The Garden of Forking Paths – Forms of Scholarship and the 'Formations' Pre-Prints System for Cultural Studies and Related Fields

DAN FLEMING
Senior Lecturer in Media Studies, School of Media & Performing Arts, University of Ulster, Coleraine, Northern Ireland BT52 1SA (e-mail: dr.fleming@ulst.ac.uk)

> *'The book is an indeterminate heap of contradictory drafts' (Borges)*

Key words: cultural studies, e-prints, electronic publishing, pre-prints, research

Abstract. This paper describes and analyses a web-based pre-prints project in the UK's Electronic Libraries Programme in order to raise issues about the forms of scholarship that are best suited to online working. Specifically, the paper describes some of the underlying processes at work in academic research and seeks to match these, where appropriate, to forms of online working. In doing so, the paper describes in detail a scholarship of integration which seems well suited to online tools such as pre-prints systems, but speculates that such forms of scholarship are too seldom explicitly identified when academics refer to research as a totality. As a consequence the potential match between working practices and emerging tools may not be obvious to academic researchers. To investigate these issues further, the paper examines the degrees of formality involved in different kinds of online communication and describes how academic working practices might be supported by adapting established 'groupware' tools such as Lotus Notes. The eLib 'Formations' project, which is using Notes to develop an integrated pre-prints and e-journal system for research in cultural studies and related fields, is described in detail, focusing on the underlying technology and the overall design.

1. Introduction

In constructing the prototype of a pre-prints system for scholarship in some specific fields of the humanities and social sciences, we found that we were assuming there the existence of a specific form of scholarship – a scholarship of integration which busies itself with accumulating, organizing and making sense out of the equivalents, in its own fields, of Borges' indeterminate heaps – but that academic work is seldom so consciously differentiated in terms of the various activities that constitute research. So the potential of online systems designed to support such

work may not be especially obvious to intended users. The form of scholarship assumed by this particular kind of online 'tool' is usually hidden somewhere in the shadow cast by a more smoothly executed, individualistic and linear (or so it seems) scholarship of discovery. And yet . . .

This paper describes an on-line system for cultural studies and related fields which was launched on 23rd September 1997 – after the paper itself was written. For that reason, unfortunately, we cannot offer here any information or discussion based on usage and user responses. If all goes as planned, however, by the time the paper is published the 'Formations' system as described will be in active use and interested readers will have the opportunity to compare the reality with the aims and intentions presented here. (In fact, the situation just before publication of this issue was that 'Formations' had run 'live' for six months and would be offline for another six months for further development based on detailed user feedback. Readers should see the website at http://formations.ulst.ac.uk for the current position, including the first evaluation results.) 'Formations', to most people's minds, will be an unfamiliar synthesis of elements from web publishing, on-line discussion groups and mailing lists, so a brief summary of what the system does will help set the scene for what follows.

'Formations' is a combined and integrated pre-prints bank and e-journal. It uses the web (although it is not a conventional website, as we will see) and is accessed with any standard frames-capable browser. We have coined this particular usage of the term 'bank' in order to evoke the associations of investment, depositing, accumulating value, interest, etc. that flow from the banking metaphor. The 'currency' in question consists of documents. Via a largely automated process, which will be described in more detail later, scholars using the 'Formations' system are able to publish issues of the e-journal from within the pre-prints bank. So the two parts are intimately linked. Indeed that linkage is at the heart of what 'Formations' is intended to achieve – a modestly new way of deploying electronic publishing that side-steps the print paradigm and uses the medium to do distinctively digital things.

The transformation, reconstruction or, more colloquially, the 'morphing' of information is a distinctively digital thing. To be able to take data and instantly re-present it in different ways is a major benefit of digitalization – nothing in conventional print publication comes close to the readily reconstructive potential of electronic publishing. The automated filtering and re-presentation of material from a pre-prints repository as 'issues' of an on-line journal is the most obvious way in which 'Formations' taps this potential, but the system also allows multiple 'views' of the documents contained in the pre-prints bank itself (say viewed by author, date or topic). The concept of a potentially very large repository of documents that can be viewed in numerous ways – including in the form of a journal – is absolutely central to how 'Formations' works. But, before describing these operations in more detail, some context is needed in order to explain where 'Formations' has come from and how its genealogy has affected its aims and working methods.

2. The Background

This is one of approximately sixty projects in the Electronic Libraries programme (eLib), funded by the Joint Information Systems Committee (JISC) of the UK's higher education funding councils. The project is based in the School of Media and Performing Arts at the University of Ulster, where it is supported by UU Educational Services (a 'converged' alignment of library, computing and media services), and has collaborators at the universities of Stirling, Aberystwyth and Luton. As some readers will know, the JISC is charged with providing a network service for the UK higher education and Research Councils community and, in this role, provides the funding for development and operation of the academic network. The JISC also provides a variety of information services and is responsible for a range of initiatives to encourage and develop the use of information systems in the university sector. The JISC established the eLib programme as a direct response to the Follett Report (1993), which made important recommendations on how the use of information technology in the 'digital library' might help alleviate some of the pressures on university scholarship in the UK today.

The eLib programme had a budget of approximately £15 million over three years to address this potential, in part by exploring different models of intellectual property management and encouraging new methods of scholarly publishing. That is where 'Formations' came in – as a proposal for a project, in the latter stages of the programme, to explore some of the 'grey' areas which established electronic journal projects were not looking into. Specifically, we wanted to investigate whether it might be possible to rework and extend the notion of 'e-prints', or electronic pre-prints, to make them of more compelling interest in areas of scholarship with no tradition of pre-print working at all. With two full-time staff on eighteen month contracts, and two part-time academic co-directors, the project took fifteen months to design and construct its first 'build' of a user-centred system with the required flexibility. It began by targeting researchers and writers in nine interconnected areas of scholarship: broadcasting; film studies; identity and culture; material culture; performance research; photography and image studies; print, publishing and journalism studies; social and cultural theory; technology and electronic culture. The first set of evaluation results, once fully analysed, may lead to fundamental re-organization of these categories. The system offers scholars working in these areas a relatively easy way to publish material instantly to the web, to organize such material into meaningful categories within on-line working areas, to find submitted material quickly and easily, to recommend and review other on-line resources, and to produce an electronic journal which abandons the fixed-length, regular schedule and related characteristics of print journals. The project team will report to the eLib executive in mid-1998 on the uptake and impact of the 'Formations' system as a whole. This paper is concerned with the thinking behind 'Formations', with the software design issues it raised and with the ways it might be used.

The concept of e-prints (electronic pre-prints) has perhaps been most interestingly explored by Paul Ginsparg. Particularly noteworthy for the range of issues it raised was his invited contribution to a conference held at UNESCO, Paris in February 1996. In a paper called 'Winners and Losers in the Global Research Village', Ginsparg described what is undoubtedly the most successful set of e-print archives to date – those serving the physics community, including Ginsparg's own initiatives at the Los Alamos laboratories. The first e-print service of note was hep-th (high energy physics – theory), launched in August 1991. Intended for use by some two hundred researchers in a narrowly defined field, hep-th boasted in fact nearly 4,000 users after only a few years of operation, by which point it had also been joined by similar physics archives reaching a total of over 35,000 researchers. Ginsparg estimated the resulting total number of daily electronic transactions at approximately 70,000. This clearly represents a dramatic uptake of a new medium for scholarly communication in one academic discipline, prompting speculation about its more general applicability. On-line pre-prints systems modelled on the physics ones have been popping up sporadically in other fields of the sciences and social sciences but, to date, comparatively little work has been done to explore alternative models or to correlate those models with a range of academic disciplines, where there might be fundamentally different approaches to publication, formal and informal communication, collaboration and the desired accessibility of work-in-progress.

3. How 'Formal' Are Pre-prints?

A 'pre-print', at least in the established model, is a research abstract – a document which formally presents some feature or features of research undertaken, without having to meet all the criteria for publication in a refereed academic print journal. That such material, at perhaps a later stage of development, would be suitable for print publication is what led to the term 'pre-print' gaining acceptance as a handy way of describing such documents. Ginsparg took pains to position the concept of e-prints carefully in relation to other forms of on-line communication, although his perspective on this may be discipline-specific and is not in fact entirely compatible with the thinking behind 'Formations':

> It is important to distinguish the form of communication facilitated by these systems from that of Usenet newsgroups or garden variety "bulletin board" systems. In "e-print archives", researchers communicate exclusively via research abstracts that describe material otherwise suitable for conventional publication. This is a very formal mode of communication in which each entry is archived and indexed for retrieval at arbitrarily later times; Usenet newsgroups and bulletin boards, on the other hand, represent an informal mode of communication, more akin to ordinary conversation, with unindexed entries that typically disappear after a short time. (1996, p. 3)

While the 'Formations' e-prints bank is not positioned in precisely the same way, it remains important to understand how Ginsparg's description identified a space for a particular form of on-line working, rubbing shoulders with other forms of electronic communication but with some meaningful distinctions to be made among them. A quick historical overview of the internet clarifies those distinctions. As we know, 1969 saw the establishment of what would become the Internet Protocol-controlled system which allows the movement of data 'packets' around wired networks. With routers and bridges linking the networks and regulating the traffic of 'packets' through them, a new medium of communication emerged. Unsurprisingly, given the comparative recency of these developments, we are still seeing the evolution of concepts for communicating effectively by such means. That evolution, to date, takes us from electronic mail to the world wide web; which may only be comparable to the development of the letter into the picture postcard and leaves a great deal of potential still to be explored. The formal-informal scale evoked by Ginsparg remains useful as a means of tracking variations in the recent evolution of networked electronic communication in general, but one quickly notices how difficult it is to position particular forms at precise points on that scale. This is an important issue in relation to understanding our project's approach to 'formal' on-line communication.

E-mail, perhaps the simplest form into which data packets can be recombined, merges the informality of conversation with the formality of the written word, leading to difficulties that are only now becoming better understood. As Pitter et al remark, 'After you use e-mail for a while, you'll see a lot of messages that should never have been sent, or ones the sender probably wishes he hadn't sent' (1995, p. 29). What is written quickly and 'informally' (e.g. with a tone or intent that would have been audible in face-to-face conversation) is often read 'formally' (e.g. at face value, as a written document, without any of the contextual modifiers that would have allowed it to be 'heard' differently). So placing email towards the informal end of the scale acknowledges how it often 'feels' as a writerly medium but allows insufficient importance to the way in which the displayed 'readerly' text on the recipient's computer screen frequently takes on a fixity, closure and one-dimensionality – a loss of the sender's 'voice' – which draws it back into the formality of all written systems. The means thus far developed for extending the communicative reach of email-based systems have run up against this central – if rather obvious – paradox in other ways.

Listservs (sometimes referred to as mailing lists, list processors or mail reflectors) have been a useful experiment in developing public or one-to-many forms of email. Topic or interest based listservs, distributing any email addressed to the list to all that list's subscribers, typically carry newsletters, discussions, announcements, requests for information or advice, etc. Interestingly, most busy listservs develop their own overall 'personality', an emergent effect of the ways in which diverse traffic tends to develop common characteristics over time, and as a consequence certain forms of etiquette frequently become clearly, if implicitly,

established on a list. Often this only becomes apparent when a breach of etiquette occurs and users become aware of their collective discomfort. In short, the list-serv often has the sort of taken-for-granted formality of a seminar room – the exchanges will frequently seem conversational but they are self-regulating accord-ing to stricter codes than might be operative in more relaxed social gatherings. An individual posting may seem essentially 'informal' but the cumulative effect, and the ways in which that effect actually does constrain individual postings, usually have a marked degree of formality. In a sense, in fact, the listserv solves many of the inherent problems found in email as a one-to-one medium by limiting users' tendency to forget that – unlike speech – tone and intention do not get automatically stamped on a message (the contrivance of 'emoticons' notwithstanding).

Where one-to-one email involves a direct mode of address (the recipient is individually addressed), a listserv disperses the mode of address or renders it indirect (recipients are a category of addressee but individually undifferentiated to all intents and purposes). This is so even though all messages are individually received by each subscriber. The Usenet newsgroup achieves much the same end but requires the user to log on with a 'reader' program to access postings rather than receiving them automatically. This is, on the surface, a subtle distinction but in practice the experience is distinctly different for the user. In effect, because a message is not actually being sent to anybody (it is not a form of e-mail), the newsgroup has become a forum for free discussion. It is important to note, though, on the evidence of fairly random sampling, that surprisingly many of the 20,000 or so newsgroups now operating often seem to be either wildly volatile or nearly impenetrable places to engage in meaningfully sustained dialogue, unless one has been closely involved with them over time. Despite often having FAQs (Frequently Asked Questions) to get new users up to speed on a particular topic, a newsgroup can be an unfriendly place for a novice participant. This is not so much the equivalent of a seminar room as of a club, where members have developed their own codes, can frequently get raucous and where a new arrival can feel distinctly uncomfortable. While undoubtedly much closer to the informal end of the scale than the typical listserv, the newsgroup's very 'freedom of expression' is itself a highly codified form of interaction, with even personal abuse formalized into 'flaming' (indeed the alt.flame newsgroup exists solely for confrontational exchanges).

What this brief detour through email, listservs and newsgroups suggests is that we have to handle very carefully the distinction between 'formal' and 'informal' modes of communication, as proposed by Ginsparg when he positioned the physics e-prints at the 'formal' end of this supposed scale. 'Informal' at the sender's end can become 'formal' at the receiver's; etiquette (or 'netiquette') quickly emerges on mailing lists, preventing users from becoming too informal in their behaviours; and even in the wilder reaches of Usenet newsgroups the freedom being exercised is often highly codified. It is necessary to think through more carefully the otherwise 'obvious' formal/informal distinction in relation to such forms of on-line commu-

nication, in order to suggest that e-prints need not be defined – perhaps cannot very usefully be defined – according to any specific degree of 'formality'. The 'Formations' pre-prints bank facilitates a range of exchanges which may mix the formal and the informal to various degrees. It is not alone in attempting this. The major internet service (as distinct from just access) providers, such as AOL, have begun to explore something similar in their various attempts to 'build communities' – AOL's Learning Zone being an interesting example. The best instance outside the academic world, however, is undoubtedly Motley Fool, an AOL area which has now spread to the web (http://www.fool.com) and deals in large volumes of financial and investment information by mixing hard facts with 'message boards' for users (averaging over 4,000 messages a day). Still, it remains important to ask whether we might too readily slip into the trap of assuming that 'informal' always means less significant, less worthy of archival respect, less important to how scholarship is pursued.

4. Off-stage Communication

There is an argument to be made for the very real significance of a certain kind of 'informality' in the pursuance of academic research and this may help us recast the distinction made by Ginsparg. Penny et al (1994) reflect in some detail on their experiences as researchers in a South African university department of education. In doing so, they also comment on how difficult it was to have their work published in a form that adequately represented some aspects which they felt to be deeply important. Instead of the formal/informal distinction, Penny and his colleagues offer the pairing of 'off-stage' detail and 'on-stage' performance in academic publishing. What they mean, in essence, is that the conventions of academic discourse and of peer reviewed print publication in established journals conspire to produce the well established genre of the academic paper. This is a public 'performance' which leaves out much that is important. It also leaves young researchers uninformed about much of the actual work that is done under the umbrella term of 'research' and about the communicative context that fosters good scholarship. The generically acceptable papers seem to exist fully-formed and independent of the debate, uncertainty, unexpectedness, false turns, excitement, dead-ends, serendipity, evolving sense of direction and general messiness of real research. (There is a parallel, perhaps, with the making of documentary films in which the apparatus of production is so invisible that a presenter appears to be standing godlike in some remote part of the world, having descended unaided from the sky, 'objectivity' unsoiled.)

> Much of the "backstage" detail is left unrevealed and thereby fails to inform the reader of the dynamics of the research process. No research is smooth and unrumpled. ... yet little of this is transmitted to the reader. What is provided is a clean, completed product devoid of the problematics of the situation being reported. (pp. 21–22)

With this question in mind – about the significance but invisibility of the 'backstage' detail – Penny and colleagues examine 'a strategy of collaborative research which was aimed at regenerating a culture of research within a university department' (p. 21). While their circumstances added genuine social weight to that pressure (educational research in South Africa is required to address some fundamental and pressing problems), the pressure itself was perhaps not so unlike that being felt within many UK university departments as they attempt to improve their rating in the national Research Assessment Exercise to which their funding is linked; especially departments in some of the ex-polytechnics and newer universities with a less developed research culture. In a fascinatingly detailed account, Penny et al describe how successful a collaborative, dialogue-based approach was for them in responding to such pressure. In twinning this account with an analysis of why, despite its success, they then had difficulty finding a place for the grain and texture of that collaborative work within the established 'dynamics of theorising and writing up', the South African group perhaps unexpectedly reveals some absolutely basic questions about academic discourse and its formalities, its protocols.

The embedding of dialogue and exchange throughout the group's research project (on how schools deal with racial integration) fostered the elusive 'culture' in the department where, by implication, continuing to work as individual researchers might not have done so. In maintaining this 'backstage' detail in how they wrote up their research, however, Penny and colleagues produced a paper which the editor of 'a well-known international journal' demanded be cut by almost half. From their account, it seems that the 'voice' behind the longer version was insufficiently univocal, authoritative and godlike in its detachment from the messy practicalities of researching and theorising. What was wanted was a more straightforwardly conventional narrative within which the research findings were presented with detached confidence and less self-reflection about how it was done. In complying, Penny et al tell us 'the team felt that the report lost much of its honesty, accountability and richness', while the published version created a not altogether truthful 'impression of clear conception and initial vision, authority, and smooth execution' (p. 29).

5. Behind the 'Smooth Executions' of Scholarship

It is not too difficult to see the inevitability of suggesting here that Penny and his colleagues may have needed access to an appropriately designed e-prints system; and how much better if that system not only allowed the writing up of the 'off-stage' detail but also supported that collaborative work while the research was actually being undertaken? But before describing how a system such as 'Formations', or its descendants, might achieve this, it is important to note the consequences for Ginsparg's use of the formal/informal distinction in relation to on-line communication. If we identify the term 'off-stage' with the 'informal' side

of that distinction, then resisting such 'informality' within an e-prints system will only serve to reproduce the censorial effect of which the South African researchers complained. In fact, in light of the detour above on the nature of email, listserv and newsgroup communication, if we substitute the term 'off-stage' for 'informal' we find that we are no longer necessarily working at the opposite end of the scale from 'formality'. There can be a great deal of implicit and explicit formality about the 'off-stage' work. The point is not that it lacks formality but that it does not fit a specific formality – that of the convention-bound final write-up, the academic paper that finds its authoritative voice and its linear narrative of research progress just in time to get into print in the expected manner. The question perhaps becomes whether we know enough about the formalities of doing good 'off-stage' work to design on-line systems to support it, as distinct from an on-line version of established print publication with its emphasis on the final 'performance'?

At this point, a step-by-step summary of how someone might use 'Formations' should help to relate the preceding necessary but abstract ruminations to some concrete detail. Logging on to the 'Formations' e-prints website at http://formations.ulst.ac.uk during its first 'live' run reveals a welcome screen with a navigation frame on the left. The left frame offers six options: How To, Journal, Library, Channels, Registration, Home, Feedback and Management. Home is the main display which is visible at this point. Feedback and Management support various housekeeping activities. How To is an area with advice and tutorials on using the system. The Journal is actually a link to a separate area where issues of the 'Formations' journal are independently accessible to readers on the web, without necessarily going through the e-prints system with its registration procedure, from where the journal is created. The Library is an area for user-generated subject gateways. There users can recommend, review or comment on other internet resources in particular subject areas. The Channels are where most of the work will be done with e-prints. (Although described here in the present tense, from the perspective of the six-month run of the first 'build', evaluation may lead to modifications in the areas described.)

Clicking on any of the Research Channels takes the user into the heart of the system's working areas. When first launched, each channel was empty except for a screen indicating 'no documents found' and a set of functional icons. One of these – 'Host New Venue' – is the means by which users set up topic areas for the submission of e-print material by themselves and others. A straightforward forms-based procedure allows a registered user to launch a 'venue', which is a topic or subject-defined area for the assemblage of documents. When venues are established on any channel, a list of their names is the first thing one sees on clicking a channel number.

Say I find a venue on film noir listed among several others on the film studies channel. It will have been started there by a self-elected host who has registered with 'Formations' and decided to set up that venue. I might decide to do the same and host a venue on Japanese cinema. Or, if I wish to contribute something on

film noir, I can select that venue, have a look at what else has been submitted to it and add my own paper, comment, report or whatever. The film noir channel may have research abstracts, full papers, conference reports, book reviews, ongoing discussions, etc. That depends entirely, as with a listserv, on how people have decided to use it. The material is not simply deposited there is one long extending column. It is 'threaded'. Each new item can either start a new thread or be added to an existing thread. A 'thread' is defined by whatever rhetorical or conceptual connections users choose in order to justify connecting the items. And everything can be sorted and viewed in various ways – by thread, by date, by author, and so on. Full-text searching allows users to find relevant material quickly and efficiently. Thread 'histories' at the bottom of every document locate that item in relation to others and, very importantly, provide hyperlinks to everything else in a thread.

So at 'venue' level, 'Formations' is a system for building, organizing and navigating around growing collections of e-prints, with the sorts of facilities for searching and filtering that users will need in order to locate what they want. So far there are similarities with sophisticated 'message board' systems such as Motley Fool. Where things begin to get unusually powerful is in the way that 'Formations' recognises categories of registered user based on what they have already done within the system and then automatically allocates specific privileges. For instance, authors of documents always have editing rights when they revisit their own documents – they can recall a document in an editing window and amend it. The system tags and dates the modified items so other readers can see that changes have occurred. Even more significantly, hosts of venues have a special privilege in relation to all material submitted to their venues – at any time they can tag items for inclusion in an issue of the journal.

The 'Formations' journal is based entirely on the idea that within the pre-prints bank some material, including – though not confined to – recognisably conventional papers, may have a different order of interest than other more transitory items; or that documents may be edited by their authors, perhaps based on posted comments, and reach a point where they have attained a more finished quality. The venue host can tag such items over time – these tags are visible to users – and when a critical mass is reached a 'publish' button will create an issue of the journal from the selected material. This is routed automatically through an 'approval' procedure which requires the system's management group to hit an 'approve' button before the issue is compiled. This is a protection against abuse of the system rather than an editorial intervention – but it does raise some interesting questions which we will have to see worked out in practice.

Journal issues will, therefore, appear at irregular intervals and with greatly varying lengths. They will be 'thematic' because published from within venues which have some thematic or topic-based unifying feature. This clearly moves away from almost all the established procedures of conventional print publishing but there is, after all, no reason why electronic publications must replicate those procedures in the first place. Instead of a ring-fenced peer review process, the relationship of

the journal to the underlying pre-prints bank embeds a general element of peer review in the whole process. It remains to be seen in practice what this means for the quality of material published in the journal. Traditional peer review can – as we know – lead as often to conservative 'playing safe' and a lowest common denominator of mediocrity in some academic journals as it does to important and intellectually adventurous writing.

It is significant for the generalised 'peer review' process that users are not anonymous. The registration procedure involves the mandatory supplying of an email address. All registered users in effect also gain an email address at 'Formations' itself. This allows users to email each other via 'Formations', the mediating role of the system offering a degree of protection against unwanted communications. But in general our expectation is that the visibility of what participants are doing will lead to the establishment of certain standards and protocols, as discussed in relation to listservs above. The system design exerts a brake on any tendency for such protocols to slide back entirely towards established print-based traditions of working. Users will be conscious that they are in a public place and are likely to behave accordingly. 'Profiles' – short biographical notes or notes about research interests – can be supplied on registration and these can be called up from the header of any document submitted by that particular user. Scholars do, after all, have personal reputations to make and sustain.

6. User-centred Design

What the small development team at Ulster has designed is, we hope, the basis for a thoroughly user-centred system where the structure lends a certain inevitable formality to user behaviour but the philosophy of the system lends itself to supporting the 'off-stage' activity which is often the large and invisible part of the research iceberg. We expect to do further work on the design described here when the first evaluation results have been interpreted. But, in the longer term, we envisage all sorts of exchanges and a wide range of document types being deposited in the pre-prints bank. The 'pull' of the journal will tend to draw many users away from 'chat' towards more polished submissions but it is vitally important to recognize that a sequence of 'conversational' exchanges might prove to be deeply interesting to others and that a host can choose to select any such items for inclusion in a journal issue. A degree of deft ring-mastering may be required from hosts to produce coherent journal issues from a diverse range of accumulating material (and users' hesitancy about this has been a focus of preliminary evaluation). To provide a degree of unification, an issue always has to have an introductory 'editorial' from the host as part of its automated creation mechanism.

The Library area of 'Formations', in the first 'build', actually uses the same basic procedures as a standard pre-prints venue but in this case things are so organised as to allow users to submit resource recommendations and reviews of other internet-based materials, instead of more diverse documents. In a sense, 'Forma-

tions' aspires to be a fairly complete on-line working environment. As the computer 'desktop' draws closer to networked resources (perhaps inevitably becoming more of a 'webtop') and browsers get integrated with the operating system and with the main graphical interface, a descendant of 'Formations' may become more convenient as a day-to-day research and communication tool. Organising venues into 'channels', for example, leaves open the option of pushing channel information onto a user's desktop at some future point – perhaps when 'smart pull' technologies have developed sufficiently to supplant conventional browsing. One we had an eye on as 'Formations' developed was the Castanet tuner from Marimba (a spin-off group from Sun's original Java development team). At some future point that kind of system could be used to set up selected 'Formations' channels directly on a user's desktop/webtop. See May (1997). But even now, in the latter stages of the project, Microsoft's Internet Explorer 4 allowed the project team to embed 'Formations' in a desktop window from where it was instantly accessible without running a separate browser and this emerging functionality began to colour the longer-term thinking about the design.

For the time being, the key point about the system as envisaged in practice, and described above, is that it combines a degree of formality with an openness to the 'off-stage' sorts of work that usually disappear behind conventional academic publication. Not least among such 'off-stage' activity is the collaborative exchange, the dialogue, the interconnection of one's own work with that of others at a formative stage. Indeed the pre-prints bank and the journal, respectively, map loosely onto the distinction between the 'off-stage' activity and the 'performance'. In practice that distinction is unlikely to be clear cut – it is not desirable that it should be – but the system design implicitly embeds that whole range of activity in the working processes which are supported.

One important proposition, therefore, is that we have re-defined the notion of the pre-print in ways that move it from the established concept in the physics archives, where the formal research abstract has a particular disciplinary role to play, towards something that might work better in the humanities and social sciences. In part this has been achieved by replacing a notion of 'informal' communication with one of 'off-stage' exchange and providing a 'stage' for that activity to be both more openly acknowledged and actually enhanced in practice. Without the link between 'stage' and 'off-stage' the potentially productive interplay between the two is severed. In designing a link, we are attempting to extract the best features of the listserv or the newsgroup and to combine them with a web publishing mechanism in order to create something distinctively robust and flexible for scholars to use as they see fit. It remains to be seen, of course, whether it is robust enough to withstand rough handling in practice. Along the way the 'informal' is less likely to disappear as mere transient ephemera and more likely to take shape as a sustaining fibre of genuine research work. The fact that 'Formations' is also based on resolutely collaborative principles is good timing at a moment when the lone scholar is starting to feel rather isolated in the face of mounting epistemological and institutional

pressures – the exponential growth of information, the fragmenting of 'subjects' and the rationalisation of how research is managed, evaluated and funded.

We have to be careful, though, about slipping into over-optimism based solely on finding technical 'solutions' which seem to realize the underlying philosophy as described here in outline. Those 'solutions' do not change the practice of scholarship just by being there. This realization requires a detour through some fundamental questions about scholarship in order to situate re-designed eprint-based working, of the kind proposed, more reliably in context.

7. Supporting the Scholarship of Integration

In a previous paper, in the course of exploring the impact of 'enterprise' on the humanities, I speculated about the distinctive features of humanities scholarship. I want to rehearse those ideas very briefly in order to add to the above description some sense of the specificity of humanities working and to consider the degree of 'fit' between 'Formations' and such work. The necessity of using pre-prints, if we use them at all, in ways that differ fundamentally from how they are used in a field such as physics, becomes even clearer in this light. It was suggested in this previous paper that one form of good scholarship in the humanities has at least the following informing characteristics:

(1) to understand complexity but not necessarily to simplify it – an assertion open to all sorts of misinterpretation, but all I mean by it is that our point is not to drive towards the elementary particles but to explore the complex structures – of ideas, of values, of narratives – that have been made from them;

(2) to find meaning in the particular instance (rather than in science's 'universal' laws) – hence Michel Foucault's dismissal of any supposedly universalised intellect in favour of 'l'intellectuel specifique';

(3) to maintain reiterative work – in other words to revisit the same problems time and again on the basis of a contention that no solutions are final;

(4) to resist an inflexible subdivision of topics of inquiry – boundaries remain permeable;

(5) to work with different depths of explanation without subsuming one within the other, from surface detail through exploration of the relationships among 'given' elements to the underlying structures that construct those elements – so linguistics might accurately describe a text, historical studies relate it to its context and literary theory unravel the ideological construction of the linguistic surface – mutually informative approaches operating at different depths of curiosity and explanation, none necessarily taking precedence over the others or making stronger truth claims.

I went on to wonder, however, whether these characteristics weren't being threatened increasingly by the fact that 'the levels of explanation have hardened into layers of academic expertise each concentrated at its own "depth" ', even while we are simultaneously discovering that 'in an increasingly complex world it is this

very movement through different depths of explanation that effectively transforms information into knowledge for the curious thinker' (1993, pp. 58–59). This notion was echoed more strongly by Middlehurst and Barnett in their important paper 'Changing the Subject: the Organization of Knowledge and Academic Culture', where they note 'the narrowing of focus and specialization required of researchers' while commenting that:

> ... the problems of the "real" world (as the world beyond academe is commonly described) do not fall neatly into subject-specific boxes, but flow over the edges of the boxes and into unexpected corners. The solutions to these problems may well be found at the boundaries between the subject boxes, emerging out of combining different ingredients in new combinations ... (1994, p. 50).

To draw this set of reflections together again in relation to the current topic, I want to suggest that the features of such 'research' – the particular but boundary-permeating, the reiterative, the movement across levels of explanation, the sensitivity to complexity for its own sake – are in fact features of one form of scholarship and that this form is something that a pre-prints system such as 'Formations', or its more developed descendants, may be especially suited to supporting. Middlehurst and Barnett describe the 'Carnegie classification' of scholarship as consisting of four kinds:

> Knowledge may be differently organised in future (with staff groups stratified accordingly) perhaps to reflect the four areas of "scholarship" represented in the Carnegie Report (Boyer, 1990): the scholarship of discovery, which is most closely related to research at the cutting-edge of a subject; the scholarship of integration which involves synthesizing the results of research within disciplines and creating new knowledge through novel conceptual formulations across subjects; the scholarship of application which involves a deeper analysis of the relationship of theory to practice and the development of a more refined conceptualization of professional practice; and the scholarship of teaching which is concerned with disseminating knowledge and promoting its understanding and its application in many different fields. The Carnegie classification is useful both in highlighting the need for a broader definition of scholarship and in offering legitimacy and recognition to different kinds of academic activity (1994, p. 54).

What has been described here in some detail – from the 'off-stage' work highlighted by Penny et al to the notion of boundary-permeating, synthesizing scholarly work – is largely a matter of the 'scholarship of integration', if such a classification is accepted. This synthesising, connective, dialogic work is often 'off-stage' in relation to the 'performances' staged, in particular, by the 'scholarship of discovery'. But if we acknowledge the former's very real importance, then working procedures such as those explored by 'Formations' take on a new potential significance. The description of the workflow processes supported by the system, as offered above, quite clearly suggests the essentially integrative, inter-connective,

discursive nature of that work. To explain now how such a workflow design was realised requires turning to matters of software design more generally and to the specific technologies that underpin 'Formations'.

8. Using Lotus Notes

First the technologies themselves. 'Formations' is a Lotus Notes application. Notes is a well established 'groupware' standard in the corporate world – widely used by many of the largest blue-chip companies, but largely ignored in academia, not least because until recently it has been both an expensive system to implement in any widespread way (each user requiring dedicated 'client' software) and a system supporting applications already highly tailored to corporate working practices. All that is changing with the ubiquity of networking, the web and Lotus' development of the 'Domino' engine which allows web browser software to be used instead of a dedicated 'client' to access Notes applications – and also, it has to be said, because something like the eLib programme has come along to allow academics to spend a year or more developing a Notes application tailored to academic working practices.

   To understand something of the deep background of Notes, one can do no better than turn to Ray Ozzie, who is often referred to as its 'creator'. We discover that, in fact, Notes began in an educational context. Ozzie was, until recently, president of Iris Associates, who do most of the core development work on Notes for Lotus, the latter now a part of the resurgent IBM empire. The following transcript by the author is from a talk about Notes' past, present and future delivered by Ray Ozzie at the European Technology Conference in Maastricht in April 1996:

> Notes has its roots in a system that was developed in the mid-70s, known as Plato, at the University of Illinois where I and several other of the founders of Iris ... went to school. Plato was a system that originally was intended for computer-assisted teaching of students. It was a centralized mainframe computer with about 10,000 users using terminals worldwide. ... While the primary purpose of Plato was computer-assisted instruction and while it was very pioneering and successful in that area, several other things emerged as side effects because of the widely dispersed and interactive nature of the system.
>
> The first one is communications. Now, basically, because these terminals were located worldwide, researchers who wanted to talk to each other about a lesson that they were preparing for a set of students would need a way of communicating. They started by using the phone and somebody had the bright idea of using the computer to help communication and ... developed a process on the system called Personal Notes. Personal Notes was what today you would refer to as electronic mail.

After Personal Notes started becoming ubiquitous on Plato, someone else ... had an idea that, instead of just sending things to people, perhaps you could share things and he implemented something called Group Notes. And Group Notes was a product that you would refer to today as a conferencing or bulletin board system. And together these programs, along with several others, enabled a tremendous amount of communications on that system; and collaboration – people really got to know each other a lot. I started to establish relationships with people that I worked with for years through these communication facilities and never got to see them. And essentially what developed was what today we would refer to as an on-line community. Again you see this in the internet today ... but in the seventies it was very unique.

It left an incredible impression on me and on the people who worked on the system. ... We graduated and went to work for computer companies and, as we were working on operating systems and eventually as PCs became a business reality in the early eighties, my friends and I began to brainstorm as to how to apply what we learned from Plato.

Since its beginnings in 1984, arising out of those brainstorming sessions, Notes has developed in stature within the software industry. IBM's acquisition of Lotus has been seen as essentially an acquisition of Notes, even though the Lotus Development Corporation has a large portfolio of other software. With its buzzwords of community, communication, collaboration, co-ordination and customization often repeated to explain Notes' distinctive character as an application development environment, the product was perhaps better prepared than any other for the startling impact of the web on notions of how people might use computers in more collaborative ways. Release 4.5 in 1997 was the fully web-enabled version of Notes and is the underlying system on which the first 'build' of 'Formations' runs. In Notes, as in 'groupware' generally, co-ordination means the automation of work processes in which multiple users can participate, usually asynchronously, although live, real-time collaborative tools such as electronic whiteboards are increasingly common as well. Notes provides a set of building blocks or components from which customized work processes can be constructed. These are all document-centred. From the beginning, the Notes developers have assumed unwaveringly that the document is the basis of most business processes. This lends itself readily to adaptation for scholarly work where the emphasis on the document is, if anything, even stronger. It is ideally suited to a pre-prints system of course.

The best part of a year of the 'Formations' project was given over to designing, building and testing a customized Notes application, based around a 'discussion database' template which is a core part of the Notes toolkit. That kit of parts centres on the grouping of documents into 'databases' (which in Notes have a particular form quite unlike the common relational databases with which readers will be familiar) and the 'viewing' of databases in multiple ways. The concept of 'views' – or filtered perspectives onto a collection of documents – is at the centre of Notes'

whole way of working. Rather than extracting separate items of information out of discrete database fields, the views in Notes sort and present a document collection in specified ways. 'Fields' are used for document elements such as author's name, date of creation and so on, but in effect the various views reintegrate the fields in whatever ways have been built into the design. This could be all the documents containing a certain piece of text, all the documents created on a specific date, all the documents by a particular author, etc. These are not unexpected ways of sorting documents but Notes allows highly customized views to be set up as well. A view can be automatically filtered, for instance, because the system 'recognizes' a pre-set category of user (so, for instance, authors can get a different view of their documents than is offered to other users) or documents can be defined in particular ways by their creators (say via a form-based input procedure), allowing user-defined categories of documents to be sifted out of a database. It is difficult to grasp abstractly what this sort of flexibility means in practice. For one thing it goes well beyond anything that can currently be achieved with standard web pages (say HTML plus CGI scripts and Java): the 'Formations' design disciplines this power around a standardised set of procedures that gets recognisably repeated throughout the system so that the user can become quickly familiar with it. Further development of the system is likely to retain this set of core procedures but may present them differently to the user, based on qualitative research into responses to the first 'build' and its six-month run. (At the time of writing, however, that qualitative evaluation is still being undertaken and it would be premature to report any findings.)

9. The Software Design Concept and the Domino Server

Before going on to explain how a web-enabled Notes application of this sort actually works, it is very important to acknowledge a particular debt that this project owes to HyperNews. While several things bear that name, we have been especially influenced by the cross between Usenet News and web publishing attempted by Daniel LaLiberte and colleagues at the National Centre for Supercomputing Applications in Illinois (this has nothing to do with the Macintosh newsreader program or the newsgroup-to-HTML extractor which both share the same name). HyperNews was evaluated in the early stages of the 'Formations' project when the team was still looking for appropriate platforms and, although rejected as insufficiently flexible, it is undeniably true that encountering HyperNews gave the 'Formations' project a basic set of ideas to be developed further.

HyperNews was set up to carry moderated document collections on the web and to allow unmoderated responses to those documents in the form of a 'response tree'. It has three simple but important characteristics: (1) unlike in newsgroup postings, articles and responses never 'expire'; (2) the integrity of the document is preserved – i.e. appended responses are not added to form one longer piece of HTML; (3) each response is numbered to allow easier identification and location.

The 'Formations' project borrows all three characteristics but with some modifications. Our use of 'hosted' areas for all submissions to the pre-prints bank in effect makes the whole system moderated – though by self-appointed moderators. Although submitted documents do not expire, we have introduced a facility which allows a host to 'freeze' a venue, at which point it becomes read-only. This allows hosts to conclude their involvement without leaving things hanging awkwardly. Where HyperNews uses hierarchical numbering to indicate structures of topics and sub-topics (i.e. there would be a document number 1 for each new main topic) we have opted instead for unique document numbering throughout the system and a visual, hyperlink 'thread history' with a 'You are here' tag to indicate structure, in addition to the various views of document sets which effectively offer different structures. Automatically giving a document a unique number provides a more precise means of indentification, and an additional way of referring to or searching for a specific item.

Where HyperNews provides a 'Respond' link at the bottom of every page, 'Formations' uses the phrase 'Add Linked Document' to indicate that linkages may be more general and conceptual than is implied by the narrower notion of a 'response'. So users may choose to thread material together based on perceived rhetorical, topic-based or conceptual connections that are as potentially varied as the material submitted. This is a significant point because it means that we have left it to users to explore what a thread might usefully be in practice. We have not imposed a conversational interpretation in advance. Users may link documents without engaging in any overt discussion via those items – the link may be perceived as a connection at the level of content and shared concerns even if one document does not make any actual reference to another. It may well be, of course, that users will take some time to settle on effective working practices which get the most benefit out of the structure of venues and threads that 'Formations' provides. Despite such marked differences and the much more elaborate web-oriented design of 'Formations', users familiar with HyperNews will quite rightly see a basic similarity in what the systems are trying to achieve – a fusion of the best newsgroup procedures with the hypermedia of the web as an electronic publishing medium. In all honesty, both may be useful experiments on the way towards a more effective long-term solution, although we certainly hope that 'Formations' will allow its users to do real work in the meantime. For its part, HyperNews has been enthusiastically adopted in some quarters as an alternative to newsgroups and remains a system in active development. (Because HyperNews is decentralized – it works across any range of servers on which users have set it up – it is not possible to refer readers directly to a central HyperNews site but the 'background' area of the Formations site will maintain links to HyperNews.)

The underlying Lotus Notes infrastructure is, of course, what sets 'Formations' apart from HyperNews or any other similar experiment. Having explained something of the 'Formations' design and the thinking that lies behind it, an outline

of how a Notes application can be accessed from the web is the last piece of this particular technical jigsaw.

Domino is the name of the technology that has made Lotus Notes into a web server. In fact the 4.5 release of Notes in 1997 adopted the Domino name for the server side of Notes, while the latter became more specifically the client side and related applications. In the most general of terms, it works like this. A Notes application is designed to structure and present a set of document-containing Notes databases. So the 'Formations' venues are essentially Notes databases to be filled with user-submitted documents. The documents are served up by the established Notes server – the technology which used to communicate only with proprietary Notes client software. Now, however, the Domino 'engine' sits alongside the established Notes technology and does some quite clever things.

Domino in effect consists of an HTTP (or web) server combined with what we might loosely think of as a translation engine. If a web browser requests something that is contained in straightforward HTML and associated files (including GIFs, JPEGs, CGI, Javascript and anything else that one might find in an ordinary web page these days), Domino serves it up as any HTTP server would. If, on the other hand, the request is for something contained in a Notes database, Domino looks there and automatically does the HTML translation on-the-fly so that the user is unaware that a Notes application is actually being accessed. This means, in short, that all the power and flexibility of a Notes application is instantly available through a web browser. Some careful interface design is needed in order to ensure that 'standard' web page elements and stuff being fed up from an underlying Notes application actually work well together from a user's point of view. But if this is properly done, most users will be entirely unaware of the delicate balancing act that is being achieved in real time between HTML files and Notes documents. Indeed, an unanticipated problem we encountered during usability testing was that users very quickly started applying criteria and expectations from desktop application software, forgetting that they were still using a web browser to access web pages – albeit pages enhanced by Lotus Notes. This upped the ante on their expectations.

'URL redirection' routes instructions received from users, via buttons, links, etc., to the relevant material in the Notes databases or to functions available in the Notes software itself. This has the disadvantage that it creates extremely long URLs, with some non-standard elements in HTTP terms, but the user does not need to handle these directly so the consequences are minimal in actual practice. One current exception is that non-standard elements such as a question mark, used in Domino URLs, form a barrier to any web search engines trying to look at the contents of pages. This makes most of the material in a Notes database invisible to search engines. In 'Formations', we have carefully implemented metadata where appropriate (based both on the emerging Dublin Core standard and the forms recognised by search engines such as Alta Vista) in order to provide as much 'findable' information as possible by automatically generating metadata information from Notes database fields. The system has its own sophisticated search engine for inter-

nal searching, once a user is actually inside the 'Formations' site. It is anticipated, in any case, that a forthcoming version of Domino will have solved the search engine difficulty.

What we are seeing with the development of Domino is a good example of a larger phenomenon on the web. The first generation of browser software, and indeed the whole concept of browsing, is being supplanted by 'back engine' systems of various kinds. These are intended to overcome the increasingly evident drawback of web publishing – the performance cost of improved functionality. Locating that functionality somewhere behind the scenes and devising systems to get it quickly in and out of the HTML environment, rather than clogging up the latter by trying to do everything there, is a sign of how the web is evolving. The 'Formations' project has built such a 'backstage' system to handle its user-centred interconnection of a pre-prints archive and an online journal. Whether web technology rapidly evolves beyond this specific solution remains to be seen. It will, in any case, have been an experiment worth doing and, in that spirit, we look forward to being able to report more fully on the results from the project's evaluation phase.

Even more important will be the question of whether a 'scholarship of integration' actually exists in the chosen fields and in a form that could take advantage of a pre-prints system such as this. At the end of the day, this may have as much to do with users' willingness or otherwise to work in wired ways as with the software design itself. What this project argues for is recognition of a potentially productive interplay between (a) developing tools and services and (b) exploring the fundamental question of how research in the humanities is determined, in part, by the available procedures for doing research. There may not be much wrong with existing procedures, in the opinion of many, but they are still social and institutional constructs and the networked computer raises alternatives which have not yet been fully explored. Not until we have pursued some of these alternatives a good deal further will we know whether we wish to construct fundamentally new ways of working around the emerging technologies, rather than simply employing them to achieve a few efficiency gains in existing practices.

References

Boyer, E. *Scholarship Revisited*. New Jersey: Carnegie Foundation, 1990.

Fleming, D. " "Enterprise" and the Humanities". *Irish Business and Administrative Research*, 14(1) (1993), 50–64.

Ginsparg, P. "Winners and Losers in the Global Research Village". Paper for UNESCO meeting, Paris, 19–23 February 1996, http://xxx.lanl.gov/blurb/pg96unesco.html.

May, M. "Tuning into Castanets". *Information Strategy*, 2(4) (1997), 42–43.

Middlehurst, R. and R. Barnett. "Changing the Subject: The Organization of Knowledge and Academic Culture". In *Managing the University Curriculum*. Ed. J. Bocock and D. Watson, Buckingham: SRHE/OU Press, 1994, pp. 48–66.

Penny, Alan J. et al. "Collaborative Research in Education: Exploring an Epigenetic Landscape". *Studies in Higher Education*, 19(1) (1994), 21–31.

Pitter, Keiko et al. *Every Student's Guide to the Internet*. San Francisco: McGraw Hill, 1995.