

## Modelling the Library Domain Using the e-Framework and addressing Library 2.0 challenges

March 2009

Web 2.0 applications offer opportunities to use library resources in learning, teaching and research in new ways and across traditional boundaries. These include enabling users to integrate digital library resources within Web environments of their choice, to share them and to personalise them.

Exploiting these opportunities poses technical and other challenges for libraries. The JISC and SCONUL Library Management Systems (LMS) Horizon Scan report<sup>1</sup> indicated that it was timely to explore current practice and to review library business requirements in the changing landscape. And indeed the JISC and British Library Google Generation report<sup>2</sup> also supported the fact that libraries needed to be more responsive to user requirements in order to supply added value services. The JISC work on personalisation<sup>3</sup> interfaces with some of these issues too.

To progress these issues, the JISC commissioned a follow-on project in 2008: Towards Implementation of Library 2.0 and the E-Framework (TILE). The TILE project<sup>4</sup> worked with libraries to:

- Capture information on the current scope and scale of Library 2.0 – library uses of Web 2.0 applications and services
- Identify the most significant challenges ('pain points') facing library systems development in this context, documenting them within the International e-Framework<sup>5</sup>
- Propose a high level library domain model that positions these challenges in the context of library business processes and assists library partners towards a shared articulation of service objectives and functions

The project produced three major deliverables: a domain model, an outline architectural proposal for aggregation and sharing of library usage and other service data and a Library 2.0 Survey Report.

1 [www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/libraryms](http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/libraryms)

2 [www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/googlegen](http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/googlegen)

3 Developing Personalisation for the Information Environment  
[www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/dpie1](http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/dpie1)

4 [www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/tile](http://www.jisc.ac.uk/whatwedo/programmes/resourcediscovery/tile) or <http://misc.jisc.ac.uk/tile>

5 [www.e-framework.org](http://www.e-framework.org)

### A Library Domain Model

#### Rationale

In a landscape transformed by pervasive internet access and emerging Web 2.0 practices, the library community needs a shared framework for articulating its purposes and visioning the domain in the context of the networked information economy and its workflows.

The value of such a framework was articulated by Lorcan Dempsey, Geneva Henry and Brian Lavoie in 2006 as 'a means of focusing attention and organizing discussion ... It does not supply the answers, but facilitates the process by which answers are sought, found, and applied.'<sup>6</sup>

#### Response

The TILE domain model offers a step in that direction as well as a context for framing the observed Web 2.0 challenges or pain points of how to tailor content to different types of user (deriving context) and how to enable and encourage user generated content (enabling contribution).

The TILE model (figure 1) defines library domain business processes in terms of the interactions within and between three realms:

#### Library Domain Ecosystem

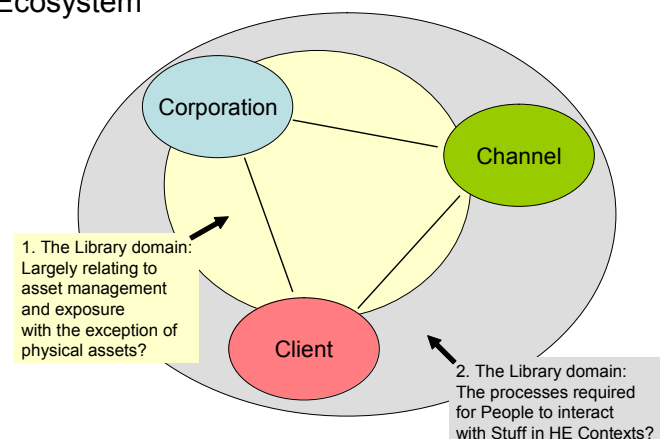


Figure 1: Scoping the Library Domain

6 Lavoie, B., Geneva, H., and Dempsey, L. (2006) A Service Framework for Libraries. D-Lib Magazine 12, [7/8]  
[www.dlib.org/dlib/july06/lavoie/07lavoie.html](http://www.dlib.org/dlib/july06/lavoie/07lavoie.html)

# Modelling the Library Domain

March 2009

## TILE & HE Services

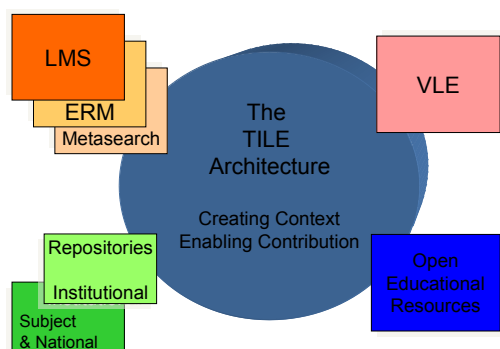


Figure 2: Concentrating HE institutional intelligence

**The Corporation:** organisations involved in the administration of knowledge assets (eg originals, copies, licences, metadata) and/or specific groups of clients (eg student records). Corporations within the library domain include all universities, some JISC services, national libraries and publishers.

**The Channel:** a means of delivering knowledge assets to Clients, not necessarily restricted to the holdings or the Clients of a particular Corporation. Channels within this model include local online public access catalogues (OPACs), JISC services and web-scale services such as Amazon and Google Scholar.

**The Client:** an individual (student, researcher, librarian, business or community partner) accessing the information environment (libraries, virtual learning environments (VLEs), repositories and the wider web world) in the context of academically motivated workflows.

## Choices

The model highlights two potentially divergent views of library business:

- **Narrowly defined** as those functions relating to the management of assets owned, curated or licensed by a library and their exposure to delivery channels. Involvement as a delivery Channel relates largely to physical assets, notably circulation and inter-lending of books and licensed electronic publications.
- **Widely defined** as the total set of Corporation, Channel and Client processes required for people to interact with 'stuff' (content, metadata, reading lists, feedback) in a higher education context.

Strengths, weaknesses, opportunities and threats exist for both perspectives:

The narrower role avoids libraries building local channels that can be more efficiently and effectively provided by the wider world (eg social services requiring network effect). However, it is dependent on the existence of appropriate

third party Channels that serve its clients (eg Copac, Serials Union Catalogue for the UK (SUNCAT), Google Scholar).

The wider role empowers libraries to provide a full set of services to meet a rich variety of locally identified user needs, potentially an institutional unique selling point. However, the downside lies in services lacking network effect and in the cost of maintenance.

TILE reinforces the LMS Study recommendation that institutional HE services (libraries, VLEs, repositories and associated content services) should excel in the narrower role in terms of metadata quality, accessible and addressable content whilst selectively fulfilling the wider role through using web-scale (global, national and consortium) services and exposing local data and services to the best Channels.

## Generating the network effect

The LMS Study emphasized the 'concentration' of users and data in web-scale Channels to generate network effect and to best serve students and researchers in a Web 2.0 world. Furthermore, this approach may generate a virtuous cycle in which library services with critical mass attract value added activity, such as recommendations and other forms of user generated content.

## Getting started

As a starting point, TILE investigations suggest there is much to be gained from the combined intelligence uniquely available to the HE community about the users of its services – notably by relating circulation and download transactions to the user's subject or course context.

Figure 2 suggests services that potentially contain the patterns of user activity. Some may already capture such activity (LMS, electronic resource management (ERM), VLE) whilst others have the potential to do so (repositories, Open Educational Resources services). The activity data sets from such systems can be logically linked through student registration data (year, UCAS code, course, module options) without compromising personal data, thus containing opportunity for aggregation of client behaviours at consortium, national or wider levels.

Based on these considerations, TILE proposes an architectural approach to concentrate the user activity intelligence. From day one, this could have critical mass of activity data from LMS circulation audit trails alone.

Figure 3 suggests there may be value added intelligence brought by some users, for example interest keywords and bookmarks. However, captured click streams rather than volunteered contributions (eg ratings, reviews, lists) are the surest source of intelligence about 'users like me'.

The running start afforded by activity data would more realistically open the opportunity to engage and curate user contribution (ratings, reviews). Furthermore, there is potential to embrace national services (consider Intute, Copac and the potential of Open Education Resources) so long as they collect activity data and have knowledge of users. The DNA of user 'academic affiliation' is the silver bullet, the knowledge about their subject, level and place of study (not name or address) which should run through all activity records.

### User Benefits

Figure 4 illustrates ways in which concentrated intelligence about user behaviour can improve system use by individual users.

Assuming that issues of quality and control can be smartly addressed, this architecture for aggregation can therefore become an architecture for participation. This would enable user contributions ranging from recommendations to tagging, from traditional publications to user generated content, from books in libraries to pages in Wikipedia.

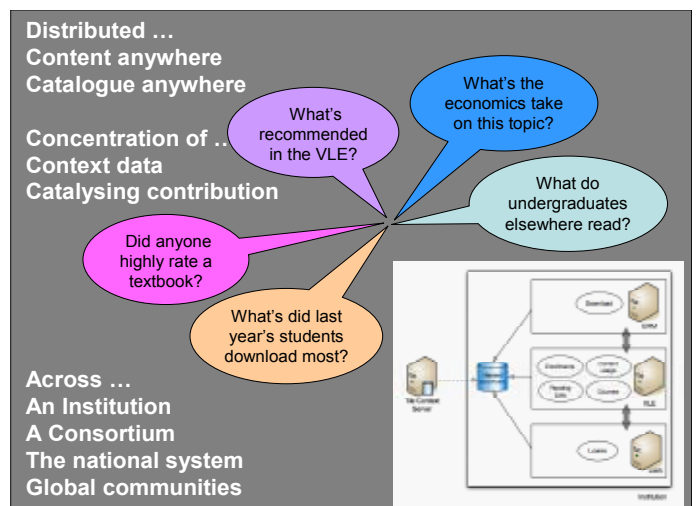


Figure 4: Benefits derived from concentrated intelligence

### Implementation

These opportunities to support students, researchers and lecturers inevitably contain challenges in terms of technology (scale, domain vendor systems, etc), ownership and protection of data, quality of metadata, and the desired levels of aggregation and value relative to global services such as Amazon and Google Scholar. For large undergraduate subject areas (such as English), benefits may accrue at a local or consortium level, whereas others (such as Anthropology) may require national or even global aggregation to generate critical mass.

Clearly, the business case for such applications is unproven. How will such capabilities address the motivations of students, lecturers or researchers? Will they serve the objectives of HE, not least the promotion of individual enquiry?

A Web 2.0 style development strategy is essential to assess sample data (bad data will not serve even enthusiastic users) and user opinion on the back of rapidly prototyped implementations (good data and especially Open Data may generate unforeseen applications, empowering innovative workflows, learnflows and researchflows).

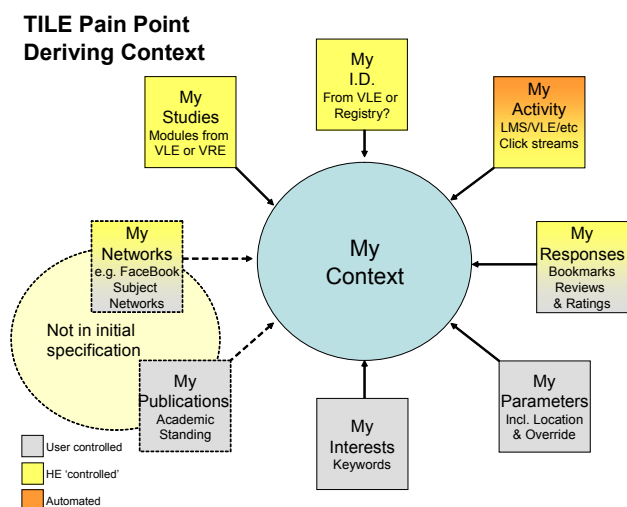


Figure 3: Deriving user context from systems activity

### Proof of concept

The University of Huddersfield broke significant ground in December 2008 by liberating its library user activity data, providing 'title' level user behaviours linked to courses. Dave Pattern emphasised the motivations and opportunities:

*At the time of writing, the library has details of just under 3 million circulation transactions spanning a period of 13 years. Mining this data has proved both beneficial to our students ... and to the library ... Involvement with the JISC TILE Project led to a decision to release a sizeable portion of the usage data in the hope that it might prove beneficial to others ... to mine the data and to exploit it, stimulate a discussion about the value of library usage data and promote collaboration and sharing of such data in order to benefit both libraries and their users<sup>1</sup>.*

<sup>1</sup> [http://library.hud.ac.uk/data/usagedata/\\_readme.html](http://library.hud.ac.uk/data/usagedata/_readme.html)

# Modelling the Library Domain

March 2009

---

Delegates at the final TILE workshop ('Sitting on a Goldmine', December 2008) agreed the time to test these possibilities is now, and recommended:

- Context data to be exposed to test interest and viability
- A mix of trials to be undertaken, ranging from national services to individual higher education institutions
- Triangulation of outcomes with both professional opinion (eg validity of search results from subject librarian perspective) and technical service considerations.

## TILE Recommendations

### For the Corporation

#### People told us

*'In many cases, library services are really only a source, not a target for users'*

*'The university knows more about "context" than anyone else'*

*'Analysis of aggregated user behaviour should deliver many opportunities'*

**Business options:** review business options for the development of context/activity based services, including an Open Data approach to expose context data so both community and commercial providers can develop services.

**Activity data:** prioritise the requirement to collect activity data (eg downloads, views, searches) in system procurements and projects. Furthermore support the community in establishing a common format and negotiating such facilities with services and vendors.

**Leveraging systems:** work with vendors and libraries to share experience, practice and knowledge in order to maximise existing systems as sources for Web 2.0/Library 2.0 services.

### For Channel Services

#### People told us

*'Back in the days of the e-lib programme libraries were leading, now they are playing catch up'*

*'Users are going to Google et al, so we need to put library services into the places where users actually are'*

*'It seems the main factor is the network effects generated by the major data hubs ... their value derives from successfully driving those network effects through wide participation, from consolidation of data and from mobilizing usage data to improve their services'*

**Context service development:** weigh up the options for medium term investment and strategic partnerships in aggregated contexts and contribution services.

**Encourage 'concentration':** the TILE Library 2.0 survey identified significant use of Web 2.0 approaches for 'diffusion' (mostly blogs, wikis, podcasts and RSS); however there was a lack of 'concentration' (providing or contributing to web-scale services to maximise network effect). JISC and its partners can back services and projects that take on this challenge and can help develop awareness of the possibilities.

**Know trusted channels:** the community needs to understand the business and quality criteria that guide the selection of trusted external Channels that might be widely leveraged to serve HE resource discovery and delivery.

### For Clients

#### People told us

*'Learners are using tools they already have rather than the institutional tools they've been given'*

*'Library services enable users to find, locate and access in context of another role, task or activity that the user is performing'*

**Demand side research:** conduct research to ascertain whether library and learning service users would be inclined to participate in the services proposed by TILE, identifying scenarios that will motivate user contribution.

**Library 2.0 evaluation:** the TILE survey indicated that work needs to be done both at the institutional level and across the sector on evaluating how current Web 2.0/Library 2.0 services are meeting their objectives.

**Web 2.0 and search:** Web 2.0 raises the bar for user expectations of and behaviours towards search and discovery to delivery (D2D) processes – in terms of inputs (search terms, system intelligence), outputs (delivery) and ease of use (interface, workflows). Services and projects should explicitly be tasked with these challenges.

**Re-use of services:** Librarians need to be empowered to re-use services at 'the point of need', which implies liberation of library data to be re-used in ways not predetermined by the library, the service or the system vendor.

This project overview was compiled by David Kay from Sero Consulting Ltd.

Alternative formats of this briefing paper can be found at:  
[www.jisc.ac.uk/publications](http://www.jisc.ac.uk/publications)