UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

BRUNO PAGNO

Everyday Visualization: Discovering More About Individuals

Thesis presented in partial fulfillment
of the requirements for the degree of
Master of Computer Science

Advisor: Profa. Dra. Luciana Nedel

Porto Alegre
June 2018

CIP – CATALOGING-IN-PUBLICATION

Pagno, Bruno

Everyday Visualization: Discovering More About Individuals / Bruno Pagno. – Porto Alegre: PPGC da UFRGS, 2018.

57 f.: il.

Thesis (Master) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2018. Advisor: Luciana Nedel.

1. Human Computer Interaction. 2. Data Visualization. 3. Personal Visualization. 4. User Monitoring. I. Nedel, Luciana. II. Título.

"Be wise, because the world needs more wisdom, and if you cannot be wise, pretend to be someone who is wise and then just behave like they would."

— NEIL GAIMAN

ACKNOWLEDGEMENTS

I believe I owe my achievements to a few people that helped me on my way to this point.

First, my family for being there since the very beginning, supporting me whenever I needed and specially believing I could make this happen.

My girlfriend Mariana that shared the tough times and kept me going on.

My friends that encouraged me to follow my ideas and wishes all the time.

My advisor Luciana who guided me on every minor and bigger achievement during this journey.

Bruno Cattelan who helped developing the software during a period.

Rosa Levandowski and Alicia Carissimi who helped with the chronobiology tests.

All the professors at Computer Graphics and Interaction lab for teaching and sharing their knowledge and experiences.

UFRGS for being my second home.

And finally PPGC for the opportunity and CNPQ for financing my scholarship.

ABSTRACT

People are becoming increasingly more interested in the use of activity monitors and self-improvement. The availability of individuals' data is also pushing the development of new applications and data visualization projects to be used at home, in science (e.g. to better understand the behavior of populations) or for governments interested in developing intelligent cities. In this work, we present an easy and intuitive set of visualizations to allow the exploration of personal data by common people. We focus on helping people to know themselves better and to make sense of their own data. Our visualizations are based on the metaphors of calendars, clocks, and maps, as well as on the use of bar charts to explore raw data. Data exploration is therefore guaranteed by the interaction between them. In order to evaluate our work we present two use cases, where few users observe and discuss the data from different points of view: the exploration of personal data for self-improvement purposes, and the use of *Everyday Visualization* by health scientists. Both use cases were ran without any training session. The resulting visualization aggregates several different data sources, going beyond many of the personal and casual visualization works. The promising results achieved demonstrated the viability of the use of such techniques for personal data visualizations and sense making.

Keywords: Human Computer Interaction. Data Visualization. Personal Visualization. User Monitoring.

Everyday Visualization:
Descobrindo Mais Sobre Indivíduos

RESUMO

As pessoas estão ficando cada vez mais interessadas no uso de monitores de atividade. A quantidade de dados de indivíduos disponível está ajudando na expansão e desenvolvimento de novas aplicações e projetos de visualizações para ser usados em casa, em ciência (e.g. para entender melhor o comportamento de populações) ou em governos interessados em desenvolver cidades inteligentes. Nesse trabalho é apresentada uma visualização simples e intuitiva que permite a exploração de dados pessoais por pessoas comuns. Com foco em ajudar as pessoas a compreenderem a si mesmas melhor e perceber coisas novas sobre seus dados. A visualização construída neste projeto é baseada em metáforas de calendários, relógios e mapas, além de utilizar gráficos de barra para explorar dados crus. A exploração desses dados se dá pela interação entre essas visualizações. Para avaliar o produto do trabalho são apresentados dois casos de uso onde alguns usuários tiveram a oportunidade de observar e discutir suas informações de dois pontos de vista diferente: exploração de dados pessoais para auto-aperfeiçoamento e o uso do *Everyday Visualization* por cientistas da saúde. Em nenhum dos casos houve treinamento. As visualizações resultantes agregam diversas fontes de dados, indo além de outros trabalhos de visualização casual e pessoal. Os resultados promissores demonstram a viabilidade de tais técnicas para visualização de dados pessoais.

Palavras-chave: Interação Homem Máquina, Visualização de Dados, Visualização de dados Pessoais, Monitoramento de Usuários.

LIST OF ABBREVIATIONS AND ACRONYMS

NBA	National Basketball Association
API	Application Programming Interface
ECG	Electrocardiography
EEG	Electroencephalography
D3js	Data-Driven Documents (javascrpt)
MVC	Model-View-Controller
USA	United States of America

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Human activity can be the source for a huge amount of information. While working, exercising, practicing sports or just living daily lives there is a vast amount of data being generated. This data can be persisted and manipulated thanks to new sensors that emerge every day. We are gradually learning how to monitor ourselves and almost everything around us. Activity monitors are becoming popular, and people are using it more and more to improve themselves. Facebook automatically remembers us that today is the birthday of some of our friends, while Google presents personalized ads based on our recent searches. These examples demonstrate that our personal information can be used for our own benefit, being just a matter of studying the best way of using that information.

By customizing the way we visualize individual's data, we can create personal visualizations. This way users can know more about themselves, set objectives, and improve in specific goals such as health, work or social life. Professionals from many areas, such as health sciences, sociology, politics and so on can also take advantage of this kind of tool to reason on the data from a single person, rather than from a community.

In this work we aim to produce casual customizable visualizations suited for individuals with or without data analysis experience. Casual visualizations are those which the end user does not need to be professional to comprehend, thus our focus on casual vis. We collected a significant amount of information from several different sources such as Jawbone Up (JAWBONE, 2015) wristband, Open Weathermap (OPENWEATHERMAP.ORG, 2015) and Foursquare (FOURSQUARE, 2015), aggregating them all. To verify the flexibility of the proposed visualization, the data was analyzed from two different perspectives: an individual looking at his own data, and health scientists analyzing the data of other individuals.

We provide to the users a consistent set of casual visualizations that allows sense making on individual's information. Our visualizations are adaptable to user's data, which means that each user can display and analyze a reasonable set of different sources of information, but not necessarily all of them. This aggregation of different sources in a unique visualization is what differs our proposal from the ones provided by the activity monitors available on the market.

1.1 Motivation and Objectives

We noticed that people are increasingly monitoring themselves. Because of that they look for activity monitors to register their actions on a daily basis. However, the way that activities are monitored can be greatly improved. Most of the monitors available to the end user focus only in a single activity and thus lack on correlations and completion. We want to expand this paradigm so that users can look at their lives and make sense of their activities on a broader level. More than that, there is a lot of information that is not user dependant, such as the temperature or the weather conditions, we believe that aggregating them to the user's information can help sensemaking and having insights.

Our main objective is to prove that collecting all this information, presenting it in a novel and interesting way, and permitting users to explore their data will allow for new and interesting insights about their lives.

1.2 Research questions and contributions

Initially we were searching for ways to improve individuals by aggregating their information. With this in mind we asked ourselves "how can we improve individuals's lifes by aggregating data?". Eventually we expanded on the subject by not only looking at individuals analyzing their information, but also allowing data professionals correlate and explore other individual's data. That lead us to the second research question: "how can we allow data professionals to interact with individuals information as they would?".

Our research lead us to casual visualizations with different levels of detail and adaptable context, which we came to believe are good measures for individuals and professionals to visualize daily individual's information.

1.3 Structure of this document

The remaining of this thesis is organized as follows. Chapter 2 presents the related works and state of the art. Chapter 3 is the design rational for the problem in hand. Chapter 4 introduces the process used for data acquisition and management, as well as the structure of the application developed. Chapter 5 presents the visualizations we developed and how they can be used. In Chapter 6 we presented the results achieved through two use cases. Finally, Chapter 7 presents the conclusions and future work.

2 RELATED WORK

Personal visualization and user monitoring have been actively discussed by researchers recently. Meanwhile, there are many challenges that can be addressed. From capturing sensors data to presenting useful visualizations and extracting interesting insights, there are many gaps to be studied. Data visualization and data collection are deeply connected as the latter strongly influences the former. We analyzed some papers related to the area in order to better understand and develop our objectives.

2.1 Sensors and data collection

Aggarwal (AGGARWAL, 2013) talks about collecting sensor data and its challenges. He describes sensor data as having numerous challenges in the context of data collection, storage and processing. It happens because sensor data processing requires efficient processing from massive volumes of uncertain data. Additionally, the data can have errors or precision problems which need to be dealt with, making the task even harder.

Sears and Jacko (SEARS; JACKO, 2007) have made a great classification and explanation of sensors. As they only focus on sensors for interaction, it is perfectly in sync with our project:

- Occupancy and motion
- Range sensing
- Position
- Movement and orientation
- Touch
- Gaze and eyetracking
- Speech
- Gesture
- Identity
- Context
- Affect
- Brain interfaces

It is possible to add “health” to the list as an important kind of sensor. Still, the ability of “sensing” all this kind of information is absolutely huge by itself. It is difficult to imagine a portion of the listed items implemented in the same context. Every and each of these capabilities require individual sensors and each has their own drawbacks such as being too big, too heavy or imprecise.

Lane et al. (LANE et al., 2010) present a very interesting survey on mobile sensors. They cover different scales, as in single individuals, groups and communities. The authors discuss that the sensor collected data can be used not only based on individuals but also in groups. It is very likely that big groups of information can help to produce useful information. The authors also explore the use of mobile phones as sensors, identifying some of the problems that they present. Generally one of the hardest problems to manage in a mobile phone is the battery. Another problem that shows up is precision of the sensors. The authors also raise the question about the bandwidth capabilities of the mobile. All these situations can be worked but there is no true global solution yet.

2.2 Personal visualization

Huang et al. (HUANG et al., 2015) did a survey exploring attributes and relations of personal visualization, and presenting its current scenario and challenges. The authors define attributes of personal visualization with a great level of detail. They identify information like the scope of who is intended to see the visualization (one or multiple users, groups and so on), the origin of the data and the effort to collect the information. They also analyze several papers on personal visualization and present the related challenges. Among other factors the authors indicate the importance to fit the data collection and visualization in the users' personal routines and the diversification of design perspectives as important challenges to be overcome. Following their work, Thudt et al. (THUDT et al., 2017) discuss goals and challenges of personal visualizations. They also indicate that these kinds of visualization have great potential to allow everyone in their daily lives to benefit from data visualizations.

Wang et al. (WANG et al., 2015) talk about the essential elements when designing personal visualizations and explore a little beyond traditional visualizations, addressing illustrative and artistic vis. They explore the idea of using visualizations that are more appealing to the user. In their example they show a bouquet to present a timeline of information. In a way it is more interesting to users to view their information, but it is very important to not let the visualization become too difficult to understand. Similarly, Pousman et al. (POUSMAN; STASKO; MATEAS, 2007) describe casual information visualization. They present the concept as being visualizations for non experts. The main idea is that end users do not need to have a professional background to visualize and understand information, which is what personal visualization is all about. Additionally, the audience for these personal visualizations is much wider than the ones dedicated to professionals.

As a consequence of this highlight in personal visualizations, several new projects have been emerging. Many new ideas for casual and personal visualization have been discussed lately. Coelho et al. (COELHO; KUMAR; MUELLER, 2015) introduce the concept of Data Memes where they present information with images that represent the data. Mah et al. (MAH et al., 2015) show nutrient information in the format of a fingerprint. It is an interesting analogy because each person will have a unique fingerprint visualization. Gou (GOU, 2015) present another flower-like visualization (Figure 2.1), where the information is abstracted in several petals. Fung and Ma (FUNG; MA, 2015) deliver a similar idea where the data is presented in a tree-like visualization. The information is divided in branches and show the individual's information in a unique manner. Flemisch et al. (FLEMISCH et al., 2015) introduce the idea of having physical visualizations representing information of a user. The idea is to have a decorative pendant which represent information. Huang et al. (HUANG; TORY; BARTRAM, 2014) show an agenda with several days of information and on the background of each day shows the amount of activity the user had that day. The idea is similar to ours as the authors composed different information in one place (agenda and activity). Larsen et al. (LARSEN; CUTTONE; JØRGENSEN, 2013) discuss the use of spiral visualization to represent data, which is quite similar to one of the visualizations we produced. On a more artistic approach, Etemad et al. (ETEMAD; SAMAVATI; CARPENDALE, 2016) discuss different aesthetics to generate graph layouts, showing weighted nodes in a view different from traditional graphs. All mentioned works present novel forms of viewing user information, focusing on displaying a concise interface with a particular purpose, and showing information of a single or many individuals. A few of the mentioned visualizations in this paragraph can be seen in Figure 2.1.

Indeed, to be able to produce such visualizations it is necessary to collect information. New sensors and trackers are actively developed, as can be seen in TicTrac (TICTRAC.COM, 2015) on Figure 2.2, a Web-based application that aggregates several personal information tracking software. The objective of TicTrac is similar to our own. The application is very strong on aggregating many different activity monitors, but the visualizations available are still generic and simple, without refining or giving detailed information.

(BUONO; CUZZOCREA, 2015) present yet another time series visualization in the form of a clock with additional information. The authors approach the subject from a slightly different point of view, where the user is not only interested in their own data, but their teammates as

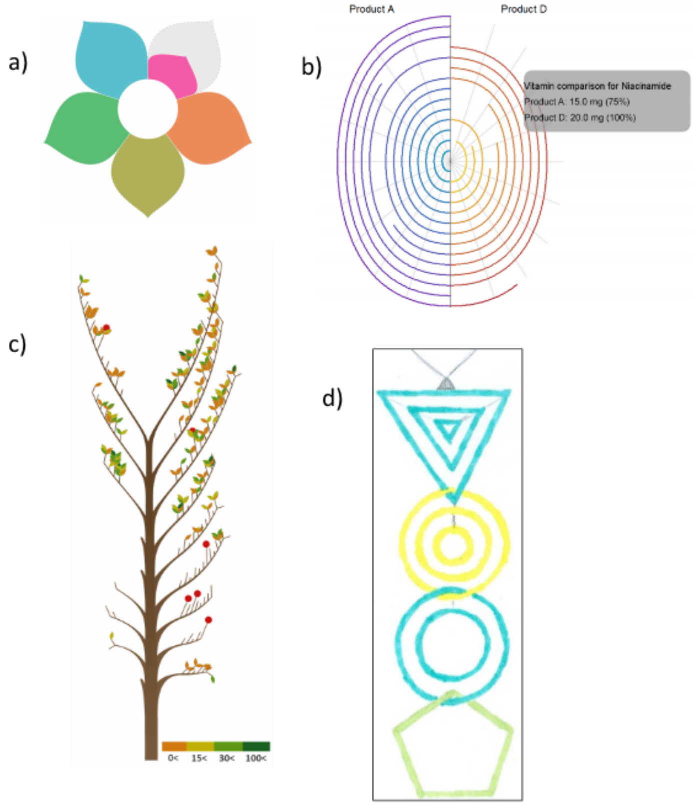

Figure 2.1 – a) Gou (GOU, 2015) flower-like visualization. b) Mah et al. (MAH et al., 2015) fingerprint visualization. c) Fung and Ma (FUNG; MA, 2015) tree-like visualization. d) Flemisch et al. (FLEMISCH et al., 2015) example of jewel physical visualization.

well. In their work the data captures comprises an office's communication activities like emails, texts and calls. The visualization groups the information in the format of a clock, allowing the users to visualize information from daily work.

Healthcare is an area that have been monitoring patients for a long time. For that reason, there are many sensors for health monitoring, as explained by Pantelopoulos and Bourbakis (PANTELOPOULOS; BOURBAKIS, 2010). The authors present an extensive list of health related sensors and classify several of their attributes and uses. Shneiderman et al. (SHNEIDERMAN; PLAISANT; HESSE, 2013) discuss the current situation of visualization on health related software and its availability. The authors also present the challenges and opportunities on the area. They indicate, between other aspects, that visualization can help clinicians with the simplification of information that would, otherwise, be time consuming. They also point that visualizations can help team decision making.

One of the most abundant data we have come across is movement. Zend et al. (ZENG et al., 2013) present a visualization for an enormous amount of movement data. The authors used a dataset from the public transportation system and presented in a concise visualization

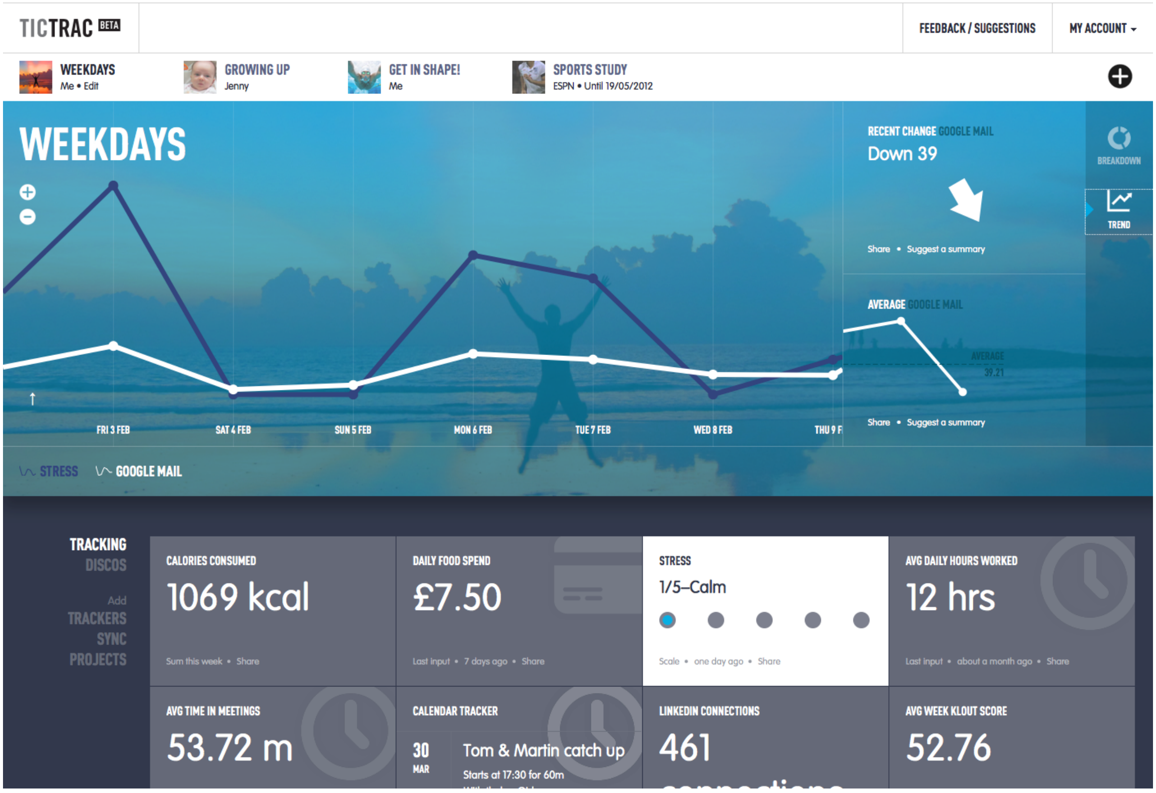

Figure 2.2 – Tictrac (TICTRAC.COM, 2015).

all the information. Their work involved a lot of visual cluttering reduction in order to produce clear visualizations. Grundy et al. (GRUNDY et al., 2009) show the movement data of animals tagged with sensors in a 3d interactive visualization. Working with animals has similarities to working with humans data because we have to collect information to be able to present it. The authors collected movement and orientation information from animals and presented it in spherical distribution. Otten et al. (OTTEN et al., 2015) provide a visualization of an individual movement data and the relation of the places the user visited. The authors present graph like visualizations based on geographic, temporal or frequent information from the visited places. As can be noticed, all the works with movement have to deal with a lot of variables like the amount of users and how to collect the data. All of this influence the final visualization.

It is also interesting to notice that most of the data collected is temporal. For that reason, many developers and researchers decide to present their data in relation to time. Kondo and Collins (KONDO; COLLINS, 2014), for instance, take advantage of this situation to show a technique to interact with timeline visualizations.

Many high-level applications show individual user visualizations. A previous work we developed, called NBAVis (PAGNO et al., 2014), had a particular view dedicated to observing information about NBA players. The information of each player could be analyzed separately or compared in rankings. Thudt et al. (THUDT et al., 2016) created an excellent example of full application on personal data visualization providing a software focused on “mementos”. A memento is an object to be kept as a reminder of people, places, and experiences. The visual design aggregates these mementos and presents it as a manner of keeping memories or telling your own story. Wood (WOOD, 2015) presents a visualization on cycling, explaining his own experience with the sport and the relation with data. His project shows visualizations based on the information of several users, but each can individually analyze his own data. These examples show the diversity of information that can be presented, and also the different interests users may have when dealing with personal data.

Bergstrom and Karahalios (BERGSTROM; KARAHALIOS, 2007) present a clock visualization to view conversations of people around a table. The clock is visualized on the table which doubles as a projection surface. Microfones around the table capture the users voice and display the streams on the visualization as can be seen in Figure 2.3. It is an interesting visualization specially because it is very similar to our development at EverydayVis.

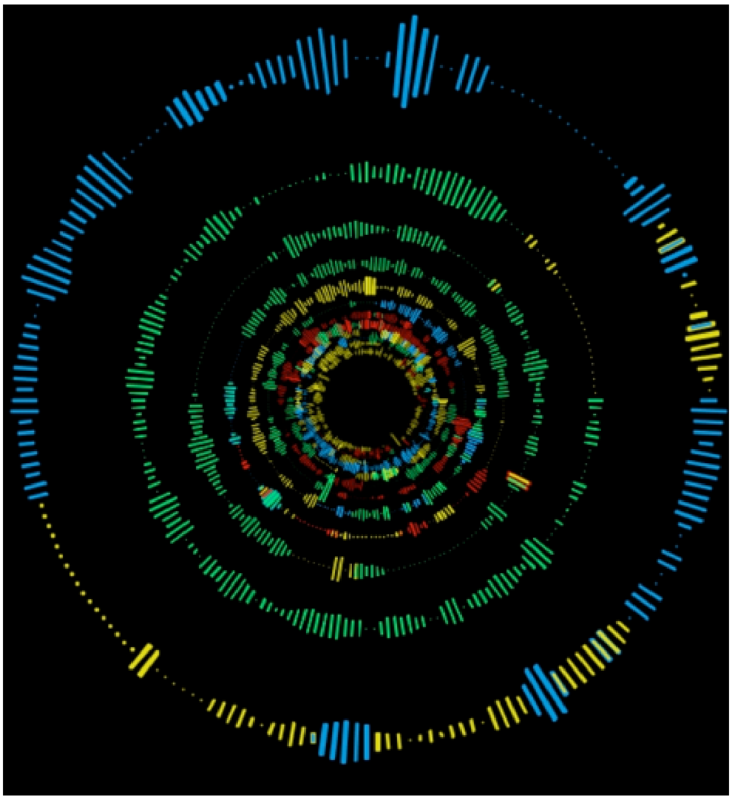

Figure 2.3 – a) *Clock* visualization by Bergstrom and Karahalios (BERGSTROM; KARAHALIOS, 2007). Each circle represents a minute of voice conversation.

Dragicevic and Huot (DRAGICEVIC; HUOT, 2002) created a spiral clock to present upcoming events. The authors discuss a bit of this kind of visualization and conclude that it has a wide acceptance.

Satyanarayan and Heer (SATYANARAYAN; HEER, 2014) present a paper about storytelling with visualizations. The authors reinforce Pousman et al. (POUSMAN; STASKO; MATEAS, 2007) on using visualizations that are not so focused on professionals. Individual user data is an excellent source to storytelling, and it can be fascinating for end users to review their memories through storytelling with data, specially with appealing visuals.

On an extra note, it is interesting to mention the physiological computing as explained by Jacucci et al. (JACUCCI; FAIRCLOUGH; SOLOVEY, 2015). Physiological computing systems are composed by collecting, analyzing and translating phases. In other words, physiological computing is done by capturing data, analyzing them and translating it into commands to an interface. The authors also mention the use of machine learning to develop these kinds of systems. In a way, their description fits quite nicely on personal visualization because of the characteristic of collecting data and returning it in the form of information to the user.

2.3 Related discussion

The works presented in the previous sections relate to our objectives at least to some extent. Sensor data collecting is not the main focus of our research, still we had a great deal of work while going through all the data collected. The experience from other authors helped greatly our decisions.

As discussed in Section 2.1, there are many different kinds of sensors that can be used to collect information. We do not aim to use *all* of the mentioned sensors, but at least a portion relevant to our interests.

A broader classification of sensor data could be done in the following manner:

- Environmental: temperature, lighting, air pressure
- Context: GPS, microphone, accelerometer
- Physiological: Respiration rate, ECG, EEG

Even though most of them could be used in some specific context to keep track of a person's information, many cannot be applied due to technical difficulties of implementation. Depending on the kind of application being developed the amount of equipment or any special necessity of a sensor may cause it to be unadequated.

Several of the visualizations mentioned in section 2.2 have goals similar to ours. Their objectives vary the perspective but still focus on visualizations and user monitoring. However, in this work, we intend to treat personal visualization from a broader point of view. We aim to aggregate much more information of a single user than seen in previous works and consistently

Author	Num. of sources	Kind of data
MAH et al.	1	Nutrients
GOU	1	Social media
FUNG; MA	1	Career (Research)
FLEMISCH et al.	1	Spoken languages
Everyday Visualization	8	Activity, location, weather status, temperature, appointments, work, luminosity, sunrise/sunset

Table 2.1 – Comparison of Personal Visualization works and the amount of data sources each uses.

present it in a highly intuitive manner, taking all the user information to himself and keeping it simple. An overview can be seen in Table 2.1

Additionally there are a few applications worth mentioning, which have similarities to our work as they present personal visualizations of tracker's data.

- Life Cycle is an app available at the AppStore which tracks user's location and activity to organize time. See Figure 2.4 a.
- Health Kit is an apple framework and application dedicated to help users with health information. See Figure 2.4 b.
- Google fit is an app for the Android which tracks user movement activity. See Figure 2.4 c.
- Polar Loop is a wristband which measures activity, similar to the Jawbone Up tracker, but with a different software for visualization. See Figure 2.4 d.

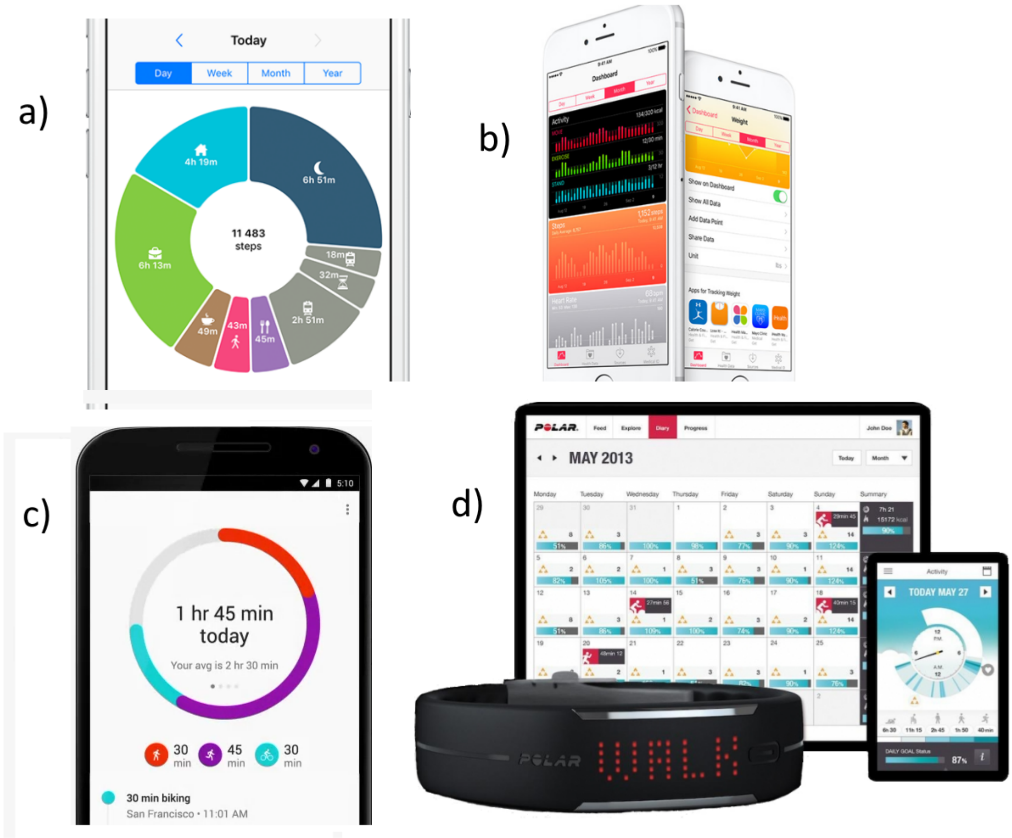

Figure 2.4 – Examples of activity monitor apps available for the consumer. a) Available at http://www.northcube.com/lifecycle/. b) Available at https://developer.apple.com/healthkit/. c) Available at https://www.polar.com/en/products/lifestyle/loop. d) Available at https://www.google.com/fit/

3 PERSONAL AND CASUAL VISUALIZATION

As pointed in Chapter 2, Personal and Casual Visualization are on the rise in the latest years. Researchers have been working on these subjects and consequently new conventions and definitions appear. We believe it is important to expose our approach towards personal and casual visualization in order to better understand the rest of the document.

The literature presents one paper which best covers personal visualization that is the one written by Huang et al. (HUANG et al., 2015). The general content of the paper was already covered in the previous chapter, so we will work on top of that.

The authors discuss the scope of personal visualizations. From their point of view, personal vis can be divided in data, context, interaction and insight. And each can be divided in several subsections. A partial transcription would be as follows:

- Data:
  - Scope: who the data is about. If the the data refers to an individual, group or community.
  - Effort: effort in data collection (none, sensor, manual, mixed)
  - Agency: how much control the person has control over the data being collected. If the user can, for example, forbid an entity of retrieving information about oneself.
- Context:
  - Design context: who is responsible for the design and development of the visualizations
  - Settings: when and how user data and the visualizations are used
  - Influence context: who will benefit from the use of the visualizations
- Interaction:
  - Attentional demand: how deeply the user needs to analyze the visualization to get information from it. For example if the user can get quick information at a glance it is considered a low attentional demand.
  - Explorability: the ability to explore
- Insight:
  - Actionability: how much the insight taken from observing the visualization can guide user's future actions

- Automated Analysis: data mining and automated analysis itself

This classification for visualizations is very useful as it easilly shows the focus of the work produced. It also becomes easier to compare with other projects, as long as they are classified using the same method.

The EverydayVis app has the following characteristics:

- Data:
  - Scope: *self*
  - Effort: *mixed*
  - Agency: *once the data is authorized it is automatically collected*
- Context:
  - Design context: *student and professor involved on this work*
  - Settings: *daily basis*
  - Influence context: *the owner of the data and a special case where a professional views patient's data*
- Interaction:
  - Attentional demand: *low*
  - Explorability: *mixed (aiming to high)*
- Insight:
  - Actionability: *mixed*
  - Automated Analysis: *no*

We have a very "individual" focus and we aim to high explorability. According to the survey made by Huang et al. (HUANG et al., 2015) we have a similar focus to the vast majority of works on personal visualization. This is a positive signal which indicates that we are working with the most popular kind of personal visualization and consequently the type with most demand. Yet we have a "lack of insight". Automatized data analysis may be an option for future work, but for now EverydayVis focuses mainly on data collecting and presenting the visualization which are very demanding tasks.

We also have an additional "influence context" that we wanted to explore. The possibility of a single professional to analyze and compare individuals data. Generally, the idea is that the user will visualize his own data, but we imagined the situation where a health professional or data scientist could want to analyze and compare subject's information. We believe that the same visualization that the user sees can be used by a professional, even if it is more casual focused, as it allows for quick insight and easy visualization.

Thinking about the audience, but still focused on individual visualization, we took a few rules for developing interfaces for personal informatics, as described by Cuttone et al. (CUTTONE; PETERSEN; LARSEN, 2014) and Shneiderman (SHNEIDERMAN, 1997). The authors make a good point showing guidelines, which we adapted in our three principles. These gave us a backbone on which we defined our application's interface.

Therefore, our method focus on the following design guidelines:

- *Keep it simple*: always aim to provide the simplest and most intuitive manner to present information, so that data can be interpreted at a glance
- *Complexity on demand*: data should be representative, not exhaustive. It is preferable to allow navigation through the data instead of overwhelming the user with excessive information
- *Access to raw data on demand*: always allow users to visualize raw data. Frequently it is the only way to find a very specific detail. In our case the raw data is represented as bar charts showing data over time.

These base guidelines focus on a simplistic, yet powerful, manner of visualizing data, which goes along with our objective of building individual focused visualization.

Still following the approach by Huang et al. we take a look at the challenges discovered by the authors. The seven main points discussed can be briefly explained as follows:

- Fit in Personal Routines and Environments: tools need to be accessible or easily usable. Any accessory needed to capture information should not hinder the user.
- Recall of Relevant Context for Reasoning: creating context that help recalling memories and literaly contextualizating to help reasoning.
- Define Appropriate Baselines: offering forms of comparison or giving some feedback about the performance of the user is very important.
- Sharing and Privacy: the information being private or shared according to users' will. When sharing, also defining the target to where they want to share.
- Diversifying Design Principles: "what information to present" and "what metaphor should convey the message", also giving control to the user over his own data/visualizations
- Integrating Computer Assisted Analysis: literally the use of computers to process the users' data. How much and when to use.
- Evaluating: verifying that the visualizations actually convey the message wanted or that

they allow the sensemaking desired by the authors.

Our project fits in a few characteristics that explore these challenges. Actually, one of the biggest motivators for the development of some features were pointing to these challenges.

The first challenge, fitting in personal routines is a pretty basic one. When we started picking the sensors we wanted to use to collect information we evaluated the possibility of using a few ones that were too complex to be used daily by the user. Sensors that provide ECG (Electrocardiography) and EEG (Electroencephalography) are still too big/complex to be used in daily life, and thus were discarded. We chose only sensors that were easilly accessible and that would not hinder the user in any way.

Recall of relevant context for reasoning is a little more tricky. According to the authors the challenge is to be able to create a context that helps users to understand their data. Our focus was to build a broader and more generic visualization that could be used by different users. We believe this challenge is better approachable to more focused data, when the designer knows exaclty what the user needs to see, which is not our case.

Defining appropriate baselines is also a challenge as it is generally difficult to measure the users. We approach this challenge by measuring users against themselves. We provide in our visualizations indicators that adapt to the users' information and challenge them to always get better results. This will be better exposed in Chapter 5.

Sharing and privacy is extremelly important. In our case we created a system that can be used by the individual himself looking at his own data, or the special case where a professional analyzes other users' information. The design of our application is supposed to never let the information leave this scope. Still we have not dealt with the situation where a user wants to share his information.

Diversifying design principles is what EverydayVis is all about. We created an adaptable visualization with different contexts. This means that the user navigates from a broader context to a narrower one. The application adapts to the users' data, being able to show to each individual only the information relevant to him/herself.

We agree that integrating computer assisted analysis is a very interesting option to most visualizations but we have not reached that level yet. Instead, it could be a future work.

Finally evaluating is a huge challenge. We faced this task by aquiring information from users with a reasonable amount of data available and having them give us a feedback on the application. It is also important to notice that we had a health professional working with data collected from patients and needed exploring it using our system.

As mentioned before, the challenges work specially well as motivation for building new forms of visualization and new solutions to problems we have, but it does not mean that taking them on is the same as solving one. We are still very confident in our approach to them as they guided us in several parts of design and development.

The last topics on this discussion we wanted to briefly touch are the concept of "casual visualization" and "physiological computing".

As described by Pousman et al (POUSMAN; STASKO; MATEAS, 2007), casual visualization is not aimed at professionals, but to "regular" individuals, or individuals without training in data analysis, statistics or any related area. On our approach we always focused in this "non professional" concept, always aiming for the most simple we could design, while still mantaining the information relevant to the user. However considering we have a special case where we want a professional to analyze several other users' data, we faced a contradiction. We believe that while the visualizations are aimed to untrained users, they still present enough information so that professionals can use. Creating casual visualizations does not limit the scope of users to non professionals, but instead expands it to them.

The physiological computing, as presented by Jacucci et al (JACUCCI; FAIRCLOUGH; SOLOVEY, 2015), is a general definition to applications/systems that use people's information and give some kind of feedback. The authors describe in a generic way the proccess where the data is collected, processed, and used to produce a result. This represents very accuratelly EverydayVis and many other visualization systems. It is a very interesting definition to something so broad.

4 DATA ACQUISITION

Acquiring data is one of the most important steps when building visualizations. The origin, management, and processing done on the data will directly influence the quality of the end product. Having the correct information makes everything easier. However, finding the right information is a significant challenge by itself. Specifically, when looking to personal visualization, it is hard to capture information for many reasons. We can assume that people produce excessive information in various forms, many on which cannot even be captured yet. The sensors that we have are frequently not precise enough or produce data with errors. Often we face situations where it is difficult to have insights because of the lack of details or information to compare with. Based on that we decided to collect as much data as possible from single users, aiming to achieve a combined visualization. We believe it expands the possibilities of insights since it allows users to make correlations between different pieces of information.

Gathering information from many sources leads to major difficulties. Each dataset has its peculiarities. Some information are captured by external APIs. Others are exported and imported through scripts. There is also the case where custom sensors are used, and the data depends on significant manual input. Getting all the sources of data working together is a difficult task that depends on many lines of code and constant maintenance. Developing the *Everyday Visualization* we also had to deal with different periods of time and frequency of data. Thereafter, we invested a considerable effort in standardizing and synchronizing all the incoming information.

Another difficulty is the fact that there are many different sensors to extract the same kind of data. The Jawbone Up (JAWBONE, 2015) is one example of tracker that collects the amount of steps of a user. But there are several other examples that do similar tracking, such as the Fitbit (FITBIT.COM, 2015) and the Google Fit (FIT.GOOGLE.COM, 2015). The tracker each user uses will depend on their preferences. The process of adjusting these different data sources to a unique representation in our application was a considerable effort. We support Fitbit, Jawbone and Actiwatch, all of them present data in a unique way and need unique treatment, yet, we have to create a single representation for all in Everyday Visualization.

4.1 Collected Data

For the *Everyday Visualization* project we collected information from a few different sources. Still as each source is very different from the other, causing for each to have a tailored treatment. An overview of the data sources can be seen in the Figure 4.1 as follows:

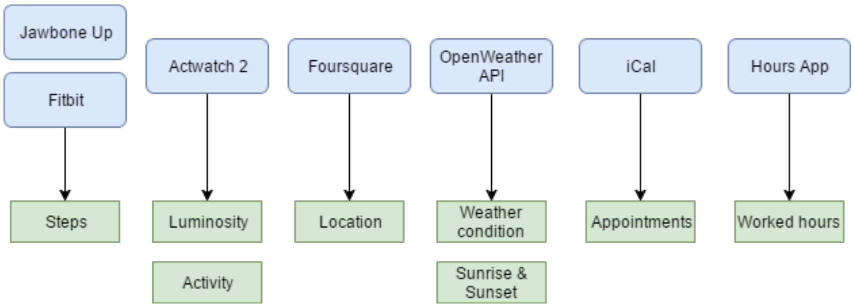

Figure 4.1 – Overview of used data sources and the information they allow us to collect.

The data is stored in a database, and each register in this database has an owner which it refers to, allowing us to desambiguate information from different users. Since most of the data is temporal, each entry in the database is registered by timestamp. This enables us to query and present information in time slices. The listing below shows roughly how data is stored in the database for the activity table.

Listing 4.1 – Simplified example of data in the Activities table.

```
TB_ACTIVITY
datetime                | amount      | owner_id
2015-03-17T17:31:00Z    | 98          | 1
2015-03-17T17:32:00Z    | 95          | 1
2015-03-17T17:33:00Z    | 97          | 1
...
```

This time separation makes it possible that with the same structure on the dataset, any time slice be queried. For example, if the system presents an hour of data, with a simple query we can find the amount of steps of given hour.

Each data source is described bellow:

Steps is the primary content offered by the commercial activity monitors on the market. Pedometer softwares have been used for some time and are continuously becoming a seamless sensor in everyday life since many are integrated to smartphones. In our case, we focused on acquiring this data from the Jawbone Up tracker. The tracker was easily available to us and already had a reasonable amount of data registration done. Initially, we were manually exporting the user data and importing it to the software, but that solution is not very interesting to any user. Based on that we started using the API offered by the company. That API allows the collection of the amount of steps and other information from the user. The communication between client and provider is based on the oauth (OAUTH.NET, 2016) authentication system where the user must authorize the application and only then the data can be collected from the provider. The general method can be seen in Figure 4.2.

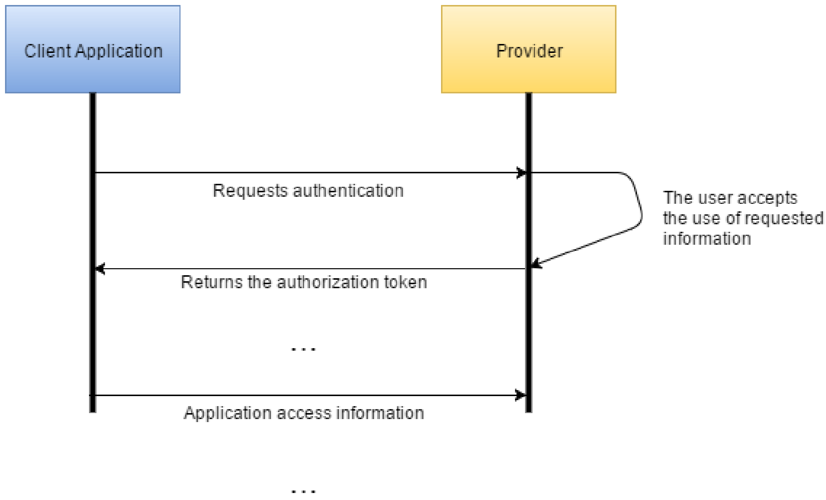

Figure 4.2 – Basic flow of authentication to collect data from a provider.

Later we also added data from the Fitbit tracker. Opposed to the Jawbone API, Fitbit does not allow the use of data per minute. For that reason we requested the developers for special authorization for research use.

Activity is similar to the steps data, but come from a specialized tracker (the Actwatch 2, from Mini Mitter Company). The data does not represent a specific action, but instead measures a generic amount of activity. The information is imported, with the help of a script, from an spreadsheet to the software and stored in the database. The raw data is similar to the steps, but its value does not translate to steps. Instead, its value is a coefficient that indicates the amount of activity of the user. In this specific case it is acceptable to manually import the data as there is a limited number of users and the use case is focused on precisely analysing that data set.

Luminosity is very similar to the *Activity* data, since this information is collected from the same tracker. It measures the intensity of day-night light exposure in lux units and convert it to an integer value. Similar to the activity data, the information is manually added to the software and stored in the database. We use the same script to import and associate the data in the database of our application.

Location is the spatial information of the user. The data is currently collected from the Foursquare/Swarm (FOURSQUARE, 2015) app with the help of their own API. We obtain the coordinates and the city and country name of each check-in given by the user. This allows us to show in spatial visualizations the user's data. A location is especially useful when correlated with the other data collected because it makes a counterpoint with all the temporal information. The API works very similarly to the Jawbone one, using an authentication token authorized by the user before collecting any information.

Weather represents the weather information. We gather the instantaneous events (e.g. rain, sun, thunderstorm, etc.), the maximum, mean and minimum temperature in real time from the OpenWeather API (OPENWEATHERMAP.ORG, 2015). From this same source we can get the sunrise and sunset time of each day. However, when we need to deal with historical data, we have to manually import the weather information to our system. This is again done by a script which reads the information collected and structured.

Daylight represents the sunrise and sunset hours of the day. It is used to show the day and night periods. This information also was collected from OpenWeather API.

Appointments are collected from Apple's iCal software and represents the appointments of the user. The user manually exports this data, and we import it into the system through scripts. Dealing with manual incoming data such as iCal could be quite complex because the application exports a custom file extension. Thankfully, there are some plugins which help us to easily do this kind of importing as shown in the code below. In the example, we record the start time, description and summary of each iCal event described in the exported file.

Listing 4.2 – Code snippet we made that parses an iCal exported file and stores appointments on the database.

```
File.open("ical.ics") do |f|
  cals = Icalendar.parse(f)
  cal = cals.first
  cal.events.each do |ev|
    Appointment.create({
      datetime: ev.dtstart,
      description: ev.description,
      summary: ev.summary
    })
  end
end
```

Work represents the number of worked hours from the user. This data is imported from the Hours App (HOURSTIMETRACKING.COM, 2015). Hours is an application where the user tracks his worktime by pointing his task during the day. The information collected consists on a label representing the type of activity the user was involved and the initial and final timestamps. We retrieved this information by exporting the user's data and importing with a custom script developed by us, in a proccess similar to the appointments data.

As mentioned before, all these information comes from different sources. So we had to adapt much of the incoming data. Yet it is very important to highlight that we did not change the content or any sensitive data. Our transformations were limited only to adjusting time formats and timezones, and dealing with different precisions. We store everything within the database of the application for organization and performance reasons.

The amount of data collected varied from source to source. Location entries were very sporadical and only had a few dozen registers on the database, while activity information had small intervals of registration, being able to store several entries a minute, and thousands each day.

5 EVERYDAY VISUALIZATION

We aim to provide an intuitive visualization that can be used by the biggest number of users possible. That means that the end user does not have to be a data analyst or professional of any area to understand the information that is being visualized. Additionally, the amount of training and effort to understand the data being displayed should be minimal. Further, we also produce highly customizable and adaptable visualization to each user's information.

Considering that most user's data are temporal, we decided to base our visualizations in the *Calendar* and *Clock* metaphors. The central idea is that the users would be able to view their information at different levels of detail. This way, the calendar provides an overview of the self-behavior along several days, while the clock gives a detailed view of a single day, hour by hour. The difference in perspective going from a broader to a more detailed one allows the user to focus on the scope of information he wishes. Additionally, some of the information collected presented a spatial characteristic besides temporal. Consequently, we added a *Map* visualization to present this point of view of the data. *Everyday Visualization* is composed by the few visualizations presented bellow.

These visualization formats such as calendar, clock and map are simple considering they are analogous to very common things. Still the clock can be considered a bit more complex as we do not use a traditional twelve hours one in EverydayVis. Instead we put an entire day of information in 24 slices separated in different arcs.

The idea of using a circular visualization is still inside our base guidelines for designing a visualization. Many authors have discussed these concepts, like Bergstrom and Karahalios (BERGSTROM; KARAHALIOS, 2007), Dragicevic and Huot(DRAGICEVIC; HUOT, 2002), and Larsen et al. (LARSEN; CUTTONE; JØRGENSEN, 2013). The later shows a similar case with the use of spiral visualizations, as can be seen in Figure 5.1. The authors argument that this format allows users to discover patterns of periodic activity which is highly desired. Even though our visualization does not allow the same depth, it has a similar effect of finding patterns, as all clocks have the same positions in each day. It is fairly easy to compare to clocks side by side.

We present a more detailed description of our visualizations in the following sections.

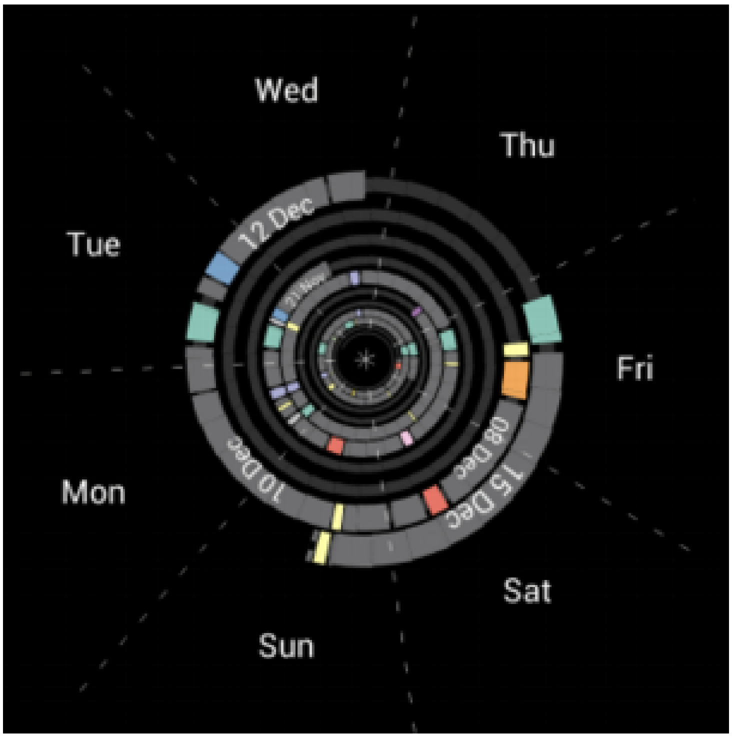

Figure 5.1 – Spiral visualization with a concept similar to EverydayVis' clock.

5.1 Calendar

The *Calendar* represents a broader visualization that contains information of overall activity, total worked hours, and weather condition of each day (see Figure 5.2). The information is presented in a clear and intuitive way without excessive details and allows the view of longer periods of time, such as a week or a month.

The calendar visualization is composed of several cells, each one containing the information of a single day. Each cell has the following attributes:

- date: the number indicating the day of the month
- weather condition: an easy to recognize symbol on the opposite side of the number, indicating if the day was sunny, rainy or cloudy.
- total activity: a subcell showing the total activity of the user. It is configured to show the number of steps or the activity value by another sensor.
- amount of worked hours: another cell on the calendar showing information of amount of the worked hours.

Each cell also presents a check box, besides the number of the day, for the user to select the respective day.

The visual of the calendar is very colorful so that users can easily see which day was more

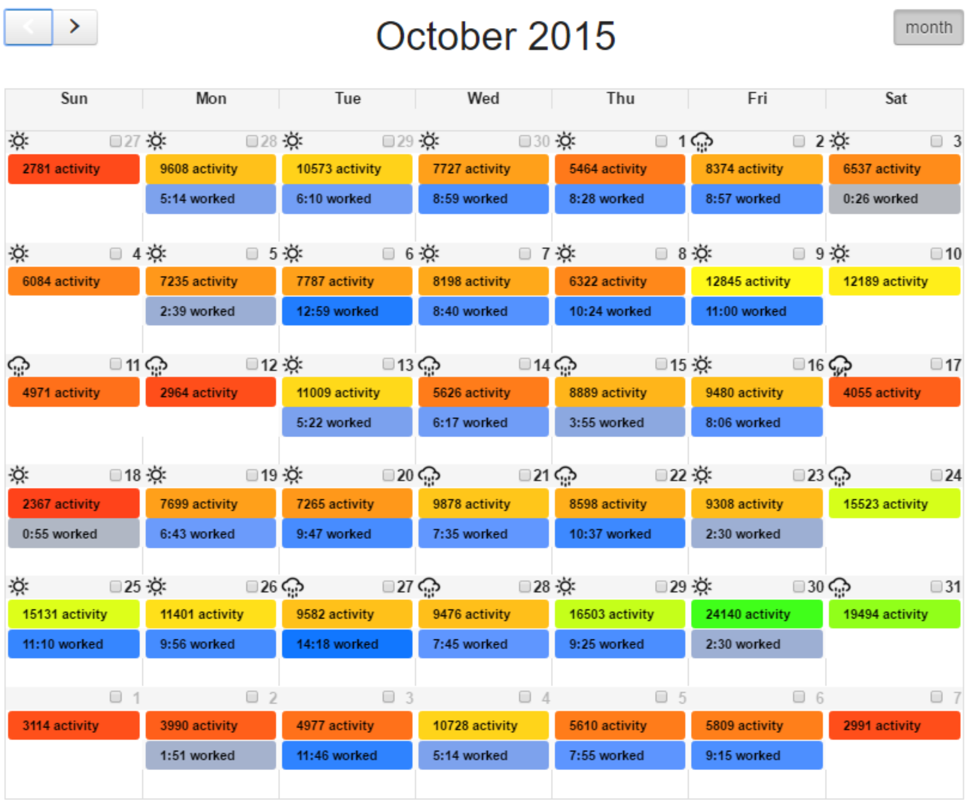

Figure 5.2 – Basic calendar visualization showing user amount of steps (activity) and the amount of work done. The color in the activity blocks goes from red (less activity) to green (more activity), while the color in the worked hours ranges from gray (less hours) to blue (more horus).

positive/negative to them. The activity cell goes from red which means less, passing through yellow, and finishing in green, which means more activity. The color varies according to the own user, being the day with most activity the most green one, and the day with the least activity the most red. This is interesting because the user will always try to beat her/himself and aim to always keep everything with higher values.

Similar to the activity, the work information goes from gray to blue, indicating less and more worked hours, respectively. We did not use a wide color range for the worked hours because the visualization already has enough colors as it is. The activity was chosen to be the most highlighted because it is the most common activity monitored by the popular sensors.

Lastly, when an activity cell is clicked, it shows a crude bar chart (see Section 5.3) with the corresponding activity data of the entire day. This goes along with our principles of always allowing users to access raw data.

5.2 Clock

The *Clock* represents a more detailed visualization that contains several arcs of information for each user. It is adaptable since it only shows information that has been collected from each particular user. Unused arcs are hidden. We made an effort to keep the clock visually simple. So, the only label presented is the hour indication. Any other text is only displayed as a tooltip and will not pop up unless the user interacts with the clock.

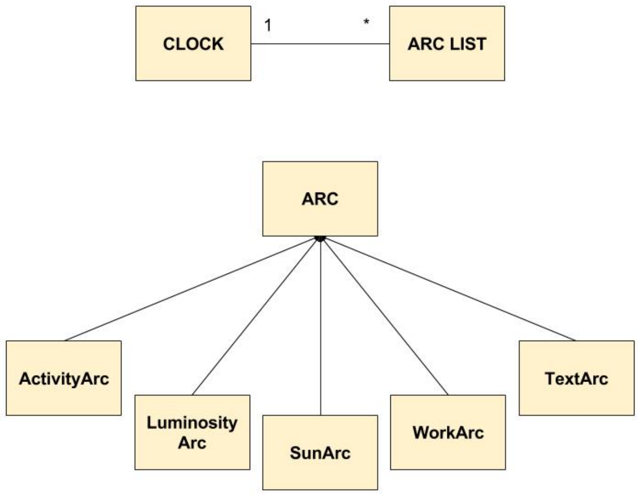

Figure 5.3 – General arc structures used to develop and present clocks on the EverydayVis application. Notice that we do not indicate any limitations to arc numbers as, even if all arcs are present, the visualization correctly displays the data. In the future, if there are too many arcs this would have to be worked.

Figure 5.3 above shows an overview of how the clock was designed to be able to show adaptable information. Each instance of a clock has a list of arcs which in case can be of different kinds. Depending on the information the user has on each day, the clock will only present the respective arcs. For example, if a user has no information on luminosity, the LuminosityArc is just skipped while rendering the interface, thus showing a more concise and better adaptable visualization. The text arc is responsible for showing the labels with the hour number.

The *Daylight* is usually the most external arc and represents the sunrise and sunset hour of the day, captured from OpenWeather API (OPENWEATHERMAP.ORG, 2015). The information is represented by a light blue color with a yellow circle (the sun) in the middle representing daylight, and a darker blue tone with a light gray circle (the moon) in the middle representing the night time.

The *Luminosity* arc represents the amount of light the user has received during the day. The arc is divided into 24 slices identified by colors. The clearer the slice, the more light the user received in that hour. Respectively, the darker means the less light received. Unfortunatelly this data is not easily accessible in any device. We only collected luminosity information from the Actwatch 2 wristband, described in Section 6.2.

The *Activity* arc is generally the most important, or at least the most popular. It is represented by 24 slices with a color filling each slice in the proportion to the amount of activity the user had in that hour. The hour of the day with most activity will be the biggest bar in the clock. This is also adapting to the user as it shows activity collected in different ways by different devices. For some users it represents an activity coefficient, while for others it shows the amount of steps taken during the day. Meanwhile, the final information is the same for all general purposes, as the number of steps is a lot related to the quantity of activity a user does.

The innermost arc represents the *Working* periods divided by tasks according to the user's personal customization. Each color represents a different task. The tooltip information reveals the name and exact amount of time spent on each task as can be seen in Figure 5.4.

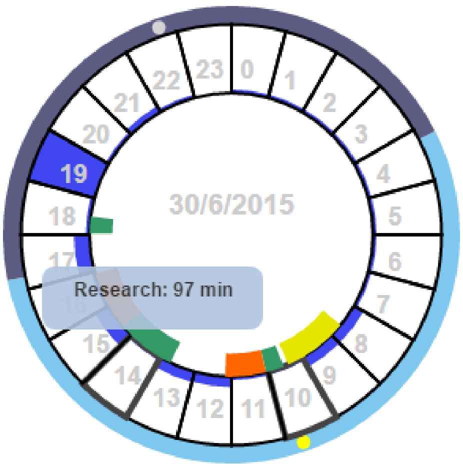

Figure 5.4 – Clock with work information showing tooltip.

In the center of each clock there is a few extra general information, such as the date, the weather, the maximum and minimum temperature, and the icon representing the weather events of the day. We also added information about *Appointments* to the clock. In the activity arc, the slices with a thicker border represent that an appointment happened during that hour.

Figure 5.5 exemplify two different uses of the *Clock* visualization.

5.3 Bar chart

The *Bar chart* presents a well-known visualization with activity information of the user. Depending on the context, the data shown corresponds to an hour (Figure 5.6), or to an entire day (Figure 5.7). It is important to notice that the bar chart data is adaptable to the context of

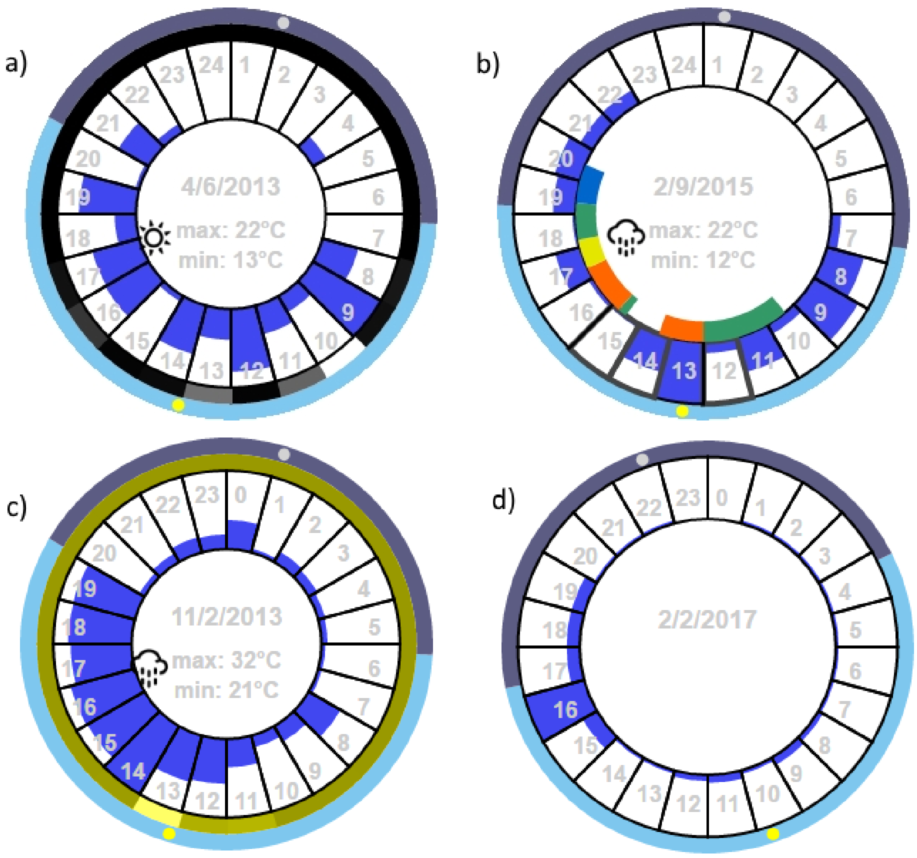

Figure 5.5 – a) *Clock* visualization including the activity, luminosity and daylight arcs. b) *Clock* visualization including the worked hours, activity (using the number of steps) and daylight arcs, as well as the appointments. c) *Clock* similar to a but with a different collor pallet for the daylight information. d) A very crude *clock* visualization only with activity and dailyght information.

the user, being able to show information independent of the amount of data collected from each user as can be seen in Figure 5.8.

The bar charts, while not a novel visualization are considered important as complimentary information. In our principles we point that access to "raw data" is necessary, so we allow users to visualize this information.

5.4 Map

The *Map* shows the locations where the user has been in the respective selected days in the calendar. This visualization was developed to offer a different view of the data – this time focused on location – instead of presenting it according to time. As the map visualization is based on Google Maps API (GOOGLE.COM, 2015), so features such as zoom and pan are already implemented and easily integrated. Users can quickly navigate through their check-ins as can be seen in Figure 5.9. Information from different dates are shown in different colors,

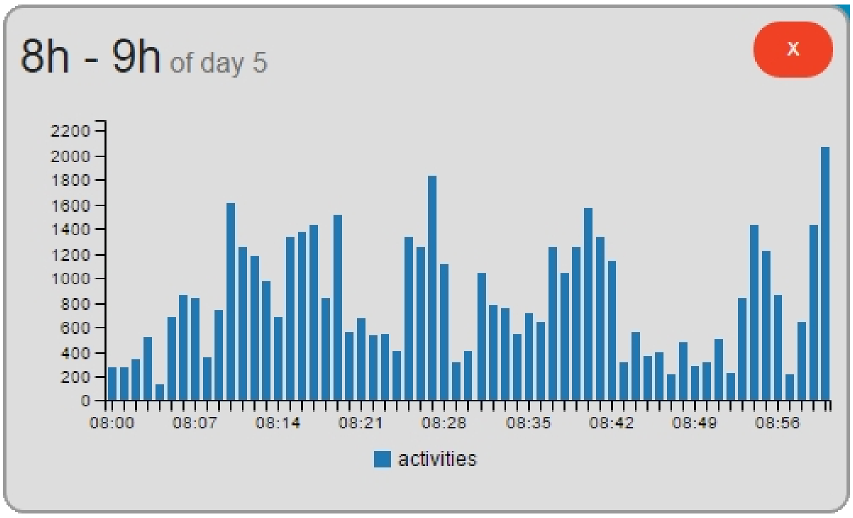

Figure 5.6 – Bar chart of one hour of activity. This box is presented when the user interacts with an activity in the clock.

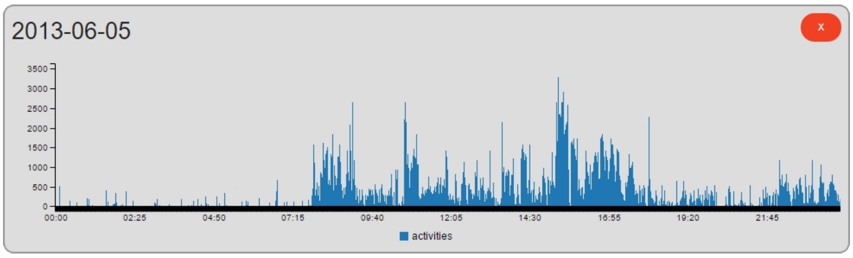

Figure 5.7 – Bar chart of one day of activity. This box is presented when the user clicks on the "H" button on a calendar cell.

according to the activity color in the calendar visualization. As the color range has consistent meaning with the calendar, it is easy to identify where the user has more or less activity. It is also important to stress, that this visualization is quite imprecise since the source of geographical locations are check-ins, made with Foursquare/Swarm App.

Following the idea being used up to now, the map is not displayed to the user that does not have checkin information.

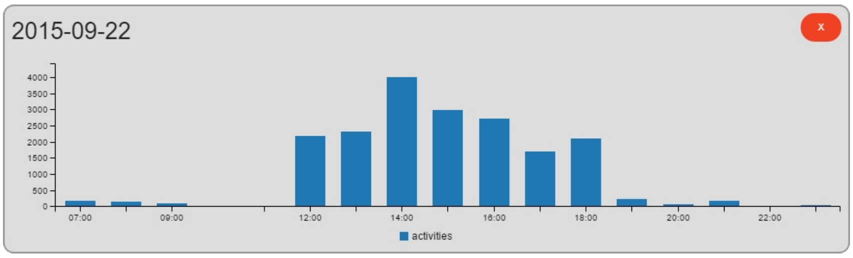

Figure 5.8 – Bar chart of one day of activity showing a user with just few data collected.

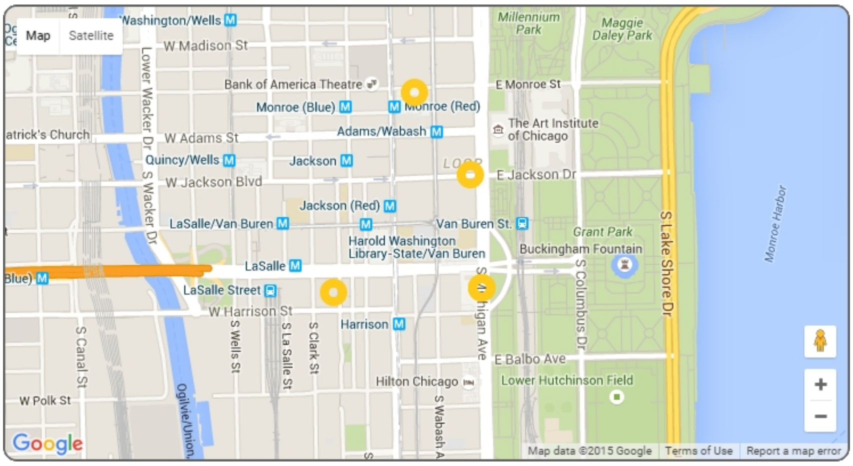

Figure 5.9 – Zoomed map with markers from a single day. Each marker is an entry on Foursquare/Swarm App.

5.5 Exploring the Data

The application shows the different levels of detail with a simple and concise, yet powerful, interaction. Initially, the user is presented to the calendar visualization, the base to all user exploration. From this point, the user can click on the activity cell of any day to visualize the bar chart of an entire day of activity. The bar chart has a tooltip to facilitate the identification of specific data. The map with the check-in locations of the user is also presented below the calendar, but only if the user has any check-in data registered. In the calendar, there are also checkboxes beside the day identifier. When the user selects any date, the map is automatically updated, showing only the check-ins regarding the selected days. The map, as mentioned before, has zoom and pan actions that allow users to explore their locations in relation to their activities.

Additionally, when the user checks a day in the callendar, it dynamically fetches the data and presents the selected clocks to the user. Each clock shows an aggregation of information of a single day and are displayed vertically one after the other. If the user hovers the mouse on each arc of the clock, there is a dynamic tooltip that shows the context information. For example, if the user hovers the luminosity arc, he will see the total luminosity of the slice. Similarly, if the user hovers a work task, he will see the work information. This allowed us to produce a cleaner visualization with fewer labels, making it easier to understand.

The clock vis offers two options to display the activity during each day. In the visualization, there is an option to visualize information relative to the same day or globally relative to the user. This means that the activity arc will fill the bars in relation to the local or global maximum value. If the user wishes to see local maximums, every clock will have a bar fully filled. On the other hand, if the global maximum is selected, there will be only one slice in one clock completely

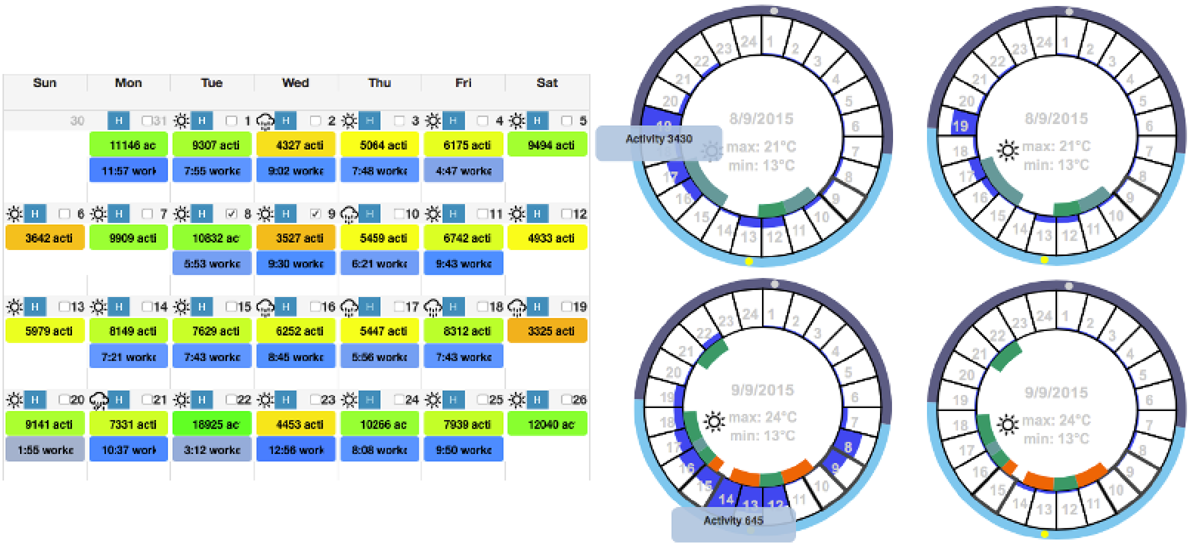

Figure 5.10 – The calendar (left) shows an overview of a single person behavior on September, 2015. The clocks (right) show in more detail how s/he spent the day in the two selected days (September, 8th and 9th). The two clocks leftmost show the activity using the local maximum, while the rightmost clocks show the same data using a global maximum, as explained in Chapter 5.

filled. Figure 5.10 shows the same data in both representations. There is also the possibility of clicking in an activity slice. This action displays a bar chart of the activity regarding the selected hour, such as shown in Figure 6.3.

It is important to highlight that these small customizations are crucial to the user, as pointed by Huang et al. (HUANG et al., 2015).

The Figures 5.11 and 5.12 represent the application flux. Even though the different visualizations are accessible in any context, the user will generally flow like the mentioned figures show. As an example the Figure 5.11 shows a user accessing clock information of two different days, then viewing a specific hour slice information for one of the selected days. The Figure 5.12 shows the user selecting the same days, but, instead, looking at a broader information on the map and the bar chart for one of the days, instead of an hour slice.

5.6 Software Solution

Since the beginning we aimed for a Web application for availability reasons. Managing or distributing this kind of software as desktop or mobile solution would be far more complex and we believe would not be as efficient as a Web application.

Regarding technology, we used the Ruby on Rails Web framework. It delivers a good "out of the box" solution for developing applications quickly. The framework also offers simple manners of importing data from files using scripts.

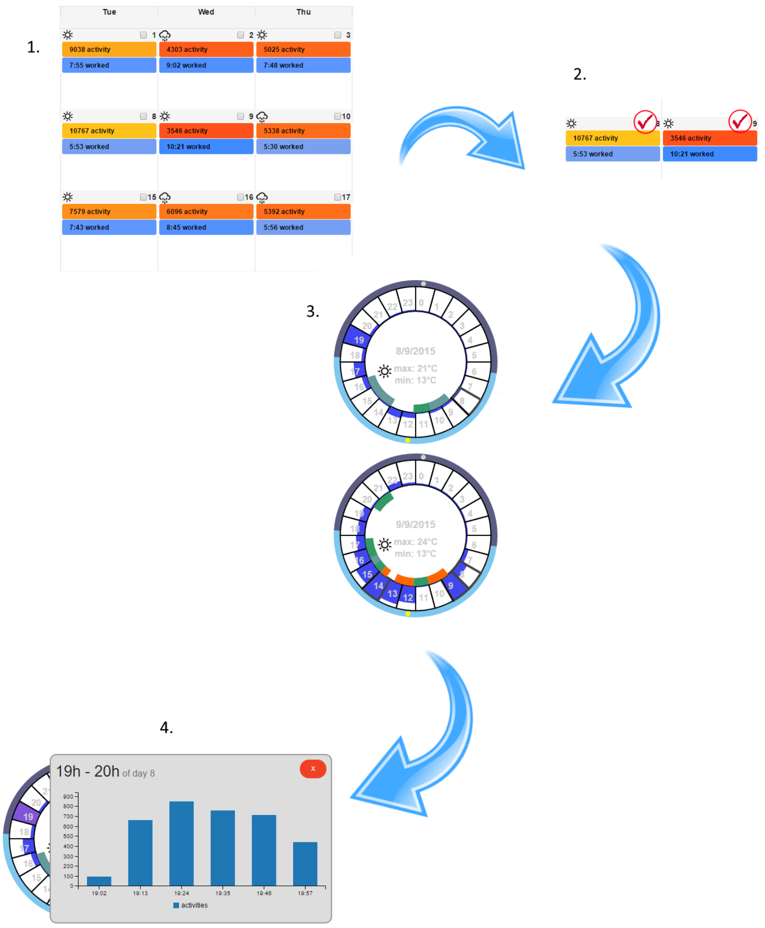

Figure 5.11 – 1. The calendar shows information of a few different days. 2. The user checks the days he wants to look deeper. 3. The user sees the clocks for the selected days and clicks on an hour slice. 4. The software shows the bar chart for the selected time slice.

We also used PostgreSQL for database as it is a highly used software and has the capabilities of storing and quickly accessing high amounts of data.

For the interface of the application we used javascript apis such as:

- JQuery: a small library that offers several features for javascript.
- D3js: a visual API that allows the development of graphics and visualizations.
- FullCalendarJs: a library which offers a customizable calendar view.
- GoogleAPI: the javascript version used for invoking the google maps view.

For hosting the application we use a nginx server running in a Microsoft Azure VPN. We configured all the software structure to run exclusivelly this application.

The EverydayVis application was developed using a MVC pattern as it is the default pattern used in Ruby on Rails. Following that structure, we created several different models, one for each kind of data that was collected. In this manner, we can individually process each information when the user requests it.

Figure 5.12 – 1. The calendar show information of different days. 2a. The user selectes two days to look into their information. 3. The user looks at the map to verify where he has been in the selected days. 2b. Shows the bar chart of activity for the day the user clicked.

For the interface we built a series of Javascript files that comprise two important structures: the calendar and the clock.

For the calendar we customized the FullCalendar (SHAW, 2015) library. We added "events" that show on each day cell related to user daily information as can be seen in Figure 5.13. Each event is filled dynamically with the user information. We also added other functionality to the calendar that are application dependent.

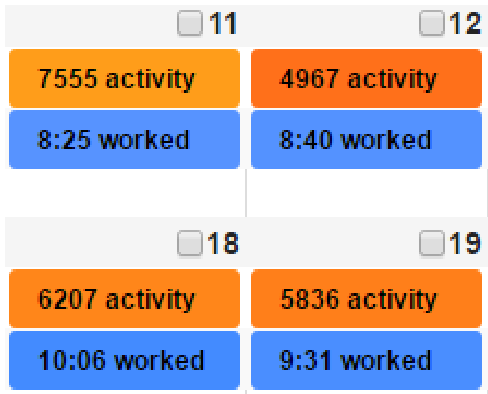

Figure 5.13 – Cells of a calendar filled dynamically

The clocks were developed using D3js (BOSTOCK, 2015) and are composed of customizable arcs. Each arc is independent and can appear in any order as the developer configures it. The arcs respond to a few default functions to be able to be displayed. This allows for new arcs to be added easily.

Figure 5.14 shows an overview of how the information flow happens on the application. User's data is collected and displayed dynamically on the calendar and clocks. We use Ajax to load the interface, this way the user can navigate between the visualizations without needing to reload the page or losing context of his task. The map uses a very similar process. The interface was build using the Google Maps API.

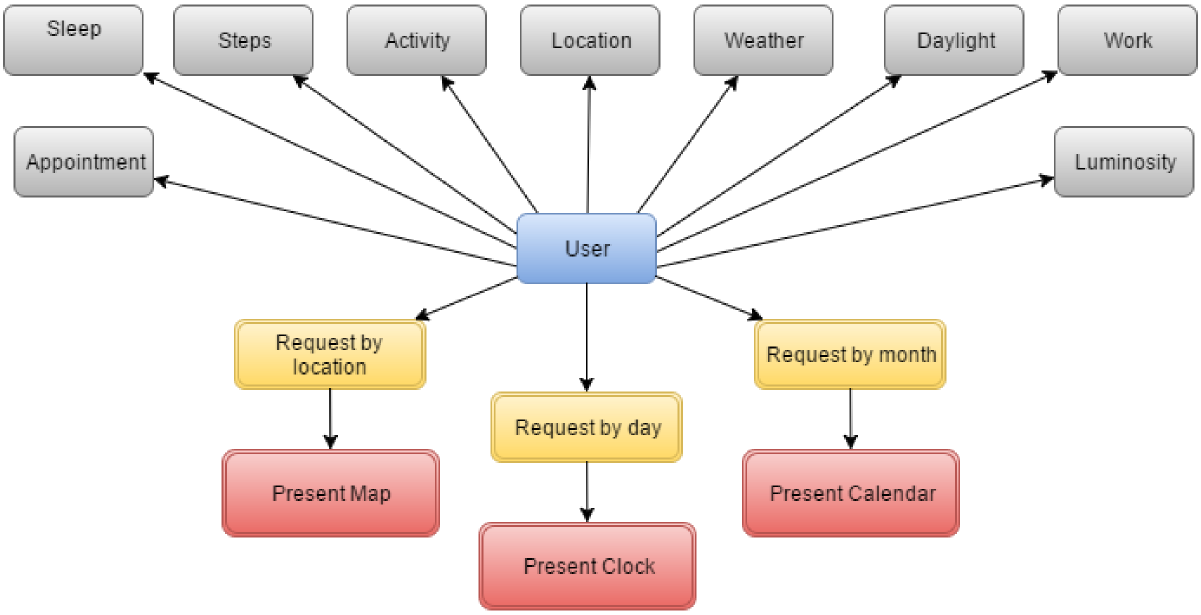

Figure 5.14 – Structure of the *Everyday Visualization* application, showing different data and the basic flow of information.

Even though we have full control of the application and host there were a few problems with performance. The speed in which each clock or calendar is shown depends on how much information the user has stored. We optimized the speed of the software through dynamic loading because, if we loaded all the information along with the page, the software would feel too slow, almost unusable. Still, because of hardware limitations our application continues to have a few performance issues. The problem is aggravated by the fact that almost all data is time sliced (divided in periods of time), which frequently requires complex queries on the database, even for minor details in the visualizations. While we could work to restructure data or optimize the code, it would not be a worthy investment, as speed is not the main focus of this work. As long as the visualizations are being loaded and presented, even if a bit slow, the users viewing it should be fine.

The server machine is a shared Microsoft Azure instance with an nginx application server running a Ruby on Rails application.

6 RESULTS

In this section, we present use case examples of *Everyday Visualization* that illustrate how people appropriated the system for distinct purposes. In the cases we used an informal approach to the users and considered their opinions on the use of the platform and visualization of their data. We had a few examples of use for self-improvement. As an extra view on the application we had a health scientist interested in chronobiological data of a specific community. For the sake of simplicity, we are considering a single dimension (the number of daily steps).

Also, we consider that the primary objective of each person is to have a good life. The concept of *good life* may vary according to one's personality, but it is often a good balance of health, work, and social activities. Each one has just a single life, and it is not easy to achieve happiness disregarding one of these aspects. However, keeping a balanced life requires personal investment, self-knowledge, and discipline. There are difficulties to manage this alone. The pursuit of a better behavior depends a lot on the knowledge of the past and, as a consequence, on the everyday monitoring.

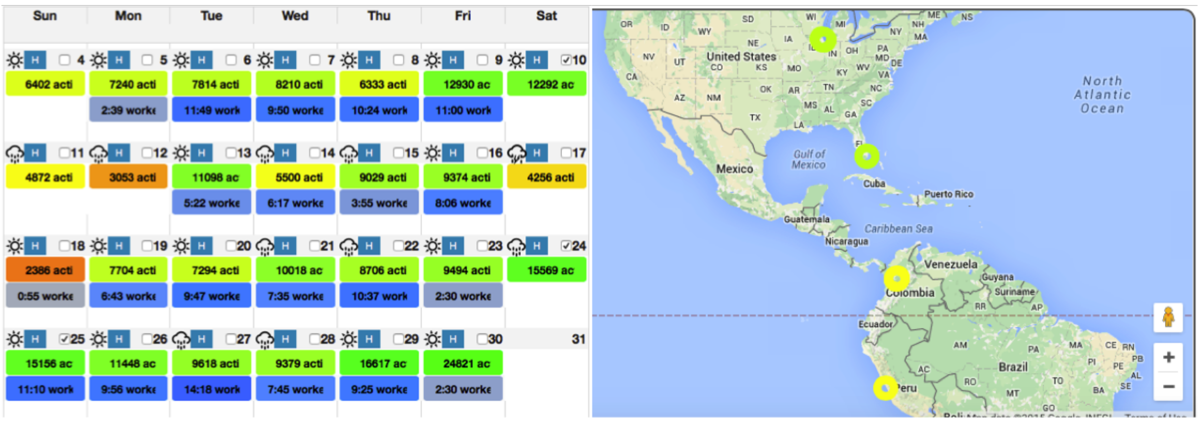

Figure 6.1 – The calendar (left) shows an overview of a single person behavior on October, 2015. The map (right) shows where this person was on the days selected in the calendar (October, 10th, 24th and 25th).

6.1 Use Case 1: Self Improvement

In this use case we analyze a single user observing his/her own personal data. We presented the application to the user and allowed him/her to associate the data through a few application hooks. In this case the user could associate their activity tracker or foursquare and have the application present them customized data.

A typical standard – but configurable – recommendation given by activity wristbands (e.g.

Fitbit and Jawbone UP models) is to walk at least 10,000 steps a day for a sustainable living. It is possible to count steps using any activity monitor, as well as to see the personal progress, the history, the mean number of steps in a week, and so on in its associated App. This information helps a lot in the self-improvement as motivation, but it is still not enough to understand why we are having such performance and what to do to improve it.

By expanding the visualizations to different contexts we allow the users to better understand and make sense out of their data. The calendar overview shows daily behaviors, and it is easy to relate the number of steps with the number of working hours and the weather. The calendar on Figure 6.1 shows a user's behavior in October 2015. We can observe that this person tends to do less activities during weekends and holidays. However, there are exceptions. That is when we can start asking the data questions. Why does this person walks much more in some weekends than in others? The application helps the user to figure out possible reasons. When these more "productive" days are selected the visualization on the map gives the user a good insight. With both visualizations side by side, it is easy to perceive that in these days this person was outside his/her hometown. Expanding on the same idea the user could notice that it is also a good explanation on why s/he has such high amount of activity on the entire last week of October, where the user was, again, away from home.

In some cases the analysis of the data can get quite complex. Figures 6.1 and 5.10 do not show a clear relationship between steps and amount of worked hours, neither an influence of the weather on the daily activity. However, considering the idea of personal visualization is that the user is looking at his/her own data a little bit of personal input can be taken into account. Based on this we looked at the first three weeks of September and found out that s/he tends to move less on Wednesdays and Thursdays. Given the question we went to the next step to try to answer it. By selecting a Tuesday (September 8th) and a Wednesday (September 9th) on the calendar and opening the clock visualization we hope to find in the differences a reasonable explanation for the given question.

- September 8th: The user reported having practiced a sports activity at 7PM and having one appointment during the day.
- September 9th: No physical activity, and three appointments reported.

With the user's information we can present a few conclusions. First we can point that the number of appointments is probably reducing the amount of activity of the user on Wednesdays. And second, practicing sports causes a reasonable impact in this specific user's activity clock. Even though the information seems pretty obvious, it is important to highlight that the visualization can help the user in this sensemaking proccess.

On an interesting extra note we can also see that the user has a regular rythm of activity on Tuesdays. On September 1st and September 15th s/he had the same "spike" of activity at 19PM as on the 8th (see Figure 6.2). This represents the reported sports activity the user reported.

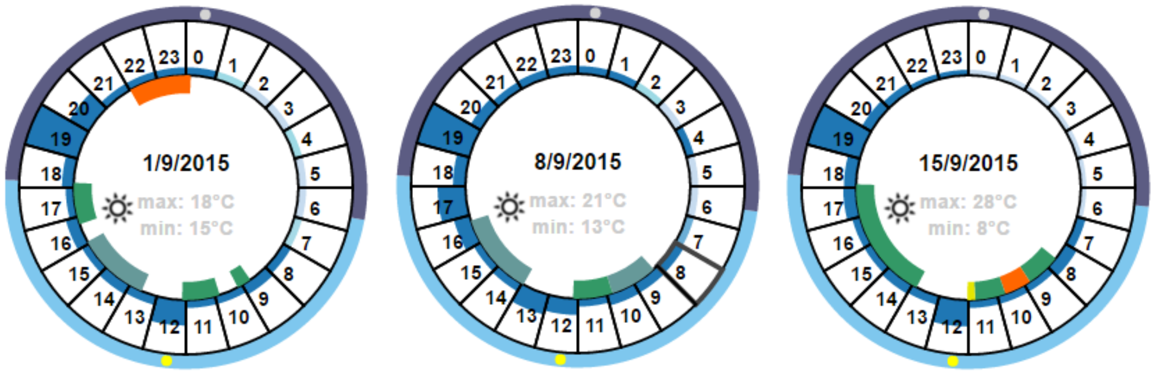

Figure 6.2 – Clocks showing the rythm of a user on the same week day.

Exploring such data described here involves a lot of interaction between the three possible visualizations (calendar, clocks, and map) and can help the user to better understand his/her own behavior, and what s/he needs to do to change it in the future. The goals may vary from person to person, but it is far easier to define and measure objectives with a visualization tool like this.

6.2 use Case 2: A Use on Chronobiology

The second use case happened with the help of a health scientist interested in analyzing individuals "biological clock". She was using a provided software for visualization when we proposed the experimentation of EverydayVis as an alternative to visualize her data. The main idea was to validate in a professional context if the visualization could help an advanced data user in some manner.

Her study describes the comparison between lifestyles of people in urban areas and people in very restricted areas with basically no access to artificial light. She explains that it is important to her to analyze daily activity and luminosity patterns of individuals. The biological clock and internal body "rythm" have been associated with a variety of health issues, including neuropsychiatric disorders, obesity and smoking (WITTMANN et al., 2006; LEVANDOVSKI et al., 2011; ROENNEBERG et al., 2013). And this is precisely the kind of information that can be observed in our visualization.

The sample data collected from ten young adults was conducted in a rural population with no access to artificial light. The data collected refers mostly to the communities living away from the cities. This study was performed according to international ethical standards, in accordance with the Declaration of Helsinki. Ethics approval was obtained from the local internal review board, and all participants signed an informed consent form.

Circadian rhythm variables were collected by actigraphy. The actigraph model used in this case was Actwatch 2 (Mini Mitter Company, Inc. Bend, USA). Actwatch is an activity monitor designed for long-term monitoring of human subjects and measures activity and ambient light. Actwatch contains an accelerometer that is capable of sensing any motion with a minimal resultant force. The data was read by the Actwatch Interface Reader and converted into an Excel file. The level of activity and light exposure were collected throughout 24 hours continuously for 2-4 weeks, with a wristband Actwatch. These devices allow us to track the subjects' rhythms in their own habitats. As a further factor, they are not invasive, allowing the capture of day and night movements in a usual behavioural life.

In shorter words, the main idea is that this professional works with biological cycles and "rythm" and was using generic visualizations like bar charts and area graphs to visualize the data from all her users. Still this kind of visualization presents a lot of difficulties as it is very crude. It would be even more painful to mine data in such large and multidimensional databases. Specially because the data scientist worked on this information for several weeks. The use of visualizations is absolutely necessary, but the main question was if our visualization were good enough.

We aggregated a subset of her collected data and presented in the EverydayVis visualization. The calendar and clock visualizations were used and the map disregarded as the location of each user did not vary with time. They stayed in the same village for the entire experiment.

After using the EverydayVis the data scientist considered the casual visualization tools (calendar, clock) very intuitive for her data. She emphasized that it allowed her to easilly verify information that was considerably harder to visualize in bar charts. The simplicity of the visualizations we offered was a highlight.

The professional pointed that while analyzing the clocks of the users it became clear that the population had earlier sleep patterns than people generally have in city life. This behaviour was expected as they rely on natural light and work mostly outdoor. It is important to notice that as these users are quite uncommon to the urban lifestyle. Their daily behaviours have slight differences. The users differentiate the week and weekend as "free and working days". Based on that we compared their free and working activities as can be seen in Figure 6.3. It is worth noting that the pattern of rest activity and sleep show a significant difference between these days. When we consider the same time slice comparing the resting day to the working day (see Figure 6.3(b) and (d)) the subject presents a far lower level of activity on the free day. It could be attributed to the fact that on the free day it was raining.

Even though there was only one professional use case of EverydayVis, the application got a positive feedback. She indicated that the proposed visualizations can be really useful to apply on chronobiology field to reveal the 24 hours pattern of cycle data and easily expose the rhythm of the date. She also pointed out that it is much more interesting to present this data to other data scientists using EverydayVis as it is easier to visualize much of the information.

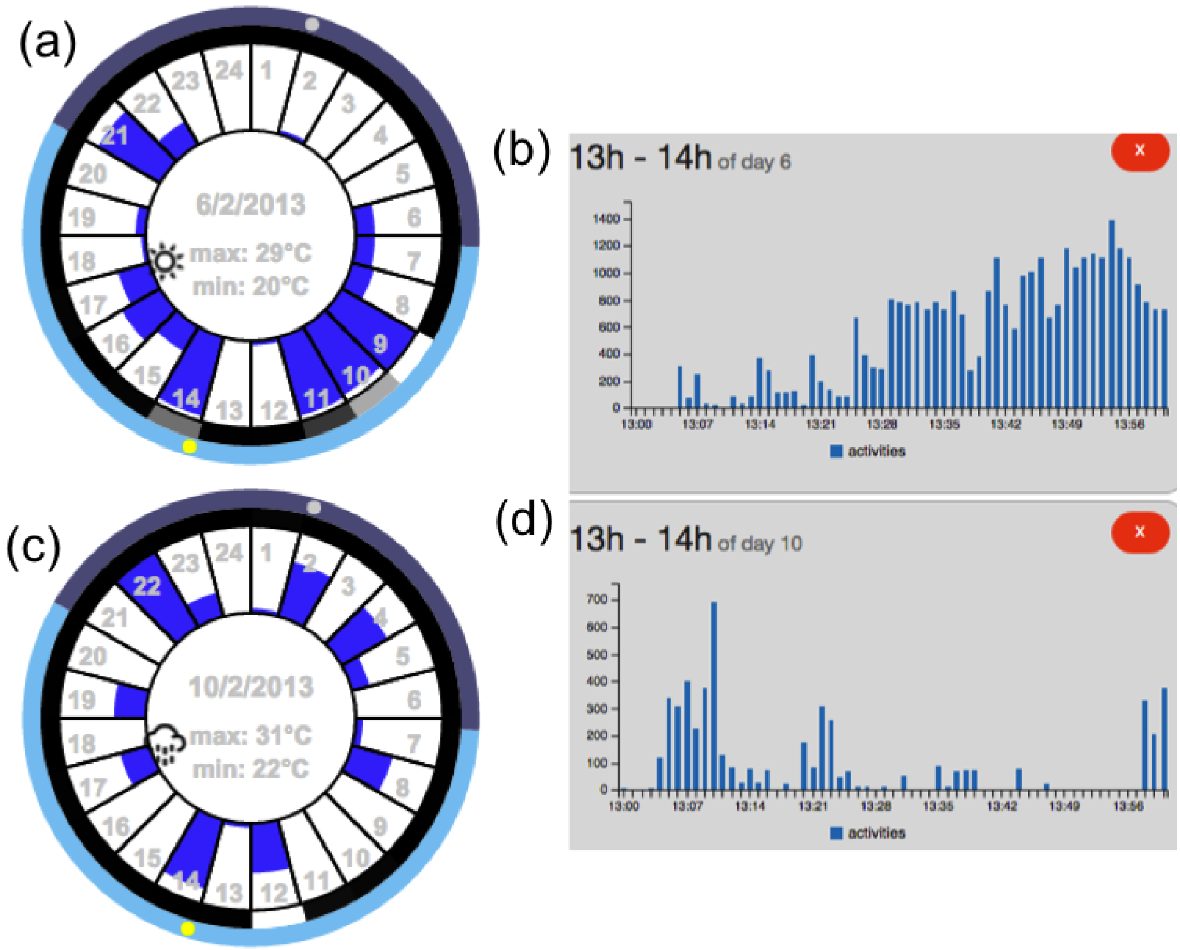

Figure 6.3 – The clock view (a) shows data from a typical working day, while the bar chart (b) depicts the movement in a specific hour. Clock (c) and bar chart (d) illustrate an individual behavior on a resting day.

7 CONCLUSIONS

In this project we presented a set of visualizations focused on individual users that aggregates data from several sources. Our visualizations are separated in levels of detail and offer a natural and seamless interaction that allows users to go through their information without much effort. We also discussed the challenges of collecting and managing user data. Furthermore, we show two different uses of our application to produce insights over the user's data. It is important to notice that both cases have analyzed individual data, but from different perspectives. In the first the user looks at his own data while in the second user data is analyzed by a professional looking at information from other individuals. In both use cases users had no negative comments regarding visualization or interaction.

The *Everyday Visualization* is a continuous project that will always have improvements and tweaks. However we are confident that the main set of visualizations produced based on simple metaphors such as calendar, clock and map are good enough to provide individuals with useful information. We are going beyond simple activity monitors and letting the end user evaluate his/her own information and to have insights about his/her own objectives and.

7.1 Contributions

We presented three main customizable visualization specifications based on analogies. They can be used in several contexts by other researchers or developers with the correct adaptations. All the code for an open source application is available in a public repository and can be used/-forked by any interested developer at https://github.com/brunopagno/everydayvis.

Figures 7.1, 7.2 and 7.3 present a exploded vision of the multiple datasources spread around the visualizations. It serves to show a broader perspective of how much information we have in small places while still maintaining consistency.

The visualizations produced in this work have a few characteristics that even though are not new, when adapted to the proposed scenario present unique positive characteristics:

- Adaptability: the visualizations automatically adapts to each user, becoming a very positive solution for different contexts of data.
- Contextualization: the different levels of visualization allow the user to keep contextualized with their information and analyze specific parts however s/he wishes.
- Minimalism: it does not force excessive information that complicates the interface for

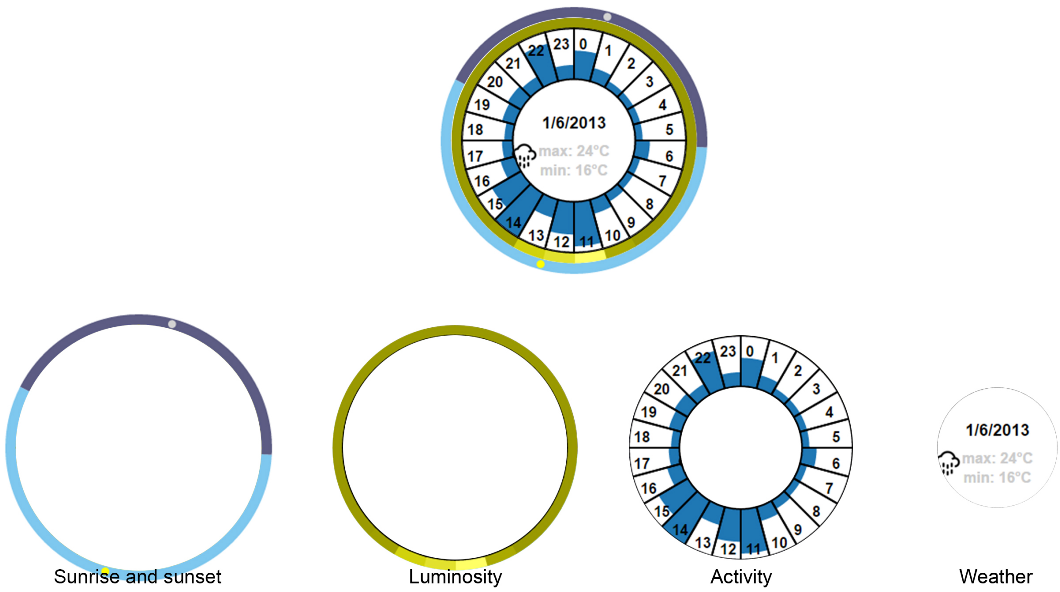

Figure 7.1 – Example of user's clock with three kinds of datasource

users that do not to want to go deep in their data.

It is also important to notice that Everyday Visualization is different from other personal and casual visualizations discussed on the related works. We apply calendar, clock and map metaphors to a new context that was not explored before. This combination presents the different levels of context necessary to view the data from broader to narrower points of view. Additionally, very few related works explore more than one type of data in their applications/visualizations, which is a huge technical difference.

7.2 Limitations

While we did a reasonable job acquiring several data sources and presenting them in our visualizations, the software still lacks in automatization. Many of the data sources are still handled manually and should somehow be seamlessly integrated in the application, otherwise it becomes very hard for users to join the EverydayVis an visualize their data. The resources that automatically collect data are still very limited. Still, considering the experimental phase of the application we believe it is acceptable that not all the information is automatized yet.

Another limitation is the fact that we have a large amount of registries in our database and present visualizations that do intense processing of this data. Considering the hardware in which we worked on during the development and deployment of the application we had to deal with a few slowdowns when loading the pages. However, for the same reason mentioned above, it was

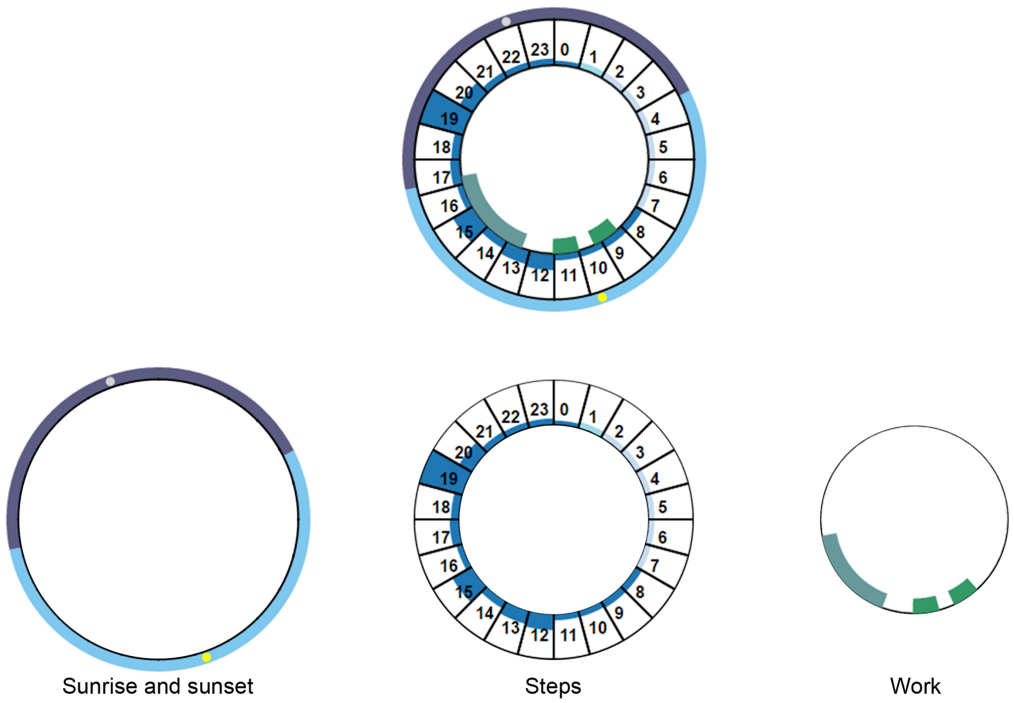

Figure 7.2 – Example of user's clock with four kinds of datasource

acceptable considering the experimental stage of the application.

7.3 Future Work

There is a clear potential to the *EverydayVis* application. The hability to visualize and relate the exposed information is very interesting to the end users. That said, we believe that expanding the variety of data sources and automaizing the data collection will expand the number of users and the quality of the insights they will be able to have. Also, a bigger number of users will

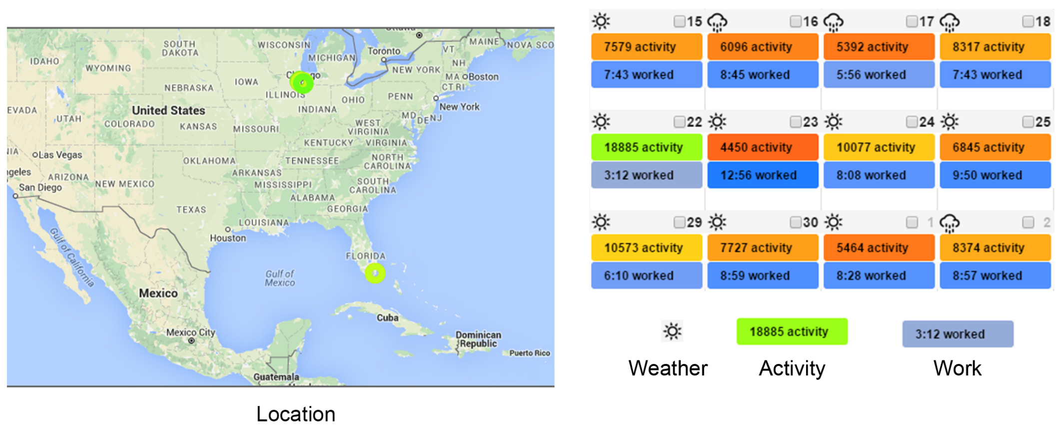

Figure 7.3 – Other datasources and their small components

allow us to make broader and better evaluation on the system, including redesigns based on users' needs. In other words, after attending to the basic needs of a user, we will expand the application to better accomodate our findings.

7.4 Publications

This work is the result of a learning experience in the field of data visualization. The research work usually takes a reasonable effort on exploring different possibilities and it was not different with EverydayVis. We worked with different datasets and contexts of data visualisation before actually developing this application. As a consequence, we published the following list of articles:

- *NBAVis: Visualizing National Basketball Association Information* (PAGNO et al., 2014): Based on an NBA dataset, we created a few visualizations and compared individual players and teams
- *Interactive Timeline Visualization of Documents* (GUEDES et al., 2014): A timeline visualisation exploring data from text documents and keywords
- *Guidelines for Designing Dynamic Applications With Second Screen* (PAGNO et al., 2015): An experiment with multiple screens in which we elaborated guidelines for the development of interfaces
- *Everyday Visualization* (PAGNO; NEDEL, 2015): The initial part of this work in which we first started exploring the clock visualization.

Each of these works allowed us to evolve and get closer to the *EverydayVis*.

REFERENCES

AGGARWAL, C. C. **Managing and mining sensor data**. [S.l.]: Springer Science & Business Media, 2013.

BERGSTROM, T.; KARAHALIOS, K. Conversation clock: Visualizing audio patterns in co-located groups. In: IEEE. **System Sciences, 2007. HICSS 2007. 40th Annual Hawaii International Conference on**. [S.l.], 2007. p. 78–78.

BOSTOCK, M. **D3.js - Data-Driven Documents**. 2015. Available from Internet: <http://d3js.org/>.

BUONO, P.; CUZZOCREA, A. A collaborative framework for supporting combined visualization of activities across time zones. In: **2015 International Conference on Collaboration Technologies and Systems (CTS)**. [S.l.: s.n.], 2015. p. 226–232.

COELHO, D.; KUMAR, A.; MUELLER, K. Data memes for personal visualization. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015.

CUTTONE, A.; PETERSEN, M. K.; LARSEN, J. E. Four data visualization heuristics to facilitate reflection in personal informatics. In: **Universal Access in Human-Computer Interaction. Design for All and Accessibility Practice**. [S.l.]: Springer, 2014. p. 541–552.

DRAGICEVIC, P.; HUOT, S. Spiraclock: A continuous and non-intrusive display for upcoming events. In: **CHI '02 Extended Abstracts on Human Factors in Computing Systems**. New York, NY, USA: ACM, 2002. (CHI EA '02), p. 604–605. ISBN 1-58113-454-1. Available from Internet: <http://doi.acm.org/10.1145/506443.506505>.

ETEMAD, K.; SAMAVATI, F.; CARPENDALE, S. Artistic inspired data visualization design process. The Eurographics Association, 2016.

FITBIT.COM. **Fitbit Official Site for Activity Trackers More**. 2015. Available from Internet: <https://www.fitbit.com/>.

FIT.GOOGLE.COM. **Sign in - Google Accounts**. 2015. Available from Internet: <https://fit.google.com/>.

FLEMISCH, T. et al. Necklan: Language as jewellery. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015.

FOURSQUARE. **foursquare for Developers**. 2015. Available from Internet: <https://developer.foursquare.com/>.

FUNG, T. L.; MA, K.-L. Visual characterization of personal bibliographic data using a botanical tree design. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015.

GOOGLE.COM. **Maps API: Application Development Analytics**. 2015. Available from Internet: <https://www.google.com/work/mapsearth/products/mapsapi.html>.

GOU, L. Visualizing personality traits derived from social media. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015.

GRUNDY, E. et al. Visualisation of sensor data from animal movement. In: WILEY ONLINE LIBRARY. **Computer Graphics Forum**. [S.l.], 2009. v. 28, n. 3, p. 815–822.

GUEDES, L. S. et al. Interactive timeline visualization of documents. **SIBGRAPI - Workshop on Visual Analytics, Information Visualization and Scientific Visualization**, 2014.

HOURSTIMETRACKING.COM. **Hours Time Tracking app**. 2015. Available from Internet: <http://hourstimetracking.com/>.

HUANG, D. et al. Personal visualization and personal visual analytics. **Visualization and Computer Graphics, IEEE Transactions on**, v. 21, n. 3, p. 420–433, March 2015. ISSN 1077-2626.

HUANG, D.; TORY, M.; BARTRAM, L. Data in everyday life: Visualizing time-varying data on a calendar. **Proc. Poster Compendium IEEE VIS**, 2014.

JACUCCI, G.; FAIRCLOUGH, S.; SOLOVEY, E. T. Physiological computing. **Computer**, IEEE, n. 10, p. 12–16, 2015.

JAWBONE. **UP by Jawbone™ | No Activity Tracker Required**. 2015. Available from Internet: <https://jawbone.com/up/app>.

KONDO, B.; COLLINS, C. Dimpvis: Exploring time-varying information visualizations by direct manipulation. **Visualization and Computer Graphics, IEEE Transactions on**, v. 20, n. 12, p. 2003–2012, Dec 2014. ISSN 1077-2626.

LANE, N. D. et al. A survey of mobile phone sensing. **Communications Magazine, IEEE**, IEEE, v. 48, n. 9, p. 140–150, 2010.

LARSEN, J. E.; CUTTONE, A.; JØRGENSEN, S. L. Qs spiral: Visualizing periodic quantified self data. In: **CHI 2013 Workshop on Personal Informatics in the Wild: Hacking Habits for Health & Happiness**. [S.l.: s.n.], 2013.

LEVANDOVSKI, R. et al. Depression scores associate with chronotype and social jetlag in a rural population. **Chronobiology International**, v. 28, n. 9, p. 771–778, 2011. Available from Internet: <http://www.tandfonline.com/doi/abs/10.3109/07420528.2011.602445>.

MAH, C. et al. A visualization fingerprint: Comparing nutrient data visually. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015.

OAUTH.NET. **OAuth Community Site**. 2016. Available from Internet: <http://oauth.net/>.

OPENWEATHERMAP.ORG. **Weather API**. 2015. Available from Internet: <http://openweathermap.org/api>.

OTTEN, H. et al. Are there networks in maps? an experimental visualization of personal movement data. In: **IEEE VIS**. [S.l.: s.n.], 2015.

PAGNO, B. et al. Guidelines for designing dynamic applications with second screen. In: **2015 XVII Symposium on Virtual and Augmented Reality**. [S.l.: s.n.], 2015. p. 42–51.

PAGNO, B. et al. Nbavis: Visualizing national basketball association information. **SIBGRAPI - Workshop on Visual Analytics, Information Visualization and Scientific Visualization**, 2014.

PAGNO, B.; NEDEL, L. Everyday visualization. In: **Electronic proceedings of the IEEE VIS 2015 workshop Personal Visualization: Exploring Data in Everyday Life. Chicago**. [S.l.: s.n.], 2015. p. 1–4.

PANTELOPOULOS, A.; BOURBAKIS, N. A survey on wearable sensor-based systems for health monitoring and prognosis. **Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on**, v. 40, n. 1, p. 1–12, Jan 2010. ISSN 1094-6977.

POUSMAN, Z.; STASKO, J.; MATEAS, M. Casual information visualization: Depictions of data in everyday life. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 13, n. 6, p. 1145–1152, nov. 2007. ISSN 1077-2626. Available from Internet: <http://dx.doi.org/10.1109/TVCG.2007.70541>.

ROENNEBERG, T. et al. Light and the human circadian clock. **Handb Exp Pharmacol**, v. 217, p. 311–31, 2013. ISSN 0171-2004. Available from Internet: <http://www.biomedsearch.com/nih/Light-human-circadian-clock/23604485.html>.

SATYANARAYAN, A.; HEER, J. Authoring narrative visualizations with ellipsis. **Comput. Graph. Forum**, The Eurographs Association & John Wiley & Sons, Ltd., Chichester, UK, v. 33, n. 3, p. 361–370, jun. 2014. ISSN 0167-7055. Available from Internet: <http://dx.doi.org/10.1111/cgf.12392>.

SEARS, A.; JACKO, J. **The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications, Second Edition**. Taylor & Francis, 2007. (Human Factors and Ergonomics). ISBN 9781410615862. Available from Internet: <http://books.google.com.br/books?id=A8TPF_O385AC>.

SHAW, A. **FullCalendar - JavaScript Event Calendar (jQuery plugin)**. 2015. Available from Internet: <http://fullcalendar.io/>.

SHNEIDERMAN, B. **Designing the User Interface: Strategies for Effective Human-Computer Interaction**. 3rd. ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1997. ISBN 0201694972.

SHNEIDERMAN, B.; PLAISANT, C.; HESSE, B. Improving healthcare with interactive visualization. **Computer**, v. 46, n. 5, p. 58–66, May 2013. ISSN 0018-9162.

THUDT, A. et al. Visual mementos: Reflecting memories with personal data. **Visualization and Computer Graphics, IEEE Transactions on**, v. 22, n. 1, p. 369–378, Jan 2016. ISSN 1077-2626.

THUDT, A. et al. Expanding research methods for a realistic understanding of personal visualization. **IEEE Computer Graphics and Applications**, IEEE, v. 37, n. 2, p. 12–18, 2017.

TICTRAC.COM. **Tictrac**. 2015. Available from Internet: <https://www.tictrac.com/>.

WANG, S. et al. Design and effects of personal visualizations. **Computer Graphics and Applications, IEEE**, IEEE, v. 35, n. 4, p. 82–93, 2015.

WITTMANN, M. et al. Social jetlag: Misaligment of biological and social time. **Chronobiology International**, v. 23, p. 497–509, 2006. ISSN 0742-0528. Available from Internet: <http://www.grp.hwz.uni-muenchen.de/pdf/wittmann_pdf/Wittmann06_Social\%20Jetlag.pdf>.

WOOD, J. Visualizing personal progress in participatory sports cycling events. **Computer Graphics and Applications, IEEE**, IEEE, v. 35, n. 4, p. 73–81, 2015.

ZENG, W. et al. Visualizing interchange patterns in massive movement data. **Computer Graphics Forum**, Blackwell Publishing Ltd, v. 32, n. 3pt3, p. 271–280, 2013. ISSN 1467-8659. Available from Internet: <http://dx.doi.org/10.1111/cgf.12114>.