

## **Encountering, keeping and organizing information; maintaining an information collection**

*Catherine C. Marshall, Harry Bruce, Brian Ross, Tiziana Catarci, Doug Gage, David Maier*

We encounter much more information that we can possibly use. We encounter information with unclear potential use or use that is sometime in the future. Despite the explosion of interest in search technologies, it is vital to remember that not all of the information that comes into our purview is actively sought. Instead, much of it is encountered in the course of our everyday activities. Sometimes we come across an interesting article or provocative photo when we're reading the morning news or when we pick up a magazine in the airport. In other cases, a directed search turns up an unexpected result, potentially useful in another context. Occasionally we record a TV show quite by accident or get caught up in a narrative we simply happen upon when we're channel surfing. A friend, colleague, or trusted media source may point us to the URL of a new resource that we intend to explore when we have time. Even in an age of increased personalization, filtering, and ranking, we still have many serendipitous encounters with information in our everyday lives. If we believe this information to have value, utility, or potential relevance to our day-to-day affairs, we may save it or share it with someone else. With that response comes the attendant responsibility: we must somehow manage what we save.

Why do we keep the material we encounter in our everyday reading? Why don't we just assume we'll look for it again when it is needed? The answer may seem obvious: there is a deeply held belief that the stuff we save will be useful at some future point. But that's an oversimplification that conceals important aspects of encountered information. What we save in this way may be emotionally evocative, reminding one of a place or event; we literally expect to re-encounter this material later in our lives and be stirred to remember our personal history. By contrast, some of the encountered information that we save has a much briefer lifespan: we want to be reminded to go to an art gallery opening; we find a review of a restaurant in a city we'll be visiting soon. But much of what the encountered information falls into a vague middle ground of both utility and permanence: we guess that we might want some guidance in building a drainage system for the backyard someday or that we might want to try cooking Kung Pao chicken.

Much of what we save represents a balancing act: the material must seem sufficiently useful, be sufficiently necessary as a reminder, or sufficiently evocative to merit the cognitive overhead of keeping it []. As William James observed more than a century ago (well before the catch phrase of information overload came into fashion), attention is a scarce resource, and cannot be overlooked as an influence in this arena: "[Attention] implies withdrawal from some things in order to deal effectively with others, and is a condition which has a real opposite in the confused, dazed, scatterbrained state which in French is called *distraktion*, and *Zerstreuung* in German." [James, 1890]. Saving encountered information may draw attention from matters at hand, and undoubtedly further scatters our already fragmented attention. Indeed, in field interviews directed at uncovering what people do (and hope to do) with the encountered material that they save, we invariably hear the term "pack rat" used as a dismissive pejorative to describe ineffective keeping strategies, including one's own; attention has been spread too thinly to manage what has been kept.

What is so different about encountered material from information that has been more actively sought is that we are less able to control the circumstances of the encounter: we are not in the task context that so often underlies PIM technologies. We may not have the appropriate filing structure to store the new content; we may find work-related information at home, and material pertaining to our personal interests when we're cruising the Web at work. In fact, encountered information may be literally orthogonal to the task at hand. Furthermore, the topics reflect our *potential* interests – what we might do in the future, hobbies we haven't yet undertaken, software we're contemplating installing, projects we see on the horizon, trips we might take – and may not adhere to our current relatively well-conceived organizational habits, structures, and systems.

A variety of technological solutions have been proposed (and in fact developed) to help us be more effective keepers of the material we encounter in the course of our everyday activities. Many of these solutions are predicated on the trend that storage no longer imposes limitations: it is possible to simply keep everything. This style of personal information management promises us that our attention won't be taxed by decisions about whether the encountered material is valuable or not and where to keep it; instead, all this information will be available to us through increasingly effective means of searching for desired items []. These methods take advantage of the context in which an item was originally encountered to recover it from an enormous personal store of familiar content.

Based on a broad variety of field experiences, both authors have observed that people often don't remember that they already have saved potentially useful or potentially meaningful material when it might be brought to bear on the problem at hand. *You can't search for something if you don't remember that you have it in the first place.* In this case, the material itself works in conjunction with the place and manner that it has been stored to stimulate memory: the material isn't sought; it gains meaning when it is re-encountered. Table 1 summarizes the relationship between re-encounter and other techniques and strategies for finding and keeping personal information.

General technique	Basic assumption	Applicability to finding and keeping
<i>Standard content analysis-based information retrieval.</i>	User has the ability to express and reformulate his/her information needs	Uses the characteristics of the materials themselves (and possibly how they are interconnected)
<i>Stuff I've Seen-style contextual search[]</i>	User knows what he/she is looking for, remembers having seen it, but not where it is	Powerful mode of access when the user remembers the context in which the information was found and kept
<i>Browsing via hierarchical file structure and hypertext links</i>	Structure of collections matches a narrowed area of interest. Multiple items may be sought	Appropriate for some types of documents, especially photos and visual material
<i>Re-encounter[] or Personal Unifying Taxonomies[]</i>	Organizing/manipulating is an essential part of keeping. User doesn't remember what he/she is looking for until he/she sees it	Appropriate for certain types of material, in particular encountered information that's saved without a particular purpose in mind

**Table 1. The relationship between techniques and strategies of finding and keeping information**

What does this mean for PIM tools? What lessons can we draw from our studies of keeping encountered information? Simply allowing encountered stuff to accumulate and relying on search to reclaim it at the right time is insufficient; the activities that surround

keeping are critical for later finding, whether they're associating material with a particular taxonomy or stable sense of place. Will we remember to look for that article about Op Art we tore out of *Life* magazine in primary school when we're trying to recall what we were interested in as a child? No, of course not. Will we remember to look through the box under the bed? Yes, that's where we keep our ancient childhood relics. We might not know exactly what's in there, but we know where it is, and what kinds of things are likely to be there. The serendipity, the utility, and the pleasure of re-encountering what we have saved cannot rely on search alone.