

---

# Intelligent tool for visual data analysis

---



Bachelor's Thesis

Jorge Oses Grijalba

Double Major in Mathematics and Computer Science

Computer Science Faculty

Complutense University of Madrid

March 2019

Documento maquetado con T<sub>E</sub>X<sub>S</sub> v.1.0.

Este documento está preparado para ser imprimido a doble cara.

# Intelligent tool for visual data analysis

*Memoria que presenta para optar al título de Doctor en Informática*

**Jorge Oses Grijalba**

*Dirigida por el Doctor*

**Tutor no definido. Usa \tutorPortada**

**Double Major in Mathematics and Computer Science  
Computer Science Faculty  
Complutense University of Madrid**

**March 2019**

Copyright © Marco Antonio y Pedro Pablo Gómez Martín

*Al duque de Béjar*  
*y*  
*a tí, lector carísimo*



*I can't go to a restaurant and  
order food because I keep looking  
at the fonts on the menu.  
Donald Knuth*





# Agradecimientos

*A todos los que la presente vieron y  
entendieron.*

Inicio de las Leyes Orgánicas. Juan  
Carlos I

Groucho Marx decía que encontraba a la televisión muy educativa porque cada vez que alguien la encendía, él se iba a otra habitación a leer un libro. Utilizando un esquema similar, nosotros queremos agradecer al Word de Microsoft el habernos forzado a utilizar  $\text{\LaTeX}$ . Cualquiera que haya intentado escribir un documento de más de 150 páginas con esta aplicación entenderá a qué nos referimos. Y lo decimos porque nuestra andadura con  $\text{\LaTeX}$  comenzó, precisamente, después de escribir un documento de algo más de 200 páginas. Una vez terminado decidimos que nunca más pasaríamos por ahí. Y entonces caímos en  $\text{\LaTeX}$ .

Es muy posible que hubiéramos llegado al mismo sitio de todas formas, ya que en el mundo académico a la hora de escribir artículos y contribuciones a congresos lo más extendido es  $\text{\LaTeX}$ . Sin embargo, también es cierto que cuando intentas escribir un documento grande en  $\text{\LaTeX}$  por tu cuenta y riesgo sin un enlace del tipo “*Author instructions*”, se hace cuesta arriba, pues uno no sabe por donde empezar.

Y ahí es donde debemos agradecer tanto a Pablo Gervás como a Miguel Palomino su ayuda. El primero nos ofreció el código fuente de una programación docente que había hecho unos años atrás y que nos sirvió de inspiración (por ejemplo, el fichero `guionado.tex` de  $\text{\TeX}$ IS tiene una estructura casi exacta a la suya e incluso puede que el nombre sea el mismo). El segundo nos dejó husmear en el código fuente de su propia tesis donde, además de otras cosas más interesantes pero menos curiosas, descubrimos que aún hay gente que escribe los acentos españoles con el `\'{\i}`.

No podemos tampoco olvidar a los numerosos autores de los libros y tutoriales de  $\text{\LaTeX}$  que no sólo permiten descargar esos manuales sin coste adicional, sino que también dejan disponible el código fuente. Estamos pensando en Tobias Oetiker, Hubert Partl, Irene Hyna y Elisabeth Schlegl, autores del famoso “The Not So Short Introduction to  $\text{\LaTeX} 2_{\epsilon}$ ” y en Tomás

Bautista, autor de la traducción al español. De ellos es, entre otras muchas cosas, el entorno **example** utilizado en algunos momentos en este manual.

También estamos en deuda con Joaquín Ataz López, autor del libro “Creación de ficheros L<sup>A</sup>T<sub>E</sub>X con GNU Emacs”. Gracias a él dejamos de lado a WinEdt y a Kile, los editores que por entonces utilizábamos en entornos Windows y Linux respectivamente, y nos pasamos a emacs. El tiempo de escritura que nos ahorramos por no mover las manos del teclado para desplazar el cursor o por no tener que escribir `\emph` una y otra vez se lo debemos a él; nuestro ocio y vida social se lo agradecen.

Por último, gracias a toda esa gente creadora de manuales, tutoriales, documentación de paquetes o respuestas en foros que hemos utilizado y seguiremos utilizando en nuestro quehacer como usuarios de L<sup>A</sup>T<sub>E</sub>X. Sabéis un montón.

Y para terminar, a Donal Knuth, Leslie Lamport y todos los que hacen y han hecho posible que hoy puedas estar leyendo estas líneas.

# Abstract

*La ciencia no se hace sola, ahí que hacerla*

Científico

This document reflects my Bachelor's Thesis corresponding to the Double Degree in Mathematics and Computer Science, developed within the area of intelligent data analytics and 'Case Based Reasoning'. During the progress of the project, the principles applicable in any environment of data processing and the science behind it are explained generally and aimed to be usable in any kind of context by any user provided the right format of data. Nowadays, highly heterogeneous data collection and processing methods are employed in all industries, however the techniques employed to get useful information out of the data usually have a generalistic aim, and the work relevant to the field itself is often done manually. In this work we aim to provide an automated way to analyze information while taking into account information and techniques relevant to the field of the analysis. The objective of this Degree's Final Project is the development of a prototype capable of carrying this analysis while being able to learn based on user input. As a Proof of Concept, we have included several medical domains with each one having developed specific methods and techniques for them. To serve as a base for this analysis, we have also developed a system for storing, loading and analyzing the information of the domain and the information provided by the user. This system will be the backbone of our architecture and enable the Case Based Reasoning analysis to function correctly in very different situations, providing the metrics and functions needed for every case.



# Contents

<b>Agradecimientos</b>	<b>ix</b>
<b>Abstract</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Program Structure: an overview . . . . .	1
Notas bibliográficas . . . . .	1
En el próximo capítulo . . . . .	1
<b>2 The program structure</b>	<b>3</b>
2.1 Introduction . . . . .	3
Notas bibliográficas . . . . .	3
En el próximo capítulo . . . . .	3
<b>3 The program structure</b>	<b>5</b>
3.1 Backend . . . . .	5
Notas bibliográficas . . . . .	5
En el próximo capítulo . . . . .	5
<b>4 The program structure</b>	<b>7</b>
4.1 Frontend . . . . .	7
Notas bibliográficas . . . . .	7
En el próximo capítulo . . . . .	7
<b>5 CBR Application</b>	<b>9</b>
5.1 CBR Applications . . . . .	9
Notas bibliográficas . . . . .	9
En el próximo capítulo . . . . .	9
<b>6 CBR Application</b>	<b>11</b>
6.1 Proof Of Concept : Medical Data . . . . .	11
Notas bibliográficas . . . . .	11

En el próximo capítulo . . . . .	11
<b>A Así se hizo...</b>	<b>13</b>
A.1 Introducción . . . . .	13
<b>Index</b>	<b>15</b>

# List of Figures





# List of Tables



# Chapter 1

## Introduction

...

**ABSTRACT:** ...

### 1.1 Program Structure: an overview

This is the introductory texttt

### Notas bibliográficas

These are the bibliographical notes (?)

### En el próximo capítulo...

This is the next chapter section



## Chapter 2

# The program structure

...

...

**ABSTRACT:** ...

### 2.1 Introduction

This is the introductory texttt

### Notas bibliográficas

These are the bibliographical notes (?)

### En el próximo capítulo...

This is the next chapter section



## Chapter 3

# The program structure

...  
...

**ABSTRACT:** ...

### 3.1 Backend

This is the introductory texttt

### Notas bibliográficas

These are the bibliographical notes (?)

### En el próximo capítulo...

This is the next chapter section





## Chapter 4

# The program structure

...

**ABSTRACT:** ...

### 4.1 Frontend

This is the introductory texttt

### Notas bibliográficas

These are the bibliographical notes (?)

### En el próximo capítulo...

This is the next chapter section



## Chapter 5

# CBR Application

...

**ABSTRACT:** ...

### 5.1 CBR Applications

This is the introductory texttt

#### Notas bibliográficas

These are the bibliographical notes (?)

#### En el próximo capítulo...

This is the next chapter section



## Chapter 6

# CBR Application

...

...

**ABSTRACT:** ...

### 6.1 Proof Of Concept : Medical Data

This is the introductory texttt

#### Notas bibliográficas

These are the bibliographical notes (?)

#### En el próximo capítulo...

This is the next chapter section



# Appendix A

## Así se hizo...

...

...

**ABSTRACT:** ...

### A.1 Introducción

...





*–¿Qué te parece desto, Sancho? – Dijo Don Quijote –  
Bien podrán los encantadores quitarme la ventura,  
pero el esfuerzo y el ánimo, será imposible.*

*Segunda parte del Ingenioso Caballero  
Don Quijote de la Mancha  
Miguel de Cervantes*

*–Buena está – dijo Sancho –; fírmela vuestra merced.  
–No es menester firmarla – dijo Don Quijote–,  
sino solamente poner mi rúbrica.*

*Primera parte del Ingenioso Caballero  
Don Quijote de la Mancha  
Miguel de Cervantes*

