An Overview of Information Management and Knowledge Work Studies: Lessons for the Semantic Desktop

Eyal Oren*
eyal.oren@deri.org

Digital Enterprise Research Institute
National University of Ireland, Galway
Galway, Ireland

Abstract Several recent initiatives aim to provide a Memex-inspired semantic desktop that would integrate with or replace our current physical and electronic desktop. For these semantic desktop initiatives to succeed, we need to consider how people organise their work and use their desktop. If we do not consider this existing work, the semantic desktop might very well suffer from the low adoption rate that is visible in other personal information management solutions [32].

The contribution of this paper is not technical, but presents an overview of relevant semantic desktop literature from the personal information management and human-computer interaction domains. We extract six practical lessons: focus on the individual, forget rigid classifications, follow the links, remember the context, value the power of paper, and keep it simple.

1   Introduction

We live in a knowledge society [57], in which people have access to tremendous amounts of information and communicate with others all over the world. Although most of this information and communication is digital, our computer desktops are still not true information portals, but rather data islands formed by individual applications. A personal information system that brings people and their information seamlessly together is now commonly referred to as the "semantic desktop" [18], an evolution and implementation of Bush's memex [12].

The memex, a visionary device foreseen by Bush, would be a mechanised private library: the memex would allow individuals to store their books, records, and communications, and to retrieve those instantly and flexibly. The memex could be consulted through an index, but more importantly, also through an

associative trail. The memex would augment our memory and help knowledge workers grasp and remember relevant information, fighting the information overload that was already visible then.

The vision of Bush and those that followed [23, 42] relies on only two crucial notions: annotations and links [26]. Annotations allow people to add their opinion to some information, and links relate pieces of information together, leading to what became known as "hypertext" [41]. Hypertext evolved into the current Web as a "portal of information" [6, p. 1], but the vision was always bigger: an information space from "mind to mind" [6, p.169].

Several initiatives aim towards a comprehensive memex-like semantic desktop from, such as Haystack [49], Gnowsis [54], IRIS [14], and MyLifeBits [25]. The semantic desktop will support individuals in their daily activities and augment their memory and intellect. To achieve such support, it is often implied that we have to abandon our wooden desks, including the piles of papers, the sticky notes and the mess of physical things: how otherwise can the machine help us?

But digitising our offices is harder than it seems: although often announced [40], we still cannot get rid of paper [37, 56]. According to Jones and Thomas [32], we have little understanding of how people organise their personal information and how they adopt digital technologies. Based on a small field study in 1997, they report that people predominantly use traditional analog techniques (80% of usage), and that digital tools are always used in combination with traditional technologies. Most common were todo-lists (used by 60% of the subjects), calendars and address books (50%), and diaries (30%). Although the study is relatively old and technology has changed in the meantime, it has been confirmed by later studies [7, 13]; and although email is increasingly used for some personal information tasks [5, 61] it is mostly unstructured [16].

In this paper we therefore analyse, inspired by earlier analyses [31, 53, 58], relevant literature in personal information management, human-computer interaction, and information retrieval. We started by investigating recent issues of relevant conferences[1] and then followed citations in both backward and forward direction. From the literature, the following six categories of challenges emerged:

- a lack of good and thorough activity models of the knowledge worker,
- a misperception of the mental model behind categorising,
- an under-utilisation of the interlinked nature of the information,
- a lack of appreciation to the context of activities,
- an unawareness of the use of physical paper documents,
- and a tendency to be overambitious in the suggested level of support.

In the following sections, we go through these problems one-by-one: we first present relevant literature and then, at the end of each section, give our personal viewpoint regarding the semantic desktop.

---

[1] particularly the ESWC, ISWC, CHI, and SIGIR conferences and the CACM journal.

2 Activity models for knowledge workers

Knowledge is "justified true belief"[44, p. 21]: it is a personal belief justified by information. Knowledge is created by individuals [44, p. 59] through social interactions [44, p. 62–73]. Organisational knowledge management, crucially important in our knowledge society [57], consists of amplifying and crystallising the knowledge of the individual.

Individuals are continuously personally committed to knowledge creation [43, 48]. This personal commitment relies on the intentions and autonomy of individuals [43]. Intention defines the understanding and actions of an individual, autonomy gives him the self-motivation and freedom to absorb and create knowledge. Personal autonomy is crucial for knowledge creation and tools should focus on, and support, this personal autonomy.

In a study into the nature of knowledge workers, Kidd [35] distinguishes three types of office workers: the knowledge worker (creating new knowledge), the communication worker (amplifying information and connecting people), and the clerical worker (managing documents). All workers perform both administrative tasks (repetitive, structured, document-driven) and research tasks (flexible, unstructured, information-driven), but in different proportions. Research tasks usually have no structured procedures and workers have "no clear idea of what the next step will be" [8].

Knowledge workers produce new information by combining an existing body of knowledge [35]. "Ideas are formed in the minds of individuals and are developed in social interactions" [43]. This process is not linear but a continuous interplay between capturing, organising, formalising, and retrieving knowledge.

Due to differences in working styles it is impossible to find a scenario that correctly describes knowledge workers' activities. Given the unstructured and creative nature of the work, document management is difficult since the role of documents is often unclear. Instead, knowledge workers often use the spatial layout of their desk as temporary holding pattern and as contextual clue after interruptions ("what was I doing"). Unstructured work is very hard to support with electronic tools [8]: knowledge workers have a high personal autonomy and they need that freedom for their creative work.

In our opinion, we should indeed focus on knowledge workers, but realise that these workers have mixed characteristics: they do some repetitive administrative work, but also creative unstructured work. Because of the mixture of their activities, and the unstructured nature of most of their work, these workers find little support in existing tools [8]. Given that the semantic desktop aims for a more radical change, it might be even harder to support the knowledge workers.

3 The mental model of categorising

People are bad at categorisation, and they do not categorise information like a librarian would, but by context [37]. The ability to retrieve information depends on clues, people remember associatively. To stimulate those associations, the

retrieval clues have to relate to the same context as the categorised information. Our mental model for categorisation is unknown, but it surely is not a strict hierarchy [24, p. 295–303], but something much more fluid and flexible.

On the other hand, most classifications in computers (e.g. folder structures) are strict hierarchies, leading to many usability problems: "classification problems underpin many usability issues [. . .] Many of the searching and sorting, 'workaround', etc. issues that are often recounted have to do with categorisation and classification systems that are relatively opaque to users. Some of this opacity at least can be attributed to the rigidity of classification schema" [50].

Categorising information is especially difficult for knowledge workers, given the unstructured nature of most of their activities: "categorisation presumes a structure – one that does not yet exist" [8]. Categorising is also strongly related to a document's intended use [36], which is often unknown to a knowledge worker.

Although people greatly prefer category-based search over keyword-based search [63], using strict hierarchies of categorisation worsens the results: strict hierarchies do not work very well for most people [28]. Also, restricting the search interfaces to predefined categories seems detrimental: it seems better to flexibly learn categories from the available data, as in [3, 17].

Bowker and Star [10, p. 10–11] define the ideal classification as complete and consistent segmentations with mutually exclusive categories. But they state that no investigated real-world classification system actually meets these requirements: categories are not exclusive (especially given disagreement or ambiguity), systems are not complete (because of practical, political, or financial reasons), and ad-hoc categories are created at will (because people ignore or misunderstand the classification, or mix together contradictory classification principles). It seems that in reality categorisations are not not rigid, hierarchical, exclusive, but flexible, ambivalent, and overlapping.

Categorisation is cognitively difficult [38, 62] and categorisation schemes are unstable and subjective [9, 10, 20]. Such results have lead to e.g. document and file systems that use metadata properties rather than folder hierarchies for retrieval [20, 27].

On the other hand, desk organisation (both physical and digital desks) serves not only for storing and retrieving documents but also for reminding one of tasks to do [4, 38]. Barreau and Nardi [4] observe that people use location-based search (directory browsing) rather than logical search (based on file metadata). Furthermore, they observe that people use files placements as reminders, placing files in location that were likely to be noticed. This reminding function of file placement would not be possible in a logical or semantic file system which would not have explicit locations and folders.

In our opinion we should, as advocated by Rohmer [53], depart from all predefined structures and not restrict people to confined categorisations or structures. On the semantic desktop users would be encouraged to enrich their documents with metadata, leading to improved search, and removing the need for explicit locations or folders. We should be careful not to restrict the annota-

tions: if metadata must conform to a single hierarchy, we seem to be back to the original problem.

4 The interlinked nature of the information

Semantic Web data is typically highly interconnected and heterogeneous without following one fixed schema [2] and, to a large extent, this will also be the case for semantic desktop data. Any technique for navigating such datasets should therefore be scalable and support graph-based navigation. It should also be generic, not dependent on a fixed schema, and allow exploration of the dataset without prior knowledge of its structure [46].

As Bush [12] noted, "our inaptitude in getting at the record is largely caused by the artificiality of indexing". Keyword search suffices for simple information lookup, but not for higher search activities such as learning and investigating [39]. An exploratory interface on the other hand allows users to find information without prior knowledge of its schema, by visually "slicing and dicing" the available information until something interesting is discovered [47, 55].

Users prefer to find information not by searching and jumping to it, but by orienteering and browsing between related items [59]. People prefer exploration not only to discover information in an unfamiliar domain [55] but also when investigating their own personal information space [15]. We often find it difficult to unambiguously specify what we are looking for and human memory can greatly benefit from context information, enabled by exploratory browsing between related items.

As Karger and Jones [34] put it, "users benefit from being able to orienteer from a document to one of its authors, to a photograph of that person, to a representation of the location where the photograph was taken, to a map of that location, and so on". Especially when users cannot exactly formulate their information need, an exploration technique is necessary that helps users to discover data [30, 60, 63]. And since people are passive in finding information, they greatly benefit from automatic presentation of related information [11, 22, 29, 52], allowing them to further explore interesting information.

In our opinion, the interlinked nature of the information is of crucial importance for people, and should be exploited to a great extent. The visions of Bush [12] and others all depend on two common factors: linking and annotations. The annotations lead to metadata, which allow for more precise searches, but the links are equally important since they enable exploratory navigation.

5 The context of activities

It is well known that context is important in information recall. According to Ranganathan [51], every element of information is classified, recalled, and remembered along the following facets: physical (what), personal (who), temporal (when), spatial (where), and energetical (why).

People categorise information based on its context and recalling therefore also benefits greatly from some contextual information [37]. In a personal classification system, the classification of a document is strongly affected by its intended use and context [36]; that means that we must capture such document context to facilitate retrieval. Dumais *et al.* [21] report that rich contextual information (people names and dates, i.e. the personal and temporal dimensions) provide beneficial cues for retrieval in their personal information system and seem even more important than standard ranking functions.

Bondarenko and Janssen [8] show that document management is strongly related to task management. By grouping and placing their documents around their desktop, people actually perform task management: they reorder priorities of work by moving documents on their desk, the spatial appearance on their desk represents the stage at which the task is. Connecting task management to their physical desktop layout helps people easily continue interrupted tasks: looking at their desktop, they are reminded of important tasks and can directly see their stage (by their spatial-visual layout).

Several results indicate that email is increasingly being used for task management [61, 62]. This can be explained by the close resemblance between email and the physical desktop [8]: email provides an overview of the communication and therefore an overview of the pending tasks, the context of conversations and subsequent tasks is kept next to the documents about the tasks, and classification of documents and emails with undecided roles is optional.

In our opinion, since documents are strongly related to task management, it is crucial to keep tasks within their context, including the communication that governs them and the information that is relevant to them.

6   The use of physical paper documents

The advent of computers in our everyday life has not made paper disappear. Even the Web has only increased the amount of printing we do [56, p. 7]. Use of paper increases because (i) we have access to more and more information, (ii) printing is becoming easier, faster, and better, and (iii) paper is extremely well-suited for authoring, reading, and reviewing documents [56, p. 13–15].

Adler *et al.* [1] found that reading occurs very often in conjunction with writing; it seems that reading stimulates writing (for example, annotating a paper with an opinion). Furthermore, 85% of the reading-writing activities were paper-based, although all subjects used computers that contained essential reading data. This corresponds with earlier findings that paper is very well-suited for reading because of its possibilities for easy annotation, quick navigation between documents, and spatial layout to form a mental picture of the document [45].

Sellen and Harper [56, p. 16] explain that "digital technologies shift the point at which paper is used rather than replacing it". Digital technology enables rather than replaces traditional technologies [32] and "understanding the way people currently work in both paper and digital domains allowed us to find basic needs in personal document management" [8].

According to Kidd [35], knowledge workers write things down not for storing them (which could easily be replaced by computers), but because the writing itself is important to them (which can still only be done on paper). Knowledge workers themselves contain the valuable information, not the documents they read or write: they write because the act of writing catalyses their thoughts. The physical marks on a document (the notes in the margin, the scribbles in the text, the diagram drawn on it) and on the desktop (the piles of papers, the little notes) are crucial, and cannot yet be adequately captured and represented digitally.

As Sellen and Harper [56, p. 22] say: "developers of new technologies need to accept and recognise the importance of paper in work-related activities and in organisational processes, and turn their attention from systems that replace paper to those that integrate paper and the electronic world."

In our opinion, more attention should be given to the importance of paper: it has some significant advantages over digital technologies. A good example is the recent work of Stäudel *et al.* [58], who combine traditional paper-and-pen with digital technology, allowing people to annotate printed papers with a stylus pen, automatically translating the annotations into the digital document.

7 A tendency to be over-ambitious

Kidd [35], after a field study of knowledge workers, warns against tools that "understand" the information being processed: only the knowledge worker himself can give meaning to the document. Instead, we should concentrate on capturing and reproducing all physical marks (scribbles, notes, arrows, drawings) on the paper. "The dream of providing an electronic encyclopedia in the sky only satisfies those cases where someone can predict what they need to know" [35].

Adoption of new technology for personal information management is low [32] although users express great needs for better solutions. The low adoption might be explained by the unawareness of many users with features of their existing tools (most people are even unaware of the automatic email filtering [8]), and by the diverse nature of people's needs. For example, notwithstanding desktop search engines[2] that can retrieve desktop information with simple keywords, people still want to categorise their data into folders and browse them occasionally [33]. People use folders not only to later retrieve their information (which could be replaced by searching) but also to remind them of tasks, to decompose problems, and to plan their work [4, 33, 61]: something that cannot be replaced by search.

We know that people use tools in a different way than we expect them to, and that most people do not understand the power of their tools. We should therefore lower our ambitions and make tools simple and accessible rather than advanced and complicated, including the absolute ban [53] of technical computer science terms from the user interface.

---

[2] such as Apple Spotlight, Google Desktop, or Beagle

8 Conclusion

Several initiatives aim towards a comprehensive memex-like semantic desktop. Without denying the technical challenges in the semantic desktop, many results in the literature indicate the severe non-technical challenges ahead. As Lansdale [37] explains, most problems in personal information management are about recall, recognition, and categorisation: psychological issues. We have discussed these challenges grouped into six categories and we summarise them as follows:

- *focus on the individual*: give individual users incentive and benefit before focusing on the social network,
- *forget rigid classifications*: leave users their freedom and do not constrain them into rigid schemas (i.e. do not use constraint-based semantics),
- *follow the links*: exploit the interlinked nature, do not rely only on search, and allow people to associate freely,
- *remember the context*: work to understand the notion of context, capture it together with the information and use it to enhance recall and understanding,
- *value the power of paper*: find ways to use the computer together with paper, or prepare for non-exclusive usage (people will not easily give up paper),
- *keep it simple*: focus on simply capturing and representing the things that the user wants to store, before doing any reasoning with it.

It will still take some time before we truly understand how people organise their personal workspace, and only then can we truly support them with a memex. In the meantime, let us beware of the myth of the paperless office [56].

References

[1] A. Adler, A. Gujar, B. L. Harrison, K. O'Hara, *et al.* A diary study of work-related reading: Design implications for digital reading devices. In *CHI*, pp. 241–248. 1998.

[2] R. Angles and C. Gutierrez. Querying RDF data from a graph database perspective. In *ESWC*, pp. 346–360. 2005.

[3] P. Anick and S. Tipirneni. Interactive document retrieval using faceted terminological feedback. In *HICSS*. 1999.

[4] D. Barreau and B. Nardi. Finding and reminding: File organization from the desktop. *SIGCHI Bulletin*, 27(3):39–43, 1995.

[5] V. Bellotti, B. Dalal, N. Good, P. Flynn, *et al.* What a to-do: Studies of task management towards the design of a personal task list manager. In *CHI*, pp. 735–742. 2004.

[6] T. Berners-Lee. *Weaving the Web – The Past, Present and Future of the World Wide Web by its Inventor*. Texere, 2000.

[7] A. Blandford and T. R. G. Green. Group and individual time management tools: What you get is not what you need. *Personal and Ubiquitous Computing*, 5(4):213–230, 2001.

[8] O. Bondarenko and R. Janssen. Documents at hand: Learning from paper to improve digital technologies. In *CHI*, pp. 121–130. 2005.

[9] G. C. Bowker and S. L. Star. Building information infrastructures for social worlds – the role of classifications and standards. In T. Ishida, (ed.) *Community Computing and Support Systems*, pp. 231–248. 1998.

[10] G. C. Bowker and S. L. Star. *Sorting Things Out – Classification and its Consequences*. MIT Press, 1999.

[11] J. Budzik, K. Hammond, and L. Birnbaum. Information access in context. *Knowledge-Based Systems*, 14(1–2):37–53, 2001.

[12] V. Bush. As we may think. *The Atlantic Monthly*, 176(1):101–108, 1945.

[13] C. Campbell and P. Maglio. Supporting notable information in office work. In *CHI*, pp. 902–903. 2003.

[14] A. Cheyer, J. Park, and R. Giuli. IRIS: Integrate. Relate. Infer. Share. In [19].

[15] E. Cutrell and S. T. Dumais. Exploring personal information management. *Commun. ACM*, 49(1):50–51, 2006.

[16] L. Dai, W. G. Lutters, and C. Bower. Why use memo for all? restructuring mobile applications to support informal note taking. In *CHI*, pp. 1320–1323. 2005.

[17] W. Dakka, P. Ipeirotis, and K. Wood. Automatic construction of multifaceted browsing interfaces. In *CIKM*. 2005.

[18] S. Decker and M. R. Frank. The networked semantic desktop. In *WWW Workshop on Application Design, Development and Implementation Issues in the Semantic Web*. 2004.

[19] S. Decker, J. Park, D. Quan, and L. Sauermann, (eds.). *Proc. of the Semantic Desktop Workshop at the ISWC*. 2005.

[20] P. Dourish, W. K. Edwards, A. LaMarca, J. Lamping, *et al.* Extending document management systems with user-specific active properties. *ACM Trans. Inf. Syst.*, 18(2):140–170, 2000.

[21] S. Dumais, E. Cutrell, J. Cadiz, G. Jancke, *et al.* Stuff I've seen: A system for personal information retrieval and re-use. In *SIGIR*, pp. 72–79. 2003.

[22] S. Dumais, E. Cutrell, R. Sarin, and E. Horvitz. Implicit queries (IQ) for contextualized search. In *SIGIR*, p. 594. 2004.

[23] D. C. Engelbart. Augmenting human intellect: A conceptual framework. Tech. Rep. AFOSR-3233, Stanford Research Institute, 1962.

[24] M. W. Eysenck and M. T. Keane. *Cognitive Psychology: A Student's Handbook*. Psychology Press, 2000.

[25] J. Gemmell, G. Bell, and R. Lueder. Mylifebits: a personal database for everything. *Commun. ACM*, 49(1):88–95, 2006.

[26] J. Gemmell, G. Bell, R. Lueder, S. Drucker, *et al.* MyLifeBits: fulfilling the Memex vision. In *MULTIMEDIA'02*, pp. 235–238. 2002.

[27] D. K. Gifford, P. Jouvelot, M. A. Sheldon, and J. W. O. Jr. Semantic file systems. In *Proc. of 13th ACM Symposium on Operating Systems Principles*, pp. 16–25. 1991.

[28] M. A. Hearst. Clustering versus faceted categories for information exploration. *Comm. of the ACM*, 46(4), 2006.

[29] M. Henzinger, B.-W. Chang, B. Milch, and S. Brin. Query-free news search. In *WWW*, pp. 1–10. 2003.

[30] E. Hyvönen, S. Saarela, and K. Viljanen. Ontogator: Combining view- and ontology-based search with semantic browsing. In *Proceedings of XML Finland*. 2003.

[31] Indratmo and J. Vassileva. Human and social aspects of decentralized knowledge communities. In [19].

[32] S. R. Jones and P. J. Thomas. Empirical assessment of individuals' 'personal information management systems'. *Beh. & Inf. Techn.*, 16(3):158–160, 1997.

[33] W. Jones, A. J. Phuwanartnurak, R. Gill, and H. Bruce. Don't take my folders away! organizing personal information to get things done. In *CHI*, pp. 1505–1508. 2005.

[34] D. R. Karger and W. Jones. Data unification in personal information management. *Commun. ACM*, 49(1):77–82, 2006.

[35] A. Kidd. The marks are on the knowledge worker. In *CHI*, pp. 186–191. 1994.

[36] B. H. Kwaśnik. How a personal document's intended use or purpose affects its classification in an office. In *SIGIR*, pp. 207–210. 1983.

[37] M. Lansdale. The psychology of personal information management. *Applied Ergonomics*, 19(1):55–66, 1988.

[38] T. W. Malone. How do people organize their desks? Implications for the design of office information systems. *ACM Trans. Inf. Syst.*, 1(1):99–112, 1983.

[39] G. Marchionini. Exploratory search: From finding to understanding. *Comm. of the ACM*, 49(4), 2006.

[40] N. Negroponte. *Being Digital*. Vintage, 1996.

[41] T. H. Nelson. A file structure for the complex, the changing and the indeterminate. In *Proc. of the ACM National Conference*. 1965.

[42] T. H. Nelson. *Dream Machines/Computer Lib*. Mindful, 1974.

[43] I. Nonaka. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1):14–37, 1994.

[44] I. Nonaka and H. Takeuchi. *The Knowledge-Creating Company*. Oxford University Press, New York, 1995.

[45] K. O'Hara and A. Sellen. A comparison of reading paper and on-line documents. In *CHI*, pp. 335–342. 1997.

[46] E. Oren, R. Delbru, and S. Decker. Extending faceted navigation for RDF data. In *ISWC*. 2006.

[47] C. Plaisant, B. Shneiderman, K. Doan, and T. Bruns. Interface and data architecture for query preview in networked information systems. *ACM Trans. Inf. Syst.*, 17(3):320–341, 1999.

[48] M. Polanyi. *Personal Knowledge: Towards a Post-Critical Philosophy*. Routledge & Kegan Paul Ltd, London, 1958.

[49] D. Quan, D. Huynh, and D. R. Karger. Haystack: A platform for authoring end user semantic web applications. In *ISWC*, pp. 738–753. 2003.

[50] D. W. Randall. Geoffrey Bowker and Susan Leigh Star, Sorting things out: Classification and its consequences – Review. *Computer Supported Cooperative Work*, 10(1):147–153, 2001.

[51] S. R. Ranganathan. *Elements of library classification*. Bombay: Asia Publishing House, 1962.

[52] B. J. Rhodes. *Just-In Time Information Retrieval*. Ph.D. thesis, MIT, 2000.

[53] J. Rohmer. Lessons for the future of semantic desktops learnt from 10 years of experience with the IDELIANCE semantic networks manager. In [19].

[54] L. Sauermann. *The Gnowsis – Using Semantic Web Technologies to build a Semantic Desktop*. Master's thesis, Vienna University of Technology, 2003.

[55] m. c. schraefel, M. Wilson, A. Russell, and D. A. Smith. mSpace: Improving information access to multimedia domains with multimodal exploratory search. *Comm. of the ACM*, 49(4):47–49, 2006.

[56] A. J. Sellen and R. H. Harper. *The Myth of the Paperless Office*. MIT Press, Cambridge, MA, USA, 2003.

[57] P. M. Senge. *The Fifth Discipline*. Currency, 1994.

[58] M. Stäudel, B. Klein, and S. Agne. Pen-based acquisition of real world annotations for document information spaces. In [19].

[59] J. Teevan, C. Alvarado, M. S. Ackerman, and D. R. Karger. The perfect search engine is not enough: A study of orienteering behavior in directed search. In *CHI*, pp. 415–422. 2004.

[60] R. W. White, B. Kules, S. M. Drucker, and mc schraefel. Supporting exploratory search. *Comm. of the ACM*, 49(4), 2006.

[61] S. Whittaker, V. Bellotti, and J. Gwizdka. Email in personal information management. *Commun. ACM*, 49(1):68–73, 2006.

[62] S. Whittaker and C. Sidner. Email overload: exploring personal information management of email. In *CHI*, pp. 276–283. 1996.

[63] K.-P. Yee, K. Swearingen, K. Li, and M. Hearst. Faceted metadata for image search and browsing. In *CHI*. 2003.