Investigating the Effects of Exploratory Semantic Search on the Use of a Museum Archive

Conference or Workshop Item

Investigating the Effects of Exploratory Semantic Search on the Use of a Museum Archive

P. Mulholland, Z. Zdrahal and T. Collins
Knowledge Media Institute and Centre for Research in Computing
The Open University, Milton Keynes, MK7 6AA, UK
Email: {P.Mulholland|Z.Zdrahal|T.D.Collins}@open.ac.uk

Abstract— **Recently, there has been a great deal of interest in how new technologies can support the more effective use of online museum content. Two particularly relevant developments are exploratory search and semantic web technologies. Exploratory search tools support a more undirected and serendipitous interaction with the content. Semantic web technology, when applied in this context, allows the exploitation of metadata and ontologies to provide more intelligent support for user interaction.**

Bletchley Park Text is a museum web application supporting a semantic driven, exploratory approach to the search and navigation of digital museum resources. Bletchley Park Text uses semantics to organise selected content (i.e. stories) into a number of composite pages that illustrate conceptual patterns in the content, and from which the content itself can be accessed.

The use made of Bletchley Park Text over an eight month period was analysed in order to understand the kinds of trajectories across the available resources that users could make with such a system. The results identified two distinct strategies of exploratory search. A risky strategy was characterised as incorporating: conceptual jumps between successive queries, a larger number of shorter queries and the use of the stories themselves to acclimatise to a new set of search results. A cautious strategy was characterised as incorporating: small conceptual shifts between queries, a smaller number of longer queries and the use of composite pages to acclimatise to a set of new search results. These findings have implications for the intelligent scaffolding of exploratory search.

Index Terms— **Exploratory search, semantic web, museum informatics, user studies**

I. Introduction

Exploratory search (White et al 2006) unlike conventional search is concerned with the more undirected retrieval and use of resources. This can be contrasted with conventional search contexts where the user has a specific goal or query in mind, and would ideally wish to identify a single digital resource that satisfies this goal.

Museum contexts provide an excellent testbed for exploratory search, where the user is finding out about a domain but not necessarily searching for specific knowledge. The user may also have little prior knowledge of the vocabulary of the domain (key people, events, locations, objects) that would generally be required for the formulation of queries.

The museum sector has also provided a key application domain for semantic web technology. This is no doubt to some extent because museum collections often have a rich source of metadata that could potentially be used to provide more intelligent and scaffolded means of search and exploration. Semantic web technology has also been found to support interoperability across collections from different museums [2].

Exploratory semantic search tools have the potential to facilitate approaches to content and use that sit between conventional search and browsing paradigms. First, they support the presentation of search results in meaningful, browsable structures rather than as a list. Second, the browsing of the structure can motivate, and in interaction terms is closely associated with, the specification of new queries.

This form of interaction with content is illustrated in figure 1 in which the user is constantly switching between the selection of new content and its navigation within provided structures. If this form of interaction is facilitated, then this begs the question as to what kinds of trajectory users will make across the available content with such exploratory search tools. For example, users may jump around the available space of content or successive queries may conform to a clear theme. The adopted style of interaction may itself may relate to other factors such as the length of the user's exploratory search episode or how the retrieved resources are browsed.

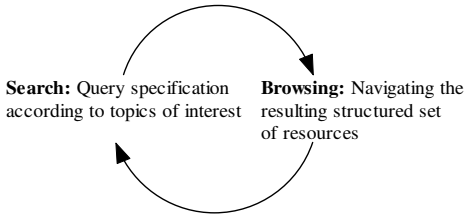

Figure 1: The mixed search/browsing form of interaction afforded by exploratory search.

In this paper we aim to shed light on these questions by investigating the use made of an exploratory semantic search tool to access an archive of museum content. The next section briefly looks at related work in exploratory search, semantic web and information foraging. We then introduce Bletchley Park Text, an exploratory semantic search application. In

section 4 we describe the methodological approach adopted in the study and in the following section present and interpret the results. Finally in the discussion section we provide an overall interpretation of the results by outlining two alternative exploratory search strategies evidenced by the results before outlining the boarder implications and opportunities for further work

II. RELATED WORK

Over the past decade web search technology has become increasingly relied upon by users and increasingly sophisticated in how it provides results to a user query. During that time, the mode of interaction with search engines has remained relatively constant in which a user enters search terms into a box and then receives a list of search results ranked by relevance. Despite the phenomenal success of existing search engines, there are contexts in which this mode interaction fails to align with the goals of the user. In a longitudinal study of search engine use, Käki [3] found that when a user's goal is more exploratory and less directed, organizing the results of a search rather than ranking by relevance facilitates their use.

Recently, technologies for exploratory search have been developed to support these less directed types of interaction with the available content. For example, mSpace [4] is a general interaction model and software framework for supporting the exploration of content. Within mSpace the user navigates the content through a series of information slices, each drawing on different properties of the content. For example, when mSpace is applied to a classical music resource, slices may relate to composer, era, form and piece. The slices scaffold the user in searching and navigating the content even if unfamiliar with the vocabulary of the classical music domain. mSpace is itself built on semantic web technologies and uses metadata of the content to construct the information slices.

Specially within the semantic web field, there has been a great deal of research recently into how the searching and browsing of museum archives can be better supported. MuseumFinland [2] is not necessarily focused on exploratory search but is one of most well known and developed semantic portals. MuseumFinland supports access to heterogeneous content from a number of museum collections and provides a number of generic search and browsing services by which the semantically enriched content can be accessed. mSpace and MuseumFinland are but two examples of how semantic web technology and new forms of user interaction are being developed to facilitate the access and use of content.

The effects of such technologies can be interpreted in terms of information foraging theory [5]. According to information foraging theory, when seeking out information, humans use similar mechanisms employed by their ancestors in search of food. Humans as information feeders (or informavores) rely on "information scent" to decide what links to follow on a website. The stronger the scent, the more likely the link will soon lead to satisfying information. Informavores must also make cost-benefit trade-offs when deciding whether move to a new feeding patch (e.g. website). The current website may have little new information to provide, however moving to a new website has an associated cost and the information value of any new website is uncertain.

Exploratory and semantic search and navigation technologies can be thought of as easing the life of the informavore. First, exploratory search reduces the gap between navigation (i.e. feeding in the current patch) and search (i.e. moving to a new patch). Secondly, semantic web technologies exploit ontologies and metadata to explicate the informational value of the current and future patches, therefore reducing uncertainty and guiding the informavore's choice of strategy.

Of potential interest to many museum website developers is the use of information foraging theory to increase website stickiness – the amount of time user spend on a website. For example, Nielsen [6] describes how information foraging theory can be used to design website that encourage information feeding with changing patch to another website.

III. EXPLORATORY SEMANTIC SEARCH: BLETCHLEY PARK TEXT

Bletchley Park Text (BPT) is a exploratory semantic search application built on our Story Fountain architecture [7]. BPT was developed for Bletchley Park, a codebreaking and computing museum in the UK. More information on BPT can be found in [8, 9, 10].

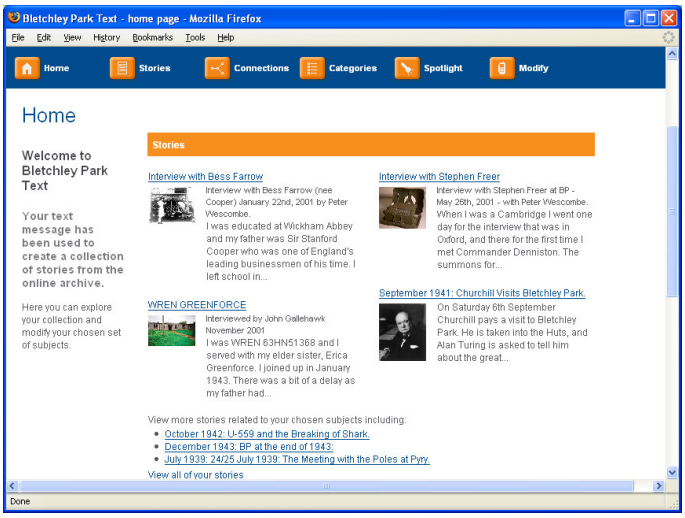

Figure 2. The newspaper-style BPT homepage.

Interaction with BPT first involves user selection of topics of interest. This can be done by SMS at the museum or via a form on the museum website. The selected topics are matched to metadata concepts which are then used to retrieve relevant stories. The retrieved collection of stories is organised and summarised in a set of *composite pages*. Composite pages impose a particular structure on a set of stories (such as pathways or set of interconnected categories) and contain story summaries from which the stories themselves can be accessed.

The organisation of the stories draws on the complete metadata description of the retrieved stories. Drawing on web design principles motivated by information foraging theory [6], a query produces a newspaper-style homepage that directly links to selected stories plus a range of composite pages that highlight different categories or pathways through the story set (see figure 2). The BPT interface provides a mechanism for reformulating the query to generate a new newspaper-style layout.

IV. METHODOLOGY

The methodological approach taken in this study focuses on server log evidence. Specifically, we looked at multi-query interaction with BPT, whereby the user has specified a query in order to generate a newspaper style website and then, after interaction with the content, has specified a further query to re-generate the content. Our definition of such multi-query episodes was operationalised as two or more queries from the same IP address with consecutive queries being no more than 20 minutes apart.

The overall aim of this study was to investigate the nature of the strategies adopted during multi-query exploratory search episodes. First, using the operational definition above, we identified the prevalence of these search episodes among users of the BPT website. Data was collected and analysed during the period 24th January to 27th September 2007.

TABLE 1. THE FIVE TYPES OF TRANSITION IDENTIFIED.

Transition	Description
Repeat	Consecutive queries contain identical terms
Subset	The next query contains a subset of the query terms.
Superset	The next query contains all the previous terms plus one or more additional terms.
Overlap	The next query contains some but not all of the previous terms plus one or more additional terms.
Jump	There is no intersection between the terms used in consecutive queries.

Once these episodes had been identified we extracted from them following set of properties. First, from the viewpoint of queries, we recorded the number of queries per episode, the number of search terms used in the episode and the mean query length. Second, in terms of navigation, the number of stories and composite pages viewed in each episode was measured. Third, we classified the types of transition made across consecutive queries within the episode. Five types of transition were identified. These are summarised in table 1. Fourth, new terms added during an episode where classified at to whether and how they were recommended in the BPT interface (see table 2). BPT presents two kind of recommendations. The terms that have the strongest statistical co-occurrence within the stories are suggested. Additionally, terms that are linked to the current query terms by a triple formally represented from the basic tour guide description (e.g. Alan Turing – was head of – Hut 8) are also recommended. These are termed *horizontal links*.

TABLE 2. THE THREE TYPES OF ORIGIN FOR A NEW TERM.

Term origin	Description
Statistical	The new term belongs to the list of top statistically co-occurring terms presented by BPT.
Horizontal link	The new term is linked by a triple from the tour guide description presented by BPT.
Unrecommended	The new term was not recommended by BPT.

The analysis of the results also looked for any correlations across the recorded properties, for example, whether a property of the query (such as query length) correlated (positively or negatively) with the number of stories and composite pages viewed, the types of transition or origin of new terms. The results are presented in the following section.

V. RESULTS

The data was taken from server logs between 24th January and 27th September 2007. The extracted server logs all relate to the forms-based entry to BPT. This is the most common method of access as it is not aimed only at physical visitors to the museum.

During this period (after the exclusion of IP addresses from the authors' university) 1154 queries were made using this mechanism. These queries originated from 777 unique IP addresses. Single query interactions (in which two queries from the same IP address were not received within 20 minutes) accounted for 643 of the 1154 queries, i.e. 55.7%. Within the server log, 166 multi-query episodes were identified, which accounted for the other 511 queries. The frequency of interactions with different numbers of queries is shown in table 3. In total 20.4% of website visits resulted in a multi-query interaction episode. This shows that BPT was able to encourage and support some significant level of multi-query interaction episodes (henceforth, for brevity referred to as episodes). The rest of the analysis will focus on these 166 episodes.

TABLE 3. FREQUENCY OF INTERACTIONS WITH DIFFERENT NUMBERS OF QUERIES.

Queries	1	2	3	4	5	>5
Frequency	643	97	30	16	8	14

The IP address was used to find out the country of origin for

each visitor. A geographical breakdown of the episodes is shown in table 4. UK IP addresses accounted for 66.3% of the episodes. The 166 episodes were from 158 unique IP addresses. Eight IP addresses each accounted for two episodes. Seven of the eight were from the UK.

Following subsections look at the properties of these episodes as outlined in the methodology section and identify any correlations between them.

TABLE 4. VISITOR LOCATION DETERMINED BY IP ADDRESS.

Location	No. of visitors
UK	110
North America	26
Europe (except UK)	19
Australia/New Zealand	6
Asia	4
South America	1

A. *Query*

Across the episodes, the mean number of queries was 3.10. The mean number of distinct query terms per episode was 5.76. The mean number of terms per query was 3.23. Using Pearson Product-Moment Correlation (see table 5), the number of terms used in a query episode correlated significantly with the number of queries ($r = 0.226$, $p < 0.01$) and the mean number of terms per query ($r = 0.748$, $p < 0.01$). Therefore, as might be expected, more query terms are found in episodes having more queries and in episodes having queries that contain more terms.

An inverse correlation was found between the number of queries and mean length of queries ($r = -0.205$, $p < 0.01$). This may suggest two distinct information foraging strategies at work. On the one hand the user may successively produce a large number of query results (i.e. from a high number of queries) but with each result associated with a small number of topics (i.e. a small number of terms in the query). On the other hand, the user may produce only a small number of query results (i.e. from a small number of queries) but with each result encompassing a large number of topics (i.e. a large number of terms in the query).

TABLE 5. CORRELATIONS BETWEEN THE QUERY PROPERTIES.

	Number of queries	Terms per episode	Mean terms per query
Number of queries		$r = 0.226$ $p < 0.01$	$r = -0.205$ $p < 0.01$
Terms per episode	$r = 0.226$ $p < 0.01$		$r = 0.748$ $p < 0.01$
Means terms per query	$r = -0.205$ $p < 0.01$	$r = 0.748$ $p < 0.01$	

When considering the number of terms in the query and its possible consequences, it should be borne in mind that as BPT was designed to support exploratory search, unlike conventional search engines, a large number of query terms does not necessarily produce a small result set. With BPT the query is initially treated as disjunctive rather than a conjunctive, and a set of stories is presented from the resulting collection.

B. *Navigation*

This section considers the number of composite and story pages read (calculated from page impressions). The mean number of stories read per episode was 3.39. The mean number of composite pages read per episode was 3.43. Including the search forms and home pages, each episode resulted on average in 13.02 page impressions. Over one fifth of website visits (i.e. those resulting in episodes) each producing an average of 13.02 page impressions would seem to indicate a reasonable level of site stickiness [6].

A correlation was found between number of stories read and number of composites read ($r = 0.542$, $p < 0.01$). As might be expected, visitors who read a larger number of stories also used a larger number of composite pages to access and understand those stories.

C. *Correlations between query and navigation*

When comparing across the query and navigation properties, the number of stories read was found to correlate with the number of queries formulated ($r = 0.297$, $p < 0.01$). Conversely, the number of composite pages read correlated with the number of query terms in the episode ($r = 0.285$, $p < 0.01$) and the mean number of terms per query across the episode ($r = 0.165$, $p < 0.05$).

It can be expected that the number of stories read would correlate with the number of queries specified. Composite pages may correlate with query terms and mean query size because larger term sets, producing a larger and less focused set of stories may require the use of more composite pages to understand the presented collection of stories.

D. *Transitions*

As described in the methodology section, five kinds of transition between consecutive queries were identified: repeat, subset, superset, overlap and jump. Table 6 shows the mean frequency of each type of transition per episode. Jumps made up 54% of all transitions. In jump transitions there are no terms shared between successive queries. This finding seems reasonable within the context of a museum website as visitors often have no specific goal in mind and are primarily interested in the breadth of available content rather than studying anything in depth. No correlations were found between the transition types. This suggests there is no pattern in the use of different transition types within episodes.

E. *Correlations between transition and query*

The number of queries within the episode was found to

correlate with repeat transitions ($r = 0.634$, $p < 0.01$) and jump transitions ($r = 0.753$, $p < 0.01$). The number of query terms within the episode correlated with overlap transitions ($r = .208$, $p < 0.01$). There was an inverse correlation between mean terms per query and jump transitions ($r = -0.179$, $p < 0.01$). Jump transitions therefore tend to be used when there are a larger number of queries and the queries contain less terms. Conversely, overlap transitions tend to be associated with episodes involving a larger number of query terms.

TABLE 6. FREQUENCY OF EACH TRANSITION TYPE PER EPISODE

Transition	Mean frequency per episode
Repeat	0.50
Subset	0.19
Superset	0.22
Overlap	0.05
Jump	1.13

F. Correlations between transition and navigation

Jump transitions were found to correlate with number of stories read ($r = 0.270$, $p < 0.01$) however, overlap transitions correlated with number of composite pages read ($r = 0.209$, $p < 0.01$). In information foraging terms, a transition from one query to the next can be thought of as moving to a new patch. This will require some acclimatization with the new patch in order to understand its information value. If the new patch is arrived at via an overlap transition then some of the terms used to structure the resources in the composite pages will already be familiar. This may allow the composites to support the acclimatization. For jump transitions, where the composites do not contain familiar terms, then the only way to become acclimatized to the content may be to read stories themselves from within the provided collection. This would explain why jump transitions correlate with reading stories whereas overlap transitions correlate with reading composite pages.

G. Term origin

New terms introduced during an episode were classified according to whether they had been recommended by BPT and whether that recommendation was based on statistical co-occurrence or horizontal linking. The mean number of concepts introduced by origin per episode is shown in table 7. Although recommend terms are occasionally used, unrecommended terms that are conceptually less close to the current story set are far more commonly taken. This suggests that overall a more speculative approach is taken when formulating follow-on queries. This can be thought of as similar to the tendency toward jump transitions, that retain no terms from the previous query. As shown in table 8, all three types of term origin were found to correlate with each other. This suggests there is no particular pattern across the origins of terms, similar to the lack of pattern in transition types.

H. Correlations between term origin and query

Correlations between term origin and query properties are shown in table 9. All three types of term origin correlate with terms per episode and mean terms per query but only unrecommended terms correlate with number of queries. This correlation between unrecommended terms and number of queries fits with the previous observation that jump transitions correlate with the number of queries per episode. Unrecommended terms and jump transitions both involve making a relatively large conceptual shift from one query to the next and both are associated with a larger number of queries within an episode.

TABLE 7. TERM ORIGIN FREQUENCY PER EPISODE

Term origin	Frequency per episode
Statistical	0.30
Horizontal link	0.27
Unrecommended	2.45

TABLE 8. CORRELATIONS BETWEEN TYPES OF TERM ORIGIN.

	Horizontal link	Unrecommended
Statistical	$r = 0.379$ $p < 0.01$	$r = 0.486$ $p < 0.01$
Horizontal link		$r = 0.216$ $p < 0.01$

TABLE 9. CORRELATIONS BETWEEN TERM ORIGIN AND QUERY.

	Number of queries	Terms per episode	Mean terms per query
Statistical	n.s.	$r = 0.425$, $p < 0.01$	$r = 0.261$, $p < 0.1$
Horizontal link	n.s.	$r = 0.361$, $p < 0.01$	$r = 0.248$, $p < 0.1$
Unrecommended	$r = 0.180$, $p < 0.05$	$r = 0.669$, $p < 0.01$	$r = 0.375$, $p < 0.1$

I. Correlations between term origin and navigation

The only significant correlation between term origin and navigation is between unrecommended terms and composite pages ($r = 0.188$, $p < 0.05$). In terms of the previously discussed findings this is difficult to interpret. This is the only finding that does not fit the with emerging model of a cautious exploration strategy involving among other things increased use of composite pages and a bolder strategy involving among other things the use of unrecommended new concepts.

J. Correlations between term origin and transition

Significant correlations between term origin and transition types are shown in table 10. The inverse correlation of both

statistical and horizontal linking recommendations with subset transitions may be because subset transitions involve focusing down on the selected content. After such focusing, the user may well wish to move to a conceptually distinct set of content having investigated that topic in some detail.

TABLE 10. CORRELATIONS BETWEEN TERM ORIGIN AND TRANSITION.

	Statistical	Horizontal link	Unrecommended
Repeats	n.s.	n.s.	n.s.
Subsets	r = -0.213, p < 0.01	r = -0.185, p < 0.05	n.s.
Supersets	n.s.	r = 0.170, p < 0.05	n.s.
Overlaps	r = 0.234, p < 0.01	n.s.	r = 0.160, p < 0.05
Jumps	n.s.	n.s.	r = 0.215, p < 0.01

Subset and overlap transitions each correlate with one or both types of recommendation. Following recommendations and retaining some or all of the previous query term are both consistent with a more cautious information foraging strategy in which consecutive queries are conceptually close. Conversely, jump transitions only correlate with unrecommended terms, both of which are consistent with a more risky information foraging strategy in which consecutive queries are conceptually distant.

VI. DISCUSSION

The interpretation of server logs over an eight month period found evidence that the approach taken in the design of BPT can support exploratory semantic search. 20.4% of website visits resulted in multi-query episodes each producing an average of 13.02 page impressions.

Analysis of the 166 multi-query episodes evidenced two distinct information foraging strategies. A more common risky strategy was characterized by the following:

- Jump transitions in which no terms are retained on consecutive queries.
- Large numbers of queries with a smaller mean number of query terms.
- The introduction of unrecommended and conceptually distant terms
- The use of stories rather than composite pages for acclimatisation to new information patches.

The second less common strategy is more cautious. This involves the following:

- Overlap transitions that retain terms from the previous query.
- A smaller number of queries but with a larger number of query terms
- Introduction of recommended and therefore conceptually close new terms
- The use of composite pages rather than stories for acclimatisation to new patches

Of all the statistical findings the only one inconsistent with this model are that unrecommended terms correlate with composite pages and overlap transitions. Further analysis of the server logs over a longer period of time may help to shed light on this.

Although the current study has provided evidence of these two information foraging strategies, server log alone do not highlight user context. A future study could triangulate server logs with interviews or protocol analyses to indicate whether choice of information foraging strategy with an exploratory search tool aligns with user goal.

Future semantic exploration tools could provide users with feedback on how cautious or risky their information foraging strategy is. This could allow them to reflect upon their strategy and possibly revise it as a result.

More ambitiously, exploratory semantic search tools could potentially automate trajectories across the available content adhering to different level of user defined riskiness, or even determining the appropriate level of riskines appropriate for the user's specified information goal. We hope to work toward addressing some of the above issues in our ongoing work.

REFERENCES

[1] R. W. White, B. Kules, S. M. Drucker and m. c. schraefel, "Supporting exploratory search," *Communications of the ACM*, vol. 49, no. 4, pp. 36-39, 2006.

[2] E. Hyvönen, E. Mäkelä, M. Salminen, A. Valo, K. Viljanen, S. Saarela, M. Junnila and S. Kettula, "MuseumFinland - Finnish Museums on the Semantic Web," *Journal of Web Semantics*, vol. 3, no. 2, pp. 224--241, 2005.

[3] M. Käki, "Findex: Search result categories help users when document ranking fails," Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI '05), Portland, Oregon, 2005.

[4] m. c. schraefel, M. Wilson, A, Russell and D. A. Smith "mSpace: Improving information access to multimedia domains with multimodal exploratory search," *Communications of the ACM*, vol. 49, no. 4, pp. 47-49, 2006

[5] P. L. Pirolli and S. K. Card, "Information foraging," *Psychological Review*, vol. 106, no.4, pp. 643-675, 1999.

[6] J. Nielsen, "Information Foraging: Why Google Makes People Leave Your Site Faster," (2003). Available: http://www.useit.com/alertbox/20030630.html.

[7] P. Mulholland, T. Collins and Z. Zdrahal, "Story Fountain: Intelligent support for Story Research And Exploration," Proceedings of the ACM Conference on Intelligent User Interfaces (IUI '04), Madeira, Portugal, 2004.

[8] T. Collins P. Mulholland and Z. Zdrahal, "Semantic browsing of digital collections,". Proceedings of the International Semantic Web Conference (ISWC '05), Galway, Ireland, 2005.

[9] P. Mulholland, T. Collins and Z. Zdrahal, "Spotlight browsing of resource archives," Proceedings of the ACM Conference on Hypertext and Hypermedia (HT '05), Salzburg, Austria, 2005.

[10] P. Mulholland, T. Collins and Z. Zdrahal, "Bletchley Park Text: Using mobile and semantic web technologies to support the post-visit use of online museum resources, " *Journal of Interactive Media in Education* (Portable Learning: Experiences with Mobile Devices. Special Issue, eds. A. Jones, A. Kukulska-Hulme and D. Mwanza)., 2005. Available: http://jime.open.ac.uk/2005/24.