EXPLORING THE BIOGRAPHY AND ARTWORKS OF PICASSO WITH INTERACTIVE CALENDARS AND TIMELINES

A Thesis

by

LUIS MENESES

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

August 2007

Major Subject: Computer Science

EXPLORING THE BIOGRAPHY AND ARTWORKS OF PICASSO WITH INTERACTIVE CALENDARS AND TIMELINES

A Thesis

by

LUIS MENESES

Submitted to the Office of Graduate Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Approved by:

Chair of Committee,	Richard Furuta
Committee Members,	John Leggett
	Enrique Mallen
Head of Department,	Valerie Taylor

August 2007

Major Subject: Computer Science

ABSTRACT

Exploring the Biography and Artworks of Picasso with

Interactive Calendars and Timelines. (August 2007)

Luis Meneses, B.S., Escuela Superior Poletecnica del Litoral

Chair of Advisory Committee: Dr. Richard Furuta

Searching for resources that are related to time periods or events can be frustrating and even problematic since it is often bound to keyword matching or prior knowledge of the exact dates of occurrence. Additionally, the ordered and itemized list that is often returned as a result is unable to provide the required affordances and constraints that users need and desire to conduct scholarly research properly. The following thesis proposes the implementation of timelines and calendar-based interfaces to browse and search through the life events and artworks documented in the Online Picasso Project.

The affordances, constraints and inherent visual nature of the proposed interfaces aid scholars and general users in answering questions regarding the relationship between life events and artworks of the famous Spanish artist. The temporal interfaces are used specifically in the context of the Online Picasso Project and provide several advantages over standard HTML interfaces.

DEDICATION

To Mari: my beautiful and wonderful wife.

ACKNOWLEDGEMENTS

I would like to specially thank my advisor, Dr. Richard Furuta for his extremely valuable advice, and for sharing his expertise and insights. Dr. Enrique Mallen gave me the opportunity to work with the Online Picasso Project, and I would like to thank him for his guidance and enthusiasm towards new ideas. I would also like to thank Dr. John Leggett for his advice and comments.

I would like to thank my friends and family. They have given me support and encouragement throughout the years. Without them, this effort would have not been possible.

I would like to deeply and emphatically thank my wife for her unconditional support. She is my inspiration when encouragement is needed, and she taught me that every obstacle can be overcome with patience and perseverance.

NOMENCLATURE

OPP	Online Picasso Project
XML	Extensible Markup Language
DTD	Document Type Definition
HTML	Hyper Text Markup Language
XSLT	Extensible Stylesheet Language Transformation
SGML	Standard Generalized Markup Language
CSS	Cascading Style Sheets

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Scholars and general users of digital editions face a difficult and problematic scenario when browsing and searching for resources that are related to time periods or events. Scrolling continuously through a long list of itemized search results does not constitute an unusual practice for users when dealing with this type of situation. The problem with this searching mechanism is that a notion of the corresponding dates or keywords associated with the event are required and constitute a precondition to a successful search.

An ordered list is unable to provide the required affordances and constraints that users need and desire to conduct scholarly research properly. It is a common practice among users to utilize the search mechanism present in most web browsers, and then perform another search among the obtained results to "narrow down" or limit the results to a smaller working set that is easier to manage. The use of an external search mechanism in a digital edition is a strong indicator that improved interfaces must be designed, conceived and implemented, just to achieve the sole purpose of facilitating scholarly research.

This thesis follows the style of *Heraclitus: A Framework for Semantic Web Adaptation.*

Aim

The purpose of this research effort is to develop interfaces that take into account the temporal metadata of the documents in a focused collection and use them to provide enhanced browsing mechanisms. The time-based interfaces are implemented in the Online Picasso Project (OPP) [1], a digital collection and repository, which document the life and artistic output of Spanish artist, Pablo Picasso.

Research Approach

To achieve this aim, Chapter II will introduce prior work and current practices. Chapter III describes the implementation and use of Calendar interfaces in a digital collection. Special interest will be given to different visualization options and metadata filtering mechanisms. Next, Chapter IV focuses on the use of Timelines in an artwork document collection. Additional metadata has been included to provide clues related to aspects that could have influenced the creation and outcome of the artist's creation. Chapter V is a discussion of the advantages derived from the use of the proposed temporal- based interfaces. Finally, Chapter VI presents an outline of the main contributions, and future directions of research.

CHAPTER II

BACKGROUND

The OPP is a digital collection and repository maintained by the Department of Hispanic Studies and the Center for the Study of Digital Libraries at Texas A&M University. As of May 2007, it contains 12691 catalogued artworks, 9284 detailed biographical entries, a list of references about Picasso's life and works, and a collection of articles from various sources regarding the renowned Spanish artist.

The OPP includes an interface that allows scholars and users in general to browse through the significant events in his life, artworks, and a list of museums and collections that hold ownership to the various art objects created by the artist during his lifetime. The implemented navigation scheme, shown in Figure 1, works well for experienced scholars who have a deep knowledge of Picasso's life and works. The amount of available information can be overwhelming to the project audience, composed primarily of art scholars and historians, because of its magnitude and painstaking detail. The deployment of calendar and timeline interfaces help categorize the huge amounts of documents and metadata related to Picasso's artworks and biographical events.

Artworks and events related to Picasso's artistic output and life are stored in the OPP have date metadata associated with them. This metadata for artworks includes title, location, medium, dimension, collection, inventory, notes, commentaries, start date and end date. Each biographical record is described by event description, commentary, start

date and end date. Start and end date metadata is especially important, since they constitute the main discriminating factors used by the temporal interfaces to display documents. For this reason, they will receive increased focus during this thesis.

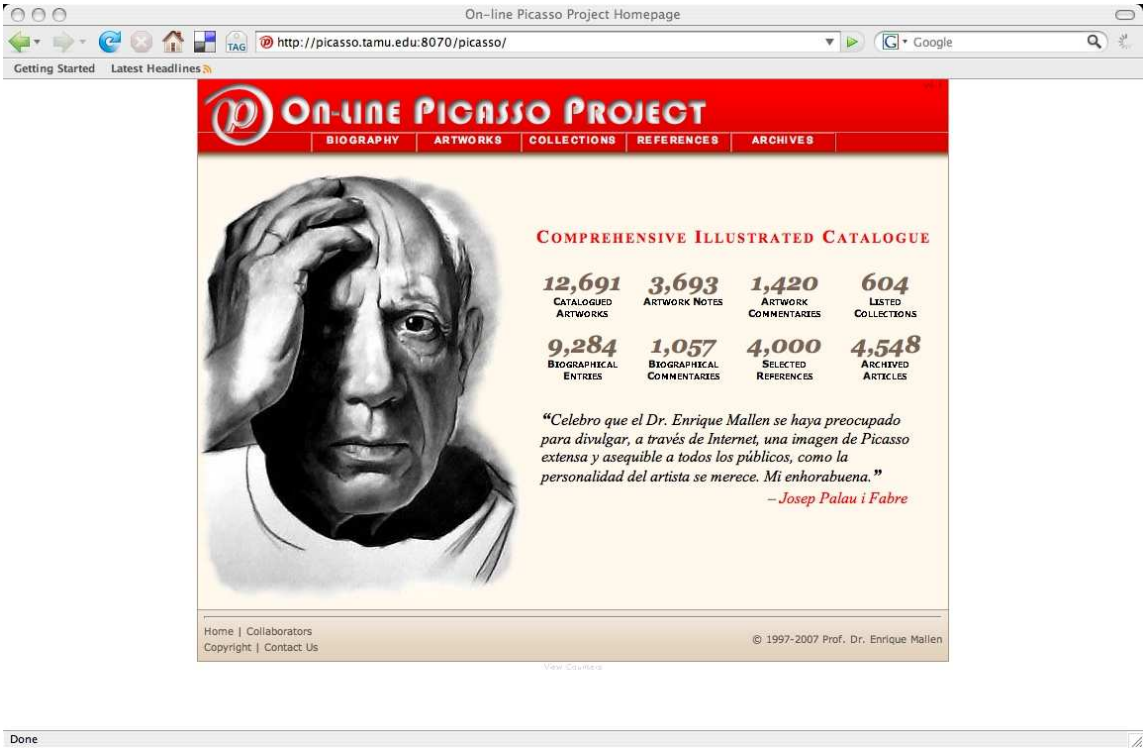

Figure 1: Online Picasso Project homepage

The temporal metadata in the OPP has a degree of uncertainty in some cases, since some of the events in Picasso's life were undocumented at their time of occurrence. The timeline interface has a mechanism for dealing with this uncertainty, and will be described in Chapter IV.

The OPP uses several different software packages to provide its content. The most crucial and important are Apache Tomcat, MySQL, Java servlets, Extensible Stylesheet Language Transformations and Cascading Style Sheets. The system architecture of the OPP is represented in Figure 2.

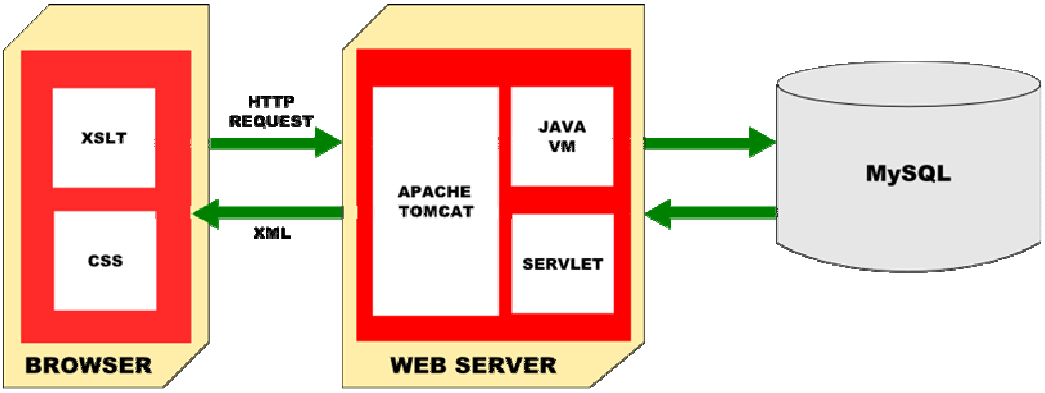

Figure 2: Online Picasso Project architecture

An Apache Tomcat web server [2] provides support for the data interfaces of the OPP. Apache Tomcat is a servlet container used for deploying Java servlet and Java Server pages. It is an open source project developed by the Apache Software Foundation.

Servlets [3] are Java [4] classes that provide a mechanism for extending the functions of a web server. Java Servlets run completely on the server side, and are platform independent. Additionally, Java Servlets can connect to databases to provide dynamic content. This means that a Servlet can output different content to web users depending upon different parameters and values.

Every Java Servlet in the OPP produces its output in Extensive Markup Language (XML) [5]. XML is defined as a subset of the SGML text format, and is commonly used for the exchange of data between web based information systems. The XML format is an open-standard, composed by user-specified tags and attributes. Additionally, the use of XML as a data exchange format is encouraged by the W3 Consortium. The structure of the XML output is not defined by a Document Type Definition (DTD) in the OPP.

The XML formatted data is transformed into Hyper Text Markup Language (HTML) [6] elements, by using Extensible Stylesheet Language Transformations (XSLT) [7]. XSLT is a language used for the transformation of XML documents. XSLT is able to transform XML into HTML, since the later is a subset of the Standard Generalized Markup Language (SGML) [8]. It is important to point out that the web browser on the client side performs this transformation.

The presentation facet of the transformed HTML is generated by Cascading Style Sheets (CSS) [9], which is a mechanism that facilitates adding styles to web documents. With the combination of XSLT and CSS, separation of content and presentation is achieved. This means that if presentation changes are required, modifications should only concern the XSLT and CSS files. This is especially useful in the OPP, since XSLT and CSS files can be used and shared by multiple web documents, giving the project consistency in its affordances and presentation.

The data stored in the OPP repository resides on a MySQL database [10]. MySQL is also an open source project, known for its speed and reliability. Minor alterations to the database structure were performed in order to guarantee the normal operation of the deployed components of the Picasso Project during development and testing and development phases.

Additionally, the OPP has a separate scenario for the testing and development of new features. The development and initial testing of the interfaces was carried out in the development server using a duplicate copy of the database. After stable versions were achieved and important features had been implemented, further testing and final deployment moved into the main production server, which serves the digital edition to its users all over the world.

CHAPTER III

CALENDARS

A calendar is a physical device used to display the events in single days, weeks, months, or years. The primary practical use of a calendar is to identify days and to specify a moment in time. Additionally, calendars are responsible for the introduction of order into modern societies [11] and hold important emotional value to their users [12].

Calendars have become ubiquitous and are present in everyday planning of activities. Reasons for their widespread use include:

- Ease of use and efficiency
- Intuitive and natural affordances
- Large amounts of information can be presented at once.

Hypothesis

Users in general carry out actions and systems interactions intuitively when using interfaces known a priori. How can this be achieved if users are not used to a new interface? The solution is to design an interface consistent with the conceptual model formulated in the mindset of the user.

The conceptual models for calendars and timelines are well established and known widely. Temporal metadata is very valuable information in digital repositories [13]. By using interfaces that take advantage of time-based metadata and bring forward the native

characteristics and usage if interfaces, scholars can take full advantage of the available documents in their research. The new affordances and constraints could potentially allow them to discover answers and possible links between events and artworks, impossible to fathom before due to the nature and limitations of the interfaces and browsing schemes.

The proposed research will be of great benefit to the humanities in general. The Humanities rely profoundly on dates to create a strong relationship between events and documents. Studies have shown that users have better recall when dealing with an image collection if temporal metadata is used to enhance the browsing mechanism [14]. A calendar-resembling interface will provide strong visual clues to answer many research-oriented questions in the humanities.

In the OPP, the time-based interfaces provide clues to the relationship between the artworks and life events of the renowned Spanish artist. It is obvious to assume that key events influenced Picasso in such a way, that they caused significant changes in artistic style and expression. The OPP contains a vast amount of information that could be used in conjunction with the proposed interfaces, in order to help answer this type of inquiry.

The research questions associated with the proposed interfaces address if visual aids and well-known browsing structures can be used as a tool to answer research-oriented inquiries. These interfaces will help find answers that have eluded scholars because of inappropriate browsing mechanisms.

Calendars

The calendar interfaces were developed to provide a timetable for the creation of artworks and occurrence of events catalogued in the OPP. Their purpose is to provide with a quick glance, relevant biographical and artistic dates. Additionally, the calendars provide means for formulating direct comparisons between dates within a single year. This is achieved through stratification of visual elements, and through surrogates. Figure 3 shows the year selection interface for the calendars, while Figure 4 shows the artwork calendar visualization for the year 1936.

Figure 3: Calendar year selection interface

Figure 4: Artwork calendar

Both calendars, artistic and biographical, were developed using Java Servlet technology. Unlike the CaTTS framework [15], they are based specifically on Gregorian time and take advantage of the GregorianCalendar class. The GregorianCalendar Java class [16] is especially useful, since it takes care of all the date calculations required. To name an

example, it is used to calculate the number of days within a given month, the day distribution within a week, the number of days in February, etc.

Given a specific year, the deployed Calendar Java Servlet runs a loop that starts with January 1st, and ends with December 31st. Inside this loop, a date iterator is used in the execution of a query which compares the dates in either the artworks or biographical table. Hits are included in the resulting XML output, which is formatted through XSL transformations. For example, April 30th 1934 has 4 artworks that were started and finished on that day. This result will be clearly present in the resulting XML file.

Besides showing information related to single days, the deployed interfaces also provide information related to events and artworks in months and seasons. This feature was included, since it increases the usability factor of calendars [17].

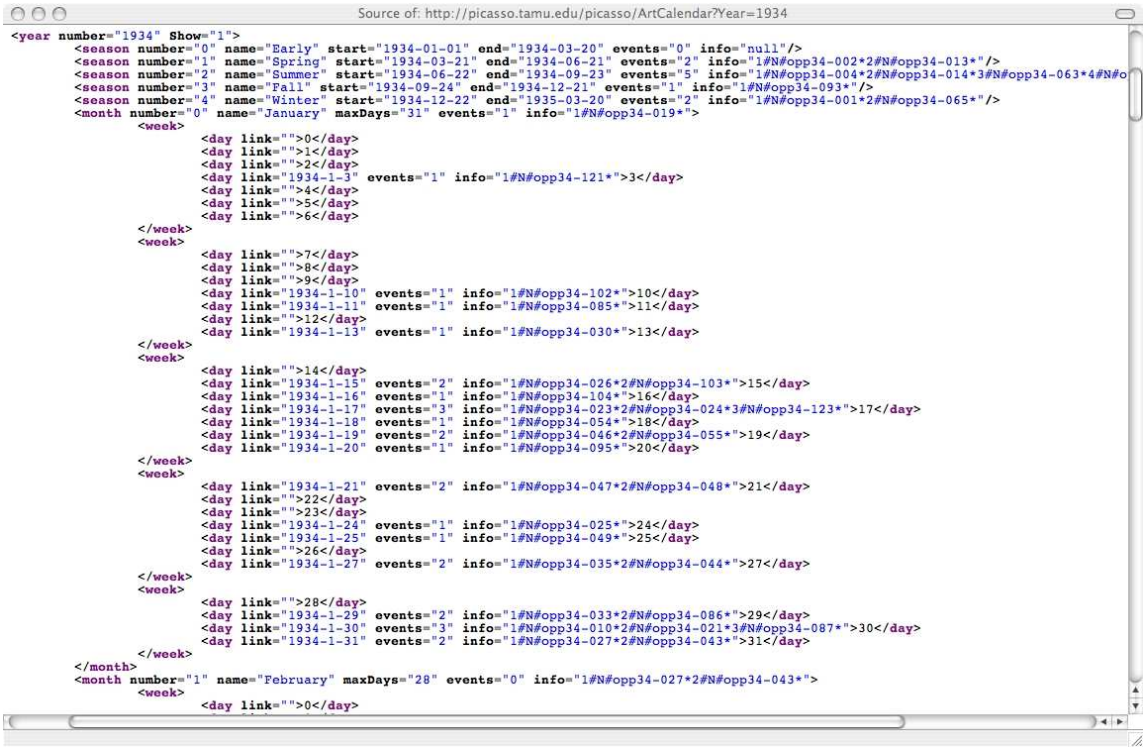

Figure 5: XML calendar output

As shown in Figure 5, the XML output divides each year hierarchically into seasons, months, weeks, and days. Except for the weeks, each of these year subdivisions has an events attribute that indicates the number of events that occurred on that lapse of time. Consequently, the XSLT produces a marker in the calendars that produces detailed information through an HTML hyperlink.

Additionally, through the use of XML and XSL transformations, total separation of content from presentation is achieved. Further modifications related to the presentation could be carried out by only modifying the XSLT, leaving the information extraction mechanism unchanged.

Query Options

The calendar interfaces have 5 display possibilities to filter results, which apply to the artworks and to the narrative:

Figure 6: Artwork surrogates – start and end dates

1.- Show start date and end date: used to display "exact matches". Exact matches refer to artworks or events that were started and terminated in the same day, month or season. In the case of months, this rule applies to the first and last day of the calendar month; while in the case of seasons, it refers to days that mark the start and end of the season. Initially, this was the only display option allowed by the interfaces, and is the default visualization option. Figure 6 provides a screen shot of the interface, along with the artwork surrogates produced by the default visualization option.

2.- Show start date or end date: refers to artworks or events that were started or terminated on any given day during the selected year. In the case of months and seasons, the filtering is applied to the first or last day respectively.

3.- Show start date only: highlights artworks or biographical events that were started on a given day during the selected year. The highlighting schema is applied to the first day of the month or season.

4.- Show end date only: highlights artworks or biographical events that were finished on a given day during the selected year. The highlighting schema is applied to the last day of the month or season.

5.- Show ranges of dates: displays selected artworks or events, which were in progress

for any given day of the year. In the case of months and season, items are displayed if they were in progress during the entire time span.

It is very important to discuss response times with the interface, as the filtering and visualization mechanism become more complicated This is a direct consequence of the complexity of the database queries. For example, the query in option 4 involves a direct comparison of end dates; while option 5 deals complicated comparisons involving ranges of dates.

Interestingly, using ranges of dates yielded the most interesting visualization from a Computer Science point of view since it displays the evolution of Picasso's artistic output through time. Following this analysis, this case should have been used as a default visualization option. However, from a humanities point of view, this visualization could potentially lead into some degree of confusion, since thumbnails of the artworks are displayed through consecutive day to denote their progress. For this reason, it was not included the option given by default even though it is valuable as a research aid.

Surrogates

In terms of output, a condensed event view is presented to the user. If multiple events occurred, or several artworks were produced simultaneously, the interface produces only

a single marker. This marker acts as a surrogate that represents multiple occurrences, from which detailed information can be accessed.

The developed calendar interfaces highlight a date if items that match the visualization criteria are found. It is often the case that multiple items are found under a specific date. How can representations be provided for the matched items?

Through the use of CSS Tooltips, surrogates are provided in the form of artwork thumbnails and textual description of the events. These surrogates are shown when a mouse over event is detected. Clicking on the specified date, month or season produces a web page, where detailed descriptions can be accessed. Three-dimensional representations [18] could not be applied to the browsing scheme of the collection, due to the fact that the interfaces are used on standard web browsers.

Figure 7: Artwork surrogates – ranges of dates

Injecting the surrogates into the developed calendar framework proved to be extremely difficult. The possibility of event and artwork surrogates was not foreseen in the design stages, so the XML format used as an output did not facilitate their inclusion. A more straightforward approach would have involved a new implementation.

To circumvent this problem, information related to the surrogates was included as an attribute in the XML tags, and each entry was limited by a separator token. However, the XSLT standard does not include a tokenizer function, which was extremely difficult to implement. An optimized XML output would have taken greater implementation effort.

Despite the challenges faced in their implementation, the retrieval of information to produce the surrogates does not impose a considerable decrease in performance and response time from the interfaces.

Stratification by Colors

Colors were added to the dates containing items, to show the distribution of the artworks and events. The design decision to include additional stratification schemes relates to the research goal of providing an enhanced browsing mechanism. The inclusion of this feature does not implicate any greater additional processing of the data, but it provides a richer environment for browsing the document collection. Figure 7 shows the color stratification in the artwork calendar interface, when ranges of dates are shown for the year 1936.

The maximum number of events is extracted individually for seasons, months and days. Colors are assigned for dates that contain items in the following percentage ranges:

(0,1/3): purple ; [1/3,2/3): orange; [2/3,3/3]: red

Figure 8: Event surrogates – start and end dates

The choices made for the color stratification are based upon the fact that the original selections must harmoniously blend with the existing framework of the OPP. More contrasting colors, such as yellow, green and red were originally chosen. The decision to

change them also follows that they do not provide an idea of progression. Purple gives visual clues related to a lower intensity when contrasted with orange and red; yellow and green fail to propagate this important characteristic.

This scheme for color assignment produces a view that resembles a statistical histogram graph, and allows pointing out the dates where more artworks were produced or events occurred, and if there is periodic nature in their occurrence [19]. Figure 8 shows color stratification produced the biographical calendar interface for the year 1936.

Even though additional modifications to the existing database were not foreseen, minor adjustments had to be made. When initially deployed, the calendar interfaces took longer than expected to render and present the data. This was a consequence of the query intensive operations that were carried in the background: for each day of the year, a query was executed on either the artworks or biography tables. To speed up the process, indexes were added to the MySQL tables. The speedup obtained in the rendering of the calendars was in the 1/100 magnitude.

CHAPTER IV

TIMELINES

A timeline is a visualization tool that facilitates the plotting of events along a time axis. In contrast to calendars, timelines are primarily used to quickly identify events according to their time occurrence and to analyze their relationship to other events. They also help visualize time that transpired between a series of events.

The OPP has a vast collection of artworks and events related to the life of the Spanish artist Picasso. Due to the dynamic nature of this electronic edition, new artworks and events are constantly being incorporated into the collection. The implementation of interfaces for this type of content is usually carried out to incorporate several features from physical pages, which are found in printed editions. However these interfaces do not provide mechanisms that aid scholars and users in the discovery of hidden relationships, or how subtle changes in his environment affected the outcome of Picasso's artworks.

This new browsing mechanism in the OPP was introduced by placing artworks as markers in a time frame. It was designed to allow users to examine the artworks produced, along with the recorded events of a given year. Certain artworks have a high degree of uncertainty related to their dates of creation. The timeline interface provides a mechanism to deal with the uncertainty in dates. This mechanism and additional

functionality will be described in future sections.

Implementation

The Simile Project [20] provides a set of open source data visualization tools. The included Timeline [21] tool, which was used, is entirely written in JavaScript. In a similar fashion as the Interactive Timeline Viewer [22] and xTG [23], the timeline produces a rectangular visualization that spans through the time span in which the artwork was created. Figure 9 shows the year selection interface for the deployed timelines.

Further modifications were necessary, since the Simile timeline was designed to support single events occurring in one day. Focusing on Picasso's artworks, many artifacts were produced on similar time spans. These artworks were grouped together, and ordered accordingly to their creation dates. Unlike the timelines produced with the ITER framework [24], the user is not allowed to select the attributes that could be displayed. Clicking on the time marker brings forth a series of image surrogates for the artifacts. Each of these surrogates can be accessed to explore in full detail the metadata associated with them.

Figure 9: Timeline year selection interface

Additional Features

Initially, the timeline was designed to focus only on Picasso's artworks. This design choice gave users great freedom to explore large amounts of information in a manipulable visual space. However, the biographical events were being excluded. These events included in the OPP, are particularly important since provide a historical framework, which is crucial to the understanding of the artist's legacy and are tightly bound to his work rhythm [25, 26]. Figure 10 displays a screen shot of the timeline interface and the stratification that differentiates artworks and life events.

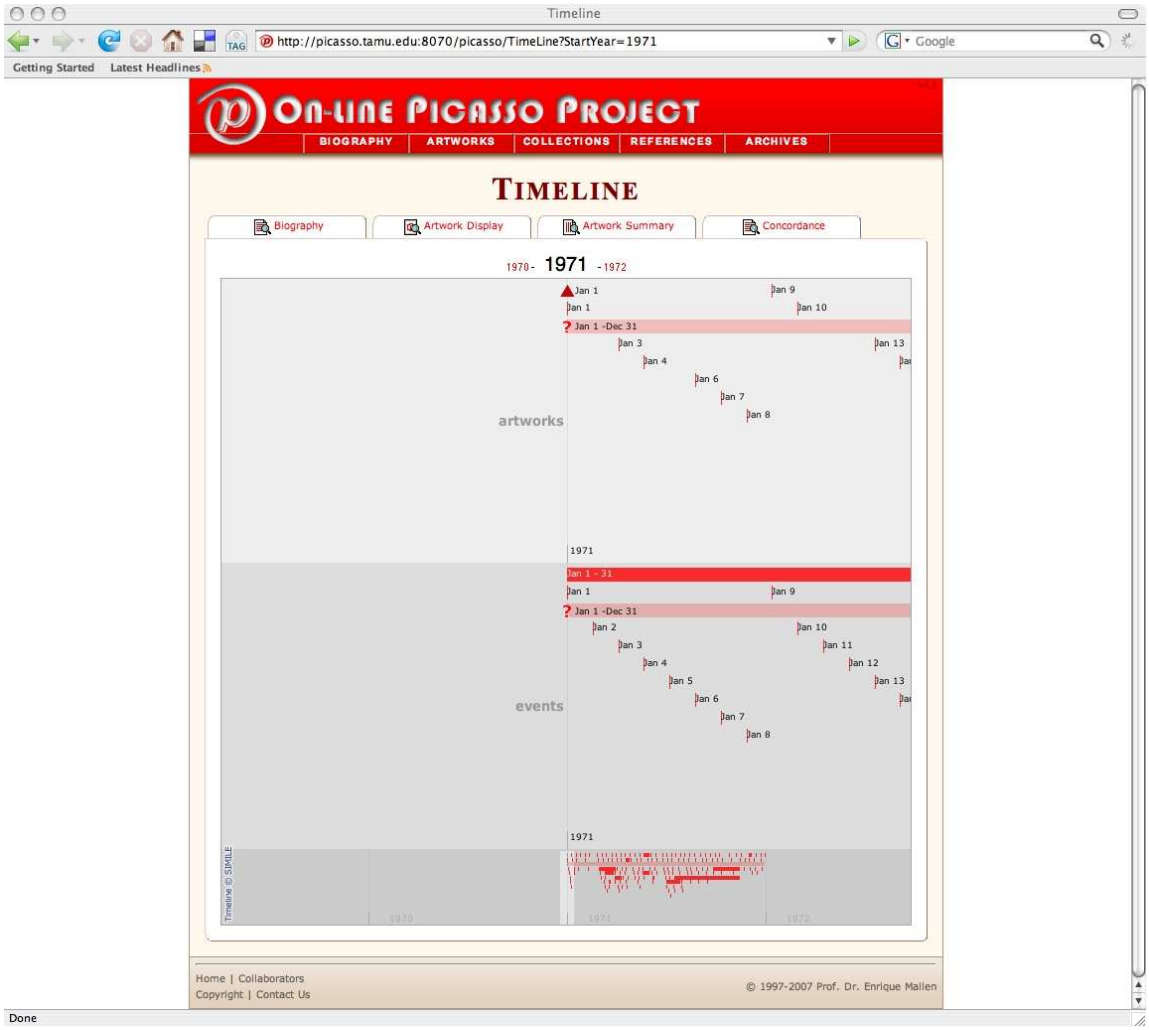

Figure 10: Timeline – year 1971

The idea of including historical events into the timeline incorporated new problems into the interface. The addition of events was a great improvement, but as a tradeoff the distinction between artworks and events became minimal. This led to confusion in the user's conceptual model.

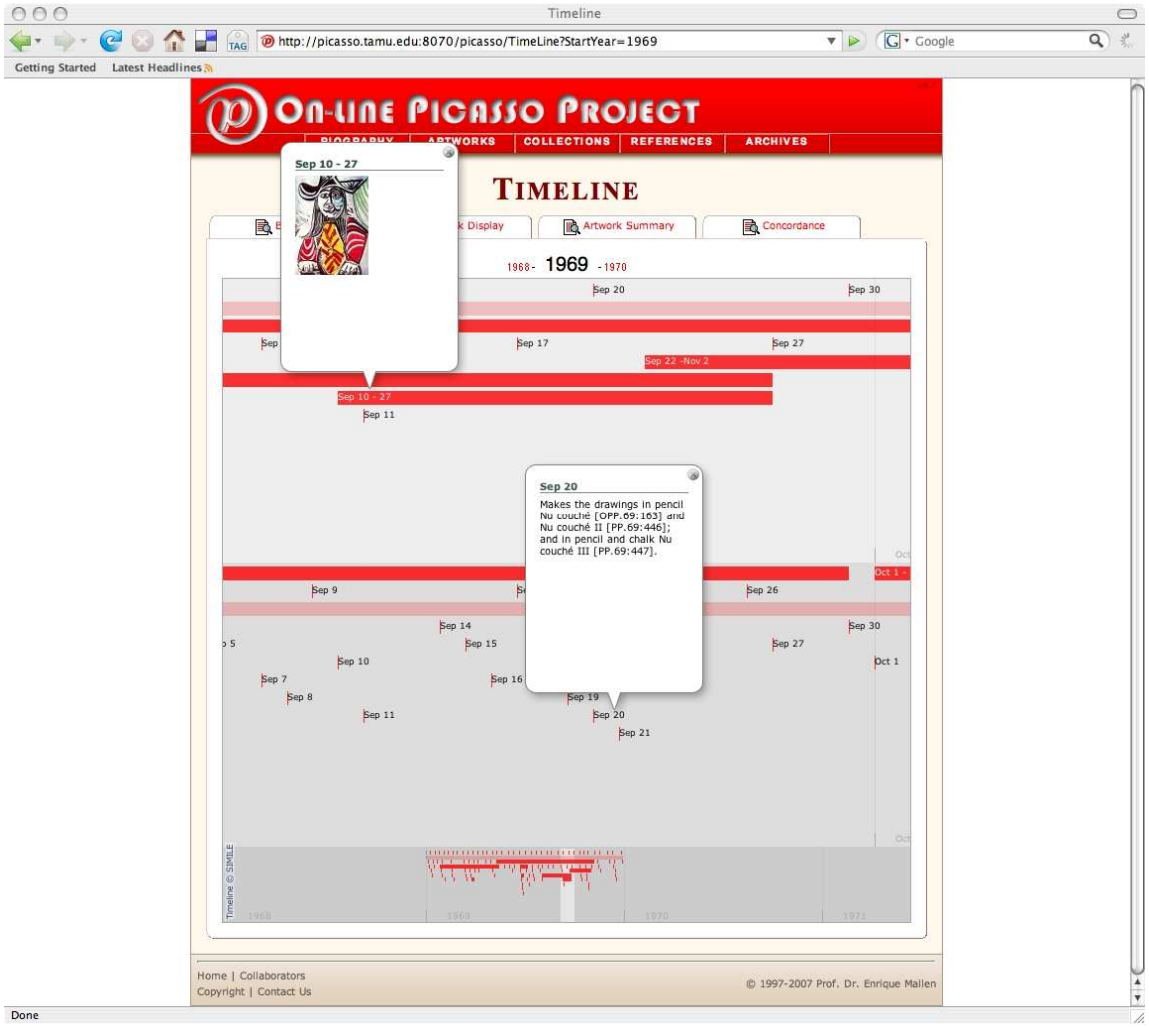

Figure 11: Exploring artwork - event correlation

To solve this important issue, there were two possible alternatives: coming up with a way of representing artworks and events on the same timeline but making them graphically distinguishable, or separating artworks and events into different horizontal bands. The later alternative was chosen due to the fact that it could be incorporated into the context provided by the Simile timeline. Figure 11 shows how the correlation between artifacts and biographical events can be explored.

Additionally, event and artwork labels were added to the time bands. These labels are positioned at the beginning and end of the year span, and are clearly visible and when the timeline is loaded. Additional labeling to the time bands was not required, because the timeline does provide clues to the user about the displayed items in the form of surrogates.

Additionally, some of the artworks produced by Picasso have a certain degree of uncertainty in their dates of creation, since their start and end dates were not documented. The timelines provide a mechanism for dealing with uncertainty, where the artworks are represented with a time bar with a lower level of saturation in their color. This gives a visual clue that the start and end dates are not fixed, and are subject to speculation. A similar color-fading scheme to handle uncertainty in dated artworks was also used in the ItLv [27].

Two servlets are used to generate the xml output for the servlet: TimeEvents and TimeLine. The TimeLine servlet is used primarily to pass the start year and year as parameters. This servlet is associated to an XSL transformation that renders the elements of the page.

The TimeEvents servlet is responsible of all the functionality and data capabilities of the timelines. This servlet is used exclusively to extract artwork and event-related metadata from the database. It uses 3 parameters: start year, end year, and source. The Source parameter indicates the source of the information, which can be from the artworks or from the narrative. Because of this separation, the TimeEvents servlet is instantiated twice, one for each time band: to extract artwork-related information, and then to retrieve biographical-related data. TimeEvents does not require any additional query strings, since it parses the entirety documents within a given year without any keyword summarization [28].

This modular separation achieved with this design, comes along with a trade-off in response time. This is a consequence of the large amount of information that must be pulled from the database. Increased response times are evident, despite the known efficiency of key technologies used: Java Servlets and MySQL databases.

Additional information such as changes in style and location were injected to the timeline, which were extracted from the artist's biography. Their purpose if to provide an additional layer of information that can be used to interpret the events that lead to the creation and mood of certain artifacts, and thus enhancing the browsing environment.

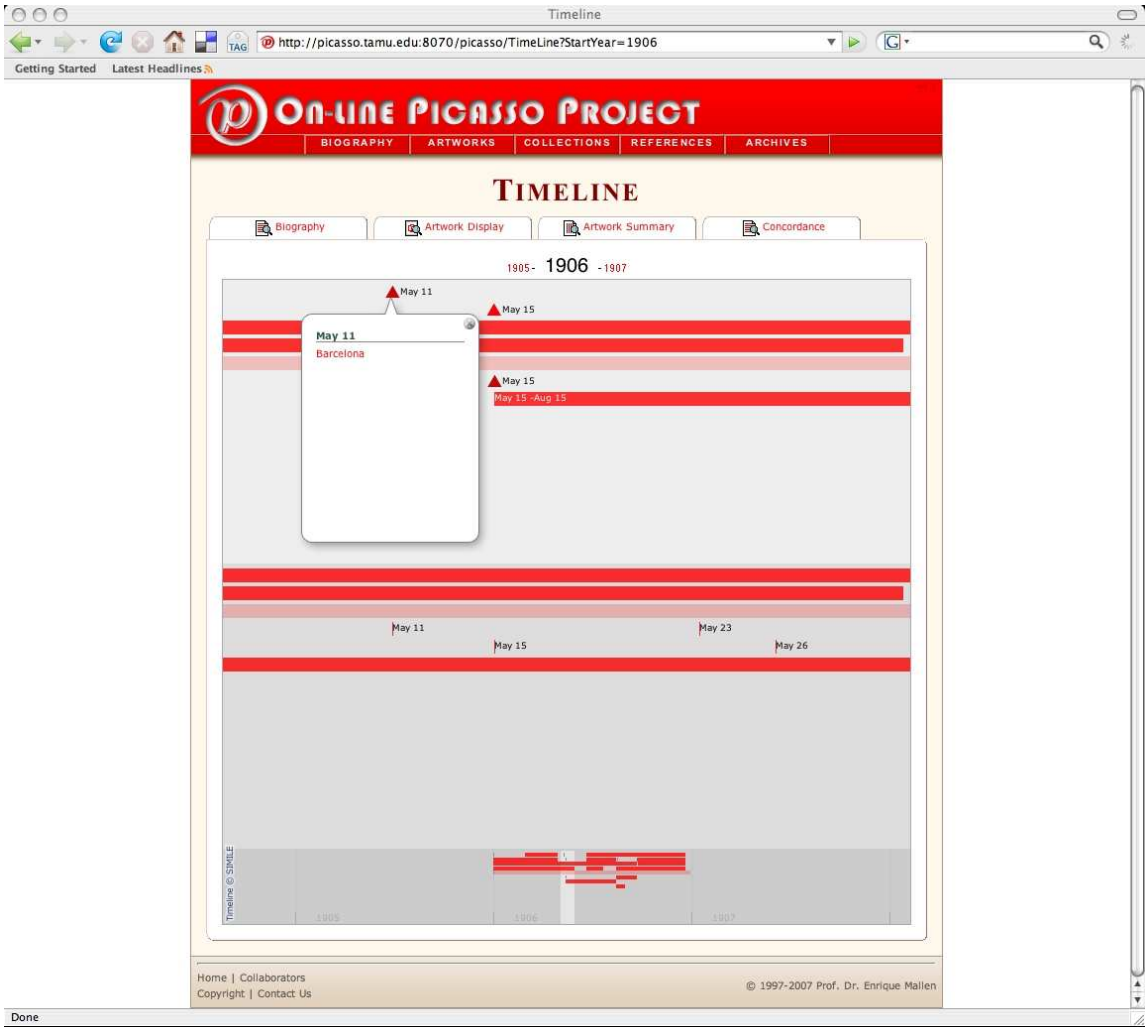

Figure 12: Changes in location

Triangular icons are used to denote changes in style and location. They are differentiated by their color intensity and different hues. Clicking on these icons produces a hyperlink that leads to a detailed description of the event. Figures 12 and 13 highlights changes in style and location in the year 1906 that occurred within a single month.

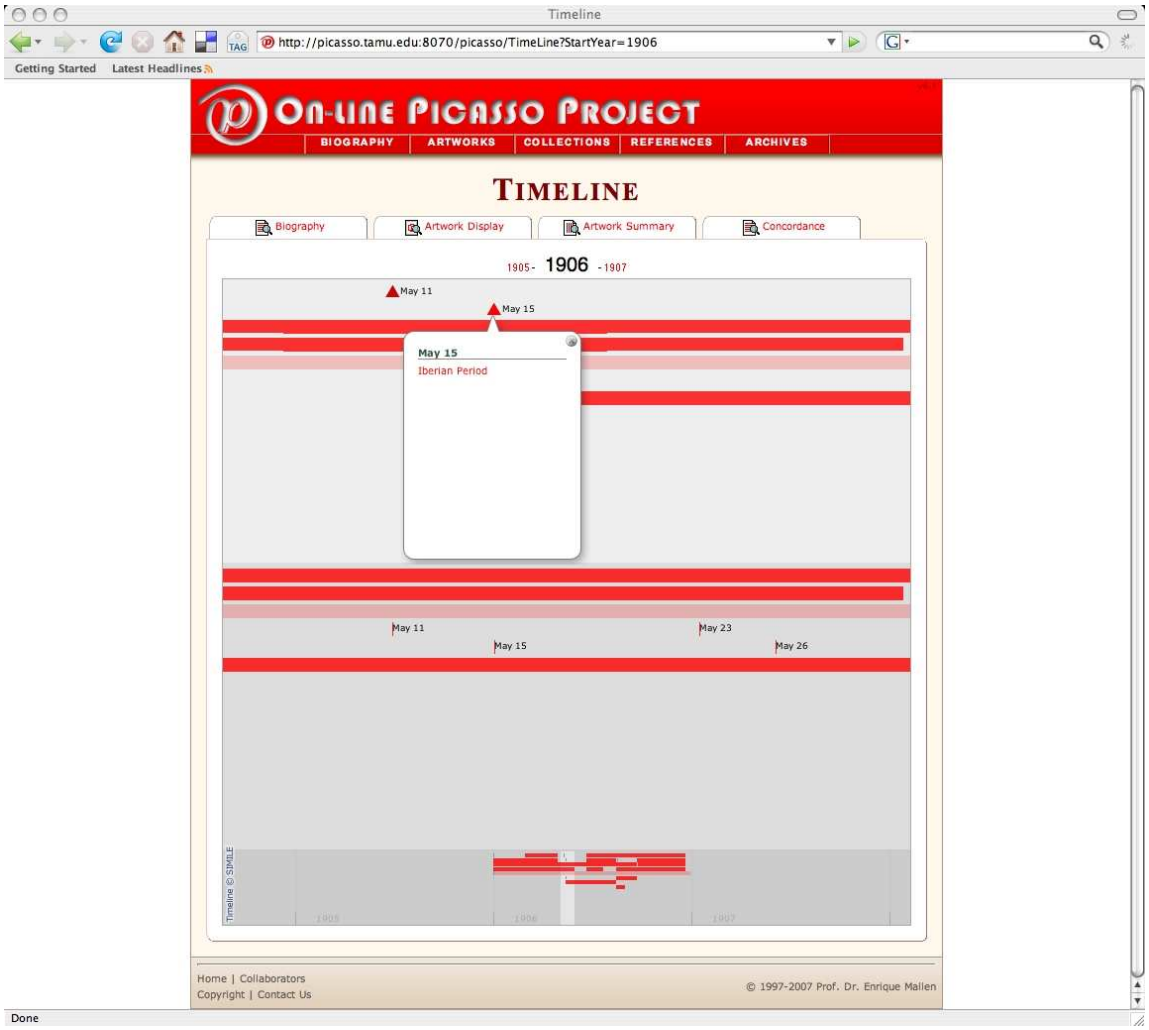

Figure 13: Changes in style

The modular mechanism, by which additional metadata is added to the timeline, facilitates the introduction of new information facets into the interface. Through this scheme, information related to Picasso's writings [29] could potentially be added to the timeline as future work.

CHAPTER V

DISCUSSION

The purpose of this chapter is to summarize the advantages that the use of time-based interfaces, specifically calendars and timelines, contributes to the browsing of focused document collection. This discussion section is aimed at the documents and developed browsing structures in the OPP. Formal conclusions will not be formulated, since they are reserved for the last chapter of this thesis.

Advantages – Calendars

The use of calendar interfaces provides new possibilities for scholars and users in general: the discovery of relationships between documents, which standard HTML interfaces do not facilitate. The main advantages derived from their use include:

1. *The possibility of visualizing an entire year in Picasso's biography and artistic career.*

Picasso's extensive artwork legacy poses a visualization challenge. Browsing his artistic output can be carried out by linearly inspecting his biography. This type of interface is effective in its task, since it facilitates the exploration of his artworks; yet it is not efficient, since it requires the user to navigate through multiple pages in order to view the complete set of artworks.

Through the use of a Calendar-based interface, artworks can be visually identified to their specific dates of creation. This provides a visualization mechanism that allows the user to navigate through a potentially large number of artworks in one screen. The number of artworks that can be accessed, depends on how esthetically prolific the artist was in that specific year.

Additionally, this interface allows the users of the project to quickly determine periods where the artist was more prolific. Dates where Picasso was "active" are clearly visible and identifiable, since they are clickable and differentiated by colors. This data stratification gives users an additional layer of information. Figure 14 shows the artwork creation distribution for March 1935, where Picasso was clearly more productive first 20 days of the month.

Figure 14: Artwork creation distribution – March 1935

For the case of biographical events, a similar scenario is created. Users can navigate through an entire year of events, and the information is presented in a way that affords quick navigation and encourages interaction. Visually, periods where more events occurred in Picasso's life are easily identifiable.

In both cases, jumping from one year to the next can be performed through the navigational hyperlinks situated at the top of the interface.

2. *The possibility of moving to specific day, month or season within a year in one single interaction with the interface.*

Users can compare his artistic output within a year jumping back and forth from one date to the next. Users can also read about specific events in his biography by visually selecting concrete dates.

Through the use of information surrogates, users have the possibility of moving to a specific day, month or season within a year with a single click. The actions produced by scrolling through multiple screens are eliminated, and users can view the artifacts produced on a specific date with ease. Consequently, comparisons between artworks can be achieved fluidly due to the enhancements in the browsing environment. Similarly, users can read about the specific events in Picasso's biography by visually selecting concrete dates.

Advantages – Timelines

The use of timelines proposes a new exploration scheme, which relies on the temporal-related metadata of documents. Its use allows users to explore the correlation between artworks, events, locations, and painting styles.

The advantages gained through the use of timelines include:

1. *The possibility of grasping visually time-extensions in Picasso's output.*

Picasso worked on several artworks at times, which share a similar theme. Even though they share common characteristics, they are not identical. Each of these artworks has variations, which differentiate them.

The OPP already offers mechanisms that allow the comparison of artworks. Artworks can be placed virtually in a "basket", to create a slide show animation similar to the one provided by the ItLV [30]. This requires that the artworks must be selected a priori. Even though artworks can be added and deleted to this context at will, interactions are not encouraged and constrained since the visualization must be halted and initiated again.

On the other hand, the timelines allows users to freely explore all the artworks and events within a given year, and point out their similitude and differences. Pointing at a

specific marked date brings forth a surrogate of the artworks produced on that specific day. All of these interactions do not require any additional intervention from the user.

Visualization tools, specifically timelines, allow users to fully explore and compare the relationships between events [31]. This browsing mechanism, which enhances the comparison and analysis process, affords further examination regarding the evolution of the common and shared theme. Since the artworks are specifically placed in time, it can be determined when the series is exhausted and a new one is taken upon.

2. *The possibility of visually comparing works ordered in chronological order.*

Pablo Picasso was a very prolific artist, who produced a large number of artworks. These artifacts are not limited to paintings, and include ceramics, sculptures, and even handwritten poems.

Due to the large number of items present in the artwork collection, navigating through them can be daunting. The artwork collection can be ordered chronologically through the Biography, but the artworks are represented exclusively by their names. The Artwork collection does provide a browsing mechanism that allows artworks to be ordered chronologically, but the interface does not give feedback about their relation to a specific time stamp.

The timelines provide a mechanism that filters artworks according to their year of creation. The enhanced navigational scheme provided, allows scholars to view artifacts in chronological order. The addition of surrogates allows users to point out a specific item, and then compare them in relation to others through and their time correlation.

3. *The possibility of seeing correlations between change of location and artwork creation.*

The frequent travels of the Spanish artist are documented in his biography. Traveling, and the discovery of new ideas and ways of thought, can change the perception of an artist towards the world. It is not unexpected that changes in Picasso's painting style are tightly bound to his travel itinerary. Interestingly, this relationship has only been formally explored on specific cubism periods [32, 33].

Changes in location are visible when navigating the artist's biography. New locations were highlighted in the text, and when activated provided information related to it through a hyperlink. This type of navigation, though effective in its purpose, does not allow further exploration of relationships and correlations.

The deployed timelines allow the exploration of correlations between location changes and the creation of specific artworks. Changes in location are marked in the timelines, and clearly denote a point in time where exposure to a new or recurring context

occurred. These location marks, can be further explored, and give a detailed description of the event that led to a location change.

4. *The possibility of comparing different stylistic periods as they relate to concrete artworks and specific locations.*

Picasso's frequent changes of style in his artworks are documented in the biographical section of the OPP. In a similar fashion to the changes in location, they are clearly identified, and give a detailed description of that stylistic period when activated. This type of interface does not provide the tools that afford the comparison of stylistic periods and how they relate to specific artworks and distinct locations.

On the other hand, timelines offer a visual mechanism to analyze these relationships. They produce a visualization that puts changes in thematic periods and in locations in a common context, along with the artworks that were elaborated in that time span. This tool is augmented with the navigational ease of clicking through a series of artworks, to compare their characteristics and perform a deeper analysis if necessary.

This gives scholars a tool that allows them to explore and discover new relationships between thematic periods and artworks. Additionally, some artworks have uncertain production dates. Through the use of this visual tool, this uncertainty can be eliminated in some cases by estimating their dates when compared alongside to other artworks. This

type of analysis could not be carried out by previously implemented browsing mechanisms in the collection.

CHAPTER VI

CONCLUSIONS

The aim of this research effort is to develop interfaces that take into account the temporal metadata of the documents in a focused collection and use them to provide enhanced browsing mechanisms.

This thesis presents how time based interfaces can be used to enhance the interactions between users and a focused digital collection. The temporal interfaces are used specifically in the context of the OPP, and provide several advantages over standard HTML interfaces.

For the calendars, these advantages include visualizing an entire year in Picasso's biography and artistic career, and the possibility of moving to specific day, month or season within a year in one single interaction with the interface.

In the case of the timelines, the advantages derived from its use with this document collection are grasping time-extensions visually in Picasso's output, comparing works ordered in chronologically, the possibility of seeing correlations between change of location and artwork creation, and the possibility of comparing different stylistic periods as they relate to concrete artworks and specific locations.

Main Contribution

The main contribution of this effort is to include browsing mechanisms that will aid scholars and users to discover new relationships between the artworks and events produced by Pablo Picasso This is accomplished through the use of interfaces that rely on temporal metadata, such as calendars and timelines.

The amount of information in the OPP is expanding and will include different facets of the artist that will give a more rich documentation of his life. Currently, work is being conducted into including his writings and poems into the document collection. Interfaces that use unexplored metadata in a certain context are required since they provide focused access to documents in a large corpus.

Before the addition of calendars and timelines, date-related metadata was practically being ignored as a filtering mechanism, and was being only primarily used as an ordering criterion. This approach relies heavily on temporal metadata, and uses it to provide mechanisms and interfaces that allow rich exploration of the artist's legacy and get a more whole and concise understanding of his life.

Future Work

The interfaces have been deployed taking into account that additional functionality could be introduced with ease. As a consequence, information regarding Picasso's writings and poems will be included into the timelines. This will allow a deeper understanding of his

legacy, since it could potentially provide a greater understanding of his artworks and biography. The writings and poems constitute a compendium of his thoughts and insights, extremely valuable because they were written by the artist himself.

Additionally, surrogates of biographical events could be included into the existing calendar of artworks. This will provide the users of the OPP, with a sense of events occurring in the background and forefront of Picasso's life. The events that occurred around the artist's environment obviously affected the outcome of his artworks. The events that will be included into the interface must give a clear idea of the atmosphere that surrounded Picasso at certain times. The level of granularity of the events that will be included into the artwork calendars is still uncertain and will be addressed in the future.

Finally, time bars could be incorporated into the calendars, to give users a visual surrogate of the position of the artwork given a timeframe of a specific year. This feature would be very beneficial for visualization 1 through 4, which do not focus on date ranges. By explicitly presenting a time frame, additional inferences could be formulated and research questions could be answered.

REFERENCES

1. The Online Picasso Project. Available at http://picasso.tamu.edu. [Accessed May 17, 2007]

2. Apache Tomcat. Available at http://tomcat.apache.org. [Accessed May 17, 2007]

3. Java Servlet Technology. Available at http://java.sun.com/products/servlet/. [Accessed May 17,2007]

4. Java Technology. Available at http://java.sun.com. [Accessed May 17, 2007]

5. Extensive Markup Language (XML). Available at http://www.w3.org/XML/. [Accessed May 17, 2007]

6. HTML 4.01 Specification. Available at http://www.w3.org/TR/html401/. [Accessed May 17, 2007]

7. XSL Transformations (XSLT). Available at http://www.w3.org/TR/xslt/. [Accessed on May 17, 2007]

8. Overview of SGML Resources. Available at http://www.w3.org/MarkUp/SGML. [Accessed on May 17, 2007]

9. Cascading Style Sheets. Available at http://www.w3.org/Style/CSS/. [Accessed on May 17, 2007]

10. MySQL AB :: The world's most popular open source database. Available at http://www.mysql.com. [Accessed on May 17, 2007]

11. E. Zerubavel, *Hidden Rhythms: Schedules and Calendars in Social Life,* University of Chicago Press, 1981.

12. M. Tomitsch, T. Grechenig, and P. Wascher, "Personal and private calendar interfaces support private patterns: diaries, relations, emotional expressions," *in Proceedings of the 4th Nordic Conference on Human-Computer interaction: Changing Roles*, 2006, ACM Press, pp. 401-404.

13. G. Crane, "Designing Documents to Enhance the Performance of Digital Libraries," *D-Lib Magazine,* 2000, www.dlib.org/dlib/july00/crane/07crane.html.

14. A. Graham, H. Garcia-Molina, A. Paepcke, and T. Winograd, "Time as essence for photo browsing through personal digital libraries," *in Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, 2002, ACM Press, pp. 326 - 335.

15. F. Bry, F. Ries, and S. Spranger, “CaTTS: calendar types and constraints for Web applications,” *in Proceedings of the 14th international Conference on World Wide Web*, 2005, ACM Press, pp. 702-711.

16. GregorianCalendar (Java 2 Platform SE 5.0). Available at http://java.sun.com/j2se/1.5.0/docs/api/java/util/GregorianCalendar.html. [Accessed on May 19, 2007]

17. C. M. Kincaid, P.B. Dupont, and A. R. Kaye, “Electronic calendars in the office: an assessment of user needs and current technology,” *ACM Transactions on Information Systems*, 1985, ACM Press, pp. 89-102.

18. J.D. Mackinlay, G. G. Robertson, and R. DeLine, “Developing calendar visualizers for the information visualizer,” *in Proceedings of the 7th Annual ACM Symposium on User interface Software and Technology*,1994, ACM Press, pp. 109-118.

19. J. V. Carlis, and J.A. Konstan, “Interactive visualization of serial periodic data,” *in Proceedings of the 11th Annual ACM Symposium on User interface Software and Technology*, 1998, ACM Press, pp. 29-38.

20. SIMILE Project. Available at http://simile.mit.edu/. [Accessed on May 24, 2007]

21. B. Martin, "Ajax timelines and the semantic web," *Linux Journal*, 2007, www.linuxjournal.com/article/9301.

22. C. Monroy, R. Furuta, and E. Mallen, "Visualizing and Exploring Picasso's World," *in Proceedings of the 3rd ACM/IEEE-CS joint conference on Digital Libraries*, 2003, IEEE Computer Society, pp. 173- 175.

23. G.M. Karam, "Visualization using timelines," *in Proceedings of the 1994 ACM SIGSOFT international Symposium on Software Testing and Analysis*, 1994, ACM Press, pp. 125-137.

24. V. Kumar, R. Furuta, and R. Allen, "Metadata visualization for digital libraries: interactive timeline editing and review," *in Proceedings of the third ACM conference on Digital Libraries*, 1998, ACM Press, pp. 126 - 133.

25. J. Begole, J. C. Tang, R. B. Smith, and N. Yankelovich, "Work rhythms: analyzing visualizations of awareness histories of distributed groups," *in Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work*, 2002, ACM Press, pp. 334-343.

26. A. Krishnan, and S. Jones, “TimeSpace: activity-based temporal visualisation of personal information spaces,” *Personal Ubiquitous Computing,* vol. 9, no. 1, 2005, Springer London, pp. 46-65.

27. C. Monroy, R. Furuta, E. Urbina, and E. Mallen, “Texts, Images, Knowledge: Visualizing Cervantes and Picasso,” *in Proceedings of the Visual Knowledges Conference,* 2003, John Frow, www.iash.ed.ac.uk/vkpublication/monroy.pdf.

28. H. L. Chieu, and Y. K. Lee, “Query based event extraction along a timeline,” *in Proceedings of the 27th Annual International ACM SIGIR Conference on Research and Development in information Retrieval*, 2004, ACM Press, pp. 425-432.

29. N. Audenaert, U. Karadkar, E. Mallen, R. Furuta, and S. Tonner, “Viewing Texts: An Art-centered Representation of Picasso's Writings,” To be published *in Proceedings of Digital Humanities 2007*, 2007.

30. C. Monroy, R. Furuta, and E. Mallen, “Creating, Visualizing, Analyzing, and Comparing Series of Artworks,” 2003, Unpublished paper.

31. R. B. Allen, “A focus-context browser for multiple timelines,” *in Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries*, 2005, ACM Press, pp. 260-261.

32. J. Palau i Fabre, *Picasso: Cubism 1907-1917*. Ediciones Polígrafa, 1990.

33. J. Richardson, *A Life of Picasso: 1907-1917*. Random House, 1996.

VITA

Name: Luis Meneses

Address: Department of Computer Science
Texas A&M University
TAMU 3112
College Station, TX 77843-3112

Email Adress: ldmm@cs.tamu.edu

Education: B.S., Computer Statistics, Escuela Superior Politecnica del Litoral, 2000
M.S., Computer Science, Texas A&M University, 2007