

Social Computation and Temporal Structure: Reflections on PepysDiary.com

Thomas Erickson
Social Computing Group
IBM T. J. Watson Research Center
snowfall@acm.org

INTRODUCTION

I am interested in how groups of people collectively work together to answer questions, solve problems or create new things. I refer to such phenomena as social computations, and I mean it quite literally: as I see it, individual members of groups are working together to compute things.

I am particularly interested in situations in which those who are working together are not really working “with” one another. That is the individuals are *not* intentionally collaborating with an eye towards achieving a collectively agreed upon goal. Perhaps they are doing small things together – talking, playing games, – or maybe they’re just carrying out individual actions that only benefit themselves. Regardless, I think such cases are interesting because the high level goal is embodied in the design of the system, and the participants’ activities are enlisted towards that end.

There are lots of examples. One set comes from Luis von Ahn’s work on CAPTCHAs [2] and “games with a purpose” [1] that label images and generate common sense facts. Another is the MATLAB open source programming contest, in which contestants effectively design and optimize programs as they compete with one another [3, 4]. Another is the Wikipedia Featured Article process, which creates a period of intense focus on improving an article so that it can be featured on Wikipedia’s front page for a day [7, 8]. A fourth example is the genre of auctions, which, as Charles Smith says in his ethnography of auctions [6], are “*social processes for establishing socially acceptable definitions of value and ownership.*”

One commonality that these examples share is that, in one way or another, they use temporal structure to focus and engage users. My interest is in understanding the ways in which temporal structure – which I will loosely define as *the implicit or explicit association of rules and behavior with time periods* – is used to structure interactions among large groups of people. I believe that such structures play a crucial role in coordinating activity, especially when groups are very large, very diverse, or simply composed of many strangers who have no desire to communicate.

I would like to have a better understanding of how to think about systems that employ this sort of web of temporally bound constraints. I suspect that there are a number of ways that temporal structures can be instantiated and used. I’d like to

have a taxonomy of such uses, and a set of examples to illustrate. But I don’t. However, I think that this would be an interesting topic to take up in the workshop, and, as my contribution, I will lay out an example that I don’t believe that anyone has described: PepysDiary.com.

PepysDiary.com is a web site started several years ago by Phil Gyford of London [5]. PepysDiary, as I will call it, is notable because it represents a unique blend of approaches to the collective production of knowledge. In its look and feel, and in the software it uses, it resembles a blog. In the way in which it is used – a large set of users focusing their collective attention on a few rapidly changing ‘items’ – it resembles SlashDot. And in its ultimate goal – to produce an online encyclopedia – it resembles Wikipedia. What makes this novel blend of approaches particularly interesting is that in the five years since it opened, it has exhibited a significant and remarkably steady rate of usage.

In what follows I’ll begin with a ‘tour’ of PepysDiary, follow with a discussion of its long term use, and wrap up with a closer look at the ways in which it is used and the ways in which it uses temporal structure.

A TOUR OF PEPYSDIARY.COM

Before beginning our tour, it’s helpful to understand a bit of background about Pepys and his diary.

About Samuel Pepys

Samuel Pepys lived in the 17th Century and rose from humble beginnings to become the Secretary of the British Navy. Although a fairly prominent figure in his own time, he is mainly remembered today because of his diary, which spans a decade that included the English Civil War, the Restoration of Charles the 2nd, the arrival of the Black Plague, and the Great Fire of London. Pepys recorded these events, intermixed with the mundane aspects of his daily life and his encounters with everyone from the rich and famous to shopkeepers and servants.

Unlike many other diarists, Pepys did not write his entries with other readers in mind. He wrote for himself, and in an obscure form of shorthand that was long mistaken for a cipher. As a consequence his entries are dense: short, telegraphic, and loaded with obscure references that require a great deal of knowledge to unpack. For example, the entry from September 25, 1665 begins: “*Found ourselves come to*

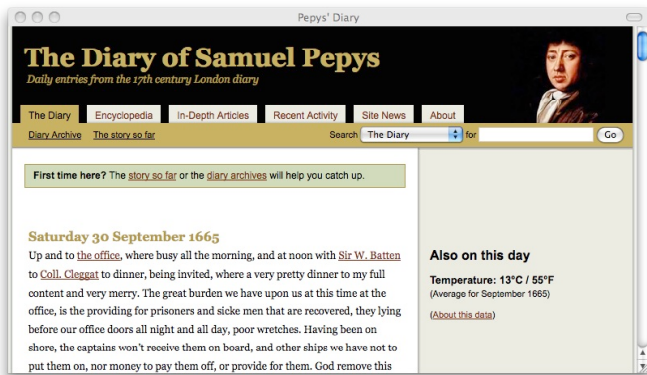


Figure 1. Entry page of PepysDiary.com

the fleete, and so aboard the Prince; and there, after a good while in discourse, we did agree a bargain of 5,000l. with Sir Roger Cuttance for my Lord Sandwich for silk, cinnamon, nutmegs, and indigo.” To make sense of an entry, the reader must be familiar with the preceding entries, with Pepys and his circles of family, friends and colleagues, and with the time, conventions and environment in which Pepys lived.

The Tour

Every day, at approximately 23:00 London time – about when Pepys typically composed his entry – a ‘new’ entry appears on the site: it is for the current day and month, minus 343 years. Figure 1 shows the basic form of the entry. In addition to the text, each entry has three notable aspects.

“Also on this day...”

First is the “Also on this day” link, in the right margin. For the entry in Figure 1, all that is shown is the average temperature, but depending on what’s available there might also be links to other sources of historical information for that day such as records from Parliament.

Hyperlinks to the Encyclopedia

The second significant aspect of the entry are its hyperlinks: these lead to the encyclopedia section, which contains collections of information about each linked item. Thus the first link in Figure 1 (“the office”) leads to a collection of information (Figure 2). Here, under the heading “Navy Office, Greenwich Palace,” a second level of tabs gives access to a map showing its location, a brief description, annotations, and a set of cross references to the entries in which the reference has previously appeared. The annotations tab allows readers to enter comments about an item; in this case, there are no annotations, but it is not uncommon for entries to have annotations (e.g., the second link, for “Sir W. Batten” has 8 annotations, some quite lengthy, produced over a period of 4 years).

The cross reference tab (shown in Figure 2) is interesting because it provides a temporal view of time and frequency of mention each item. Thus one can see that “the office” is

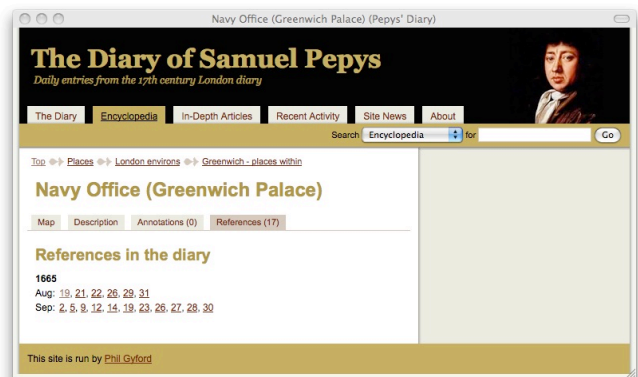


Figure 2. Part of an encyclopedia information collection

mentioned quite frequently, but only from the 19th of the preceding month – as the description explains, it was moved to its current site for safety after the outbreak of the Plague. (This example also makes the point that the hyperlinks are manually maintained: e.g., earlier references to “the office” point to the entry for its previous Tower Hill location.)

Entry Annotations

The most interesting aspect of the PepysDiary site comes at the end of each entry: the annotations link. It is behind the annotations link that most of the activity of diary lies. As noted earlier, the entries are thick with obscure references that, if any sense is to be made of them, require unpacking. What happens is that readers, in a flurry of comments, engage in that unpacking. Thus, for the entry shown in Figure 3, we see a discussion of a reference to a play as well as sympathetic comments about “Sam’s” (as he is familiarly referred to) financial worries. Most of this

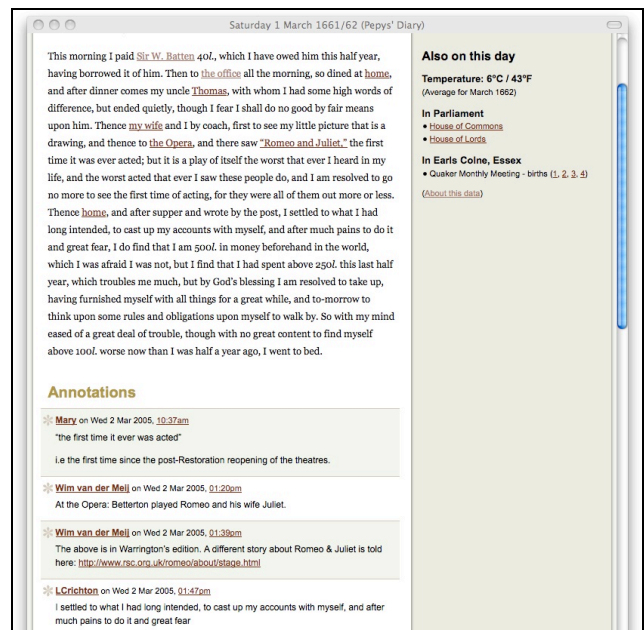


Figure 3. Diary entry and annotations

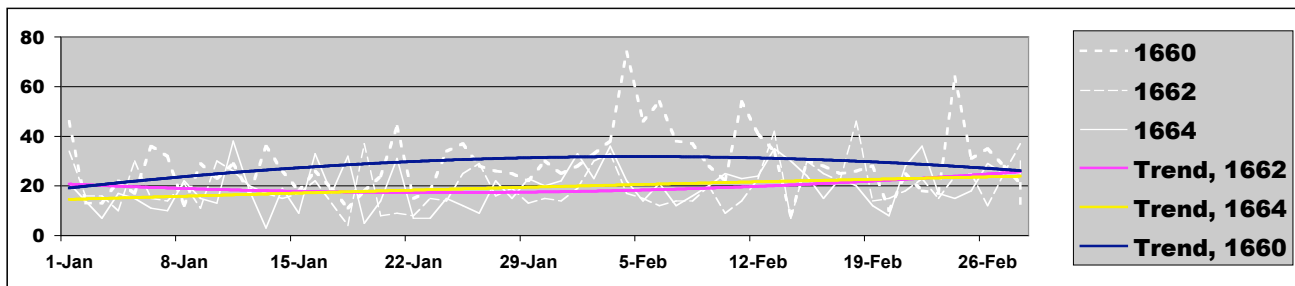


Figure 4. Annotations (absolute numbers and trend lines) per day for January and February of 2003, 2005 and 2007.

discussion happens in the 24 hours following the entry’s publication, after which the collective focus of attention shifts to the next entry.

THE LONG TERM USE OF PEPYSDIARY.COM

PepysDiary has been active for over five years. Currently the site is getting about 25,000 unique visitors per month, making about 55,000 visits. Most of the visitors simply read, but those that do participate contribute about 600 annotations per month. To provide a sense for the activity over the life of the site, Figure 4 shows the number of annotations for the first 2 months of the first, third and fifth years of the site. As is evident, after the initial burst of popularity the first year, the site shows a remarkably constant rate of annotation – about 20 per day.

Over time this activity has resulted in a considerable amount of content. As of late summer, 2008, PepysDiary has a bit over 2,000 of Pepys’ entries, which have in turn garnered about 41,000 annotations, an average of 20 per entry. There are over 30,000 links from the entries to encyclopedia topics. The encyclopedia itself has over 3,000 topics (though many of them are quite brief or are simply stubs); topics have an average of just over 2 annotations per topic (though I suspect the median is substantially higher).

HOW DOES PEPYSDIARY WORK?

So, to summarize thus far, PepysDiary operates like this. Every day at 23:00 London time, a new diary entry is posted. Items in the entry are linked to encyclopedia topics (existing or newly created stubs) by the site’s owner. Over the next 24 hours, the site gets about 1,800 visitors, most of whom just read, but a small fraction of whom annotate the current entry (the large majority), or previous entries or encyclopedia topics (a small minority on any given day).

If one reads the site on a daily basis, what is most striking about it is the nature of the annotations. The annotations play a number of roles. Sometimes they are just annotations: someone may, for example, explain the meaning of an archaic term like “scallop whisk,” or provide a reference to something that happened previously (e.g., that the men working in the basement are repairing a problem reported several weeks ago).

However, much of what is going on is sense-making of a higher order: commenter are trying to puzzle out the social

logics behind what is going on. In the annotation excerpt shown in Figure 5, the participants are puzzling over Sarah’s (a now-ex maid) report that Sam’s wife “borrowed 50s for Will”: Why the loan? Is it “for Will” or “from Will” (as another edition has it)? Does will have enough money that it could be from him? Why did Sarah report it to Sam? She has been ‘let go,’ but it is known that Sam has had a ‘dalliance’ with her. Elsewhere in the stream, other’s discuss Sam’s astonishing comment that Sarah would make a suitable wife for Tom. How does it all fit together?

Another interesting motif visible in this stream of annotations is an interest in differences between the “now” of the diary and the “now” of the commentators. Earlier in this stream of annotations, one commenter has noted that it is eight days before Christmas, but Sam hasn’t mentioned it. Others take up the question, venturing explanations ranging from that it was “not the deal we make it today” (citing similar lack of anticipation prior to previous Christmas’s) to the celebration of Christmas having been banned before the Restoration, currently only in its second year. Another participant reminds others to check the background information in the encyclopedia (which consists of 14 annotations, most created during previous Decembers).

Yet another feature of the annotations is the tendency to cast Pepys – almost always familiarly referred to as “Sam” – as a personal figure. He is sympathized with when ill or worried, chided when he errs, and mocked, at times, for hypocrisy. He is treated rather like a beloved but eccentric uncle. Indeed, some participants go so far as to invent episodes complete with plot and dialog to further imagine Sam, as Gertz is beginning to do in the last annotation (only partly shown) of Figure 5.

There are many other features of interest in the annotations – a complete absence of flaming, a core of participants who know and refer to one another, the assumption of idiosyncratic personas – but these suffice to give a sense of the conversation. All of this is of interest because, from my observations, it is the conversation in the embedded annotations that drives PepysDiary and is responsible for its sustained use and the gradual accretion of its impressive body of content.

DISCUSSION

PepysDiary is an interesting case for the workshop to consider for a number of reasons. First, as I've already

<p>* In Aqua Scripto on Sat 17 Dec 2005, 03:13am It dothe bother me too, the meaning of 50s, Will would not have that much, he only receives less 15s a week [50 Quid/yr] That would be using up most of his savings unless he be very thrifty with his cash flow and if he be, then would not lend it out any way unless the Mistress be promising a nice reward. So this might be a tale out of spite/ along with that other porkie of La chanteuse.</p>
<p>* The Mollusc on Sat 17 Dec 2005, 03:21am Sarah's surveillance It could be quite damaging for the reputation of our Pepys family (Samuel and Elizabeth) if word got out that Madam was 'in hock to the servants' while Sir was on his way to 'the top' in society etc. Revealing unspoken or unresolved conflicts between them would make raise doubts about how solid that relationship is...</p>
<p>* In Aqua Scripto on Sat 17 Dec 2005, 04:02am "I was mystified about why he would think Sarah a suitable wife for Tom " A.S. Man will be enamoured by a peck on the kisser,..."...indeed she seemed very well-favoured to me to-night, as she is always...." Mans brain be a Yo Yo, up and down the spine it dothe go. Re: The 25th: not the deal we make it today: last year, it was not an awe inspiring occasion. He dothe record. "...Dined at home all alone, and taking occasion from some fault in the meat to complain of my maid's sluttury, my wife and I fell out, and I up to my chamber in a discontent...." He gets around to mentioning it only 5 days before, the 20th '60 and 21st '61. No sleighs.</p>
<p>* Paul Chaplin on Sat 17 Dec 2005, 04:55am Christmas It's only the second Christmas since the Restoration. During the Commonwealth, as I understand it, Christmas celebrations were forbidden or strongly discouraged. It may be taking people a while to get back into the swing of it.</p>
<p>* Mary on Sat 17 Dec 2005, 08:22am the borrowed 50s. Indeed the L&M edition gives a reading here of "from Will", but notes that this is an editorial emendation of "for Will" in the original text.</p>
<p>* Australian Susan on Sat 17 Dec 2005, 09:50am I was just being peevish really. The average 17th century housekeeper had far more problems than the 21st century equivalent. At least I don't have to chase and kill pigeons as Elizabeth has to at one point. Although my grandmother's cook was expected to cope with a live turkey one year (it was a present from a late paying customer of my grandfather's). Twelfth Night was much more important for festivities in the 16th and 17th centuries than Christmas Day. Remember last year's 12th Night Pie???</p>
<p>* Ruben on Sat 17 Dec 2005, 12:34pm all those interested in Christmas Day should check the Background information (including Australian Susan that contributed to the place). Have a nice Christmas!</p>
<p>* Robert Gertz on Sat 17 Dec 2005, 02:36pm Back again to the Tower treasure quest. I don't know Lee, I'd consider making sure a third man is along, Sam seeming so fixed on the hunt. Perhaps a little Fred C. Dobbs creeping in... One am...The Tower's dankest basement. "Lee! There!"</p>

Figure 5. Excerpt from a day's annotations

noted, it seems to be a unique amalgam of better known digital forms that is successful in sustaining collective activity and producing a coherent result. Second, I think it's interesting to note the ways in which the structure of the site shapes its mores. This is not Wikipedia, with its neutral point of view; nor is it Slashdot with its fractious and only-sporadically-coherent comment streams; nor, of course, is it a blog in the traditional sense of the word. However, in the familiarity with which readers position themselves relative to Pepys, PepysDiary seems to partake more of the blog than anything else.

What I find most interesting about PepysDiary is that way that it uses time to structure the activity within it. PepysDiary uses time in at least three ways. First of all, the daily posting of a single entry serves as a powerful mechanism to focus readers' attention. The second way in which time plays a role is that – due to the nature of the diary – the entries are connected to one another and play out a larger narrative, albeit one that isn't always clearly laid out. A third aspect of time has to do with the historical nature of the content of the diary: many of the events, people and places are significant, and can provide an entrée to a larger historical context that some readers find extremely engaging. A fourth aspect of the temporal structure of PepysDiary is that it is synchronized with the calendar: December 25, 1665 will be published on December 25th, 2008, and that correspondence will be remarked upon and used to explore the similarities between Pepys' time and today.

REFERENCES

1. Ahn, L. v. Games with a Purpose. *Computer* 39, 6 (2006), 92-94.
2. Ahn, L. v., Blum, M., and Langford, J. Telling humans and computers apart automatically. *Commun. ACM* 47, 2 (2004), 56-60.
3. Gulley, N. 2001. Patterns of innovation: a web-based MATLAB programming contest. CHI 2001 Extended Abstracts. ACM Press (2001), 337-338.
4. MATLab contest.
<http://www.mathworks.com/contest/overview.html>
Accessed September 25, 2008.
5. PepysDiary.com. <http://www.pepysdiary.com/> Accessed September 26, 2008.
6. Smith, C. W. *Auctions: The Social Construction of Value*. Berekely, CA: University of California Press, 1989.
7. Viegas, F., Watteberg, M., McKeon, M. M. The Hidden Order of Wikipedia. *Online Communities and Social Computing*, Springer (2007), 445-454. Also available at: http://www.research.ibm.com/visual/papers/hidden_order_wikipedia.pdf
8. Wikipedia Featured Articles.
en.wikipedia.org/wiki/Wikipedia:Featured_articles
Accessed September 26, 2008.