

Liquid Mirò: Semantic Softlinking to Support Cooperative Document Exploration

Aaron Quigley

Nathan Lee

Smart Internet Technology Research Group
School of Information Technologies
Sydney University, Australia

CustomWare
200 George St
Sydney, Australia

aquigley@it.usyd.edu.au

nathan@cutomware.net

Abstract

In this position paper we present a non-intrusive mechanism for evolving the overall quality of semantic relationships between elements of information in hypermedia document systems. The evolutionary aspect of this work is an application framework that includes a combination of hard links and temporal soft links between existing documents. A hard link completes the binding between two documents in the system upon creation, whereas a soft link delays the binding until some later time. The ability to monitor, weight, integrate, delay, and then order the soft links is what offers the power in Liquid Mirò document systems. Here we focus on the use of hypermedia document systems which support existing online communities, ranging from social to professional groups..

Keywords Personalised Documents, Document Management

1 Introduction

On-line distributed hypermedia environments for targeted communities, continue to evolve at an unfettered pace across the web. In support of these growing environments, numerous navigation and exploration mechanisms have been employed. These mechanisms include, structuring the information in a hierarchical manner, site-maps, navigational patterns, link previews, textual and visual search engines and associated expressive query languages, end-user centric structure and review, along with expert manual intervention. Due to the volume of information, and the speed at which it both changes and grows, any structuring mechanism incorporating repeated human intervention becomes less efficient [1,2,7] and hence less useful.

Here we present a novel community based mechanism, to improve the quality of the overall hyper-linking structure, within a given community using hypermedia documents. This mechanism is

further encompassed in the *Liquid Mirò* web application framework to couple existing document repositories to emerging client device types, both desktop and mobile. The goal is to allow specific communities to evolve abstract, un-predicted semantic associations, while presenting the information in a fluid form that can be interacted with via any class of device type.

Online community environments range from governments providing logistical information and support, to educational institutions providing distance education support to both educators and students. The typical model is for the few owners of the online environment to publish information as hypermedia documents which are then read by regular users of the environment.

In contrast to this, numerous on-line portals offer features, which more closely mirror the notion of a community as a group with "common rights and privileges". On-line portals that offer community-building features include Yahoo groups and CNN interactive communities. Due to financial considerations such portals are turning to increasingly sophisticated software and interaction paradigms to reduce the manual effort involved in supporting, maintaining, and growing such on-line communities.

This paper presents an application framework for supporting communities using a variety of desktop and mobile access devices, as shown in Figure 1. This framework incorporates all the navigation and exploration mechanisms noted, along with the evolutionary formation and structuring of semantically relevant soft links. Instead of relying on manual effort to create lists or relationships we propose to let the actual usage of the system create and evolve the inter-connections between elements of information.

The rest of this paper is organised as follows, Section 1.1 describes the background of soft linking and semantic web structuring. Section 1.2 describes related work. Section 2 describes the Liquid Mirò framework. Section 3 outlines four possible Liquid Mirò web applications.

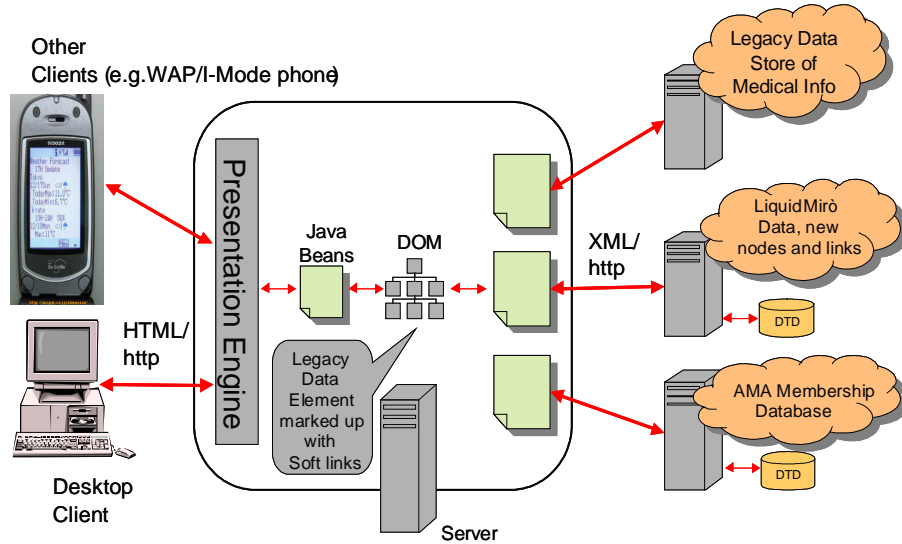


Figure 1: Example Liquid Mirò application overview for coupling a large legacy data source, data from a professional user organization with semantically relevant linking

2 Background

Many online communities are now centred around hypermedia systems, which superimpose an external structure on data [2,6]. Experience has shown that large-scale hypertext systems requiring extensive manual support are costly, error prone, and difficult to maintain [1,2,7]. These problems stem from the fact that even with support tools, maintenance is a laborious and often painstaking task.

The concept of hypertext *soft linking* between elements of information is not new and has been explored in file system research [8], software engineering [1], digital libraries [3], video streams [5], conceptual hypermedia [6], via dedicated services [3,4], and in online services such as Everything2. In the Liquid Mirò application framework we present a more generalised and powerful soft linking mechanism. Liquid Mirò based applications include the ability to monitor, weight, integrate, delay, and order the soft links before they are incorporated into elements of information.

Figure 2 indicates a soft link formation and

ordering, between non-connected legacy information elements that is typical of Liquid Mirò applications. Here a registered user starts exploring the information at node X via existing structures, search engines, hard links or other mechanisms. Once the user arrives at another elements of information (a,b,c), Liquid Mirò creates a bi-directional soft link between the now semantically related elements. Based on other users usage patterns, the relative *strengths* of these links change. Depending on the application and user type, if other users don't follow such soft links, they eventually disappear, to be replaced by new soft links.

Dedicated soft link services provide a remote link lookup facility [3]. These services rely on augmenting web documents with soft links prior to delivery to the client. The end point of these links are determined via a dynamic search mechanism when the link is followed. The remote link service provides the end point of the link i.e. it determines the "late-binding" of the link.

Soft link structures such as conceptual indices, are automatically formulated based on an analysis of the overall hypermedia structure. Concept indices provide a temporal cross-linking mechanism, which represent

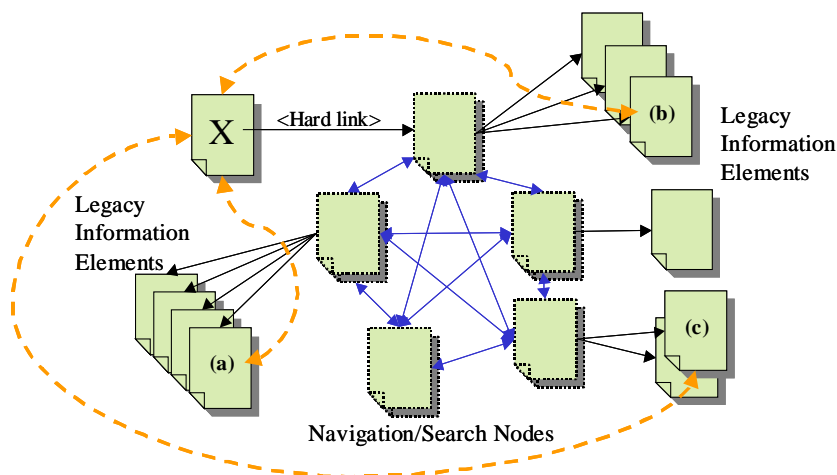


Figure 2: Example semantic soft link creation by user navigation

semantic relationships rather than pre-determined links.

2.1 Related Work

Slashdot is a prototypical online community, which is focussed on “*News for Nerds and Stuff*”. This site allows its members to submit stories that are effectively moderated before they are posted on the web site. In the long term, a regular user can become a moderator. A variety of factors go into deciding which users become moderators and for how long. The end result is a pool of eligible moderators that represent the average positive Slashdot contributor. This ad-hoc approach has many limitations and still requires a great deal of manual intervention on behalf of the moderators.

Everything2 is social experiment, which is trying to form a large peer reviewed information web. Everything2 incorporates a soft linking scheme and aims to allow users to read, write, and share quality information. Within Everything2 it would not be possible to create a special purpose e-community to cater for example, to medical professionals or educators with specific needs. The vast quantity of irrelevant, incorrect and ambiguous information would soon render such a system useless. In contrast Liquid Mirò applications incorporate soft linking with an *existing* high quality document repository Systems based on this approach also allow registered users with different levels of access, and makes the information easily accessible across device types.

3 Liquid Mirò

The Liquid Mirò application framework is built around a mechanism of extending existing legacy data to incorporate a dynamic record of usage-based semantic associations within the data. The first step is

to evaluate which existing data sources are suitable for integration. The basic criteria are that the legacy information elements can be embedded in a hypermedia system. Often the legacy data is stored in a relational database or flat file system. In the case of legacy data in a special format it can be fully re-engineered into a presentation independent format such as XML, or wrappers can be developed to access the data. Once each element of information can be uniquely distinguished by the system, then the information can be incorporated into a Liquid Mirò application.

Next a soft linking scheme must be decided upon to seed the legacy information units with soft links. Textual information is suitable for such a soft linking treatment, with soft link selection based on partitioning along words or phrases using a simple scheme. For a more sophisticated approach it is possible to integrate with one or more legacy information sources to determine where the seeded soft links may be found.

Development of a mechanism for the presentation of soft links and tracking of their usage requires knowledge of the presentation medium. This part of the Liquid Mirò framework is responsible for altering the community usage information associated with the existing legacy data. The emphasis is on creating a mechanism that enables the users to find the most appropriate links by themselves. The seeding process is merely to enable the most useful semantic associations to be discovered quickly. This mechanism permits a user viewing a particular piece of information to travel to what others or the system regards as the next relevant piece of information. This traversal creates a soft link if not already in existence or strengthens an existing soft link. The strength of the soft link can be represented in a number of ways by simple ordering, colour, size, or the use of glyphs.

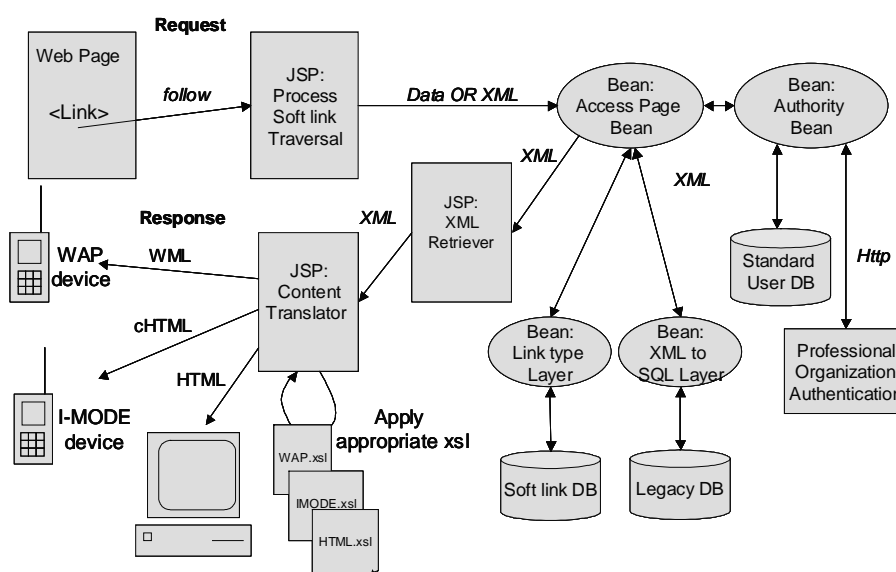
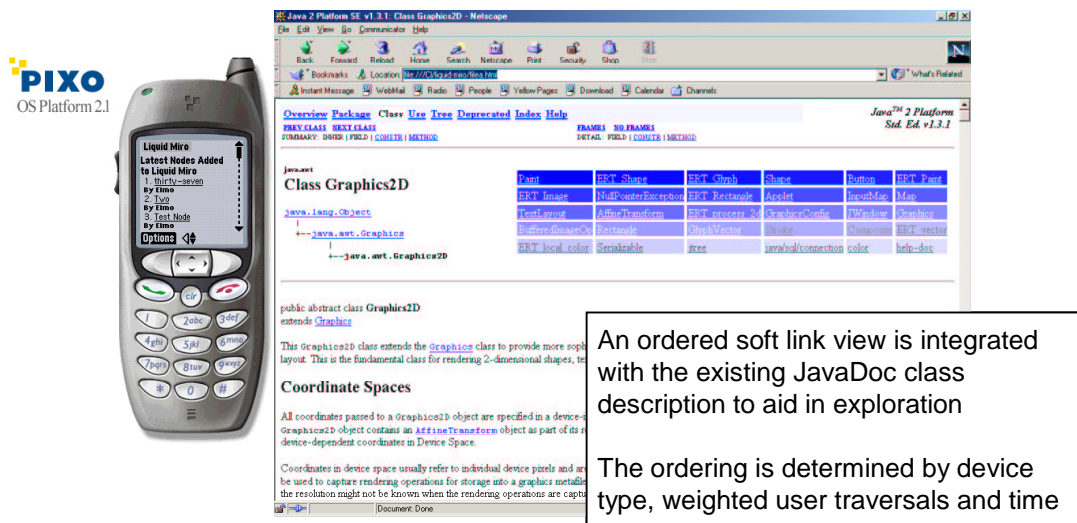


Figure 3: Liquid Mirò implementation (using JSP) supporting multiple client types, with client specific soft linking, internal and external user levels coupled to an existing information web

For more sophisticated implementations of Liquid Mirò, particularly in community based applications, a

creating a community determined “most relevant” set of associations.



weighting of both the influence a particular community member has and of the relative importance of the Soft Links when generating the presentation format might be integrated into the tracking and presentation mechanism. This can be achieved by integrating a separate user modelling system. For example, senior members of a community or organization might have far greater influence over junior members in terms of the affect their actions have upon the soft linking within the system. In terms of influence over presentation of soft links, weighting of soft links frequently traversed by community members using a similar type of browser might be higher than those used by members viewing with another type of browser (small screen mobile devices versus traditional web browsers). Figure 3 shows an example application permitting assignment of priority via integration with an authentication server and through separation of presentation logic from the main application through the use of XML and XSL to easily tailor the output format. The emphasis is on determination of soft links at request time rather than *a-priori*.

The example software engineering application of Liquid Mirò shown in Figure 4 demonstrates how an existing information service (the automatically generated API documentation of applications developed in the Java programming language) can be enhanced with the addition of soft links. The standard JavaDocs are a set of HTML pages that provide information on the software application. This takes the form of a page for each class in the system with in-line hyperlinks to any associated classes. With a Liquid Mirò framework built around this information source, the users of this system will reveal further relationships between the classes that are then available for use by others. The soft links are created as the software engineers navigate the documentation,

Acknowledgements

Particular thanks to Brett Anderson, Kersten Fernandes, David Hemingway, John Hinton, and Jeffrey Pond all whom allowed us to use screen grabs of their respective Liquid Mirò applications.

References

- [1] Oli Kai Paulus, *et. al.*, *Adding Softlinks to the Web*, 8th Joint European Networking Conference, <http://www.terena.nl/conf/jenc8/proceedings.html>
- [2] Monika R. Henzinger, *Hyperlink analysis for the web*. IEEE Internet Computing, 5(1) pp. 45-50, January/February 2001.
- [3] Leslie Carr, *et. al.*, *The distributed link service: A tool for publishers, authors, and readers*. In Proc. Fourth International World Wide Web Conference. O'Reilly Associates, 1995.
- [4] Eytan Adar and Jeremy Hylton, *On-the-fly Hyperlink Creation for Page Images*, in Proceedings of the Second Annual Conference on the Theory and Practice of Digital Libraries, 1995
- [5] Jason W. Smith, *et. al.*, *An orthogonal taxonomy for hyperlink anchor generation in video streams using ovaltime*. in Proceedings of Hypertext 2000, pp. 11-18.
- [6] Mark d'Inverno and Michael Hu, *A Z Specification of the Soft link Hypertext model*, In ZUM'97: Proceedings of the Tenth International Conference on Z for Users Bowen, Hinchey and Till (eds.), Lecture Notes in CS, 1212, 225-240, Springer-Verlag, 1997.
- [7] Leslie Carr, *et. al.*, *Conceptual Linking: Ontology-based Open Hypermedia*, In Proceedings of WWW10, May 2-5, 2001, Hong Kong.