"Tell me a Story"
Issues on the Design of Document Retrieval Systems

Daniel Gonçalves, Joaquim Jorge

Computer Science Department, Instituto Superior Técnico, Av. Rovisco Pais
1049-001 Lisboa Portugal
djvg@gia.ist.utl.pt, jorgej@acm.org

Abstract. Despite the growing numbers and diversity of electronic documents, the ways in which they are cataloged and retrieved remain largely unchanged. Storing a document requires classifying it, usually into a hierarchic file system. Such classification schemes aren't easy to use, causing undue cognitive loads. The shortcomings of current approaches are mostly felt when retrieving documents. Indeed, how a document was classified often provides the main clue to its whereabouts. However, place is seldom what is most readily remembered by users. We argue that the use of narratives, whereby users 'tell the story' of a document, not only in terms of previous interactions with the computer but also relating to a wider "real world" context, will allow for a more natural and efficient retrieval of documents. In support of this, we describe a study where 60 stories about documents were collected and analyzed. The most common narrative elements were identified (time, storage and purpose), and we gained insights on the elements themselves, discovering several probable transitions. From those results, we extract important guidelines for the design of narrative-based document retrieval interfaces. Those guidelines were then validated with the help of two low-fidelity prototypes designed from experimental data. This paper presents these guidelines whilst discussing their relevance to design issues.

1 Introduction

In recent years, computer hardware has become increasingly cheap. As a consequence people tend to use computers not only at work, but also at home. Furthermore, PCs are losing their dominance and laptops or PDAs are ever more commonly used in all settings. Moreover, the advent of ubiquitous, pervasive computing will only increase the number of devices available from which documents can be handled. Because of this trend, more and more often users edit and store related documents in different locations. Thus, new tools that allow users to more easily find a specific piece of information, regardless of where they are, or to visualize the Personal Document Space (PDS) as a whole will soon become imperative. One of the major challenges of HCI in the upcoming years will revolve around these issues, as pervasive computing becomes a reality [1] [2] [13].

The biggest problem with current hierarchic organization schemes is that they continuously require users to classify their documents, both when they are named and when they are saved somewhere in the file system. Such approaches force users to fit their documents into specific categories. Also, since users know that a good classification determines their ability to later retrieve the documents, classifying ever increasing numbers of documents becomes a painful task, causing undue cognitive loads while choosing the category in which each document should be placed.

This was first recognized by Thomas Malone [12] on his groundbreaking work where two main document organization strategies were identified: *files* and *piles*. On files documents are classified according to some criteria, whereas Piles are *ad-hoc* collections of documents. The latter were shown to be more common due to the difficulties inherent to the classification task. Nowadays, similar results are found not only for documents on computers but also for other applications in which hierarchic classification has become the primary information organization strategy. Such is the case of email, where it was found [4] that most users' inboxes are often filled with large numbers of messages, given the difficulty and reluctance in classifying them into other folders. However, despite the apparent lack of classification, the same study found that the users think it easier to find email messages in the inbox than finding a document on the file system. This is because email messages are associated to useful information elements, ranging from the sender of a message to when it was sent and what messages were received at about the same time. This causes some people to overload their email tools to work as To Do lists or to maintain sets of unread documents [14]. Even considering that email tools were not designed with those ends in mind, the trade-off in relation to traditional applications seems to be positive.

This shows the importance of information other than a name or classification for retrieving documents. Users more readily remember other contextual, real world, information, rather than some arbitrary classification made months or years ago. Several works try to make use of such additional information to help users retrieve their documents. One of the first was Gifford's *Semantic File Systems* [7], where properties are associated to documents, either automatically inferred (from email headers, for instance), or explicitly created by users. Documents can then be found in 'virtual-folders', whose contents are determined by queries on the defined properties. This work inspired others such as Dourish et al's *Placeless Documents* [4] and Baeza-Yates et al's *PACO* [3], where enhancements for features such as support for multiple document locations and management of shared documents can be found. Other works, such as Freeman and Gelernter's *Lifestreams* [6] recognize the importance of temporal information, presenting all documents in an ordered stream.

Although alleviating some of the problems users must face, new problems appear with those approaches. Property-based systems require users to handle (and remember) arbitrary sets of properties. Furthermore, each property is an isolated piece of information with no apparent relation to the others. Temporal-based approaches disregard other kinds of information. An integration of the several relevant information elements that could help users in finding their documents is lacking. The most natural way in which users can convey that information to someone is in the form of stories or narratives. Humans are natural-born storytellers. From early times have stories been told, first in oral tradition and later in written form. Elements in a story do not appear

separately but as part of a coherent whole. The relations between those elements make the story easier to remember. An interface that takes advantage of those abilities and allows users to tell a story describing a document in order to retrieve it will allow for a more natural and efficient interaction.

The design of such an interface should take into account not only the most common and expected elements in a narrative, but also how they inter-relate. This will allow it to know what shape the stories might have, what will come up next at any given point in the narrative, and what information users might remember even if it wasn't volunteered in the first place, resulting in a dialogue that is natural, informative and not awkward. Thus, it is important to find out exactly what document-describing stories are like.

To correctly address the aforementioned challenges, we performed a set of interviews where several stories describing documents were analyzed. This allowed us to extract patterns for common narrative elements and ways in which they are used. Some recurrent story structures were found. From those, we extracted a set of guidelines that systems for narrative-based document retrieval should follow to correctly address the users' needs. Ultimately, we envision the design of a system that continuously gathers information about the users' interactions with their documents and whose narrative-based interface is able to extract vital information about the documents from the users, allowing the documents to be retrieved.

We'll start by describing how the study was conducted. Next, we'll analyze the results thus obtained. Then we will present the design guidelines, and how they were validated. Finally, we'll discuss the main conclusions and possible future work on the area.

2 Procedure

With this study, we tried to answer two main research questions: (1) in document-describing stories, *what are the most common elements?* (2): *how do they relate to form the story?* To find the answers, we conducted 20 semi-structured interviews. The volunteers were interviewed at a time and place of their choice (previously arranged), often in their own offices or other familiar environments, to set them at ease. We asked for their consent in recording the interviews.

Of the 20 subjects we interviewed, 55% were male and 45% female, with ages ranging from 24 to 56. Academic qualifications spanned all levels, from high-school to PhDs. Their professions were also fairly diversified: Computer Science Engineers, High-School Teachers, Law Students, economist, social sciences professor, etc. This accounts for the wide range of computer expertise we found, from programming skills to sporadic use of common applications (such as Microsoft Word). Overall, we feel we collected data from a diverse sample that won't unduly bias the results.

After explaining the study to the subjects, they were asked to remember specific documents from three different classes and to tell stories describing them. Those classes were: Recent Documents on which the user worked on in the past few days or weeks; Old Documents, worked on at least a year ago; and Other Documents, not

created by the user. They were chosen to allow us to evaluate the effect that time might have on the nature and accuracy of the stories (regardless of their correctness, since real documents were not available to validate them), and to find if stories are remembered differently for documents not created by the users themselves, since their interaction with those documents was different. We didn't provide actual documents to be described because that would require the interviewer to have access to the subject's computer in order to choose those documents. Previous experiments [8] showed that users are reluctant to allow that kind of intrusion. Also, preliminary test interviews demonstrated computers to be distractive elements during the interviews, resulting in stories of poor quality. Furthermore, asking interviewees to remember the documents to be described better mimics the situations in which they might want to find a document in everyday life.

For each document, the interviewees were instructed to "tell the story of the document", and to recall all information they remembered about it. It was specifically recommended that information besides the one resulting from the interaction with the computer itself was important. Additional questions regarding several expected elements were posed in the course of the interview. They were asked only when the interviewees seemed at a loss of anything else to say, to see if some other information could still be elicited from them, or whenever they had started talking about some unrelated subject and we needed to make them go back to describing the document at hand. Three test interviews were conducted to tune and validate this procedure

Stories usually took five minutes to be told. Their transcripts averaged two to three plain text pages, although some users told longer stories. A typical story might start like this translated excerpt from a real interview:

Interviewer: So, now that you have thought of a document, please tell me its story…
Interviewee: It's a paper I had sent to my supervisor. We had sent it to a conference some time ago. It was rejected… meanwhile I had placed the document on my UNIX account…

3 Interview Analysis

All interviews were subjected to a Contents Analysis [15]. We coded for several elements we expected to find in the stories (Table 1). New elements could be considered if required during the analysis process. As it turned out, no new elements were necessary after the initial encoding. Since the users were free to tell their stories as they chose, we're fairly confident that we considered all relevant elements.

Table 1. Story Elements

Time	Place	Co-Authors	Purpose
Author	Subject	Other Docs.	Personal Life
World Events	Doc Exchanges	Doc Type	Tasks
Storage	Versions	Contents	Events
Name			

Contents analysis is often performed by defining a coding dictionary which contains, for each specific word or expression that might occur in the interviews, the class to which it belongs [11]. In our domain such a dictionary could contain an entry stating that the occurrence of the word "hours" is a reference to a "Time" element. This approach would allow the encoding to be made automatically. However, it requires the researcher to anticipate all relevant words or expressions that might appear. This was impossible in our experiment since the subjects were free to say whatever they chose about documents previously unknown to us. Hence, no coding dictionary was used. Instead, we conducted the coding manually with the help of a set of heuristic rules that clearly define what should belong to each category, considering not only specific words or expressions but also their meanings. We coded for frequency rather than for occurrence, since frequency can give us an estimate of the relative importance of the elements in terms of the amount of information of each kind in the stories. Also, we took notice of what elements were *spontaneous* (proposed by the interviewees) and *induced* (promptly remembered by the interviewee after a question or suggestion from the interviewer). We also considered that not knowing something is different from knowing something not to have happened. An element was recorded only in the latter case. For instance, some users remembered that a document had no co-authors, while others couldn't remember if that was the case or not.

We also performed a Relational Analysis [15] to estimate how the several elements relate in the story. We considered the strength of all relationships to be the same. The direction of the relationships was given by the order in which the elements appear in the story. The signal of a relationship (whether two concepts reinforce or oppose each other) wasn't considered since it isn't relevant in this case. This allowed us to create a directed graph whose nodes are story elements, arcs represent the relationships between those elements, and arc labels contain the number of times the corresponding transition was found. No transition was considered when the destination element was induced, since in that case no real connection between the elements existed in the interviewee's mind.

4 Results

Overall, we collected and analyzed 60 different stories, 20 for each document type. We produced not only quantitative results relating to the relative frequencies of the different story elements and transitions between those elements, but also qualitatively analyzed the stories' contents. We took care to compare stories for different document kinds. Finally, we were able to infer archetypical stories about documents. Several statistical tests were used whenever relevant. In what follows, all quantitative values are statistically significant to 95% confidence. More results can be found in the experiment's technical report [9].

4.1 Story Length

We found stories to be 15.85 elements long, on average (std. dev.=5.97). The fairly large standard deviation accounts for the difference between stories relating to documents created by the user and those of others, with average lengths of 17.7 and 12.15, respectively. From this we conclude it is easier to remember information about your own documents. There is no significant correlation between story length and subject age. Although the interviewees were relatively young, this is a surprising result. Cognitive problems arise with age and some trend could already be visible. As to gender, we observed that women tend to tell longer stories than men (16.81 vs. 14.67 elements), suggesting it is easier for them to remember potentially relevant information.

4.2 Transition Numbers

Since no transition is recorded between two elements if the second is induced, the ratio between the numbers of transitions and story elements provides a good estimate of how in control of their stories the interviewees were. On average, 47% of stories were spontaneous, regardless of document type and interviewee gender. A significant but weak (0.22) correlation was found in relation to age: older users are marginally more in control of their stories, allowing for less interference from the interviewer.

4.3 Story Elements

The most common overall story elements were **Time**, **Place**, **Co-Author**, **Purpose**, **Subject**, **Other Documents**, **Exchanges**, **Type**, **Tasks**, **Storage** and **Content** (Figure 1). Some elements appear more than once in a story, showing that users sometimes provide additional information to reinforce or clarify them. The least mentioned elements were those pertaining information about **Authors**, **Personal Events**, **World Events**, **Versions**, **Events**, and **Names**. This shows how those elements are harder to remember or considered less important by the users.

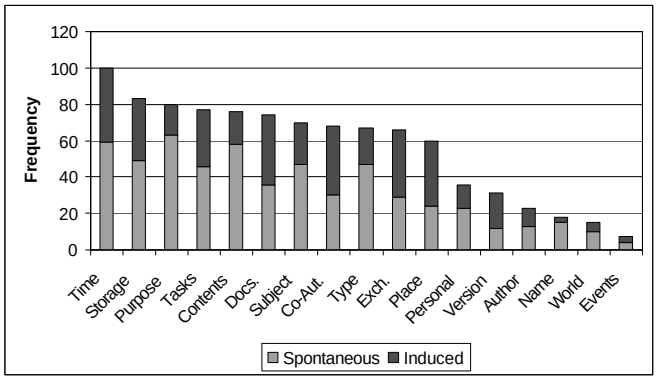

Figure 1 – Overall Element Frequencies

Figure 2 shows that element frequencies for Recent and Old Documents seem to follow similar distributions. Statistically, we found significant differences only for the **Subject** element. When a document is recent, users tend to reiterate it on their narratives, since they easily remember more relevant details.

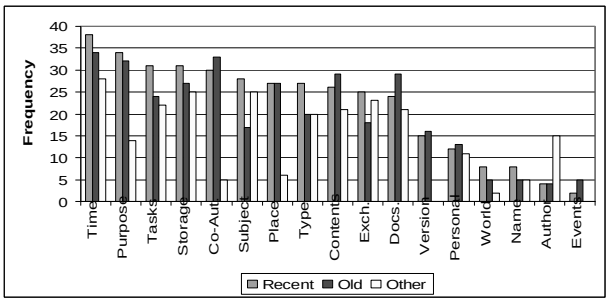

Figure 2 – Element Frequencies by Document Kind

Larger differences could be found among documents created by the user and those of others. The most noteworthy differences are related to the frequencies of **Place**, **Co-Authors**, **Purpose, Author**, and **Version**. The differences in Author and Version are easy to explain: when the user itself is the author of a document, he will take the fact for granted, and it is hard if not impossible for a person to know if a document someone else wrote had different versions. Co-Authors are also harder to remember. Only the author, if anything, is remembered. As to the Place where the document was handled, reading a document is less prone to memorable interactions than actively writing it, making it harder to remember where it happened. Finally, and regarding the document's Purpose, the reason for the difference seems once again to be the ease in which it is possible to remember what a document was for when we were its author.

We found little difference in the amount of times an element was induced, given its total number of occurrences, for the different document types. The only significant differences occurred between documents created by the users and those of others, for **Place**, **Co-Author** and **Version**, as was to be expected from the different element frequencies we described above.

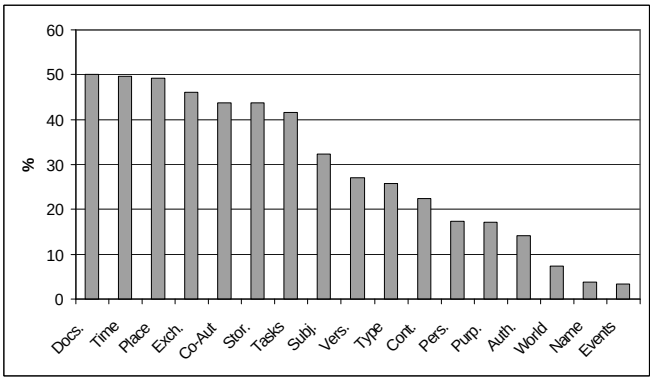

Figure 3 – Overall Percentages of Induced Elements

Overall (Figure 3), we found that the less often induced elements are **Purpose**, **Author**, **Personal Events**, **World Events**, **Events** and **Name**. With the exception of Purpose, these are the least frequent element categories. Keeping in mind that induced elements are those subjects remembered after a question, the fact that these elements were rarely mentioned and, when they were, they appeared spontaneously, means that either they are so important they are remembered without need for external aid, or no amount of suggestion can make the users remember them. Purpose's case is different. It is an element that is seldom induced but that appears fairly often in the narratives. This shows it to be something users consider important and easy to remember.

The more often induced elements are **Time**, **Place**, **Co-Author**, **Other Documents**, **Exchanges**, **Tasks** and **Storage**. All of these appear fairly often in stories, at least once, on average. They are important elements, but hard to remember: mentioned often but only after something triggered the subject's memories about them. Even so, no element is, on average, induced more than 50% of its occurrences in the stories, showing that, even if it is hard to remember, there is a fair chance it might come up spontaneously after all.

The Nature of Story Elements

A closer look at the elements themselves allowed us to find exactly what form the phrases where they are described actually takes.

The level of accuracy for references to **Time** tends to vary. For Recent Documents it is fairly specific: *"(…) about one hour and a half ago (…)"*. For Old Documents it is only roughly remembered: *"(…)I delivered it around April (…)"*. In stories about Other Documents, the references to Time vary in accuracy, depending solely on how long ago the document was handled. References to **Place**, on the other hand, are very accurate (*"At home"*; *"It was updated here"*), as are those about the document's **Purpose**, which include information on where and for what the document was used: *"(…) it will be used in the school's newspaper (…)"*.

References to **Co-Authors** are seldom actual names. Often, the subjects only remember if they existed or not. The mentioned **Subjects** were of very diverse natures: *"(…) the subscription to a magazine (…)"*; *"(…) the weekly results of my work"*; *"(…) an analysis of the company's communications infrastructure"*.

The **Other Documents** that were mentioned sometimes included actual paper documents, and not electronic ones. It was common for users to mention the existence of other documents without actually specifying what documents they were talking about (but apparently knowing it themselves). Finally, sometimes the reference to another document was enough to cause a 'short story' about that document to be told. Information about the document **Exchanges** usually described email exchanges, but also other forms, such as posting it on a web site. References to a document's **Type**, included not only the mention of specific formats ("text", "image"), but also to applications commonly used to handle documents of a given kind ("Word", "Excel", "PowerPoint").

We found references to computer-related and 'real world' **Tasks**: "(…) went to the library to find some references (…)"; "(…) downloaded and selected the photos."; "(…) I printed the document (…)". References to where the document was **Stored**

often mention entire computers, but also removable media and specific (unnamed) locations in a hard drive or local networks. In the case of online documents, the site is often mentioned.

As to **Content**, it was common to find mentions to specific information about the document's structure. References to specific contents were rare: *"It had a sentence that started by 'And to those persons that…' "*; *"(…) it was divided into tables (…) It had lots of graphics (…)"*.

It is not always possible to remember a document's **Author,** especially for foreign, hard to pronounce names. **Personal Events** usually happened to the interviewees themselves or to someone directly related to them. Often it is something that could be found on someone's agenda, but not always: *"It was the day my car's battery went dead."*; *"(…) I finished it before my vacations."*; *"(…) my son had a serious asthma crisis (…)"*.

Almost completely absent were references to **World Events**, often not directly associated to the users but directly relating to their jobs or co-workers. Only once was some important news event mentioned. Also rare were references to **Versions**, normally to state that they didn't exist. The least mentioned story element, **Events** that might have occurred when the subject was interacting with the document, often described actions done by the users and unrelated to the documents, rather than events outside their control. It seems that such incidents are unimportant and quickly forgotten: *"(…) I prepared instant soups (…)"*; *"Someone arrived at my home (…)"*. Finally, there were some references to **Names**, either of the document files themselves or of folders where those files are stored. Sometimes, no specific names were uttered, but it was clear the user had a specific, well identified, folder in mind.

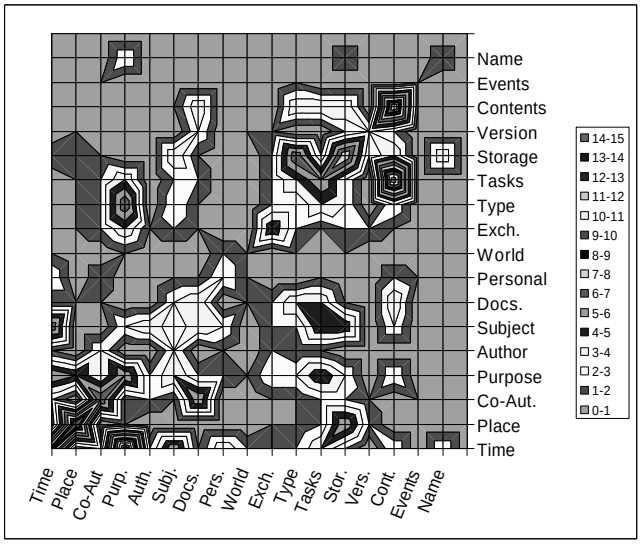

Figure 4 – Transition Frequencies

Element Transitions

Only 36.7% of all possible transitions occurred more than once, reinforcing our assumption that there are indeed especially relevant transitions underlying the stories. The most common transitions were **Time-Purpose**, **Tasks-Content**, **Subject-Time**, **Type-Purpose**, and **Storage-Type** (Figure 4). Reflexive transitions such as those involving **Content, Place,** and **Time,** are also common, whenever the user feels the need to refine or clarify something.

A situation could arise in which a transition between two frequently-occurring elements would itself have a high absolute frequency while happening (for instance) only 50% of the times those elements were present in a story. This could make it seem more important that a transition that occurs 100% of times among rarer elements. Normalized transition frequency values accounting for the frequencies of the involved elements were calculated and no significant bias was detected.

We calculated, for each story element, the probabilities that another of a particular kind might follow. For the most common transitions (for the others, the data is not trustworthy), we found the most probable to be **Place-Place** (0.417), **Content-Content** (0.344), **Tasks-Content** (0.316), and **Time-Purpose** (0.25). Also with a fairly high transition probability we found **Co-Author-Co-Author** (0.259), **Author-Co-Author** (0.25), **Author-Subject** (0.25), and **Place-Storage** (0.25). These probabilities are enough to build some expectations but not to have any certainties.

Finally, we found little symmetry in the transitions. For instance, the Time-Purpose transition occurs over three times as often as Purpose-Time.

5. Discussion

The thorough description of document-describing stories we obtained provides important insights on what the designer of interfaces that make use to those stories should consider. We collected those insights in the form of guidelines we will now describe.

5.1 Customization

We found little relevance of personal factors such as gender and age to the way stories are told. The only exceptions were that women tend to tell longer stories than man, and that older persons are marginally more in control of their stories than younger ones. Apart from those aspects, the stories remain the same. Hence, *little user customization will be necessary in relation to what to expect from a story*. This does not preclude other customizations, such as adapting the interface to the particular subjects users usually work on, or to better visualize a particular Personal Document Space.

5.2 Memory

We expected to find that a user's memory about a document would fade with time, allowing them to remember less information. However, except for Subject (more com-

common for Recent documents), no significant time-related difference was found for the remaining elements, story length, or transition numbers. Likewise, no differences were recorded in the percentages of induced elements stories: nearly half of the narratives were spontaneously told by the subjects. Differences in information correctness might exist, but were not addressed by this study.

What does seem to affect the information a user can remember about documents is their origin. Stories about documents created by the user, regardless of when, are longer. Some elements such as Place or Purpose are mentioned more often, suggesting they are easier to remember. In short, *some differences in the story structures and accuracy can be expected according to the age of the document being described*. However, the biggest differences derive from the document's origin. *It is important to determine it early in the narrative, to correctly form expectations about what can be found ahead in the story*.

5.3 The Importance of Dialogues

For some story elements, a significant number of occurrences were induced by questions posed by the interviewer. Elements such as Time, Place, and Other Documents are among them. They are also some of the most frequent elements, suggesting that users consider them important and can actually remember them, if asked.

It is important to establish dialogues with users in order to obtain all information they can actually remember. Some care should be taken about thematic shifts. However, they are fairly rare and should pose no significant problem.

On the other hand, *the dialogues should not waste time and resources trying to discover certain elements*, such as Author, Personal Events, World Events, Events and Names. They are rarely mentioned but generally spontaneously, showing that if they are remembered at all, they will most likely be volunteered with no need for inducement.

5.4 Context-Dependent Information

It is common for stories to include indirect references to elements that are taken for granted by the storyteller. For instance, references to the Place where a document was produced and its Author are based on assumptions or contextual information. Often, no specific places or names are mentioned because they seem obvious to the person telling the story. This happens, for instance, if a document arrived by email and the user only has email access at work. *It is important to take the context in which the story is told into consideration, comparing it to a model of the users' world and of users themselves*.

5.5 Ambiguity

Some level of ambiguity is common in stories. For instance, references to time become more inaccurate for older documents. Something similar occurs when trying to

remember names of authors or co-authors. The user can remember what the name sounded like, or that it had some co-authors, but not their actual names.

Some level of ambiguity must be tolerated by narrative-based interfaces. Techniques to automatically disambiguate stories with the help of context and user and world models are to be considered. Users themselves often try to help, providing information about the same element more than once in the same story. That willingness to help should be encouraged and used.

5.6 World and User Models

When referring to such elements as Purpose, World Events or Personal Events, a wide range of information can be conveyed. It is probably impossible to just use keywords extracted from the stories to effectively gain some insight on what document is being talked about. Trying to understand those elements just by looking at what was said is also insufficient, due to great numbers of things that would be important to understand them but are taken for granted and not explicitly mentioned. To aid in that understanding, *a model of the world around the users and of the users themselves (including typical activities, co-workers, etc.) should be used.* Important information can also be found on the user's agenda, and also in that of his friends or co-workers. Some facts from the 'wider world', such as important news could also helpful, albeit rarely.

5.7 Overall Document Structure

Users remember more easily overall document structures than actual keywords or phrases in that document. Some technique that identifies the overall structure or visual appearance of a document and can use that information to differentiate among several documents would be useful.

5.8 Events Arising During Interactions With the Document

In short, *these are not relevant*. It was extremely rare for any such events (someone entering the office, a phone call, etc) to be remembered.

5.9 Recursive Stories

When describing related documents, it is common for several information elements pertaining those documents to be told. They can constitute small recursive stories (stories within a story). *Special care should be taken to capture those elements, which provide important information, while keeping in mind they relate to a document different than the one the story is about.* Also, those stories should somehow be controlled in order to prevent the storyteller from loosing himself in them, sidetracking from the document he really wants to find.

5.10 Expected Elements and Structure

The stories we analyzed share, up to a point, similar structures. Designers of narrative-based interfaces should take advantage of those similarities. They will allow the system to know what to expect from the stories, help guide the user towards providing useful information, and collect that information.

Some story elements are more frequent than others, and should be expected more often. Several will be mentioned only if prompted by some external factor. This information is useful, helping decide if some more information should be expected (if some frequent elements weren't yet mentioned) or not. It will help decide whether it's worthy to invest some time and effort to discover more elements.

5.11 Probable Transitions

Of all possible transitions between different story elements, only 37% have some credible probability of showing up. Of those, five are to be expected fairly often. Combining this information with the probabilities of what will be the next element, given the current point in the narrative, it will be possible to build expectations of what the next element in the story will be. This will help recognize it and extract all relevant information, facilitating disambiguation.

6. Validating the Guidelines

The guidelines we just described are based solely on stories told to human interviewers. To validate them, it is necessary to verify if stories told to computers, no longer free-form but in a more structured environment, are similar to those in which the guidelines were based. We designed two low-fidelity prototypes that embody the guidelines. In both, time plays a special role, as does determining the documents' authors, allowing the use of the different expected story structures. Several story elements are suggested to the users in the order found to be the most likely in the previous study, but any of them can be referred to at any time, if the users so wish. Specialized dialogue boxes are used to enter the elements. Prototype A allows the direct manipulation of the elements, graphically represented on the interface as little boxes, and Prototype B presents those elements as natural language sentences (Figure 5 and Figure 6). More details on the prototypes' design can be found in the experiment's technical report [10]. Ten users where asked to tell document-describing stories using Prototype A, and ten others using Prototype B. We used a Wizard-of-Oz methodology, in which the researcher simulates the workings of the prototypes.

Comparing the stories told using the prototypes to those previously collected immediately showed them to be similar. The relative frequencies and importance of the several story elements is analogous to those found for stories told to humans, as is the nature of the information. The stories were actually longer than those told to humans (20%), thus conveying more information. Prototype B was clearly better, allowing for longer stories to be told, with fewer differences to the ones in the previous study. For

instance, in only 3% of stories did the users of that prototype deviate from the proposed story order, whereas this happened on 43% of the stories told using Prototype A. Also, the qualitative evaluation of the prototypes (using a questionnaire), showed that the users found Prototype A to be more confusing. We attribute the differences between the two prototypes to the fact that on Prototype B, the users were able to see the entire story as a whole, in textual form, and Prototype A dispels the illusion of telling a story by dividing the narratives into discrete elements.

This shows that, despite the validity of the guidelines (using them, we were able to come up with an interface that allows stories similar to those told to humans to be told), the judicious design of the interface is crucial for the quality of the stories.

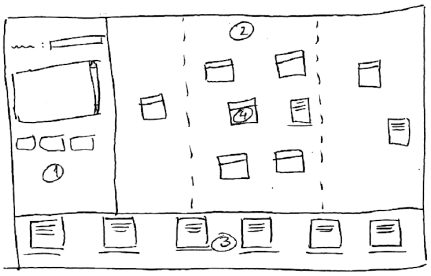

Figure 5. Prototype A

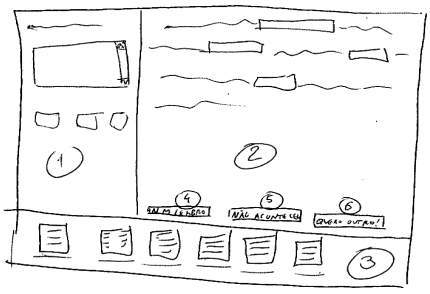

Figure 6. Prototype B

7. Conclusions and Future Work

With the growing numbers of documents users must deal with on a daily basis, new techniques to help finding them are imperative. One such technique involves taking advantage of our innate ability to tell stories. We verified that stories about documents provide a wealth of information about them, helping the users to remember more details than they would otherwise, as shown by the existence of induced elements. We found that dialogues are important to allow those elements to come up. The stories shared several common properties and structure, including the most common elements. This will allow for narrative-based interfaces to build expectations on what shapes the stories might take, helping to understand and disambiguate them. In short, several important guidelines could be extracted that will allow future research in the area to be developed on a sound basis. Those guidelines were validated with the help of low-fidelity prototypes.

One factor we didn't take into account in this study and that might constitute interesting future research is to ascertain to what extent the information users tell in their stories is accurate. In the present study, when someone said that a document was written four months ago, we had no way of verifying that assertion. Such verifications would require access to the users' documents. However, such extended access leads to important privacy concerns that will have to be dealt with. This would be something better tested by resorting to a story-gathering prototype which is able to gather story details and verify their accuracy without the intervention of a human interviewer.

References

[1] Abowd, G. Software Engineering Issues for Ubiquitous Computing. *Proceedings of the 21st international conference on Software engineering*, pp 75-84, ACM Press, 1999.

[2] Abowd, G. and Mynatt, E. Charting Past, Present, and Future Research in Ubiquitous Computing. *ACM Transactions on Computer-Human Interaction*, 7(1), pp 29-58, ACM Press 2000.

[3] Baeza-Yates, R., Jones, T. and Rawlins, G. A New Data Model: Persistent Attribute-Centric Objects, Technical Report, University of Chile, 1996

[4] Bälter, O., Sidner, C.. Bifrost inbox organizer: giving users control over the inbox. In *Proceedings of the second Nordic conference on Human-computer interaction*, pages 111-118, ACM Press, 2002.

[5] Dourish, P. *et al*. Extending Document Management Systems with User-Specific Active Properties. *ACM Transactions on Information Systems*, 18(2), pp 140-170, ACM Press 2000.

[6] Freeman, E. and Gelernter, D. Lifestreams: A Storage Model for Personal Data, *ACM SIGMOD Record*,25(1), pp 80-86, ACM Press 1996.

[7] Gifford, D., Jouvelot, P., Sheldon, M. and O'Toole, J. Semantic File Systems. *13th ACM Symposium on Principles of Programming Languages*, October 1991.

[8] Gonçalves, D. and Jorge, J. An Empirical Study of Personal Document Spaces. In *Proceedings DSV-IS 2003*, Lecture Notes on Computer Science, Springer-Verlag, vol. 2844, pp. 47-60, June 2003, Funchal, Portugal.

[9] Gonçalves, D. Telling Stories About Documents, Technical Report, Instituto Superior Técnico, 2003 (http://narrative.shorturl.com/files/telling_stories.zip).

[10] Gonçalves, D. 'Telling Stories to Computers'. Technical Report, Instituto Superior Técnico, December 2003.
http://narrative.shorturl.com/files/telling_stories_to_computers.zip.

[11] Huberman, M. and Miles, M. Analyse des données qualitatives. Recueil de nouvelles méthodes. Bruxelles, De Boeck, 1991.

[12] Malone, T. How do People Organize their Desks? Implications for the Design of Office Information Systems, *ACM Transactions on Office Information Systems*, 1(1), pp 99-112, ACM Press 1983.

[13] Myers, B, Hudson, S and Pausch, R.. Past, present, and future of user interface software tools. *ACM Transactions on Computer-Human Interaction*, 7(1), pp 453-469, ACM Press 2000.

[14] Whittaker, S., Sidner, C. Email overload exploring personal information management of email. In *Conference proceedings on Human factors in computing systems*, pages 276-283, ACM Press, 1996.

[15] Yin. R. Case Study. Design and Methods. London, Sage Publications, 1989.