Mycrocosm: Visual microblogging

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1109/HICSS.2009.836">http://dx.doi.org/10.1109/HICSS.2009.836</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Institute of Electrical and Electronics Engineers</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://hdl.handle.net/1721.1/58849">http://hdl.handle.net/1721.1/58849</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher’s policy and may be subject to US copyright law. Please refer to the publisher’s site for terms of use.</td>
</tr>
</tbody>
</table>
Mycrocosm: Visual Microblogging

Yannick Assogba
MIT Media Lab
20 Ames Street, Cambridge MA
yannick@media.mit.edu

Judith Donath
MIT Media Lab
20 Ames Street, Cambridge MA
judith@media.mit.edu

Abstract

This paper introduces Mycrocosm, a microblogging site in which users communicate via statistical graphics, rather than the usual short text statements. Users of Mycrocosm can record everyday "personal statistics" and share this information in the form of simple graphs and charts. People have used it to track how they got to work and how long it took, what color clothes they wore, how often they checked their Facebook updates, how often a colleague shows up late, and many other trends and events. Mycrocosm raises awareness of the strong yet subtle effect a medium has upon the messages it conveys. Most computer-based communication interfaces are text-based, so while the conduit is computational, the medium is not far different from a typewriter. Mycrocosm, by using easily updated statistical graphs, creates an easy to use and understand medium that nevertheless is quite different from other forms of communication, both direct and computer mediated.

1. Introduction

The primary language of the web is, and has been for a long time, text. Recently, we have seen the delivery and adoption of other forms of media on the web mature greatly. Sites such as Youtube\(^1\) have made the consumption and creation of video widely accessible. Photo-sharing sites play a similar role for static images while contemporary social network sites blend image, sound, video and text to create rich spaces for social interaction. The landscape for media-rich interpersonal communication has greatly broadened, and it is within this landscape that we situate our system, a microblogging site that appropriates the visual language of statistics.

Mycrocosm is a web service that allows users to track and share everyday "personal statistics" using simple graphs and charts to display their data. We describe Mycrocosm as a visual microblog because, like other microblogging services, users add data in short but frequent entries; like other microblogging and status-updating systems, our goal is for people to use this tool as a medium for self-expression. Our goal in creating Mycrocosm is to explore more graphical approaches to human interpersonal communication than existing tools.

Graphics are becoming part of everyday communication for ordinary individuals. While we are familiar with sites like Flickr\(^2\), Zoom\(^3\) and YouTube, Mcdonald [8] has done a deeper study on the use of image (primarily photographic) as a conversation tool in web communities and identified six broad categories of visual conversation styles in those communities; reaffirming our belief that more graphical alternatives to text conversation are indeed practical.

Ordinary individuals are also increasingly exposed to info-graphics as a means of communication. Publications such as the New York Times increasingly use info-graphics to support their stories [12] and we observe an increasing use of info-graphic style in pop culture and advertising [9],[13]. Web sites such as ManyEyes\(^4\), Swivel\(^5\) and TrackNGraph\(^6\) put tools for data analysis and information visualization in the hands of ordinary web users.

Although information graphics are becoming increasingly common, their use is still primarily utilitarian: when we think of information graphics, it is usually in the context of "serious" data. Pousman et al [11], have identified casual information visualization systems as a distinct subgroup of tradition information visualization systems noting four core differences between the two, particularly with respect to their "user population, usage pattern, data type and insight". Mycrocosm is better situated in this model of casual infovis systems as it is entirely designed to be a social space where visualization techniques are used as graphical tools for self-expression.

\(^1\) http://youtube.com
\(^2\) http://flickr.com
\(^3\) http://www.zoomr.com
\(^4\) http://services.alphaworks.ibm.com/manyeyes/home
\(^5\) http://www.swivel.com
\(^6\) http://www.trackngraph.com/www/
We also live in a world where issues of personal data collection, control, and use are becoming increasingly important. Our representation of people as their data seems appropriate to the current technological climate in encouraging people to think to the people of how data collected about themselves can be used.

This paper will describe the function and design of the Mycrocosm web application, giving some examples of how it is used. We shall also situate our work with respect to current social and artistic practices that engage with data as representation and aesthetic. Finally we shall discuss our motivations in creating this application and future directions that this work could take.

2. Related Work

2.1. Visualization in popular culture

Sites such as ManyEyes and Swivel place visualization tools in the hands of ordinary web users, users are able to upload data sets to these sites and collectively analyze and annotate them. While the primary focus on these sites is on collective data analysis of complex or large data sets, Viegas et al [14] concludethat, “users are reinventing visualization technology for social and personal purposes”. Mycrocosm aims to create a sociable visualization tool that takes this newer use case as its goal.

Another area from which our work is informed our work is artistic approaches to the visualization of personal data. Artists and designers are using personal data and a data driven aesthetic to explore our individual and societal relationships to creating, recording and publishing personal ‘data’. At the same time these artists suggest new ways at which we can look at notions of identity. Ellie Harrison is an artist whose practice engages directly with personal data collection. Her works include ‘Eat 22’ in which she photographed everything she ate for a year, and projects where she counted how often she sneezed or wrote down what she was thinking every time she had a cup of tea [4]. Craig Robinson created “Personal Pies” a set of pie-charts the recount various aspects of his life from birth to adulthood [7] which provides an excellent example of the use of statistical visual language in an expressive manner. Nicholas Felton is a graphic designer who for the past three years has compiled yearly “Annual Reports” detailing numerous minutiae from the past year in a richly designed booklet [1]. More broadly we can observe the appropriation of visualization tools on sites like GraphJam[3] where users upload graphical interpretations of songs and pop culture memes, made using tools like Microsoft Excel, Microsoft Visio or Google docs. These examples give strong evidence that this particular visual form is adaptable to non-traditional, social and communicative uses.

Figure 1- Graph produced by user of graphjam

2.2. Microblogging

Mycrocosm also relates to the practice of microblogging. Through microblogging, and in blogging at large, people are presented with the opportunity to “say what is on their mind” and engage in discourse over the Internet. According to the Pew Internet & American life project, 44% of US Internet users have created content online [6]. Nardi, Schiano and Gumbrecht [10] have examined peoples’ use of and motivations for blogging, while more recently Java et al. [5] have looked at usage patterns within Twitter. Much blogging is done primarily as a social activity. From simply keeping people aware of ones activities (exciting or banal) to publishing thoughts and opinions to ones’ small but attentive audience, to everyday chatter, these tools function as spaces for social interaction for a large and ever increasing number of people.

Microblogging is a relatively recent form of blogging. Here, users publish very short entries about what they are doing or thinking at a given moment. Mycrocosm is similar to text microblogs in that each entry takes a minimal amount of effort, thus making it easy to frequently update, and also being of a scale that encourages posting about mundane activities. However, text microblogs focus on the “now” – they are about what you are doing now. Mycrocosm is also about trends over time, accumulations of in-

---

7 http://www.graphjam.com
8 http://twitter.com - One of the most popular text based microblogging sites.
3. How it Works

Mycrocosm users create and update data charts. The charts can be about anything – what they ate for dinner, what time they got up, whether they wore eye shadow that day and so on. Charts are publicly displayed on the Mycrocosm site unless the user specifies that chart to be private. Each user has a home page that displays the charts they have created in reverse chronological order (fig 2), currently users cannot re-order their charts but this is certainly a planned feature to enable users to more effectively use multiple graphs together to communicate a message. The Mycrocosm front page also displays recently updated publicly visible charts, as well as a list of users sorted by the amount of content they have on the site (fig 3). Mycrocosm charts (the front page and user’s pages) are visible to visitors of the site without requiring login, this is aimed at making it easier to share graphs once created.

Users create and update charts from the site itself or from their mobile phone, all via short text entries. Given the audience and use pattern Mycrocosm was conceived for, its design strives towards simplicity in both data entry and options for visualization.

3.1. Data Entry

Once users have signed onto the Mycrocosm website, they can start creating datasets for things they want to track. A dataset is a collection of entries – each entry may be about now, but the display extends into the past.
on some subject that are grouped together to make a graph. The main mode of data entry is a single line text-field on the users home page. Text was chosen as entry modality to allow easy interaction via a mobile device. Mycrocosm uses a very simple syntax to enable data entry, a plus symbol (+) followed by the name of the dataset is used to indicate that the user wants to create a new dataset, thus if a user wanted to begin tracking what colour shirt they wear day to day they would type in “+Shirt Colour” (the quotes are not included in the user input).

To enter data into this newly created dataset the user types in the name of the dataset followed by the data they want to put into it, thus if the user wore an orange shirt today they would type in “Shirt Colour. Orange”. Notice the use of the full stop in the entered string; this separates the name of the dataset from what is being entered into it, this full stop is optional if the name of the dataset does not contain spaces. The plus symbol (+) can also be used to associate a particular cardinality with an entry. Thus if a user wanted to indicate that they have seven pairs of sneakers they may enter the following string “shoes + 7” (assuming they already have a dataset entitled “shoes”), if no cardinality is given by the user it is assumed to be one. The plus and full stop are the only symbols used apart from the data that the user is trying to enter. The user can also assign a short code to any dataset; this is usually a short string to save typing the full name of the dataset. Thus “Shirt Colour” in the above examples could be replaced by “s”.

If a syntax error is made during data entry an error message is displayed to the user on their user page, directly below where they enter data. This error reporting does not extend to the mobile text interface, which is certainly a limitation, however we could easily provide feedback via SMS or email in future iterations.

Datasets can be edited by their creator at any time. Apart from changing the datasets options such as the datasets name, the short code associated with the dataset or the type of graphic used to visualize the data, each entry can be modified after its creation. This enables corrections to be made to entries and the dates associated with entries to be changed, this latter use case also enables a user to enter data for a number of different dates at once (i.e. retroactive data entry). The editing of datasets is done using a web based form interface rather than the command line style text entry. This prevents edits to entries and datasets from being done over the mobile text interface. While it would be possible to create a messaging interface to enable these features over text messages, it would likely be extremely cumbersome and not necessarily all that useful as the user would still not be able to see the graphic without access to the web.

As described above, the following sequence of commands, “+Rambo Kills”, “Rambo Kills. Shirt on + 129” and “Rambo Kills. Shirt off + 91” would produce the following graph.

![Figure 5 - Pie chart](image)

We will discuss the creation of visualizations further in this paper. In general Mycrocosm’s understanding of user data is as strings that are grouped together and counted in order to create the graphs (the different ways in which entries are grouped together will be discussed further on in this paper). However Mycrocosm can differentiate regular strings from strings representing numbers and strings representing time and treat these data differently (e.g. numbers can be summed up rather than counting the number of instances of that number in the dataset).

Another feature supported by Mycrocosm is the tracking of time. We were interested in allowing people to capture aspects of the daily rhythms of their lives, and the system has particular features to enable this. When a user creates a dataset with the word ‘time’ in its title it is initially assumed to be a dataset in which a user wishes to record time. For example if a user wanted to track when they took breaks during the day, they can first set this up by entering “+Break time” into the system, then whenever they take a break, they need only enter “Break Time.” and the system will record the current time as the time of their break, they can also enter a time in natural language to specify when the event happened (e.g. “Break Time: 3:45pm”). Time tracking in Mycrocosm is also extended to recording durations, when working with a time dataset, one can input the keyword “start” to start a timer and enter the keyword “stop” to end the timer on that dataset and record that

---

9 The caption below the image “Rambo Kills under various conditions” is added separately on the edit dataset options page.
duration. These events and durations are also displayed in a simple visualization.

Mycrocosm also provides another way to create a new dataset, by copying an old one. Next to each dataset, logged in users can see a link to copy a dataset’s template; this is visually represented as a light grey plus (+) sign next to title of the dataset. Copying a datasets template creates a new dataset with the same title, captions and options for the user that did the copy. It does not copy the entries existing in that dataset. Apart from providing an easy way for a user browsing the site to start tracking something they come across for themselves, providing this shortcut also helps reduce the possible number of different names for very similar graphs that may exist. We plan to augment this feature with a tagging interface to make it easier to browse through graphs that may be related in some way.

3.2. Mobile data entry

One of the reasons the command line style, text input interface was chosen was to enable easy mobile entry even for those without sophisticated phones with web browsing capabilities. Once a user has registered his phone number or email with the system the user can use the exact same syntax described above to create and populate datasets. Users can send SMS messages or emails from their cell phones to add data to their collection. We believe enabling the use of Mycrocosm on the go is important in enabling people to capture events as they happen as easily as possible. The other advantage of the text-based interface is that it is very easy to create other services and interfaces that can talk to the Mycrocosm web site.

3.3. Visualization

Mycrocosm currently uses a small number of simple statistical graphs to visualize user data; currently it supports pie charts (fig 5.), bar graphs (fig 7), stacked bar graphs (fig 8), floating column graphs for showing time and duration (fig 6), and line graphs (fig 9). For datasets with a single entry the user can also use formatted text to display the data (fig 10, aesthetic considerations led us to provide this option). For each dataset users can select which visualization to use, the exception to this is that the floating column chart is the only one used with datasets that explicitly track time. Additionally users can select different ways of organizing the data. By default entries are simply counted and the magnitude of each value is displayed. However the user can organize the data by ‘date’ and get a view of the distribution of entries since they started entering them, or they can organize the data by ‘day of the week’, here entries are grouped and counted according to the day of the week they were entered. Each visualization is restricted to an area of about 300x300 pixels, this enables us to display each user’s page as a set of small visualizations in the style of an information dashboard. Thus visitors to a users page can quickly get a look at the variety of what that user is tracking and any interesting data that appears.

Figure 6 – Two time tracking charts, the user wanted to compare coffee breaks to commits to a version control system

Figure 7 - Bar chart

Figure 8 – Stacked bar chart
3.4. Social and Collaborative features

Mycrocosm also has features to ease interaction between users of the site as well as those that allow for direct collaboration between users.

Like many web sites featuring user-generated content, registered users can leave comments on other users datasets. We also support more direct collaboration in the form of shared datasets, these are datasets that allow multiple people to contribute to or edit a dataset. To collaborate on a dataset, a user has to create a dataset and share it with other users. We added the concept of “followers” to restrict number of the people one can share datasets with or from whom one can receive “dataset invitations”. Followers in Mycrocosm are similar to followers on Twitter; they are people who have decided to follow your activity. To do this users click on a ‘follow’ button that is available on all users home pages. Once a user follows another, that user is added to their ‘following’ list and they are added to that users ‘followers’ list. These lists are also displayed on every users home page. Graphs can only be shared between mutual followers, we place no limit on how many people a user can follow, so in effect we are not limiting how many people one can collaborate on graphs with, the purpose of requiring mutual following was to avoid situations where a malicious user may try and ‘spam’ other users with numerous invitations to collaborate on datasets. While this restriction was the main reason we introduced followers, we realize that having ones followers displayed can act as an indicator of who ones audience may be (at least within the mycrocosm community) and may act as encouragement to continue posting.

Shared datasets allow multiple people to contribute and edit entries on a particular dataset, however they also allow each user to choose a different visualization for their ‘copy’ of the dataset, each user can also create their own caption for the dataset. We do however prevent the name of a dataset from being edited once it has been shared. Shared datasets are also displayed using a different colour palette from other graphs on the site in order to immediately differentiate them.

4. Discussion

4.1. Why graphical communication?

Our motivation in creating Mycrocosm was to enable a different form of communication than the plain text we are all accustomed to. New communication media afford opportunities for new kinds of communication. We believe that using the powerful yet constrained form provided by Mycrocosm can enable and encourage different kind of message passing than commonly seen at present.

Two salient affordances of this visual form are its compactness and the cumulative nature of the graphs. Its compact nature allows a graph to act as an easily understandable overview of what may be a large amount of data. We are not referring to data that is large in the traditional info-vis sense but rather in the sense that if it were to be expressed in prose it would likely take up a lot of space or otherwise be overly verbose. The cumulative nature of the graphs over time allow an ongoing narrative of events to be represented in a single graphic capturing the temporal
rhythms of day to day activity and allowing the communication of histories and patterns of behaviour.

The ability to capture and store vast amounts of history is a hallmark of our digital communication tools, yet that capability is not equally expressed in all media. Contrasting text with the medium we propose, we can see how our proposed style of ‘writing’ is more adept at simultaneously displaying information from past to present due to its compact and cumulative nature.

As a form for online identity the above factors enable this medium to be a way to quickly get a sense of who a person is, or at least what a person chooses to represent about themselves. One of our initial goals for Mycrocosm was that it could enable a form of open, dynamic, and temporal ‘profile’. The ability of the graphs to represent a long period of effort in collecting data can act as a strong signal of the importance placed on what the user chooses to track.

In addition to the above affordances, there is very much a qualitative difference in the ‘feel’ of communicating with a graph, this difference in feel is the crux of the ability of a medium to affect the content of its message. For example the difference in saying (or writing) that one’s highlight of this year’s Olympics was the performance of Michael Phelps and maintaining a graph such as the one below.

Its distinctiveness from plain text also allows it to communicate things that would be difficult or awkward to communicate in prose. The graph in fig. 4 has a wonderful ambiguity about how the user may actually feel about getting most of their birthday wishes through Facebook, on one hand it is possible that the user is excited that Facebook has allowed more people to remember their special day and wish them well while on the other it may act as a commentary about the nature of their different friendships as she contrasts those who do not need the memory prosthesis provided by this social networking site in order to wish her a happy birthday to those who do. In fact this particular user created the graph as “‘a sign of the times, and thought it would be fun to show how much we depend on Facebook nowadays’”. While it has little to do with capturing history, this graph as well as others such as fig 12 and fig 7 are more about making statements, graphically.

It is this difference in medium that offers a change in the mode of expression that we are interested in exploring through Mycrocosm.

4.2. Preliminary user observations

At the time of this writing Mycrocosm has been fully publicly available for just over a week. We will first describe what kind of usage we are observing amongst our new users then do a more in depth discussion of a much smaller set of users that acted as testers while we were developing Mycrocosm.

We started to advertise Mycrocosm mainly through word of mouth via email and twitter. We were also fortunate to have it featured on a popular information visualization blog, which helped drive traffic to the site. Over the past 10 days we have received over 4000 hundred unique visitors, 235 of these visitors have signed up and made 578 datasets (127 of these datasets are currently empty) with a total of 2980 entries. Of the 578 datasets created, 55 are private and 523 are publicly visible. Common themes for graphs among new users include tracking habits relating to eating and sleeping, traveling/commuting, mood and music. We also see graphs indicating particular events and special occurrences in user’s lives, while another class of graphs are those describing some aspect of their creator. Figure 16 also shows a pair of graphs (belonging to a pair of users that presumably know each other) that in fact reference a common subject. We shall show a small selection of these graphs here\textsuperscript{10}.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure12.png}
\caption{"Olympics" graph}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure13.png}
\caption{Has Someone Irritated Me Today?}
\end{figure}

\textsuperscript{10} Usernames have been blurred out.
The graphs above were created by users of Mycrocosm from the general public; earlier on in the development process we invited a small set of people to use Mycrocosm. We shall discuss the experiences of three of these users in this next section. These early users come from the authors’ academic community; two are members of the authors’ research group though they are not directly involved in developing Mycrocosm. The third was an undergraduate student working for another research group over the summer. She came to know of mycrocosm via word of mouth and was interested in trying it out. The length of time they spent using mycrocosm varied, going from approximately one week for two of the users up to 47 days for the third. None of these users were given any specific instruction when directed to the Mycrocosm site. A little while after their activity on the site reduced they were sent a questionnaire to gauge their experience. While this is quite a small sample it does shed some light on possible motivations to using mycrocosm.

The most commonly described use of mycrocosm amongst these users was introspection. Users tracked sports scores in personal competition, eating destinations, number of dinner companions, commute time, meals consumed and other “eccentricities of life” such as teeth brushing behaviour and fashion choices. This is not surprising as Mycrocosm provides a way to record and quantify one’s behaviour. One user, Jane, says “I always felt that my commute was long and out of personal curiosity I collected numbers just to see a) how long it is, b) how big a chunk of my time every day is eaten up by commuting ... although I had a general sense of what the pattern was, I thought it would be interesting to find out precisely.”. Peter (our longest running user) writes “I’m primarily interested in introspection - it’s fun to see what the variety of places I eat is ... Because eating is

---

11 Names used here are pseudonyms.
something that happens every day, it's easy to make updating mycrocosm a part of my routine. I keep it on my bookmark bar and when I'm looking for ways to procrastinate, I'll log on and add the previous day's information."

Related to the above use is that of communicating something to others about yourself, using data. Jane recounts that she created a graph to track what she eats in order to rebuff a complaint that her roommate had that she ate too skewed a diet. However, this use can take an interesting turn, as users are careful of what they put in their graphs and how it makes them look. For example Peter says, "For all my graphs, I'm always excited to enter in new data that increases the diversity of my chart, or grows a category that's rare. For instance, if I eat dinner somewhere I don't usually eat, I like to enter that because it increases the diversity of places it looks like I've been." He continues, "I also think a little bit about how other people will interpret it. For instance, I don't want to over-represent the times when I eat alone. Since that's already a reasonably frequent category, I probably end up under-reporting those instances because they're not unique and I don't really want it to seem like I'm often eating alone." Peter engages in the practice of impression management [2] commonto online social spaces, and indeed almost any social space in which we are perceived by others.

There is also of Mycrocosm for more general non-egocentric communication. We have already described the graph commenting on Facebook as memory prosthesis, but we also observed graphs made to comment on others' behaviour and habits, for example one of our users made a graph to track/comment on how often one of their acquaintances was late for their meetings as a gentle reminder to that acquaintance.

Two of our users also commented on how they felt the investment of adding data to the graphs over time gave them a certain weight and credibility that differed from more the more ephemeral text-based microblogging systems. Peter says, "I really like my charts on mycrocosm have some depth to them. I can show to people that I really care about something by how much I update it, as opposed to facebook where they kind of have to take my word that I've always liked a certain band or book. In mycrocosm, I could have a chart that shows how many pages of certain authors I've read, or how many movies I watch a month".

While it is still too early to draw strong conclusions from the above user accounts, they provide promising indicators that Mycrocosm provides a flexible, expressive and unique communication tool.

4.3. Truth, Truthiness and Statistical Graphs

Mycrocosm can be, and often has been, used as a tool to keep track of factual data—the length of one's commute, one's waking time, diet info, exercise records, etc. Yet it can also be used for more subjective narratives, and there is a certain irony in using statistical graphs to represent thoughts, ideas and fictional data. Statistical representations are often seen as representations of truth or fact and in our context this association provides an interesting juxtaposition to what may actually be communicated; that feel of truth and authority adds a certain irony when what is being represented is, say, the amount of "Haterade" consumed amongst a group of friends. Just as we can express truth, opinions and non-truths in text, Mycrocosm presents users with the opportunity to use statistics' association with credibility in representing who they are but also allows one to express partial/non-truths. This odd juxtaposition actually adds to the humanity of the graphs, as users push on these associations of charts with facts, they turn data into stories, and collections of graphs into portraits.

5. Conclusion & Future Work

In this paper we have introduced Mycrocosm, a new web service for social communication and identity representation that allows users to use simple statistical graphs as an expressive communicative medium.

Some of our future work includes further exploration of what this visual vocabulary should contain, as we scale the web site up in use we will have a better chance to evaluate what the different kind of visualizations that people need are as well as examine what types of communication emerge within this medium. We also want to expand and examine the usefulness of Mycrocosm in the context of users other lifestreaming/personal publishing data. We plan to enable the embedding of live Mycrocosm graphs on external web pages (such as a user's blog), and to hopefully integrate Mycrocosm with social data aggregation systems such as FriendFeed13 or Facebook and see whether it can augment these existing forms in a meaningful way. Additionally we also plan to provide APIs to allow easy automation of data entry on the site.

12 A slang term that rhymes with “Gatorade” (a popular sports drink), for a drink purportedly drunk by and fuelling those who ‘hate on’ (are jealous of) others. http://www.urbandictionary.com/define.php?term=haterade
13 http://friendfeed.com
5. References


