

# 1

---

## **Encountering on the road to Serendip? Browsing in new information environments**

David Bawden

### **Introduction**

This chapter considers the continuing relevance of the ideas of browsing, serendipity, information encountering and literature discovery in the context of the information retrieval (IR) environment of 2010, though its scope extends to the ideas in the broader contexts of information seeking and information-related behaviour. It is based around a selective review of the literature since 1990 and reflection and speculation on the results. The central focus is on questions of how the concept of browsing, serendipity and related ideas have changed in the new IR environment of the web and whether, indeed, they are still meaningful concepts.

By the early 1990s, computerized retrieval was well established in practice and much investigated academically, but the internet was not widely used and the web was not developed. Browsing was a ubiquitous way of finding information, regarded as reasonably well understood. The literature on browsing, serendipity, creative use of information and associated topics up to that point has been thoroughly reviewed (Bawden, 1993a; Chang and Rice, 1993) and this older material will not be systematically covered in this chapter. More recent reviews and literature summaries are provided by Rice, McCreddie and Chang (2001), Foster and Ford (2003), Bates (2007), McBirnie (2008), Anderson (2010), Makri and Warwick (2010), Nutefall and Ryder (2010) and Case (2007), who analyses the concept of browsing and gives numerous examples of studies where its importance as a behaviour has been shown.

Browsing, as a concept, did not necessarily have a positive image among information specialists. I can recall being introduced, while studying for my master's degree, to the views of some senior practitioners who held that if users of a library were found to be browsing to any great extent, then this was a reflection on the library and its staff. Either its catalogues and indexes were

inadequate, or the users had not been properly instructed in their use. Dr Urquhart, a luminary of the British library world at the time, expressed this view well: 'Browsing is no doubt a useful activity for an undergraduate who does not know what he wants but who must have a book immediately. Is it really a sensible activity in a research library? Is the browsing habit but the survival of an ancient custom or does it survive owing to bibliographic laziness or ignorance?' (Urquhart, 1976, 9). It is only fair to say that this view was challenged at the time by several writers (Bawden, 1993a).

It was also held that browsing was something that one did, indeed could only do, in printed materials. Computer searching was good for finding specific information, but was ill-suited and too expensive to allow for a browsing approach.

All these ideas now seem rather quaintly old fashioned. Developments in retrieval systems have meant that searching is now so effective that it is unlikely that anyone would adopt a browsing approach simply because there was no other way to find the desired information. On the other hand, the availability of large quantities of essentially free information, particularly on the internet, has led to the emergence of some new forms of information behaviour, which are often thought of as browsing; as Nicholas et al. (2004, 36) comment of their studies of web users, 'Browsing, time and time again in our studies, has proved to be the main method of obtaining information.' With other forms of media, browsing seems also to have grown, rather than diminished, in importance: it has, for example, been called 'a central search tactic in image retrieval' (Westman, 2009, 74).

Since the early 1990s, the major development in the practice of IR has been the advent of web-based digital information and the ubiquity of the web search engine as the retrieval tool. Web 2.0 tools and practices, such as social tagging, or folksonomy, and the widespread adoption of social media have had a 'second wave' impact on information seeking and retrieval, including browsing (De Meo et al., 2009). They have greatly expanded the amount of searching for information and simply 'coming across it' that takes place. The majority of this is clearly much more informal and more unstructured than heretofore and it may be reasonable to categorize much of it, roughly, as browsing. Indeed access to the web is invariably through a piece of software termed a browser. But does this mean that browsing and encountering cease to be distinct or useful specific concepts; are we, in fact, all browsers now? To answer this question, we need to consider the wider contexts of information seeking and information-related behaviour.

Studies of web behaviour have shown unusual and previously unsuspected patterns of use (Nicholas et al., 2004; Heinström, 2005), and the concepts of 'information encountering' (Erdelez, 2004) and 'serendipitous seeking' (Foster

and Ford, 2003) have been put forward. Although clearly browsing of a kind, they are rather different from the concept as it was understood 20 years ago.

The link between browsing and the use of information systems and services for stimulating innovation and creativity has long been recognized, and for an obvious reason. Search implies, in a way that browsing does not, a clearly defined idea or concept in mind, with information being sought to confirm, support or refute it, or to set it in context. It cannot, as browsing has always been held to do, throw up new ideas or new connections between ideas, which are the *sine qua non* of creativity and innovation.

The idea that IR systems could be relevant as stimuli and aids to creativity had been put forward well before our starting point of the early 1990s; see reviews by Bawden (1986) and Swanson (1990). It has been considered since, again in the new context of a largely digital information environment; see, for example, Schneiderman (2001), Eaglestone et al. (2007) and Swanson, Smalheiser and Torvik (2006). It seems reasonable, therefore, to consider these developments alongside those for browsing, encountering and serendipitous information.

This chapter deals with the central question of how the concept of browsing and the related ideas noted above have changed in the new IR environment of the web, and to what extent they are still meaningful concepts. It is based on a selective review of a very wide literature since 1990, focusing on contributions from information science rather than the other disciplines that have an interest in these topics.

Seven more specific questions are considered:

- Is browsing still a meaningful description of a style of IR and a type of information behaviour?
- How do the newer forms of information behaviour in web environments relate to traditional ideas of browsing?
- How do newer forms of IR system enable serendipity, information encountering, literature discovery and creativity stimulation, compared with earlier forms?
- Do tagging, folksonomy and other Web 2.0 and social media influences on information seeking promote browsing, and how may they best be applied?
- How can we best understand, explain and categorize browsing in the current IR environment?
- Can we helpfully identify individual 'information styles' associated with browsing, encountering and serendipity, and creativity?
- How can IR systems best support browsing, encountering and creativity?

## Browsing

Browsing has generally been distinguished from searching by having a less well-defined end-point, in terms of the information being sought, and a less structured approach to finding items of interest and relevance. Where searching is formal and analytical, browsing is informal and heuristic (Marchionini, 1995).

While it seems clear what browsing is, in practice, precise definitions are few and not generally agreed, perhaps because the term can cover a variety of activities and purposes. Chu (2010) describes it as 'seeking and selecting information by skimming, scanning and other similar activities'. Cove and Walsh (1988) gave perhaps the best general explanation, saying simply that it is 'the art of not knowing what you want until you find it'. The finding may be in itself a Eureka moment, again emphasizing the link between browsing and innovation/creativity.

Equally, a number of terms are used to mean browsing and similar activities. A nice example of this is the index entry from Case (2007, 415):

Browsing

see also Discovering, Encountering, Foraging, Grazing, Navigating, Scanning, Zapping

It has been usual to distinguish between various forms of browsing. *Active* browsing, for example, where some definite information was sought, may be distinguished from *passive* browsing, a general scanning with no particular end in mind. A distinction has also often been made between *directed*, *semi-directed* and *undirected* browsing, again depending on the extent to which some defined information is being sought through browsing, as against a situation of simply scanning sources in the hope of noticing something interesting. This latter categorization was first made in one of the earliest reviews of browsing in the literature (Herner, 1970) and was revised by Vickery (1977) as *purposive*, *semi-purposive* and *capricious*. A somewhat similar categorization of browsing as *systematic*, *exploratory* and *casual* has been made by Marchionini (1995) and as *search/directed*, *general purpose* and *serendipitous/random* (Cove and Walsh, 1987; Catledge and Pitkow, 1995). A number of rather similar categorizations, generally three-way, have been proposed (Bawden, 1993a).

A variation was introduced by Bates (1986), who divided information-seeking behaviour by two criteria – active/passive and directed/undirected – giving a four-way categorization, with the active-directed category corresponding to conventional IR. Bates regarded the active-undirected category as browsing – approximating the purposive form of other categorizations – but not the two passive categories, as no active attempt was being made to find information. It seems more realistic to regard Bates's passive-directed and passive-undirected

as being other forms of browsing, akin to what would later be termed encountering. More recently, Bates (2007) has analysed studies of browsing to suggest that it typically has a four-stage nature: glimpsing a 'field of vision'; selecting an object within it; examining the object; acquiring or abandoning it.

It seems that, looking beneath the variant terminology, a fairly consistent understanding of different forms of browsing had emerged by the early 1990s and that later work has not substantially revised it. The only major theoretical extension has been an analysis presented in a 1995 thesis and reported by Chang (2005), which identified nine patterns of browsing within five general purposes. From outside the information sciences, the 'information foraging' model has gained some interest as relevant to browsing approaches (Pirolli, 2007; Pirolli and Card, 1999; Cronin and Hert, 1995; Jacoby, 2005).

Browsing has, arguably, been given greater credibility by its being included as a component of a number of models of information-seeking behaviour, although, as Foster and Ford (2003) point out, these do not usually extend to unstructured serendipitous encountering. Examples are:

- Wilson's (1999, 2005) model, which includes *passive attention* and *passive search* as components
- Ellis's (1989, 2005) model, which has *browsing* (understood as semi-directed searching) as one component
- Foster's (2004, 2005a, 2005b) non-linear model, which has a process of opening, including activities of *browsing* and *serendipity*
- McKenzie's (2003) model of information practices in everyday-life information seeking, which involves *active scanning* (including semi-directed browsing) and *non-directed monitoring* (serendipitous encountering)
- Chu's (2009) analysis of environmental scanning, which includes *undirected viewing* as one of four modes, involving the scanning of large amounts of information from many and varied sources of information
- Makri and Warwick's (2010) model for 'information for inspiration', which, in addition to *browsing*, includes *exploration* ('finding information either without a predefined goal or to address a vaguely-defined goal').

Further examples are noted by Yuan and Belkin (2010).

Makri and Warwick (2010) also relate stages or processes in information-behaviour models to a 'creativity framework' with associated information tasks (Schneiderman, 2001). The four stages of this framework are:

- collect - essentially the use of conventional information systems, with browsing and visualizing emphasized
- relate - communicating with peers and mentors

- create – which may include serendipitous association of concepts and reflective reviewing of information
- donate – communication of insights gained.

As noted above, browsing – once largely a method of finding information in printed sources which had to be adopted because of the lack of adequate indexes and other tools to assist searching – has not disappeared in the digital environment, despite advances in search functionality. Its continuing popularity is unsurprising, since, as Chu (2010) points out, it is a natural and simple way of finding information, less intellectually demanding than alternatives and needing little in the way of training and practice.

Furthermore, browsing can achieve things which alternative tactics cannot. There are numerous purposes for browsing, some of which are enumerated by Marchionini (1995) and Chu (2010). They include:

- finding information in a context where browsing is the only feasible method
- finding information on topics which are not clearly defined or which are hard to specify exactly, i.e. where the information need is broad and poorly specified
- getting an overview or sample of the information in a collection
- finding items which are similar to, or dissimilar from, those which one has identified
- finding one's bearing in a subject of which one knows little
- selecting the 'right' information from a large collection of 'relevant' material
- looking for inspiration, new ideas, or just something interesting, i.e. allowing for serendipity.

Crucially, even in fully digital environments, browsing may lead to unexpected, serendipitous information discovery precisely because it is much more likely to lead to the finding of unanticipated material. It does this because of its unstructured nature, compared with conventional search, and hence its openness to material of very varied – at the extreme, random – nature.

In addition to categorization by purpose – *why* people browse – there have been categorizations of browsing tactics – *how* browsing is done.

I put forward the idea, in a review more than 15 years ago, that browsing, in practice, usually meant one of three distinct things, though all have been called browsing (Bawden, 1993a):

- finding items similar to one or more items already known, where the

nature of the similarity remains fuzzily defined

- following a predefined categorization, usually hierarchal, through an information space, to identify interesting items
- obtaining an overview of the variation of items in an information space, so as to identify interesting areas of that space.

Marchionini (1995) distinguishes four browsing strategies:

- scanning – looking sequentially through lists of information
- observing – looking in an unstructured way at whatever information presents itself
- navigating – following routes provided by the system
- monitoring – a multi-tasking approach, examining various information sources simultaneously.

Chang's (2005) model presents nine *patterns of browsing* within five *general themes*:

- looking for a specific item
  - situational browsing
  - opportunistic browsing
- looking for things with common characteristics
  - systematic browsing
  - evaluative browsing
  - focus browsing
- keeping up to date
  - monitoring browsing
- learning or finding out
  - indicative browsing
  - preparatory browsing
- goal free
  - invitational browsing.

While all of these strategies and patterns could, in principle, be used in a print-on-paper environment, they have been supported and enhanced by features of digital information systems. These are reviewed by Kowalski (1997), Ruthven (2008) and Chu (2010), and only a few additional example references are given below. These features include:

- ranked results lists and other 'browsable' lists (Jacso, 2005; Hoare and Sorensen, 2005)

- clustering of results (Crestani and Wu, 2006)
- faceted browsing (Perugini, 2010; Fagan, 2010)
- automatic classification and categorization (Tang, 2007; Golub and Lykke, 2009)
- hyperlinks (Mobrand and Spyridakis, 2007)
- directory and classification structures (Perugini, 2010; Koch, Golub and Ardo, 2006)
- highlighting of information elements (Toms, 2002)
- graphical visualization of information (Hoare and Sorensen, 2005; Westerman, Collins and Cribbin, 2005)
- support of integrated searching and browsing (Tang, 2007).

These seem, on the whole, to support the more purposive/directed styles of browsing and navigation, rather than serendipitous/undirected styles.

We should also mention developments in the handling of chemical information. With the advantage of dealing with complete and unambiguous representations of chemical structures, techniques were developed for calculating similarity and dissimilarity between substances and diversity of substance collections (Bawden, 1988, 1993b, Willett, 2008, 2009). These methods are transferable to other forms of information (Wade, Willett and Bawden, 1989; Willett, 2000) and are highly relevant to the browsing process, where what is often wanted is items similar to a given, or dissimilar from it within constraints, or representative items from a collection, spanning its diversity as much as possible.

The development of folksonomies and social tagging, as a complement or alternative both to conventional indexing by controlled vocabularies and to full-text searching, has the potential to improve browsability and also to offer a better chance of serendipitous discovery because of the multiple, often unconventional, perspectives offered (Peters, 2009, ch. 4). Systems relying on social tagging often have interfaces offering specific support for exploratory searching and browsing (Shiri, 2009). However, the sheer volume and diversity of tags can be a problem, and faceting and clustering of tags can aid the browsing process (De Meo et al., 2009; Ding et al., 2009; Spiteri, 2010).

Despite the current overwhelming emphasis on digital materials, support for browsing is still an important feature of printed collections, as witness the appeals by Kirk (2010) for 'browsing collections' in academic libraries and by Hoeflich (2007) for the support of serendipity in legal libraries and archives, the argument of Boyd (2000) for the serendipity-provoking arrangement of physical collections, and the success of the London Library's 'idiosyncratic' classification schema 'based firmly on the principle of serendipity' (Flood, 2007, 51).

Browsing therefore still seems to be a realistic description of an important, and arguably increasingly important, way of finding information that may be

categorized and examined. However, the findings of studies of web behaviour, based mainly on detailed analysis of web-logs combined with more traditional survey methods showing exactly what users are doing, shows a somewhat different picture. Although some aspects of these behaviours may be described as browsing, it is a rather different style of browsing from that discussed above. Nicholas et al. (2004) identified 'bouncing' and 'flicking', a very rapid movement from site to site without any in-depth examination of their content, in contrast to other search styles involving, for example, repeated returning to the same pages or checking of the same information on different sites. Heinström (2005) identified 'fast surfing', a similar style of rapid movement from site to site with only limited examination of information possible, and 'broad scanning', a wide-ranging and generally unstructured examination of a variety of material (discussed further below), by contrast with styles such as 'deep diving', in-depth investigation of a few resources. These results suggest that new forms of browsing have developed to deal with the web environment and that browsing is an important, though not universal, approach.

### **Encountering and serendipity**

Browsing, as discussed above, may be used in a wide spectrum of information seeking, in different ways and for different purposes. One aspect or style of browsing is particularly associated with serendipitous or accidental information access: the random, exploratory, passive scanning aspect, which is commonly associated with the use of information to stimulate innovation and creativity. Indeed, the two things – browsing and creative use of information – have, for many years, often been treated as synonymous (Bawden, 1986). Regarded in earlier years as something of a mystical art, inaccessible to analysis, this style of information acquisition has been examined recently from a variety of viewpoints. These have been as much concerned with the wider information-related behaviour context as with the more specific contexts of information seeking and IR.

The most thorough analysis has taken the form of the concept of 'information encountering', introduced by Erdelez (1997, 1999, 2004, 2005). This alludes to the finding of useful information by accident, also termed serendipitous information acquisition, and to means employed by information users to maximize the chance of this. On an initial basis of interview and questionnaire studies with university students, Erdelez identified four categories of information user, depending on the extent to which they relied on accidental encounters to satisfy their information needs: non-encounterers, occasional encounterers, encounterers and super-encounterers. For the last group, this was the preferred means of finding information and they took active steps to

promote it, going well beyond use of formal information systems, for example, systematically examining reading materials left behind by others. They were also typically active in identifying encountered information of use to others and communicating it to them, strongly emphasizing the personalized nature of this kind of information interaction (Erdelez and Rioux, 2000).

The idea has been developed by Erdelez (2005) into a more detailed model of accidental encountering of interesting information during some other task, involving steps of noticing, stopping, examining, capturing and returning. She also positions it within a more general category of 'opportunistic acquisition of information', allowing for the possibility that encountering is not the only mechanism for such acquisition.

Interestingly, the negative image of browsing mentioned above has persisted, to an extent, to the present day, since Erdelez (1999) comments that super-encounterers were unwilling to speak about it, as they felt theirs was not an 'approved' method of finding information and they might even be ridiculed for admitting they relied on it. Echoing the points made above about the value of printed materials for browsing, many super-encounterers preferred these resources, and in particular avoided the internet as a source. For others who rely less on encountering, the internet is a primary environment for encountering information, along with libraries and bookshops.

This encountering idea has been used in other studies; Pálsdóttir (2010), for example, identifies four categories of information-seeking behaviour in a large-scale study of the ways in which Icelanders find information on health and lifestyle and finds encountering to occur in all of them, but in active or passive ways in different categories. In a small-scale study of health-information seeking among Latino immigrants to the USA, Courtright (2005) found that serendipitous encountering figured strongly in information interactions with positive outcomes; these were generally personal interactions. Essentially the same phenomenon, though termed 'incidental information acquisition', was identified in a study of the everyday-life information seeking of older people (Williamson, 1998). Ross (1999) described 'finding without seeking', whereby those who read for pleasure encountered information that made a significant difference to their lives - a good example of encountering in unconventional sources. Foster and Ford (2003) found that serendipity was widely experienced in the information behaviour of inter-disciplinary academic researchers. Makri and Warwick (2010) found that serendipitous encountering was a significant factor in the information behaviour of architecture students, particularly when seeking inspiration and often in the context of examining images.

Encountering also seems rather similar to the idea of 'broad scanning', noted above as a style of information behaviour identified by Heinström, as an aspect emerging from a factor analysis of detailed data on information behaviour. This

was defined as 'a search pattern characterized by wide and thorough information seeking . . . [practitioners] sought information from many different sources, retrieved information by chance and found it easy to judge information critically' (Heinström, 2005, 237). Accidental information acquisition was much more common in this style than in others.

We should also mention a rather different form of creative use of information, quite distinct from browsing and encountering. This is the specific and formal use of a procedure or algorithm for 'information discovery', finding analogies and relations within information collections, which had not been previously recognized; for a thorough review, see Kostoff et al. (2009). This idea has been mainly promoted and developed for the medical information domain (Swanson, 1990), with a variety of systems developed to identify unappreciated ideas implicit in the literature, generally on the basis of joint occurrence of thesaurus terms in metadata records; see, for example Swanson and Smalheiser (1999), Weeber et al. (2001), Swanson, Smalheiser and Torvik (2006) and Petric et al. (2009). This is a very different concept, in practice if not in principle, from the browsing ideas discussed above, showing the variety of solutions necessary for this diverse area.

### **The personal factor**

A prevalent theme in writings on information for creativity has been the extent to which this has a strongly personal and individual dimension. The kind of information needed, the way it is identified and the means used for browsing and encountering all differ considerably between individuals; some examples have been given above. This is one specific example of the general finding that personality traits can make people more prone to serendipity (Merton and Barber, 2004).

This 'individual differences' aspect has been analysed more precisely in a number of studies of the relation between information-related behaviour and personality factors, learning styles and cognitive styles; for recent reviews, see Heinström (2010) and Bawden and Robinson (2010). A number of such studies have specifically investigated browsing and serendipitous approaches to information, finding links between preferences for such behaviours and ways of implementing them, and personality factors.

Graff (2005), for example, using Riding's Cognitive Styles Analysis (Riding and Rayner, 1998; Ford, 2000), found a distinction between the web-browsing strategies of verbalizer and imager style, the former using more pages set up in a hierarchical arrangement, the latter using more relationally linked pages.

Heinström (2006) has related a tendency for 'incidental information acquisition' to personality factors, showing it to be favoured by those with high

motivation, an energetic personality and positive emotionality, and avoided those inclined to stress, insecurity and low motivation.

The same author (Heinström, 2010) has given a detailed analysis of many aspects of information behaviour in terms of the widely used 'five factor' model of personality (Wiggins, 1996), which considers five personality traits or dimensions: neuroticism, extroversion, openness, agreeableness and conscientiousness. A high level of openness, for example, might be expected to lead to a propensity for encountering and for creative use of information, and there is some evidence to support this. This was not established, however, in one study which specifically examined such factors (Heinström, 2006), suggesting that other factors may play a significant part. A high degree of extroversion, leading to an active and spontaneous style of information seeking, might be expected to lead to a large extent of incidental encountering, and this has been confirmed in one study (Heinström, 2006).

On the basis of such considerations, Heinström (2010) developed a fivefold categorization of 'information attitude'. Three of the attitudes involved a preference for some form of browsing:

- *invitational attitude*: an open personality with an intuitive searching style, leading to a high degree of serendipity
- *exploring attitude*: again an open personality, but relying more on a broad scanning style
- *passive attitude*: low on conscientiousness and relying on a fast surfing style of browsing to find the necessary information with the minimum of input.

The Honey–Mumford conception of learning styles (Honey and Mumford, 1986), which identifies four preferred styles – Activist, Pragmatist, Theorist and Reflector – is held to have a relation to information behaviour because of the strong links between learning and information acquisition (Bawden and Robinson, 2010). They have been used explicitly, as a surrogate information style, to plan training and personalized service delivery in a multinational company (Donnelly and Craddock, 2007). An 'information activist', with a Honey–Mumford activist style, might be expected to be an enthusiastic networker, keen to try many and diverse sources in a rather unstructured way.

Although many of the results of such studies are indicative rather than conclusive, we can see that it is feasible to define and identify information styles as they relate to browsing and creativity.

### **Serendipity systems?**

Nearly 25 years ago, I suggested a number of desirable criteria for information systems focusing on the support of serendipitous information discovery (Bawden, 1986). Of course, innovation and creativity depend on the 'prepared mind', which will generally rely on the input from more conventional information systems. But there may be some specific features which can be recommended to meet this particular purpose. My suggestions at that time were:

- inclusion of peripheral and speculative material
- provision of interdisciplinary information
- representations of information to bring out analogies, patterns, exceptions etc.
- emphasis on browsing facilities
- encouragement of informal channels
- information geared to individual preferences and requirements
- direct involvement of the information user
- appropriate use of new information technologies
- an overall information-rich environment.

The last three have been effectively assured by technological developments in the intervening years; the rest I believe to be still valid and supported by more recent developments and studies.

Since then, there have been a relatively small number of studies specifically addressing this issue: these include those of Ford (1999), Schneiderman (2001), Eaglestone et al. (2007), Makri and Warwick (2010) and Anderson (2010). Many other studies have addressed these issues in part, as noted above.

So what would such a system be like?

First and foremost, it would have to be a personal system with a customizable interface and an adjustable balance between focus and diversity. We have noted above the strong element of individual and personal style in this area, and this must be catered for if systems are to be accepted and successful.

It would have to be based around promoting the possibility of serendipitous information discovery. But, as Ford (1999, 532 and 537) points out, 'such flashes of insight seem far removed from the focused and systematic search for information . . . there would seem to be a considerable mismatch between the element of serendipity that often characterizes creative thinking and what IR systems are essentially designed to deliver'. It is not clear, despite the various suggestions noted above, that we really know what such a system would look like. Some would say that it would look like the browsable shelves of a library of books: either a physical library or its virtual equivalent (see, for example, Beheshti, Large and Bialek, 1996).

As we have seen, browsing is frequently based on analysis of similarity and its converse, dissimilarity. As Ford (1999) points out, similarity recognition is also central to creative thinking, another link between browsing and creative use of information. Any system for these purposes would have to be based around similarity and dissimilarity assessments, in a deeper and more sophisticated way than the widely used probabilistic retrieval systems.

It has been suggested for many years that the kind of information that is most valuable in supporting creativity and innovation may be rather different from that needed for more usual purposes. I suggested some long while ago (Bawden, 1986) that four kinds of information were of most value: interdisciplinary, peripheral, speculative and that focusing on exceptions and inconsistencies.

This amounts to saying that relevance criteria should be relaxed or modified in systems focusing on this purpose, and this suggestion has received support in the literature. For example, Spink and Greisdorf (1997) found that it was partially relevant, rather than highly relevant, items which were most helpful in promoting new ideas, while Ford (1999) argued for a balance between relevance and diversity in support of creativity. Interestingly, Bodoff (2006) found that very different judgements about the relevance of an item were made by searchers who were browsing around an area, as compared with those who were carrying out a focused search, emphasizing the different kind of material valued by each.

Going even further, it has been suggested that random, or even incorrect, information may be valuable in provoking creativity. The former is catered for by undirected forms of browsing and has frequently appeared in internet search systems, seemingly as a 'fun' option rather than anything more serious. The latter is still beyond the pale for most information systems and services, though it is accepted as a brainstorming tool. (An intriguing fictional account of the value of providing deliberately incorrect information [Jones, 1971] was one of the – serendipitous – stimuli for my interest in this topic.)

A system of the kind we are envisaging would need novel search facilities, so distinct from the norm that we might not wish to call them 'search', or even 'browse', at all.

As long ago as 1979, Bates proposed a set of 17 'idea tactics' for searchers, to help generate new ideas and promote creativity. They include such heuristics as:

- *wander* – move among one's resources, being receptive to alternative sources and new search ideas which come into view
- *jolt* and *breach* – think laterally, redefine a problem completely differently, revise ideas of the boundaries of the area being considered

- *focus* and *dilate* – look at an issue or query more narrowly or more broadly
- *change* and *break* – alter an established search pattern, search in a very different way, try different sources etc.
- *notice* – watch for clues in information retrieved that may change one's view of the problem.

These clearly bear strong a relation to browsing approaches discussed earlier and typify the sorts of unconventional facilities needed if information systems are to support creativity.

Ford (1999) argued that such systems must respond to two general types of query: those which find a pattern of relationships, which will integrate seemingly disparate information, and those which find instances of a concept or relationship similar to a starting-point in a variety of contexts and subject areas. As Ford notes, this would need higher-order knowledge representations, in order to allow abstraction across different subjects and contexts and fuzzy matching, considerably beyond what is the norm.

Makri and Warwick (2010) give a detailed set of recommendations for 'creative' electronic resources, with features including: augmented searching; browsing by various criteria; supporting serendipity by integrating materials for diverse sources, categorized and shared in personal ways; and visualization techniques to show a rapid snapshot of diverse information. They emphasize the importance of browsing of visual material for serendipity and inspiration at least for some user groups, as do André et al. (2009). Yuan and Belkin (2010) have shown how a system may be designed to support specific information-seeking strategies, including scanning of resources and results.

To conclude this section, we might consider what basic set of facilities a system would have to offer in order to be regarded as adequate for this purpose. I suggest the following, based on the considerations presented above:

*Show examples of what the system has*

- 1 show a random document  
*given:* a statement of the content of the collection
- 2 show representatives of document clusters  
*given:* a statement of the content of the collection  
*and:* an option to select the number of clusters required.

*Show the structure of what the system has*

- 3 show an alphabetically ordered list of subject terms  
*given:* a statement of the content of the collection  
*and:* an option to choose the level of generality of the terms

- 4 show a classified list of subject terms  
*given:* a statement of the content of the collection.

*Show what the system has, related to what I already know about*

- 5 show a ordered list of documents in descending order of similarity to the input  
*given:* an initial specified document or query
- 6 show an ordered list of documents in descending order of dissimilarity to the input  
*given:* an initial specified document or query.

This requires that the system have:

- a collection of structured records, e.g. bibliographic records
- a field for indexing terms, these terms being taken from a vocabulary with a term hierarchy, i.e. a thesaurus or taxonomy
- a search function, to identify initial documents for functions 5 and 6 (from the above list)
- the capability to display documents indexed with terms selected from a list
- the capability to cluster documents, in such a way that the user can be offered alternative numbers of clusters.

This is a rather basic set of functions and abilities, with an arguably rather uninteresting collection of material. It could be extended, based on the same general functions, to allow for multiple types of materials and varying representations, including those of higher abstraction.

## **Conclusion**

This review has shown that increases in understanding of the issues and developments in practice bring closer the possibility of information systems that will be able to make direct and genuine contributions to innovation and creativity.

But we should remember that the individual and idiosyncratic natures of the use of information for innovation and creativity are likely to elude the best-motivated attempts of system designers to cater for them in any formal sense. Twenty-five years ago, a staff member of the then College of Librarianship Wales related a telling anecdote in a letter to a newspaper (Roe, 1985):

One university lecturer has stated that he makes a practice of noting down the registration numbers of the cars parked outside the library, and consulting the books shelved at those numbers in the library. Rarely does he fail to find something of interest and instruction in this way. Information technology whiz-kids and the new librarians . . . ignore such quirks of human behaviour at their peril.

Building systems and services in the light of the considerations reviewed in this chapter should help whiz-kids and librarians alike to avoid this danger.

## References

- Anderson, T. D. (2010) Kickstarting Creativity: supporting the productive faces of uncertainty in information practice, *Information Research*, in press.
- André, P., Cutrell, E., Tan, D. S. and Smith, G. (2009) Designing Novel Image Search Interfaces by Understanding Unique Characteristics and Usage: paper presented at *INTERACT 2009, Uppsala, Sweden, August 2009*, <http://research.microsoft.com/en-us/um/people/desney/publications/interact2009-tendrils.pdf>.
- Bates, M. J. (1979) Idea Tactics, *Journal of the American Society for Information Science*, **30** (5), 280–9.
- Bates, M. J. (1986) An Exploratory Paradigm for Online Information Retrieval. In Brookes, B. C. (ed.), *Intelligent Information Systems for the Information Society*, North-Holland.
- Bates, M. J. (2007) What is Browsing – Really? A model drawing from behavioural science research, *Information Research*, **12** (4), Paper 330, <http://InformationR.net/ir/12-4/paper330.html>.
- Bawden, D. (1986) Information Systems and the Stimulation of Creativity, *Journal of Information Science*, **12** (5), 203–16.
- Bawden, D. (1988) Browsing and Clustering of Chemical Structures. In Warr, W. A. (ed.), *Chemical Structures; the international language of chemistry*, Springer Verlag, 145–50.
- Bawden, D. (1993a) Browsing: theory and practice, *Perspectives in Information Management*, **3** (1) 71–85.
- Bawden, D. (1993b) Molecular Dissimilarity in Chemical Information Systems. In Warr, W. A. (ed.), *Chemical Structures 2*, Springer Verlag, 383–8.
- Bawden, D. and Robinson, L. (2010) Individual Differences in Information-related Behaviour: towards identifying information styles. In Spink, A. and Heinström, J. (eds), *New Directions in Information Behaviour*, Emerald, forthcoming.
- Beheshti, J., Large, V. and Bialek, M. (1996) PACE: a browsable graphical interface, *Information Technology and Libraries*, **1594**, 231–40.
- Bodoff, D. (2006) Relevance for Browsing, Relevance for Searching, *Journal of the*

- American Society for Information Science and Technology*, **57** (1), 69–86.
- Boyd, B. (2000) Serendipity of the New, *Journal of Rare Books, Manuscripts and Cultural Heritage*, **1** (1), 36–7.
- Case, D. O. (2007) *Looking for Information: a survey of research on information seeking, needs, and behavior*, 2nd edn, Academic Press.
- Catledge, L. D. and Pitkow, J. E. (1995) Characterizing Browsing Strategies in the World-Wide Web, *Computer Networks and ISDN Systems*, **27** (6), 1065–73.
- Chang, S.-J. (2005) Chang's Browsing. In Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behavior*, Information Today, 69–74.
- Chang, S.-J. and Rice, R. E. (1993) Browsing: a multi-dimensional framework, *Annual Review of Information Science and Technology*, **28**, 231–76.
- Chu, C. W. (2009) The Art of Scanning the Environment, *Bulletin of the American Society for Information Science*, **25** (3), available at [www.asis.org/Bulletin/Feb-99/choo.html](http://www.asis.org/Bulletin/Feb-99/choo.html).
- Chu, H. (2010) *Information Representation and Retrieval in the Digital Age* (2nd edn), Information Today.
- Courtright, C. (2005) Health Information-seeking among Latino Newcomers: an exploratory study, *Information Research*, **10** (2), Paper 224, <http://InformationR.net/ir/10-2/paper224.html>.
- Cove, J. F. and Walsh, B. C. (1987) Browsing as a Means of Online Text Retrieval, *Information Services and Use*, **7** (6), 183–8.
- Cove, J. F. and Walsh, B. C. (1988) Online Text Retrieval via Browsing, *Information Processing and Management*, **24** (1), 31–7.
- Crestani, F. and Wu, S. (2006) Testing the Cluster Hypothesis in Distributed Information Retrieval, *Information Processing and Management*, **42** (5), 1137–50.
- Cronin, B. and Hert, C. A. (1995) Scholarly Foraging and Network Discovery Tools, *Journal of Documentation*, **51** (4), 388–403.
- De Meo, P. et al. (2009) Exploitation of Semantic Relationships and Hierarchical Data Structures to Support a User in his Annotation and Browsing Activities in Folksonomies, *Information Systems*, **34** (6), 511–35.
- Ding, Y. et al. (2009) Perspectives on Social Tagging, *Journal of the American Society for Information Science and Technology*, **60** (12), 2388–401.
- Donnelly, A. and Craddock, C. (2007) Information Discovery Stir-fry: information literacy in the commercial sector. In Secker, J., Boden, D. and Price, G. (eds), *The Information Literacy Cookbook*, Chandos, 45–70.
- Eaglestone, B. et al. (2007) Information Systems and Creativity: an empirical study, *Journal of Documentation*, **63** (4), 443–64.
- Ellis, D. (1989) A Behavioural Approach to Information Retrieval System Design, *Journal of Documentation*, **45** (3), 171–212.
- Ellis, D. (2005) Ellis's Model of Information-seeking Behaviour. In Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behavior*,

- Information Today, 138–42.
- Erdelez, S. (1997) Information Encountering: a conceptual framework for accidental information discovery. In Vakkari, P., Savolainen, R. and Dervin, B. (eds), *Information Seeking in Context: proceedings of the international conference on research in information needs, seeking and use in different contexts (August 1996, Tampere)*, Taylor Graham, 412–21.
- Erdelez, S. (1999) Information Encountering: it's more than just bumping into information, *Bulletin of the American Society for Information Science*, **25** (3), [www.asis.org/Bulletin/Feb-99/erdelez.html](http://www.asis.org/Bulletin/Feb-99/erdelez.html).
- Erdelez, S. (2004) Investigation of Information Encountering in the Controlled Research Environment, *Information Processing and Management*, **40** (6), 1013–25.
- Erdelez, S. (2005) Information Encountering. In Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behaviour*, *Information Today*, 179–84.
- Erdelez, S. and Rioux, K. (2000) Sharing Information Encountered for Others on the Web, *New Review of Information Behaviour Research*, **1**, 219–33.
- Fagan, J. C. (2010) Usability Studies of Faceted Browsing: a literature review, *Information Technology and Libraries*, **29** (2), 58–66.
- Flood, G. (2007) The London Library: portrait of a lady, *Information Today*, **24** (1), 1, 51, 54.
- Ford, N. (1999) Information Retrieval and Creativity: towards support for the original thinker, *Journal of Documentation*, **55** (5), 528–42.
- Ford, N. (2000) Cognitive Styles and Virtual Environments, *Journal of the American Society for Information Science*, **51** (6), 543–57.
- Foster, A. (2005a) Nonlinear Information Seeking. In Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behavior*, *Information Today*, 254–8.
- Foster, A. (2005b) A Non-linear Model of Information Seeking Behaviour, *Information Research*, **10** (2), Paper 222, <http://informationr.net/ir/10-2/paper222.html>.
- Foster, A. and Ford, N. (2003) Serendipity and Information Seeking: an empirical study, *Journal of Documentation*, **59** (3), 321–40.
- Foster, A. E. (2004) A Nonlinear Model of Information Seeking Behaviour, *Journal of the American Society for Information Science and Technology*, **55** (3), 228–37.
- Golub, K. and Lykke, M. (2009) Automated Classification of Web Pages in Hierarchical Browsing, *Journal of Documentation*, **65** (6), 901–25.
- Graff, M. (2005) Individual Differences in Hypertext Browsing Strategies, *Behaviour and Information Technology*, **24** (2), 93–9.
- Heinström, J. (2005) Fast Surfing, Broad Scanning and Deep Diving: the influence of personality and study approach on students' information-seeking behaviour, *Journal of Documentation*, **61** (2), 228–47.
- Heinström, J. (2006) Psychological Factors behind Incidental Information Acquisition,

- Library and Information Science Research*, **28** (4), 579–94.
- Heinström, J. (2010) *From Fear to Flow: personality and information interaction*, Chandos.
- Herner, S. (1970) Browsing. In Kent, A., H., Lancour and Nasri, W. Z. (eds), *Encyclopedia of Library and Information Science*, Vol. 3, Marcel Dekker, 408–15.
- Hoare, C. and Sorensen, H. (2005) Information Foraging with a Proximity-based Browsing Tool, *Artificial Intelligence Review*, **24** (3–4), 233–52.
- Hoeflich, M. H. (2007) Serendipity in the Stacks, Fortuity in the Archives, *Law Library Journal*, **99** (4), 813–27.
- Honey, P. and Mumford, A. (1986) *The Manual of Learning Styles*, 2nd edn, Peter Honey.
- Jacoby, J. (2005) Optimal Foraging. In Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behaviour*, *Information Today*, 259–64.
- Jacso, P. (2005) Browsing Indexes of Cited References, *Online Information Review*, **2991**, 107–12.
- Jones, R. F. (1971) Noise Level. In Crispin, E. (ed.), *Best SF Five*, Faber.
- Kirk, T. (2010) What Has Happened to Browsing Collections in Academic Libraries?, *Library Issues: Briefings for Faculty and Administrators*, **30** (5), May 2010.
- Koch, T., Golub, K. and Ardo, A. (2006) Users' Browsing Behaviour in a DDC-based Web Service: a log analysis, *Cataloguing and Classification Quarterly*, **42** (3/4), 163–86.
- Kostoff, R. N. et al. (2009) Literature-based Discovery, *Annual Review of Information Science and Technology*, **43**, 241–85.
- Kowalski, G. (1997) *Information Retrieval Systems: theory and implementation*, Kluwer Academic.
- Makri, S. and Warwick, C. (2010) Information for Inspiration: understanding architects' information seeking and use behaviors to inform design, *Journal of the American Society for Information Science and Technology*, **61** (9), 1745–70.
- Marchionini, G. (1995) *Information Seeking in Electronic Environments*, Cambridge University Press.
- McBirnie, A. (2008) Seeking Serendipity: the paradox of control, *Aslib Proceedings*, **60** (6), 600–18.
- McKenzie, P. J. (2003) A Model of Information Practices in Accounts of Everyday Life Information Seeking, *Journal of Documentation*, **59** (1), 19–40.
- Merton, R. K. and Barber, E. (2004) *The Travels and Adventures of Serendipity*, Princeton University Press.
- Mobrand, K. A. and Spyridakis, J. H. (2007) Explicitness of Local Navigational Links: comprehension, perceptions of use and browsing behaviour, *Journal of Information Science*, **3391**, 41–61.
- Nicholas, D. et al. (2004) Reappraising Information Seeking Behaviour in a Digital Environment: bouncers, checkers, returners, and the like, *Journal of Documentation*, **60** (1), 24–43.
- Nutefall, J. E. and Ryder, P. M. (2010) The Serendipitous Research Process, *Journal of*

- Academic Librarianship*, **3693**, 228–34.
- Pálsdóttir, A. (2010) The Connection between Purposive Information Seeking and Information Encountering: a study of Icelanders' health and lifestyle information seeking, *Journal of Documentation*, **66** (2), 224–44.
- Perugini, S. (2010) Supporting Multiple Paths to Objects in Information Hierarchies: faceted classification, faceted search and symbolic links, *Information Processing and Management*, **46** (1), 22–43.
- Peters, I. (2009) *Folksonomies, Indexing and Retrieval in Web 2.0*, de Gruyter.
- Petric, I., Urbancic, T., Cestnik, B. and Macedoni-Luksic, M. (2009) Literature Mining Method: RaJaLink for uncovering relations between biomedical concepts, *Journal of Biomedical Informatics*, **42** (2), 219–27.
- Pirolli, P. (2007) *Information Foraging Theory: adaptive interaction with information*, Oxford University Press.
- Pirolli, P. and Card, S. K. (1999) Information Foraging, *Psychological Review*, **106** (4), 643–75.
- Rice, R. E., McCreddie, M. M. and Chang, S. L. (2001) *Accessing and Browsing Information and Communication*, MIT Press.
- Riding, R. and Rayner, S. G. (1998) *Cognitive Styles and Learning Strategies*, David Fulton.
- Roe, B. J. (1985) On Line but Out of Mind, *Guardian*, 22 July, readers' letters.
- Ross, C. S. (1999) Finding without Seeking: the information encounter in the context of reading for pleasure, *Information Processing and Management*, **35** (6), 783–99.
- Ruthven, I. (2008) Interactive Information Retrieval, *Annual Review of Information Science and Technology*, **42**, 43–91.
- Schneiderman, B. (2001) *Leonardo's Laptop: human needs and the new computing technologies*, MIT Press.
- Shiri, A. (2009) An Examination of Social Tagging Interface Features and Functionalities: an analytical comparison, *Online Information Review*, **33** (5), 901–19.
- Spink, A. and Greisdorf, H. (1997) Users' Partial Relevance Judgements during Online Searching, *Online and CD-ROM Review*, **21** (5), 271–80.
- Spiteri, L. F. (2010) Incorporating Facets into Social Tagging: an analysis of current trends, *Cataloguing and Classification Quarterly*, **48** (1), 94–109.
- Swanson, D. (1990) Medical Literature as a Potential Source of New Knowledge, *Bulletin of the Medical Library Association*, **78** (10), 29–37.
- Swanson, D. R. and Smalheiser, N. R. (1999) Implicit Text Linkages between Medline Records: using Arrowsmith as an aid to scientific discovery, *Library Trends*, **4891**, 48–59.
- Swanson, D. R., Smalheiser, N. R. and Torvik, V. I. (2006) Ranking Indirect Connections in Literature-based Discovery: the role of medical subject headings, *Journal of the American Society for Information Science and Technology*, **57** (11), 1427–39.
- Tang, M. (2007) Browsing and Searching in a Faceted Information Space: a naturalistic study of PubMed users' interaction with a display tool, *Journal of the American*

- Society for Information Science and Technology*, **58** (13), 1998–2006.
- Toms, E. G. (2002) Information Interaction: providing a framework for information architecture, *Journal of the American Society for Information Science and Technology*, **53**(9), 855–62.
- Urquhart, D. J. (1976) National Lending/Reference Libraries or Libraries of First Resort, *BLL Review*, **4** (1), 7–10.
- Vickery, J.E. (1977), A Note in Defence of Browsing, *BLL Review*, **5** (3), 110.
- Wade, S. J., Willett, P. and Bawden, D. (1989) SIBRIS: the Sandwich Interactive Browsing and Ranking Information System, *Journal of Information Science*, **15** (4/5), 249–60.
- Weeber, M. et al. (2001) Using Concepts in Literature-based Discovery: simulating Swanson's Raynaud-fish oil and migraine-magnesium discoveries, *Journal of the American Society for Information Science and Technology*, **52** (7), 548–57.
- Westerman, S. J., Collins, J. and Cribbin, T. (2005) Browsing a Document Collection Represented in Two- and Three-dimensional Virtual Information Space, *International Journal of Human-Computer Studies*, **62** (6), 713–36.
- Westman, S. (2009) Image Users' Needs and Searching Behaviour. In Göker, A. and Davies, J. (eds), *Information Retrieval: searching in the 21st century*, Wiley, 63–83.
- Wiggins, J. S. (ed.) (1996) *The Five-factor Model of Personality: theoretical perspectives*, Guilford Press.
- Willett, P. (2000) Textual and Chemical Information Processing: different domains but similar algorithms, *Information Research*, **5** (2), Paper 69, <http://InformationR.net/ir/5-2/paper69.html>.
- Willett, P. (2008) From Chemical Documentation to Chemoinformatics: 50 years of chemical information science, *Journal of Information Science*, **34** (4), 477–99.
- Willett, P. (2009) Similarity Methods in Chemoinformatics, *Annual Review of Information Science and Technology*, **43**, 3–71.
- Williamson, K. (1998) Discovered by Chance: the role of incidental information acquisition in an ecological model of information use, *Library and Information Science Research*, **20** (1), 23–40.
- Wilson, T. D. (1999) Models in Information Behaviour Research, *Journal of Documentation*, **37** (1), 3–15.
- Wilson, T. D. (2005) Evolution in Information Behavior Modelling: Wilson's model, in Fisher, K. E., Erdelez, S. and McKechnie, L. E. F. (eds), *Theories of Information Behavior*, Information Today, 31–6.
- Yuan, X. and Belkin, N. J. (2010) Investigating Information Retrieval Support Techniques for Different Information-seeking Strategies, *Journal of the American Society for Information Science and Technology*, **61** (8), 1543–63.