display a side bar of results from outside services along side its own.

Figure 3: a search for "maori battalion" on Horowhenua Library Trust's Koha catalog showing results from Koha along side "Results from Horowhenua Kete".

In addition, by providing a Z39.50 interface to its OAI-PMH DC records, Kete's data could be accessed from Koha searches using the companion pazpar2 federated searching software.

# Conclusion

As Katipo and HLT's second collaboration after Koha, Kete benefited from a number of lessons learned in writing Open Source Software. The most important lesson perhaps was that Kete should use Open Standards where possible.

As a new way for libraries to engage their patrons in a collaborative manner, the Kete project was a logical companion to Koha. Sharing information between the two programs made sense. Like many organizations, HLT wanted to make the most of their data by creating as many relevant pathways to items in their collections as possible. Kete and Koha have both evolved towards being a part of an ensemble of software tools that can work together by using Open Standards rather than only providing their own data exclusively in their own context.

In the end Koha and Kete's focus on supporting Open Standards and general interoperability not only makes it simple to integrate instances of the two software projects, but also creates larger possibilities for data sharing with a number of web sites and services.

### **Biographies**

#### **Walter McGinnis**

Walter McGinnis has been an open source and online community developer for more than a decade. He has contributed to the ArsDigita Community System, OpenACS, and numerous other software projects.

In 2005 he moved to Wellington, New Zealand from the United States. That year he also began working for Katipo Communications. He is the lead technical developer for the Kete open source project and has been since its inception. He currently specializes in Ruby and Ruby on Rails applications, especially those with emphasis on community based functionality.

#### Joann Ransom

Joann Ransom is currently Deputy Head of Libraries at Horowhenua Library Trust, Levin, New Zealand. She has worked on two award winning open source projects: as one of the team who developed Koha back in 1999-2000, and more recently leading development of Kete <a href="http://kete.net.nz">http://kete.net.nz</a>, an open source project for building a community digital archive of articles, images, audio, video and documents. Joann is a professional librarian with qualifications in English literature and computer network administration. She regularly blogs <a href="http://library-matters.blogspot.com/">http://library-matters.blogspot.com/</a> and tweets jransom <a href="http://twitter.com/jransom">http://twitter.com/jransom</a>, and leads a quiet life in a small coastal village with her menagerie of children and pets.