Data driven drawings

González Jiménez, Juan Camilo

2019

González Jiménez, J. C. (2019). Data driven drawings. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/106573>

<https://doi.org/10.32657/10220/47992>

DATA DRIVEN DRAWINGS

by

JUAN CAMILO GONZÁLEZ JIMÉNEZ

Vibeke Sorensen

Professor, Chair

School of Art Design & Media, Nanyang Technological University

Thesis Supervisor

Marsha Kinder

University Professor Emerita

Professor Emerita of Critical Studies

University of Southern California

Thesis Co-Supervisor

School of Art, Design & Media

A thesis submitted to the Nanyang Technological University in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Singapore, 2019

Statement of Originality

I certify that all work submitted for this thesis is my original work. I declare that no other person's work has been used without due acknowledgement. Except where it is clearly stated that I have used some of this material elsewhere, this work has not been presented by me for assessment in any other institution or University. I certify that the data collected for this project are authentic and the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

March 10, 2019

.....
Date

.....
Juan Camilo González Jiménez

Supervisor Declaration Statement

I have reviewed the content of this thesis and to the best of my knowledge, it does not contain plagiarised materials. The presentation style is also consistent with what is expected of the degree awarded. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accordance with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

March 10, 2019

.....

Date

.....

Vibeke Sorensen

Authorship Attribution Statement

This thesis **does not** contain any materials from papers published in peer-reviewed journals or from papers accepted at conferences in which I am listed as an author.

March 10, 2019

.....
Date

.....
Juan Camilo González Jiménez

Acknowledgments

First to my mother, whose love and support are immeasurable and has led the way for my brother, sister and I to follow by example of resilience under the most difficult circumstances.

Special thanks to my supervisors Vibeke Sorensen and Marsha Kinder for providing clarity, guidance, and support. Equally important to this research has been the accurate proposed revisions by the thesis committee: Michael Walsh, Cristina Venegas, and Andreas Kratky.

To my friends and colleagues that embarked with me on the adventure of a newly created program: Javad Khajavi and Cinzia Bottini.

My friends that made life in Singapore a blast: Kathrin Albers, Gallina Mihaleva, Ross Williams and Lea Vidakovic.

To my teacher, mentor, and friend Giannalberto Bendazzi.

And all the amazing staff at ADM: Hong Bee Kuen, Aw Lee Ting, Kumar Chockanathan, Cheng Lim and Sam Yeong.

Table of Contents

Acknowledgments	ix
Table of Figures	xiii
Abstract.....	xv
Introduction.....	17
Research Objectives	18
Literature Review.....	22
Methodology.....	28
The Need for Practice-Led Research.....	30
Conceptual Framework.....	32
Data Drawings	33
Data Driven Drawings	44
Data and Animation	52
Personal Motivation.....	54
Precedent: From the stream of consciousness to stream of data	58
Chapter 1: Data Representation.....	63
Artistic data representation	76
Evocative data representation.....	82
Example 1: “Rite of Passage”	88
The Sublime	96
Chapter 2: Data visualization as graphic notations	110
Graphic Notations	113
The proposed methodology for studying graphic notations.....	122
Custom software for studying graphic notations	128

Results of the case studies	130
Case Study A: Fleischer Studio - "Mr. Bug Goes to Town" (1941).....	130
Case Study B: Zbigniew Rybczyński - "Tango" (1980).....	133
Case Study C: Norman McLaren - "A Chairy Tale" (1957).....	136
Case Study D: Juan Camilo González - "SiSiSiSiSiSiSiSiSiSi" (2011).....	139
From graphic notations to data visualization	143
Chapter 3: The Artist Dilemma	145
Functional Data Visualization, the case against subjectivity	146
Jaques Bertin - Formalization of a visual grammar	147
John W. Tukey - Exploratory vs confirmatory data analysis	152
Edward Tufte - Graphical Excellence and Moral Responsibility	154
Evocative representation of reality in Animated Documentary.....	159
Mimetic substitution	161
Non-mimetic Substitution.....	164
Example 2: "Eterna".....	166
Evocation.....	171
Chapter 4: Production of Data Driven Drawings	174
Project A: "Vista al Mar"	184
Project B: "Dibujo Defunciones"	194
Project C: "N.N."	200
Conclusions	207
Future Research	208
Bibliography	211

Table of Figures

<i>Figure 1.</i> Three points of view of Data Driven Drawing projects, from “Vista al Mar” & “Dibujo Defunciones” & “N. N.”, by González J., J. C., 2018.	31
<i>Figure 2.</i> Cholera epidemic map emphasizing the high density of cases near the Broadwick Street water pump, from “On the Mode of Communication of Cholera”, by Snow, J., 1854.	37
<i>Figure 3.</i> The complete diagram of the Brookes ship, from “Stowage of the British Slave Ship Brookes Under the Regulated Slave Trade Act of 1788”, by Alfred, W., 1788.	39
<i>Figure 4.</i> Detail of the note describing the Regulation Act of 1788, from “Stowage of the British Slave Ship Brookes Under the Regulated Slave Trade Act of 1788”, by Alfred, W., 1788.	42
<i>Figure 5.</i> Sprite Sheet for animation experiment, from “Tireless Feather”, by González J., J. C., 2012.	49
<i>Figure 6.</i> Screenshot of the interactive application that modularly renders drawings based on seismic data of Ingeominas, from “Circle Variations - Pencil”, by González J., J. C., 2012.	51
<i>Figure 7.</i> Data Driven Drawing using seismic data and a sprite sheet of if strokes, from “Circle Variations - Ink”, by González J., J. C., 2012.	52
<i>Figure 8.</i> Structure of steps in data representation, from “Computational Information Design”, by Fry, B., 2004.	86
<i>Figure 9.</i> Detail section of the graphic notation used in the film “Tango”, from “Tango – 1980 Lodz, Planning the Action ´s Progress in Time”, by Rybczyński, Z., 1980.	124
<i>Figure 10.</i> Detail section of the graphic notation used in the film “A Chairy Tale”, from “Visual Outline of A Chairy Tale, prepared by McLaren for Ravi Shankar”, by McLaren, N., 1957.	124
<i>Figure 11.</i> Detail section of the graphic notation used in the film “Mr. Bug Goes to Town”, from “Story Mood Chart”, by Fleischer Studio, 1941.	128

<i>Figure 12.</i> Circular graphic notation for short-film, from “SiSiSiSiSiSiSiSiSiSi”, by González J., J. C., 2011.	140
<i>Figure 13.</i> Graphic notation and final frame superposed, from “SiSiSiSiSiSiSiSiSiSi”, by González, J, J. C., 2011	141
<i>Figure 14.</i> Graphic notation from data visualization of 2003 data of seismic activity and terrorist attacks, from “Rite of Passage”, by González J., J. C., 2013.	149
<i>Figure 15.</i> Frames corresponding to the sentimental analysis score, from “Eterna”, by González J., J. C., 2015.	169
<i>Figure 16.</i> Animation sequences contained in a sprite sheet, from “Eterna”, by González J., J. C., 2015.	169
<i>Figure 17.</i> Data Driven Drawing introductory interface delineating the neighborhood <i>El Gallo</i> , from “Vista al Mar”, by González J., J. C., 2018.....	185
<i>Figure 18.</i> Data Driven Drawing final visual composition of <i>Google Street View</i> and superposed drawing, from “Vista al Mar”, by González J., J. C., 2018.	186
<i>Figure 19.</i> Diagram of simplified cropping technique, from “Dibujo Defunciones”, by González J., J. C., 2018.	195
<i>Figure 20.</i> My father’s signature, from “Dibujo Defunciones”, by González J., J. C., 2018.	196
<i>Figure 21.</i> Data Driven Drawing state of final visual composition, from “Dibujo Defunciones”, by González J., J. C., 2018.....	198
<i>Figure 22.</i> – Data Driven Drawing resulting state of the year 2008, from “N. N.”, by González J., J. C., 2018.	202
<i>Figure 23.</i> Graphic notation of violence data, from “Monitor Info”, by González J., J. C., 2017.	203

Abstract

This practice led doctorate investigates the potential of artistic representations that express subjective views derived from traumatic personal experiences, using a series of impersonal datasets relating to acts of violence in Colombia as the point of origin. This investigation is led by a series of animated projects of my own creation in which I utilized data that contextually relates to deeply personal events that affected my life, seeking to build a defense for the subjective experience to permeate and inform the representation of data. These artistic projects seek to interrogate complex experiences that indicate unresolved states of mind in constant states of flux, leveraging the potential of data to be an index of phenomena with these characteristics. Relevant theoretical research that contextualizes these works has been conducted to elucidate contrasting arguments for the potential of representing data either as an expression of these complexities to allow for clearer reasoning or the opposing view that investigates the re-presentation of the complex and the limitless as states where reason becomes forestalled. The first chapter contextualizes artistic data representation and examines the role of subjectivity in it. The second chapter proposes a connection between data visualization and the tradition of graphic notations in music and animation to study a point of intersection between functional data visualization and artistic practice. The third chapter examines in more detail the contrasting arguments for approaching data representation. Once the theoretical proposal for subjectivity in data representation is established, chapter 4 discussed my own practical work as an embodiment of it. Formally and aesthetically Data Driven Drawings investigate the possibilities and affordances of animation to develop temporal articulations of data, in order to explore the mutating flux of underlying themes that motivate artistic projects. The practice and theory are interwoven to create a single unified argument for the notion of Data Driven Drawings presented here.

Introduction

This practice led doctorate investigates the potential of artistic representations that express subjective views derived from traumatic personal experiences, using a series of impersonal datasets relating to acts of violence in Colombia as the point of origin. Despite the prominent emphasis on empirical data, the way in which this information is represented does not attempt to separate the subjective feelings from the “real” events, for these representations are deeply informed by an emotional connection to the data.¹ By using datasets that describe complex and wide encompassing issues of violence in Colombia, the author seeks to examine the role of subjectivity within the overpowering phenomena that contextually relates to his personal life, but does not singularly describe it. While a dataset is, by virtue of its form, a reduction in scale of an otherwise limitless phenomenon, its potential representations, or ways in which it is presented to an audience, can either fulfill the goal of providing an object that allows for clear reasoning and understanding of such phenomenon, which this in itself is a valuable objective or the approach that is investigated here in which the representation denotes back the unresolved and abstract nature of the events captured in the dataset. This artistic attitude towards data representation is as an approach to animation proposed here as Data Driven Drawings in which the animations are tangible audiovisual manifestations of the abstract and complex nature of the phenomena that data denotes and the emotions that influence their interpretation. These artistic

¹ For instance, some of the datasets contain information about the history of violence in Colombia. This information refers not only to the context in which the author grew up but also to the environment that led to violent events in which people close to him lost their lives.

projects seek to interrogate complex experiences that indicate unresolved states of mind in constant states of flux, leveraging the potential of data to be an index of phenomena with these characteristics.

Research Objectives

The main objective in Data Driven Drawings is to build a defense for subjectivity in the general practice of data representation.

Considering that the current debate in data representation presents two seemingly opposing approaches; on one side of the spectrum the proposed argument is that the goal of representing statistical information has to be “functional”. In it, the goal is to expose “evidence” of a hidden reality embedded in the data. This approach prescribes to the rigor of the scientific method and reasonably the academics in the field contest that in order for the data to be faithful to the phenomena of study, the artist should not interfere with his/her own subjective view in the process of representing the information.

The favored conceptual premise found in the leading academics of the field (Franceschet, 2014; Friendly, 2008; Fry, 2004, 2008; Halpern, 2014; Lupi, Popova, & Posavec, 2016; Manovich, 2002; Tufte, 2001; Tukey, 1962; Yau, 2011), argue that the goal of representing data should not be to bring forward an artist point of view, but to allow for the information itself to decide the perspective from which the information is communicated (Chapter 3).

Furthermore, even in cases where the representation of data is framed within artistic practices (Husbands, 2013; Manovich, 2010; Viégas & Wattenberg, 2007), the relationship between representing

reality and the subjective point of view of the artist has not been clarified.

This thesis focuses on this theoretical conflict. The goal is to develop a conceptual framework where data representation and artistic intention overlap. At the same time, discuss how this union is less problematic than it initially seems. It is argued that data representation and animated documentary share a fundamentally similar intent, this is the representation of realities that are inaccessible to unaided human perception. For this reason, they also share theoretical questions about the relationship between the artist subjective views in the process of articulating an image of reality. The theoretical work adheres to the tradition of animated documentary research (Honess Roe, 2013; Sofian, 2005; Ward, 2005; Wells, 1997) to initially disambiguate the subjectivity against representation of reality conflict. Particularly referring to Honnes Roe's proposition of "evocative" animated documentary (Chapter 3).

Regardless of these similarities, the author does not claim that Data Driven Drawings are only a form of animated documentary. As examined in Chapter 4, the work developed here is a hybrid between the cinematic tradition of animation and the digital paradigms under which data representation commonly operates. For this reason, the present dissertation debates the position of evocative representations of reality from within both domains under which the practical work has been created: animated documentary and data representation.

For the practical work, the central premise is located within the authors subjective and personal experience of an internal phenomenon, states of the mind in a state of flux triggered by violence. Despite the highly personal, the referred state of mind is

one in which full self-awareness is not reached, nor the goal is to contain it for its comprehension. Tor Nørretranders develops a lucid articulation of this problem in his book *The User Illusion: Cutting Consciousness Down to Size* (1999) in which he makes a clear distinction between the *I* and the *Me*². Being the *I*, the version of the self that is aware itself, its history and how it presents to others, while the *Me* takes care of all the bodily and mental processes that the *I* is not to be bothered with, the *Me* is the general person while the *I* is the version we are conscious of. Within this framework, *I* am aware of the context and biographical story that I might relate to you about the death of my father, but really what I want to articulate and present to you is the *Me*, that I have no access to. The personal opinions *I* have of the history of violence in Colombia, how they affected me personally are of a secondary order, but the flux of the experience that inhabits *Me*, is intuitively what an artistic endeavor becomes a medium to access that unrecognizable state of mind.

The *I* can communicate and control its communication with others. The *Me* can also communicate, but it does not have the same awareness of what it communicates. The *I* is social and can enter into agreements with other *I*'s. The problem is whether the *Me* will abide by those agreements. (Nørretranders, 1999, p. 269)

Nørretranders' proposition of the *I* that enters into agreements is where the type of data representation that reaches consensus and describes the world clearly makes sense. But it is not the one this thesis pursues, the central problem is the one stated by

² The reader might also consider expanding on this idea from the proposal of neuroscientist Antonio Damasio in his book *The feeling of what happens* (2000). where he articulates his three layers' theory of consciousness. He proposes the idea of "Protoself" as the most fundamental and relates closely to Nørretranders' *Me*, the "Core consciousness" that is a state of awareness concerned with the present moment and functional operations, and lastly the "extended consciousness" that provides the self a more complex awareness of itself in historical context.

Nørretranders, that “strange” way in which the *Me* seeks communication with others but does not necessarily “abide by those agreements”. Those agreements pertain to an approach to data representation that removes the *I*, it strives for the world to reveal on its own because the ability to record data empowers the chaos of life to be contained and understood logically, in this process *I* have nothing to do with its representation. The same procedure that allows data representation to be functional is examined here as the way to delegate to data the full control that the *I* has in the process of communicating with others. The difference is that setting it within an artistic context permits the exploration of new agreements under which data can be communicated to others. The practical work developed for this thesis is a series of efforts in which the *Me*, is trying to configure ways to present itself and share the experience of violence.

The theoretical research that contextualizes these practical projects, is examined under contrasting philosophies that argue for the potential of representing data either as the fulfillment of bounding complex phenomena to allow for clear reasoning, and the opposing view that investigates the re-presentation of the complex and unbound where reason is forestalled. The first examination of these opposing views is presented in Chapter 1 in which a proposal is put forward to define two categories that abide by each of these schools of thought: the “functional” and the “evocative”, these are then paralleled with the philosophical tradition of the beautiful and the sublime respectively as proposed by several authors in the context of artistic data representation (De Landa, 2000; Heinrich, 2015; Manovich, 2002; McCosker & Wilken, 2014). The argument proposed in Chapter 1 is the first defense towards validating subjective views on data representation as defined by the intent of

the author (either evocative or functional) and the potential aesthetic experiences that can be denoted through the use of data (beautiful or sublime).

In Chapter 2, a proposition is put forth that argues that functional data visualization can fulfill an important role in the preparation of artistic work and study this through the history of graphic notations and by analyzing 4 case studies. In Chapter 3, I continue with the contrasting philosophical views in discussion with authors that reject the idea of subjectivity from data representation and in contrast the subjectivity is validated in the theories of animated documentary as the field that studies the imaged representation of reality. Once the theoretical defense for subjectivity in data representation is established, the practical work is discussed in Chapter 4 in relation to it.

Literature Review

Considering subjectivity in the fields of data representation presents a contentious debate. The following summarizes the key authors discussed in this thesis and their ideas relating to the research objective put forth earlier, in order to introduce the logical flow of the argument.

The ability to create an artistic image that is able to remove any personal trace of the author is in itself a complicated task. In terms of the form and not necessarily the intent, the seminal work of Jaques Bertin (2011)³ sets the groundwork for the development of a

³ The edition referenced here is of 2011 but the book *Semiology of Graphics* was originally published in 1967.

homogenous graphics language for data representation. His ideas have been widely tested and implemented in computer graphics today as they conceptually provide a rigorous method for translating data types to specific graphical form that best describes them. This is not to say that human decisions can be completely removed from the representation by delegating to computer Software the “drawing” of graphics based on the data structure, but our relationship to reading data is heavily influenced by the standardization of data graphics as they evolve from Bertin initial proposition into automation in modern computers. A contrasting approach to this is to depart from codified data graphics into the creation of new visual forms that respond to specific research questions and needs.

We find new forms of visually codifying data particularly in the arts, but the seed for this approach can be found in the ideas of John Tukey, a mathematician who developed an eloquent plea to his peers in the scientific community to establish an “attitude” towards data that is “exploratory” vs “confirmatory” analysis (Tukey, 1962, 1977, 1990). In an exploratory attitude toward data, Tukey is proposing that we should allow for the early stages of visualization to be in constant iterations and as the term suggests “explore” these images in various forms to allow for questions and hypothesis to emerge from the patterns that emerge, and not limit the data a priori to answer a specific question before it is explored. His ideas have also influenced the development of computer software for allowing precisely the automation of large calculations and mappings of the data, to allow for agile exploratory analysis. This attitude towards data is particularly relevant for artists as this is where the aesthetic forms that embody a dataset require a lot of exploration and fluid iteration. Gaviria (2008) proposes a clear

distinction between data artists that seek to appropriate and articulate their ideas based on previously codified forms of representation “Genre Art” and those who adhere to Tukey’s proposal and explore, interrogate or create new forms based on the particular questions posed to the data “Research Art”. While Gaviria is proposing these categories of Genre and Research Art in response to contemporary practices of data visualization, it is important to identify and draw connections to the seminal ideas that crossover from the scientific community into data art practices.

Even in the most automated forms of data visualization, the author is always intrinsically connected with the end result, bias can easily permeate the representation and in certain fields of research, this is strongly rejected. The most influential voice in the rejection of subjectivity in data representation is Edward Tufte (2004; Tufte, 1990, 1997, 2001, 2006, 2017) who is referred to in almost all the literature on data representation either to defend and reject his deep-seated views. His conceptual rejection of subjectivity is very well argued, through a long list of examples and historical revisions, he is able to articulate a unified defense that data is able to contain evidence of the complex “world out there” and being able to allow for this evidence to emerge is vesting the data of its intrinsic power to tell us fascinating stories. This stories and clear manifestations of the complex are buried in the data waiting for a representation to bring it forth to us. In this sense, he forcefully rejects the authors' opinion and subjectivity since these are elements that compromise the manifestation of the phenomenon for what it is, what is able to reveal and excites us. While Tufte’s arguments are compelling, they leave no room for researching ways in which the subjective experience of the author may produce forms for interpreting the “world out there” artistically. In other words, Tufte puts all his

confidence in the world manifesting itself through data, but this research investigates ways in which we can provide views of that world as it manifests on data. This operation is relevant in the arts, but can also be found in studies of information philosophy.

Two key authors in this field are Rosenberg (Gitelman, 2013) and Floridi (Floridi, 2011, 2016) who provide clear conceptual counter-arguments to Tufte. Rosenberg's etymological study of the word data, illuminates a critical idea for artistic endeavors which is that data is "specifically rhetorical" (2013, p. 18), in other words, data is articulated into an argument to persuade rather than assert, without the rest of the parts that configure how data is presented to others, data in itself is voided of meaning. Drawing a parallel back to Tufte, this is not to say that Tufte is unaware of this since he is always considering the role of the author as one that builds that context for data to have meaning. The difference between the two is that Rosenberg's argument is wide-encompassing, no opinion is made as to how data should preferably be given meaning in the way that Tufte does to reject subjectivity. Aligned with Rosenberg are the ideas presented by information philosopher Luciano Floridi who argues that there is "no information without data representation" (Floridi, 2011). Developing further Rosenberg's idea of data as a rhetorical device, Floridi argues that information is the sum of data and meaning, but he goes further to say that there is no singular meaning corresponding to a dataset, but that datasets are transient objects that can be interpreted and answer multiple questions depending on who studies them. Floridi's proposal is crucial for identifying the role of artistic interpretations of data. His argument sets the premise that anyone one, from their own domain, is able to singularly pose questions to a dataset and articulate its meaning to produce a data representation. The type of information that

emerges from these data representations is directly connected to how the author seeks to communicate with others the meaning inferred from the data. If the artist intends to evoke an emotional experience, the ways in which the information is presented may be convoluted and challenging to favor what emerges from the effort rather than providing a clear space for reasoning.

The artistic intent as to how data is represented and communicated to others is also related to how the artist understand the datasets in itself, what is it accounting for. For media theorist Lev Manovich (2001, 2002, 2010), the meanings that artistic data representation affords to articulate are a new order of artistic abstraction and claims that they are able to fulfill an artistic and philosophical impossibility of representing the ineffable, he proposes for this that artistic data visualization are “anti-sublime” (Manovich, 2002). This premise is debated as it is seen here as a contradiction. What Manovich claims is that phenomenon that before was impossible to represent, data visualization can now make perceptible to the scales in which clear reasoning can operate. The problem with this claim is that by omitting emotions and the subjective feelings produced by the ineffable, his definition of the sublime is only considering one type of human knowledge which is that in the scales of clear reasoning, leaving aside the other forms of knowledge that emerge from the tension of not knowing and the human drive to reason. If a phenomenon, as Manovich claims, is reduced to the scale of reason, what it is argued here is that he is simply referring back to the nature of the beautiful, not the anti-sublime. To further develop this rejection to Manovich’s argument and articulate who precisely it is in the sublime where artistic data visualization has its most potential, a lengthy section of Chapter 1 is dedicated to revision key authors of the sublime (Acosta, 2012; Burke, 1757, 1796; Heinrich,

2015; Kant & Goldthwait, 2003; McCosker & Wilken, 2014; Morley, 2010; Pseudo-Longinus, 1890) and drawing connections to artistic data representations.

Having set this theoretical backdrop contrasting the functional and the evocative categories of data representation let's discuss two more precedents with which Data Driven Drawings finds more direct alignment. Firstly, considering data as an index of reality, the ways in which artistic representation refer or point back to that provenance of the data is not as direct as other artistic practices like photography (Bazin, 2005). Rosalind Krauss (1986) examines a key moment in Art history in which photography and the index begin to permeate other forms of artistic practices producing a wide array of original works in the 70's avant-garde. Through the analysis of a series of case studies, Krauss is able to deconstruct the notion of the index (Peirce, 1932) in photography and draw connections of its mechanism to artistic works that explore the idea of signing back to reality in way in which the physical correspondence is not as evident as it is in photography. For instance, in Deborah Hay's piece were she reconfigures the codes of a dance performance by using her speech to indicate to our imagination that the molecules in her body are performing the dance she alludes to. The expansive view on how art is able to experiment and produce new forms of understanding the index is directly connected to Data Driven Drawing's goal to provide the audience with information that challenges the notions of an apprehension of the world out there. Data as an index for the arts require exploration (Tukey, 1990), research (Gaviria R., 2008) and experimentation with new forms of representation (Krauss, 1986).

Lastly, Data Driven Drawings finds its closest ally in the practices and scholarship of animated documentary (Honess Roe, 2013; Willaim Moritz, 1997; Patrick, 2004; Sofian, 2005; Ward, 2005; Wells, 1997). This research adheres to the tradition of this field as the one exploring the subjective experience of reality through animation, and also brings forth to the scholarly discussion a new index of reality which has not been considered before which is data.

Methodology

Data Driven Drawings is practice-led research for which two fundamental questions arise for the creator of this artworks. The first has to do with the intent of the artist and the second to the film form: this includes the type of narrative and methods of production of an animated piece.

Research question 1: What is the artist's intent when data is the index of reality?

The first problem defines how the artists will interpret the data and present it to the audience, what is the artist intent? Will the representation be about how "I" see the phenomena? Or is it about what the phenomena are in itself? For instance, in the first case, the goal could be to communicate what the phenomena means to me based on personal experience, or an effort to express an opinion of what the data signifies. In the second case, the artist ought to separate their own subjective experience and bias from the representation in order to allow for the phenomena to be studied and clearly understood by others. These scenarios are not proposed here as mutually exclusive, but rather as a conceptual range in which multiple possibilities for representing data will fit.

Research question 2: What is at stake for the animated film form when data is introduced to the production workflow?

The first basic problem that has been identified regarding this question is that we do not have a clear naming convention for animations that use data as an index to what the animation represents. Even though artists have been creating animated works based on data arguably since the '40s and before digital computers⁴, but more consciously since early digital animation (Youngblood, 1970) under the broad definition of “Expanded Cinema” proposed by Youngblood.

Presently, the relationship between arts and data is even more omnipresent in approaches and techniques using digital technologies. Artists are able to work seamlessly with the computer as an expressive tool, it is not uncommon anymore to find artists who are able to write code and create their own artwork directly on the computer using a widely available source of data openly on the web. Despite this, no clear precedents and examples were found in which traditional animation using empirical data from social events were combined to propose a subjective view of the data, as this research attempts to examine. For this reason, the methodology established here was to develop practice-led research for which the practical work could serve as the primary source of examination to develop a single unified argument for Data Driven Drawings.

⁴ One of the earliest approximations to data driven animation can be identified in the experiments that were developed in the Fleischer studio during the 1940's during the development of the film “Mr Bug Goes to Town”. In it, the artists devised a technique they called “Story Mood Chart” in which they plotted into a graph the emotional journey of the film. Even though it was not a formal acquisition of data in order to develop these charts, an empirical approach to attaining information about some complex phenomena such as emotions of a narrative and visualizing this information, is in principle not any different to modern digital approaches to representing data.

The Need for Practice-Led Research

As mentioned above, this practice led research stems from the lack of examples of animations using data and an index to evoke and explore subjective themes. For this reason, the following projects were developed to support and to be tested by the theory.

The main conceptual premises established in the outset of the practice was that the datasets used should be directly connected to a personally traumatic event, this way, the ability to separate oneself emotionally from interpreting the data was deliberately complicated from the beginning. Also, the resulting artworks ought to articulate a subjective view through metaphorical associations and present back the data in a tangible audiovisual form that embodies the abstract nature of the subjective experience and complexity of the phenomenon in its natural states. This forces the artworks to embrace ambivalent presentations of data that result in contingencies rather than assertions of the information contained in the datasets.

To do this, the three main projects described in Chapter 4 and the backbone of this thesis are conceived as three points of view of the same event. The traumatic experience they embody relate to the death of my father as seen through data that relates to his violent passing. The first project “Vista al Mar” is the closest to the event and uses data collected personally (the author’s own heartbeats and information extracted from documents in the family archives). The second project “Dibujo Defunciones” steps further away from the specificity of the event to see it through statistics on homicide in

Colombia at large. For the third and furthest away point of view, the project “N. N.” again makes this same exercise but by relating the personal emotion of the trauma to the wider encompassing dataset of all available data of violence categories in Colombia.

Figure 1. Three points of view of Data Driven Drawing projects, from “Vista al Mar” & “Dibujo Defunciones” & “N. N.”, by González J., J. C., 2018.

The way these points of view are then articulated in the theory has to do with the conceptual range between functional and evocative data representation. While they all seek to be evocative, the internal negotiations for each piece vary. In “Vista al Mar” the event is alluded to more directly, the animation uses Google Street View to locate the personal drawings in the place where my father was murdered. The specificity of the data, the heartbeats that set the rhythm for the image and the drawings of how I remember him, all are at its most subjective and personal. The data points to our family, yet the representation tries to aesthetically evoke the emotions of the author. The second level has both a direct allusion to him through photographs of him hidden behind a white veil that slowly get uncovered and covered by mapping a dataset of all the homicides reported in Colombia. A tension here grows between the functional operation to map these homicides and reveal the

overwhelming frequency, and the imposition of the singular allusion to my own father within them. The personal information is interwoven with the functional display of data. In the third and more distant level, the reference to the personal life is removed from within the image, yet the way that the lines that are drawn based on all of the types of violence available in Colombia, are charged by what they sign point to the author. The overwhelming amount of violence, seen from afar, beholding a vast dataset into a single image that shows itself calm and “beautiful” in its appearance, is still seeking to provoke an emotional experience despite its ability to reduce a large and complex phenomenon like all registered cases of any type of violence in a country.

Conceptual Framework

Animation is not commonly associated with data representation since the underlying sources that operate in the creation of each of these objects are often understood as distant apart. The former relies on the author’s own ability to provide a trace of either an observation of the world or propose an imagined one, while the latter depends directly on the world to offer that trace. In animation there is always a proposed articulation of the source that denotes reality, passing through the subjective view of the artist, while data representation may or may not rely on a personal view and can be delegated to systematic rules that have the potential of distancing subjectivity from the end result. This fundamental difference between these two forms of representation, which is contained in their internal structures of production, has conceptually separated them as distinct modes of representing reality. Yet, a close revision of these structures will provide clarity as to how these modes of representation may be combined, blurring the gaps between

subjectivity and data representation and therefore providing a pathway to the proposition of Data Driven Drawings where this hybridization is investigated.

Data Drawings

The first examination pertains to how drawing operates in the creation of a static data visualization, being this the transversal form of representation found in other more complex structures that use animation or interactivity. Drawing with data, or more precisely, allowing data to be operative in the process of “observing” the world and render it in a diagrammatic disposition of lines, introduces new challenges to the traditional understanding of drawing as a creative activity. Certain aspects of the drawing operation are delegated to the data. This delegation is not fashionable or arbitrary, but a necessity when our physical senses cannot access the versions of the world that data is able to yield. Delegating to data does not mean to simply separate the self from the act of drawing, but to engage in a creative negotiation with data as a material.

A type of data drawing that separates the self is often a drawing that seeks to be an index in itself and pertains to a category of drawings that has purely functional objectives. For example, a seismograph is a device that transforms the impulses of seismic activity into electrical impulses that the mechanism is designed to translate into lines that progressively denote the intensity of the tremors. This visual codification is functional because of general consensus as to how this drawing should be interpreted. The device is precisely calibrated to produce lines with a specific meaning and human interpretation only arrives after the lines are drawn. The technological development allows for a full delegation of the

drawing to a standardized machine that has no judgment over how the drawing is created and therefore fulfilling its task. Yet drawings that delegate to data do not have to exclusively distance the self if a deliberate thinking process is a position between the raw data shaped by the phenomenon and the rules that define how the drawing should be executed.

An artist that investigates this deliberate imposition of the self in what seems like drawings that respond directly to raw data is Tim Knowles. In his series “Tree Drawings”⁵ he delegates the drawing to a tree by attaching pens and brushes to the extremities of the branches, then carefully places white paper near it and waits for the wind to move the branches naturally. Similar to the seismograph, this system translates the impulses of the wind into lines on the paper, but these drawings do not produce standardized lines that can be interpreted under general rules of analysis. The lines in Knowles work are not only the traces left by the phenomenon that activates them, but they also make reference to the artist who decided how these traces are registered and presented. In the exhibitions of this work, Knowles displays both the final drawings and pictures of the process, highlighting that the process of how it was made is as important as what is registered in the paper.

Knowles’ work highlights that the process of delegating certain aspects of the drawing to data for artistic purposes, implies that the author cannot rely solely on the consensus of how to codify the information into lines (like a seismograph), but propose those forms of abstraction in the invention of his own device. Knowles does this by designing his own analog “machine” (the tree with attached

⁵ See the project in the artist’s website:
<http://www.timknowles.co.uk/Work/TreeDrawings/tabid/265/Default.aspx>

utensils) for registering raw data that manifests from wind and even though he cannot control the outcome, his own presence as an artist emerges from the articulation of how the data is drawn by the device. This strategy works for natural data that has a physical presence that can be captured as impulses as they happen, but when the source of the data is not so directly connected to a physical phenomenon, the author needs to get more involved in the decision-making process of translating the data into visuals. As this happens, it turns more complicated to separate the self from the representation and often the delegation turns not only to let a machine decide how to draw the data, but also aesthetic decisions that standardize the graphical elements that codify the information in the image.

Two early examples of data visualization are examined for their contrasting use of standardized graphics. The John Snow's map of the cholera epidemic in London 1854 and the Brookes slave ship plan of 1788 by William Alfred. Snow's map shows the utility of primary geometric shapes to concentrate solely on communicating with a line a specific piece of information, and in the Brookes diagram, the use of figurative and idiosyncratic drawing that suggest personal opinions from the maker⁶.

The city of London experienced a catastrophic cholera epidemic between 1849 and 1854 with a major outbreak on August 31 of

⁶ Further reading on data drawings can be found in the excellent essay "Index, Diagram, Graphic Trace" by Margaret Iversen in which the author proposes the idea of graphical trace in the context of the philosophical tradition of the index and contextualizes data drawings as used by a series of contemporary artists. (Iversen, 2012)

1854 where 616 people died⁷. One of the main difficulties to contain the epidemic was that at the time, the leading hypothesis suggested that cholera was initially transmitted through particles in the air that surfaced from the contaminated soil that surrounded the Thames river. This assertion known as “miasmatic theory”⁸ was led by the well-respected epidemiologist William Farr who drew his hypothesis from studies that correlated mortality with elevation and therefore inferred that the disease was transmitted through the air. A counter-hypothesis was being proposed by physician John Snow who argued that the disease was contracted by ingestion of a contaminated germ and then transmitted between people. At the time there was no theorem for germ disease and therefore Snow had no way to clearly demonstrate that the outbreak was caused by contaminated water and not foul air. Snow attempted to examine samples of the water distributed to residents through pumps around the city, but these experiments were not conclusive. He then approached the issue by using data visualization as a method that at the time was novel, but ultimately persuaded local authorities and his peers in the scientific community. Snow’s visualization was a map that first located the public pumps from where people usually

⁷ An exhaustive account of this event can be found in Sandra Hempel’s book *The Strange Case of the Broad Street Pump: John Snow and the Mystery of Cholera*. Hempel describes in great detail the life story of English physician John Snow, the African migration to Europe with the arrival of new diseases, London’s infrastructure, the issues with water companies that provided contaminated water to their users, and provides a clear picture of London during the mid-nineteen centuries which indicates the context where Snow struggled and ultimately succeeded. (Hempel, 2007).

⁸ Miasmatic theory is now considered obsolete by the scientific community and this is due, to great extent, by Farr’s own investigations after the cholera outbreak of 1854. This transition miasmatic theory to the now accepted conclusions that the disease is transmitted through ingested agents “germ theory” is clearly articulated by historian of modern medicine John M. Eyler in *Victorian Social Medicine: The Ideas and Methods of William Farr* (Eyler, 1979) and specifically addressing the case of the cholera outbreak in *The changing assessments of John Snow's and William Farr's cholera studies*. (Eyler, 2001)

obtained their drinking water and placed them as nodes. He then drew rectangular lines that represented the geographical places associated with those who died of the disease. The main cluster of lines that emerged from this method was located around the pump in Broadwick Street and further investigations showed that a nearby sewer had started leaking into the well and it contained fecal material of cholera contaminated diaper.

Figure 2. Cholera epidemic map emphasizing the high density of cases near the Broadwick Street water pump, from “On the Mode of Communication of Cholera”, by Snow, J., 1854.

This now iconic example of data visualization as evidence producing document used simple geometric shapes to denote each of the cases in the data, and therefore each line carries a singular expressive significance that distances from the personal expression

of Snow's own hand that draws them. The lines do not denote an emotion or opinion of how he felt about the person but are rendered carefully equal as if done by a machine. Each line removes away traces of personality that describe the victim, we don't know their gender, occupation or any trace of personal history to concentrate exclusively on the geographical location where the case happened. In this way, the image that emerges is allowed to denote what Snow was trying to find which is a map of the phenomenon as indicated by simple geometric shapes with black ink and not an internal exploration of a subjective view that a more intuitive line may suggest. Snow's personality is hidden away from the drawing so the perceivable object distances from the author formally⁹. This is not so clearly the case in the diagram of Brooke's slave ship and our reading is influenced emotionally by the author in a conspicuously but effective way.

⁹ This view is challenged by Tom Koch in his article "The Map as Intent: Variations on the Theme of John Snow. *Cartographica*" (Koch, 2004) where he revisits the original data used by Snow and uses digital technologies to remap it. In his experiments he is able to conclude that Snow left out a lot of data variables that could provide a much clearer view of the epidemic, but Snow was so focused on his objective to prove that the Broadwick Street pump was contaminated that he makes personal decisions on leaving out data he deemed unnecessary for his hypothesis. Even though he did not purposefully lie, Koch uses this to argue that our relationship with images that suggest evidence are powerful in their ability to remove the author, yet internally there is no map or visualization that structurally is completely removed from the maker's personal view and opinions.

Figure 3. The complete diagram of the Brookes ship, from "Stowage of the British Slave Ship Brookes Under the Regulated Slave Trade Act of 1788", by Alfred, W., 1788.

The Brook ship diagram by William Alfred is one of the earliest data visualization images dating 1788, furthermore, it also proved to be an image that could enact social outrage by emotionally appealing

to the viewers and was widely printed throughout the UK¹⁰ at the time. The diagram denounces the inhumane conditions under which slaves were transported inside the Brookes ship by abstracting the average height of a man, woman, and child and then diagrammatically render each possible person that could fit inside the dimensions and empty spaces of the ship. Different from Snow's image, each instance of the data was not abstracted into geometric shapes, but into a figurative depiction of a human being. The plan of the lower deck shows how they barely had any personal space between them if they laid down on the floor at the same time, which suggests that the space of the ship was optimized to its maximum capacity regardless any possibility of dignity and responsibility for these human beings. The diagrams render each human figure with very small differences between one body to the next and give prominence to the difference in height to indicate which are a man, woman or child and their gender-specific placements inside the ship. The "average" drawing depicts each human as a figure with dark skin and clothing with no more than short underpants. In the men's section we see that each one of them is attached to each other's feet and arms with what could be a rope or a chain, intensifying the deplorable conditions by suggesting that not only there was little or no space between them, but also chained to each other to reduce the chances of moving around.

The "longitudinal section" or side view of the lower deck proposes a few particular characterizations that are not found in the other sections. We see two men sitting in the floor with their legs shrugged, holding their own arms towards their chest and looking

¹⁰ See Thomas Clarkson *"History of the Rise, Progress, and Accomplishment of the Abolition of the African Slave Trade"* (1839)

towards the floor. Another man in the same position is shown from a front view and near him another man crawling towards him. At the other end of the deck, we can see two men walking toward the gun room in hunchbacked positions and one holding his arms towards the chest and intensely looking towards the floor. The man next to him is holding his arm on his leg as if pausing to rest. In the context of what the image narrates, this side notes of what the author imagines how it could be to be inside the ship, turns these men personable and human, even if the renders are not hyper-realistic or detailed, their depiction is very expressive and emotionally captivating. Is the author lying by allowing his imagination to take part in the visualization? Reading it today we can see all the diagrammatic codes of scale legends, textual notes and careful abstraction of the bodies as nodes that indicate each instance of the data, yet the figurative approach to drawing and the distinctive expressive characters imagined by the maker, enhances the complication that it was not an image purely delegated to data, but is intervened deliberately by the author. The introductory note contained in the top-right side of the image also indicates rhetorically that what we are seeing is not merely evidence of an event, but a denunciation of a catastrophe.

Figure 4. Detail of the note describing the Regulation Act of 1788, from "Stowage of the British Slave Ship Brookes Under the Regulated Slave Trade Act of 1788", by Alfred, W., 1788.

The note in Figure 3 reads: "*The Brookes after the regulation act of 1788, was allowed to carry 454 slaves. She could stow this number by following the rule adopted in this plate, namely allowing a space of 6ft. by 1ft. 4in. to each man; 5ft. 10in. by 1ft. 4in. to each woman, & 5ft. by 1ft. 2in. to each boy, but so much space as this was seldom allowed even after the regulation act. It was proved by the confession of the salve merchant that before the above act, the Brookes had at one time carried as many as 609 slaves...*"

The note not only describes the law and suggest the methodology applied to render this data, but also indicates that the image we see is the abstraction of what the Brooker could hold if it carried 454 slaves and even then, it overwhelms as inhumane. The note also allows for us to imagine how it could be when they actually reached the 609 slaves in one journey and delegating the creation of that image to our minds, to our imagination, and in this way the author effectively appeals to our emotional reading of the event. The

representation may be bias but it doesn't necessarily lie about the event and the data that upholds it. Its articulation supports the information obtained and overtly exposes its subjective point of view to the observer. This shows the underlying complications when the author takes presence in a data drawing and assumes a moral responsibility that may be challenged, yet the data pulls our attention towards the veracity of the information and the depiction and rhetoric towards an emotional reading. The tension that emerges between these two allows for data art to find a place in these types of representation, where an author positions his/herself inside the codes provided to us for reading the image.

Connecting this idea back to the particular field of animation that is of concern to this investigation, even when animation operates as a representation of reality and not entirely from the imagination, it is contingent on the author's subjective view of the phenomenon as it never provides an unmediated index of reality. As discussed here data representation or data drawings, on the other hand, is not exclusively dependent on personal opinions and has the potential of operating itself as an index of the reality it reveals, as shown in the examples above. This particular affordance of data representation to be in such close proximity to the phenomenon it indicates has traditionally positioned this practice as a privileged form of giving us access to the world out there that is inaccessible to our physical perception. As such, subjectivity, or the world inside, is not strictly required, and in some cases, theoretically rejected (Tufte, 1990, 1997, 2001, 2006). The combination of animation and data drawings open new possibilities for an animation to redefine its affordances with an index that is not normally used in its production process. Data Driven Drawings explore this combination as the process of animation is rearticulated to include ways in which

aspects of its constructions are delegated to the impulses contained in the data. As we have seen in the previous examples, a range can be established between how much effort is put into designing the systems that allow for the data to manifest without much of the authors opinion (seismograph and Snow's map) to other systems in which the authors' own views are calibrated into the operation (Knowles and Alfred) without removing the indexicality of the data. In any case, the procedure of animating or drawing with data involves the personal calculation as to how much of the self ends up becoming part of the representation. This negotiation is at the heart of the theoretical discussion presented in this thesis.

Data Driven Drawings

The conceptual premise of Data Driven Drawings is to explore the artistic representation of data in which the subjective point of view of the artist takes part in the articulation of the information presented to the audience. Like in the Brookes ship diagram described before, the goal is to enunciate and propose rather than completely delegate the emergence of the images to the properties of the data. Formally and aesthetically Data Driven Drawings investigate the possibilities and affordances of animation to develop temporal articulations of data, in order to explore the mutating flux of underlying themes that motivate artistic projects. The practice and theory are interwoven to create a single unified argument for the notion of Data Driven Drawings presented here. In order to expand the previous examination of drawing in static images, let's examine the structural characteristics of animating with data.

Two types of drawings are used for the animations produced during this investigation, analog hand-drawn images and digital ones. Both

visually articulate in some form the data using different algorithms that parallel the information in the datasets with how the drawing is either composed on screen or generated in “real-time”. Some of the early experiments discussed below, constitute small steps into understanding the different formal stages of production in a Data Driven Animation and even though they do not constitute finished pieces that denote the intended themes discussed earlier (and in more detail in Chapter 4), they served as important milestones in the learning curve of producing animation with data. At the same time, using either analog or digital drawings respond to an aesthetic exploration which is to denote in the drawing that it is made by a human, reveal that the articulation of the data is being proposed by someone and not solely delegated to the data or a machine. Conceptually, this premise responds to the theoretical proposition of investigating subjectivity as an operative element in the communication of the data to the viewer, and this is done by overtly exposing the images as “handmade” or digital images that simulate organic systems.

The underlying algorithms that compose the animations on screen share one basic function which is to read the dataset line by line and based on the datum, define how the images are drawn on screen at that particular instance. The flow of the animation is then defined by how one instance connects with the next in order to suggest a verisimilar percept of motion¹¹. This way of constructing an animation differs from classical forms of animated cinema where

¹¹ A verisimilar motion is understood in animation as something that is believable yet not necessarily “realistic”. A common use of this is to anthropomorphize an animal to denote human-like attributes in a character. For this, the motions given to the animal may not be “realistic”, but closer to the idiosyncrasies of how humans move and behave, turning the character believable as to having the emotions and unique personalities identified in humans.

the relation between each frame is positioned within a linear timeline. For instance, animating a walking cycle can be easily abstracted into a linear timeline because one has control as to how one drawing relates to the one before and the one after, facilitating the construction of a believable walk by having the anticipated understanding that the motion flow will be contained in a linear structure. With data, on the other hand, the relationship between each frame is not solely structured within a linear timeline but in other forms that emerge depending on how the variables are visually interpreted.

For example, consider the following dataset of geographical locations to be hypothetically translated into an animation of an object moving around the screen:

Table 1
Small Sample of Geographical Data

Latitude	Longitude
6.795	-73.052
5.258	-72.78
2.012	-53.11

The first step of the process is to translate these values into screen coordinates so the program knows where to position the object and then translate its view to the next location. The algorithm defines certain rules as to how that transition will happen: the speed, how to draw the in-betweens, what happens when it arrives at the new location, how much time to wait before it steps into the next row of the dataset, and any other aesthetic and formal decision that one would like to enact on this particular motion. Considering this example's dataset, with just three instances, analog drawings could

be created easily to move the object to these three specific positions on the screen. But datasets normally have hundreds or thousands of instances and producing the drawings by reading line by line to infer the structure of the data becomes a monumental task that usually defies normal conditions of production. One solution is to draw the images digitally with code¹², a method explored here as well, but analog drawings demanded an alternative form of production.

The method for approaching analog data drawings to turn animation production viable was to use modularity to blindly anticipate the structures that emerge in the datasets. This method was developed for early video games in which “sprite sheets”¹³ that contain all possible actions are contained in a single image, saved in the memory buffers of the system and then a scene can be rendered modularly from the sprite depending on the user's input. For example, if the player wants to move the character from left to right and send this signal through the device controller, the motion of the

¹² See conference presentation “*Stop Drawing Dead Fish*” (Bret, 2012) who argues that classical animation uses “dead” or static images to represent dynamic and complex human stories. Presenting a case study of a computer software that allows for real-time digital animation that renders the images dynamically and in response to live performance that changes on every presentation. Also, the book “*New Challenges for Data Design*” (Bihanic, 2015) which contains a series of essays and interviews with contemporary data artists and graphic designers who expose their working methodologies for creating digital data visualizations.

¹³ A sprite sheet or image atlas is a large image that contains all the frames of the animation in a single file. This method was used in early computers and game consoles due to memory limitations and today this limitation is still prevalent in animation produced for the internet. The sprite sheets solve two formal obstacles for animations with data: it allows for modularity so the structures under which the motions are rendered are blindly anticipated, and allows for faster loading times since all the images need to be sent to users’ every time a page is loaded and wants to render the animation. Since the animation is rendered in real-time based on a dataset and an algorithm that contains the rules of composition, all frames need to be loaded first on the user’s computer to allow for a fluid reproduction of the animation.

character in profile is contained in the sprite and rendered in real-time to accordingly respond to the players' command. A basic example of this was tested in the experiment contained in the project's website as "Tireless Feather"¹⁴. In it, the mouse position generates two data points that represent the coordinates on screen and based on these values, a particular image of a bird is rendered in the center as if it was looking at the mouse. The structure of the sprite sheet blindly anticipates any position of the mouse and in order to create the effect of the bird "looking" at that specific point, screen coordinates data is algorithmically translated into the position in the sprite sheet that contains the required image.

¹⁴ See experiment at: <http://www.dddrawings.com/lab/tireless-feathers/>

Figure 5. Sprite Sheet for animation experiment, from “Tireless Feather”, by González J., J. C., 2012.

This method allows for an optimization of the number of drawings required to render the scene in any of its iterations and in turn drawing the images by hand is deemed reasonably easy. But the particular data in this experiment is a purely digital one that has no particular meaning other than a reference to a user’s interaction.

Also, the blind anticipation is minimal because the mouse cannot jump abruptly from one coordinate to the next because the computer mice function as such. When using social or natural datasets (similar to the hypothetical example described in the sample data table above), the structures are not so linear, yet the method still applies for them as described in the next example.

In the experiment “Circle Variations” (Figure 5) which is also contained in the project’s website¹⁵, the dataset used is of seismic activity in Colombia and extracted from the official government institution *Ingeominas* that publishes them online on a daily basis. The algorithm that composes these images reads the dataset line by line and renders the scene where each instance is contained within a circular structure. The data is sorted in chronological order and the corresponding image is placed clockwise in relation to the date. Different sprite sheets were tested in order to identify different types of results and the images in them correspond to a visual abstraction of the Richter scale values. For example, in the pencil version, the sprite sheet has soft pencil strokes that gradually increment their intensity. When the image is rendered, a continuous line emerges with subtle variations of the stroke’s intensity that refer to the modulation of the tremors in the data. Changing the year in the application, yields a new variation of the line with another complex pattern that constructs the drawing. The application also allows for an animated version in which the process of reading the data line by line is paralleled with the frame rate so the process of construction can be viewed in motion.

¹⁵ See experiment at: <http://www.dddrawings.com/lab/circle-variations/>

Figure 6. Screenshot of the interactive application that modularly renders drawings based on seismic data of Ingeominas, from “Circle Variations - Pencil”, by González J., J. C., 2012.

Another variation of this interprets the Richter scale in different lengths of ink strokes (Figure 6), the same algorithm and data are applied to the new sprite sheet and a very different image emerges from it. Due to the complexity of the internal structures in the data, a very small variation in the underlying analog drawing can produce vastly different results and each highlights different aspects of the information. This iterative process of producing the sprite sheets, testing it with different datasets, adjusting the algorithms and source images, becomes viable due to the economy of images required to produce.

Figure 7. Data Driven Drawing using seismic data and a sprite sheet of ink strokes, from “Circle Variations - Ink”, by González J., J. C., 2012.

Data and Animation

The artistic propositions in fields such as data visualization and interactive art have expanded the epistemological relations between Art and Data, exposing new fields of artistic inquiry in which data provides a view to the “world out there” and the way of representing it is empowered by accessible digital technologies. This dissertation investigates the relationship between data and animation in light of these new digital technologies. It focuses on the conceptual problem of representing reality as an evocative experience, which is rooted in the traditions of animated documentary.

As such, the work presented here constitutes a further exploration of the traditions of animated documentary (Honesty Roe, 2013; Patrick, 2004; Sofian, 2005; Ward, 2005; Wells, 1997) but in this case the relation to reality is not invoked through the documentary

sound¹⁶ or narrative devices that reference reality¹⁷ but by the use of statistical data that reduces complex phenomena into structured numbers and text. Regardless of this formal difference, the central arguments contested in animated documentary theory remain the same here.

Rather than questioning the viability of knowledge-through-documentary, animated documentaries offer us an enhanced perspective on reality by presenting the world in a breadth and depth that live action alone cannot. (Honesty Roe, 2011, p. 229)

The point of departure of Data Driven Drawings from the traditions of animated documentary is in the type of index that refers back to reality. In traditional animated documentary data as an index is not considered into the academic discussion due to the lack of examples and artworks¹⁸, to develop a relationship with the current tradition. Having said this, all of the academic research around animated documentary develops a mode of understanding subjectivity and internal experiences of the authors as valid positions from which to articulate a view of the world out there. This validation of subjectivity and the types of representation that animated documentary affords is conceptually the same as the one pursued artistically here, but the technical and practical methods of

¹⁶ For instance, the work based on interviews is arguably the most widely used approaches to animated documentary today. See Sheila Sofia, Richard Linklater, Michel Gondry, Chris Landreth.

¹⁷ A more permissive understanding of documenting reality in animation is to accept that what is being portrayed exist in real life, this allows for films such as "Waltz with Bashir", "Persepolis" and even more experimental approaches like "The Death of Stalinism in Bohemia", to be regarded as documentary even if fiction permeates and deliberately complicates the representation of such events.

¹⁸ One brief exception is the work of the British Duo Semiconductor who work with different natural phenomena data to create animated films. Their work is discussed in an excellent article by Lilly Husbunds "The meta-physics of data: Philosophical science in Semiconductor's animated videos" (2013)

data driven drawings provide a new type of index to animated documentary not articulated before.

Personal Motivation

The union of data representation and hand-drawn animation stemmed from a recurrent inquiry in my own animation artwork. Over the past decade, my central goal has been to investigate the ways in which I can represent the complex nature of my experience after the death of my father. This momentous event in my personal life has deeply affected the way in which I see violence and the effects it has on the human experience.

In the past, I have explored similar themes in the creation of short animated films, but these made no direct reference to the realities that motivated them or made use of data as a formal component. The primary reason for investigating the use of data in combination with animation is that now I seek to represent complex experiences that indicate unbound states of mind in constant states of flux, which in turn I can leverage on the potential of data to abstract and signify phenomena with these characteristics. These internal experiences are personal memories that resist conforming to a single “image” in my mind and relate to the context of violence that led to my father’s violent death. In a series of three Data Driven Drawings projects, I explore different possibilities for embodying the mutating memories I have of my father as activated when confronted with the context of violence in Colombia indicated in the data.

It has always been very clear to me that my work is not aimed to represent my own biographical story and the events that have taken

place in it. Yet, a series of events that were triggered by the context of violence in Colombia and affected my personal life, have created a kind of ubiquitous lens through which I observe and infer meaning from the realities of my life. My artwork is about that view of the world and the relationship I establish with the viewer is an invitation to experience through my own point of view, it is not about what is it like to be me as in “Juan Camilo González Jiménez”. As such, the work that I do develops from an autobiographical reflection of events that shape who I am and the way I experience my surroundings, but the intention is to communicate to others something that emerges only after reflecting on my autobiography. It serves the purpose of embodying in an artwork the emotional experience and subjective inferences that have remained with me after experiencing violence and death in my personal life.

The subject matter of the practical work I present here is something that I have no rational clarity of, and in it I make an effort to create a complex system accepting that what emerges from it is something that I don't know and would like to experience myself, very much as if I am the audience of the work I make. The starting point for me was a series of personal observations that summarize as the following: I hate and love my father, I am sad and relieved that he is dead, he has passed away but remains present in my life, I feel emotionally handicapped due to the way he was with me, but I am satisfied with the person I have become despite or due to his influence on me. I see this as a complex system in itself and the parts that constitute the system are seemingly contradictory. The goal I have established for myself in the practical work is to portray these kinds of contradictory events as if they are naturally merged together. The personal experience I ought to communicate to others exists in defying the contradiction of opposing feelings such as love

and hate, violence and life as if they are singular entities in which these conflicts coexist and manifest as naturally blended together.

The decision to not express my own biographical story, but at the same time to allow for my personal experiences to guide the point of view from which I represent “reality”, has implications for the two modes of investigation developed in this research.

Firstly, it outlines the conceptual approach to representing data in the form of hand-drawn animation that is developed in the practical work. The data, in this case, is the indexical reference to “reality” in Colombia, and the animation is created with an attitude in which I allow for my own intuition to guide the way in which the information is finally represented. The goal with this approach is to evoke in the audience a sense of what is it like to be human in violent places like Colombia and not intended as the presentation of “evidence”¹⁹ of the information itself.

This particular framing allows for the representation of reality from a defined and subjective point of view. The audience is provided with clues to the fact that the artist is making the decisions as to how the information is communicated to them. This is the case due to the formal characteristics of the representation: the data serves as an index of reality, while at the same time a human is making an

¹⁹ Clarification on the differences between representations that are intended to evoke based on reality, as opposed to provide evidence of that reality, will be further discussed on Chapter 2.

interpretation of it through the act of drawing²⁰. I have decided to use handmade drawings as a case study for creating a conflict between the mathematical realm of quantitative data and digital software, and the instinctive nature of my own drawings. Since the goal of this research is to make a case for evocative approaches to representing reality, this conflicting method has provided a useful scenario for testing this hypothesis (Chapters 4).

Before moving forward, it is important to clarify that the ideas that motivated my personal work may or may not apply to other artists and have been just mentioned in order to provide the reader a clear context of my intentions as an artist. But, the theoretical contribution of this dissertation is not the practical work²¹ in itself,

²⁰ The opposite formal approach to this would be to allow for computer-generated images to represent the information. But, computer-generated or hand drawn images can ultimately serve the same purpose for an artist to communicate his/her subjective point of view. The difference between them is initially naïve, in the sense that audiences might think that computer-generated images can be more “truthful” to the data since they both operate within the same mathematical realm. In other words, computer-generated data representations might give the impression that the artist is acting at a lesser extent on how the data informs or “drives” the representation. This idea will be challenged on Chapter 2.1 in order to focus on the artist intent and not on the form characteristics of the representation. The goal of this research is to make a case for evocative approaches for representing reality and this is regardless of the tools and techniques that the artist decides to employ.

²¹ The work I have done in the past using animation where already engaging with the idea of evoking the experience of complex relationships and identity. These subject matters have been successfully approached by other artists in cinema as well: Aleksander Sokurov cycle “Mother and Son” (1997) and “Father and Son” (2003), Thomas Vinterberg “Festen” (1998), and in the case of animation, Jan Svankmajer “Dimensions of Dialogue” (1982), Yuri Norstein “The Heron and Crane” (1974) John Canemaker “The Moon and the Son” (2005), to name a few. As an artistic proposition, the theme is not new, what is different and constitutes the theoretical contribution, is the formal approach of using digital technologies that facilitate the exploration of new expressive potentials to old conceptual ideas. The dissertation will focus on a novel approach to digital animation. The contribution of the thesis is the arguments describing this approach to animation, the contextualization of this work within the tradition of cinema and the formal proposition of a workflow that makes it possible for other artist to investigate the expressive potential of combining data and animation.

but a conceptual framework developed throughout the process of creating these artistic works that combine hand-drawn animation with data representation.

Precedent: From the stream of consciousness to stream of data

In the past, I've investigated the problem of representing the complex nature of the relationship with my father through animated short films. For instance, in *Mecanismo Olvidador* "Forgetter Mechanism" (2009)²² I developed the film using as a premise the image of a pendulum that depicts the oscillation of a man's state of mind inspired by Schopenhauer's proposition that "life is like a pendulum".

Willing and striving are its whole essence, and can be fully compared to an unquenchable thirst. The basis of all willing, however, is need, lack, and hence pain, and by its very nature and origin it is therefore destined to pain. If, on the other hand, it lacks objects of willing, because it is at once deprived of them again by too easy a satisfaction, a fearful emptiness and boredom come over it; in other words, its being and its existence itself become an intolerable burden for it. Hence its life swings like a pendulum to and fro between pain and boredom, and these two are in fact its ultimate constituent. (Schopenhauer, 1969, p. 312)

I related this metaphor of the pendulum to my personal oscillating feelings towards my father. Transitions in a state of mind moving from love to hate, from repressing difficult memories to confronting them, feeling a sense of "emptiness and boredom" as proposed by Schopenhauer when pushing difficult memories away and then swinging back to pain when confronting those memories. In the

²² Available online at: <http://juancgonzalez.com/films/mecanismo-olvidador/>

film, I wanted to portray this vertiginous feeling by placing a man inside a pendulum. When I was making the film, I established this conceptual idea as the premise, and no script or storyboard was produced beforehand. The animation was made by allowing a stream of consciousness to take place. In short, the subject matter was established beforehand and then directly transferred to production without “knowing” clearly what will be the end result.

There are some elements that transition seamlessly from this approach to Data Driven Drawings as developed here. First, a clear conceptual theme is established and then allow the animation to manifest despite my rational self. What changes in Data Driven Drawings is that instead of a personal stream of consciousness to be in command of the final result, this is replaced this with a “stream of data” that refers to the world outside of the self, I still have control over the images that get rendered: I draw and animate them in the traditional hand-drawn technique, but once these animations are placed in the system that renders them²³, absolute control is lots over the temporal structure in which they ultimately get rendered on the screen.

By shifting from “stream of consciousness” to “stream of data”, the traditional techniques that allow us to make an animated film turn out to be at stake. The link between one frame and the next are not fixed into the timeline anymore but become interchangeable and no longer predefined. At the same time, the final animation is no longer contained into one single object (film stock or digital file)

²³ See Chapter 6 for more details. The real-time rendering is happening on an Internet browser and the rendering engine is written in JavaScript. The user is able to select the data to be rendered and the system feeds the drawings into the rendering screen as it goes line by line through the events present in the data.

that can be viewed over and over again without it “changing”, but the result is a mutable object that varies depending on the data that is fed into the system at any given time. In principle, data driven animation is the investigation of this shift from a fixed object into an ever-mutable one.

As referred to earlier in this chapter, the idea proposed by Tor Nørretranders can be useful again to elaborate on the stream of data issue. In Nørretranders terms, the stream of consciousness process can be understood as the suspension of the *I* to be in full control of how ideas are being articulated, the consequence is often non-conventional or modes of communication that do not correspond to a consensus of how we communicate with one and other. The stream of consciousness is giving “voice” to the *Me* since it is still me who is trying to communicate something that *I* am not aware of. A simple process of delegation is taking place here and is required for interrupting the *I* to be in full control. In data driven drawings the process has a variation and the “voice” is delegated to the world outside myself, how can it then relate to my experience if it is not coming from within? I would argue it doesn’t have to as long as the data that indicates the world out there passes through our own personal experience of it. In other words, an artistic articulation of data is a process of designing a system in which the personal experience of an event can reproduce our observation of it.

Let’s consider as an example a piece by Colombian artist Doris Salcedo called “Noviembre 6 y 7” from 2002. This site-specific artwork was presented in the 17th anniversary of the siege by the then guerilla group M-19 to the justice palace in the center of Bogotá. The group entered the building and took hostage the over 200 people including members of the supreme court, causing an extreme reaction from the government of then-president Belisario

Betancur. The violent reaction of the military forces provoked the death of 98 people, including 11 supreme court justices and 8 people disappeared in still today unresolved circumstances. The event remains one of the most iconic cases of brutality and abuse of power by the government in Colombia. The way in which Salcedo configures her work is, in my opinion, an interesting use of data representation in contemporary arts. Salcedo collects data surrounding the event as her starting point, the duration of the event was in total 27 hours on November 6 and 7 of 1985. During this period of time, she is able to trace back the times in which people were killed. Her work takes place in the façade of the justice palace where the massacre took place and chooses the exact same dates of the year to produce the work. Following the dataset collected before, a single wooden chair descends from the rooftop in the exact same time a person was murdered. Each chair remains hanging outside the building for the duration of the piece and slowly, for 27 hours, the justice palace starts to fill with these floating chairs.

Even though this work is not often associated with data representation, I argue that the procedure is exactly what artistic data representation is capable of. Once she has the data and being faithful to it, she activated the highly deliberate system under which the stream of data is turned visible. She does not decide when a chair should descend, the data indicates it. Salcedo designs a system for the stream of data to go through, not to inform us the times in which people died, but to force an experience on the audience that relates to that real event that took place there 17 years before. She could have decided arbitrarily a rhythm for the chairs to descend and create the same image since the objects hanging from this particular building can already carry the history of

the tragedy, but her decision to delegate to the data part of the procedure is what makes it so singular. She removes her / without removing her emotional experience and opinion over the event. This particular procedure of designing a system of artistic expression for a stream of data to operate on is the basic mechanics of artistic data representation as discussed in the next chapter.

Chapter 1: Data Representation

Obviously, there is no classification of the universe that is not arbitrary and conjectural. The reason is very simple: we do not know what the universe is [...] But the impossibility of penetrating the divine scheme of the universe cannot dissuade us from outlining human schemes, even though we are aware that they are provisional. (Borges, 1964, p. 104)

Data in itself does not impose a medium or form of representation, it is an index of a particular event in the world that has been captured in a structured set of quantitative values, waiting to be given meaning through its representation. Daniel Rosenberg develops an important etymological study of the word “data” in his introductory chapter to *Raw data is an Oxymoron* (Gitelman, 2013) aptly titled “Data Before Fact” (Gitelman, 2013, pp. 15 - 40). His central argument proposes that we understand the term data as “specifically rhetorical” (2013, p. 18) in the context of its historical use which usually locates the word in discussions of “knowledge and argumentation” (2013, p. 15). For Rosenberg, data is rhetorical because it is there in support of the argument, but in itself does not contain the facts or evidence.

A “datum” in English, then, is something given in an argument, something taken for granted. This is in contrast to “fact,” which derives from the neuter past participle of the Latin verb *facere*, to do, whence we have the English word “fact,” for that which was done, occurred, or exists. The etymology of “data” also contrasts with that of “evidence,” from the Latin verb *videre*, to see. There are important distinctions here: facts are ontological, evidence is epistemological, data is rhetorical. A datum may also be a fact, just as a fact may be evidence. But, from its first vernacular formulation, the existence of a datum has been independent of any consideration of corresponding ontological truth. When a fact is proven false, it ceases to be a fact. False data is data nonetheless. (Gitelman, 2013, p. 18)

Rosenberg’s clear distinction between data, facts, and evidence is critical to understanding data representation as the construction of

an argument where data in itself serves as a rhetorical agent in support of the argument, data is not the argument but exists prior to its articulation.

When a singular datum is presented to us without any context, it has no resemblance or connection to its source, this is a particular benefit and constraint of the process that governs any use of data. If we compare the process of capturing data with that of a photographic image, both operate under the same premise of abstracting the sensible impulses of phenomena “out there” and the technology used to capture these impulses is designed to contain it in a form that allows us to revise it later. Both are ways of capturing and both intrinsically create a reference to the phenomena being contained. The camera that produces a photograph is designed as an extension of the eye, it may construct its image by mimicking our own perception or even enhance it through the use of specialized lenses that allow the camera to “see” far away, deep close, expand our normal range of view, but in all of these cases it is producing an image of what is in front of the lens. As such, the photographic image is inescapably a reproduction of the phenomenon being captured, or what André Bazin describes in his essay *The Ontology of the Photographic Image* as the image being “the model”:

The photographic image is the object itself, the object freed from the conditions of time and space that govern it. No matter how fuzzy, distorted, or discolored, no matter how lacking in documentary value the image may be, it shares, by virtue of the very process of its becoming, the being of the model of which it is the reproduction; it is the model. (Bazin, 2005, p. 14)

By the photographic image is the model, Bazin is making a distinction between what the image may be able to signify and the fundamental process of it being an index of what was in front of the camera. Highlighting that the resulting object already carries that

reference because the primary result cannot escape it being a reproduction. Borrowing Rosenberg's ideas, the primary result of the photographic process is also rhetorical. The light that travels from the subject to the camera, imprints a trace of what it was at that instance in the sensible support of the camera (the emulsion in analog photography and the digital sensor in modern devices), this light or "data" being captured is rhetorical, precedes the creation of an image and is dependent of the technology used. Even if the result is, as Bazin puts it "fuzzy, distorted, or discolored". The way in which the camera retains the information of what is being captured is an image that reproduces this preceding technical event of light traveling from the subject to the device. This reproduction is initially also rhetorical, what Bazin describes as it being "the object itself...freed from the conditions of time and space" (2005, p. 14) or paraphrasing Rosenberg, a false photograph is a photograph nonetheless. The main difference between the photographic process and data representation is that the primary result of the capture in photography is an image and not the unconstructed data of the light that defines the subject. As such, the end result of photography, with its ontological and epistemological constructs,

can be this same image²⁴. In data representation, the process forces a form of representation of the primary source, the data. The photograph has a resemblance to its source, but data is an object in itself with no such properties until given meaning through its representation.

Luciano Floridi proposes a unifying principle in his treaty on information philosophy summarized as “no information without data representation” (Floridi, 2011) in which information is understood as the sum of data with meaning. For Floridi, data itself is a transient object which can support an argument in a particular context and equally move on to be interpreted differently as new questions are posed to it, he defines the datum as:

x being different from y, where the x and they are two uninterpreted variables and the domain is left open to further interpretation. (Floridi, 2011, p. 85)

²⁴ For further elaboration of this idea, read Rob Kitchin’s proposal that differentiates quantitative and qualitative data in chapters 1 and 6 of “The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences” (Kitchin, 2014). Kitchin argues that photographs, among other objects operating as qualitative data, are currently being transformed back into data for new forms of analysis, but this process has certain limitations due to its reduction of the information of the original source: “In contrast, qualitative data are non-numeric, such as texts, pictures, art, video, sounds, and music. While qualitative data can be converted into quantitative data, the translation involves significant reduction and abstraction and much of the richness of the original data is lost by such a process. Consequently, qualitative data analysis is generally practiced on the original materials, seeking to tease out and build up meaning and understanding rather than subjecting the data to rote, computational techniques. However, significant progress is being made with respect to processing and analyzing qualitative data computationally through techniques such as machine learning and data mining” (2014, p. 5) Such efforts can be seen in the work developed in the “Cultural Analytics Lab” led by Lev Manovich, which has developed a series of projects using social media pictures that are treated quantitatively and qualitative to produce interactive interfaces that allow for new modes of studying big social phenomena through qualitative objects. Notably “Selfiecity” (2014-2015) that studies patterns of self-representation manifested in the practice of the “selfie” and “The Exceptional and the Everyday: 144 Hours in Kyiv” (2015) that uses digital visualization techniques to analyze the reactions of regular people during the social upheaval of 2014 in Kiev.

The importance of Floridi's argument exposes a powerful virtue of data not forcing a specific definition of the phenomenon where it comes from, data may serve as evidence of its existence and structure, but its absence of uniformity allows for multiple articulations that provide us information of that phenomenon. When the research objective is to better understand a phenomenon, this transient property of data is useful for replicating experiments, update the statistical models to further corroborate or disprove previous findings, or pose new questions previously not explored with the data²⁵. But of relevance to the present study, this flexibility is particularly important for data art practices because data not only precedes argument, but its articulation is not strictly attached to the domain in which it was originally captured. All of the practical projects created and discussed in this research elaborate precisely on Floridi's proposal that "the domain is left open to further interpretation." Not only to further interpretation, but ones that do not necessarily serve the purpose of understanding the phenomenon for what it is, but as metaphorical supports to new associations.

In these terms, the representation in "data representation" builds the epistemological and ontological constructs around the data and is when the particular ways of understanding the data are given to us, what Floridi proposes as information "data + meaning". A lucid demonstration of how data without context is voided of meaning can be seen in the video works "Opus" (2005) by Cuban artist José Luis Toirac and "5'116.137" (2010) by Colombian artist Andrés

²⁵ A detailed study of the transient nature of data within the scientific community and further explanations of Floridi's contribution can be found in the excellent essay by Sabina Leonelli "The Philosophy of Data" chapter 17 of the book edited by Luciano Floridi "The Routledge handbook of philosophy of information" (Floridi, 2016)

Felipe Uribe Cárdenas. Both artists use the same strategy, select archival recordings of the presidents Fidel Castro and Alvaro Uribe Vélez respectively, and meticulously cut away everything they say in the speech except for numerical data. The results are a stream of dates, percentages, and singular numbers that could be referring to anything. Once the context of the number is removed, both artist acutely highlights the rhetorical nature of the datum, and in consequence, expose how its naturalized use in political speeches should not amount for evidence or facts on their own. Toirac and Uribe understand data in the same way that Rosenberg proposes theoretically:

Data means — and has meant for a very long time — that which is given prior to argument. As a consequence, the meaning of data must always shift with argumentative strategy and context — and with the history of both. (Gitelman, 2013, p. 36)

Toirac's and Uribe's artworks problematize any predisposition to equate the term data to evidence or facts, by removing precisely what Rosenberg refers to as "argumentative strategy and context". This context is projected back to these numbers from what we may know of these political figures, their history, their past behaviors, but this context is subjective as it only takes form based on our bias and previous information of the politicians. The numbers they refer to produce no information because their meaning has been taken away and in consequence are given to us in that transient nature they carry in the raw state. What might have been a percentage of unemployment can be interpreted as something else and the frustration and ambivalence of data that produces no information is what the artists exploit for an argument that relates to political rhetoric and not the original phenomenon the datum refers to. The discussion that follows understands the term representation as the place where those complex strategies and context are given to the

data, turning particularly contentious in the type of questions that art practices pose to data.

The particular process of representing data establishes a unique relationship between the object or phenomenon captured and its embodiment in an image, sound, interactive application, animation and potentially a rich array of possible forms yet being explored by artists and industries alike. The distinctive nature of the process is that the instruments and methods used to capture the information do not necessarily carry through an “opinion” as to what the data signifies and how it should be represented. This idea applies only in the cases where the instrument is designed to capture what will be referred to here as raw data and not when the instrument is designed as a representational tool in parallel to capturing the phenomenon. This distinction is critical for identifying the specific role of representation in its own right and why this thesis draws attention primarily to representation in the overall complex process of working with data.

An illuminating case for understanding this was discussed in the previous chapter in reference to the seismograph and the work of the artist Tim Knowles. In both the seismograph and Knowles’ drawing trees, the instrument for capturing the data is the same as the one that represents it, therefore collapsing the distance between these two distinct processes. The result is a deliberate impediment for the raw data to be captured and saved in a state voided of signification, very similar to the previous discussion of the photographic image. For example, the drawings produced by the seismograph contain in themselves the consensus of the scientific community. Their opinion as to how this phenomenon operates, its structure and how it should be represented for clarity in its study is designed into the instrument that captures/represents. In Knowles’

trees, the same principle applies but for very different reasons, he is not working under consensus of opinion but proposing a personal one. Yet the instrument (a tree with drawing tips attached to its branches drawing on canvas) is designed to reveal his personal view in the context of an artistic exercise of drawing what cannot be seen. In both of these cases, the representation is in charge of providing the path to how we are able to read the data, extract meaning from it and engage in the intellectual and emotional process of its significance. This is particularly clear in Knowles' work for two reasons that empower the overarching affordances of artistic data representation: firstly, by revealing the unconventional device that produces the work, the piece complicates any predisposition or assurance that the data operating in the work contains the "truth", yet by delegating the drawing to the nature of the phenomenon and the constructed device, the result suggests that the representation beholds a "truth". Secondly, the ambiguity of the resulting images refers back to the process of capturing the data and use the phenomenon symbolically and not literally. The representation that Knowles' articulates is not a direct path to understand wind patterns over a tree, but a metaphor to question authorship, traditional understanding of drawing from nature and challenge the dominance of the eye as the primary source for accessing information of nature so it can be transformed into a drawing on a piece of paper. Yet another set of issues and affordances arise when the instruments that capture the data are not intended to represent, but exclusively to collect raw data.

The term "raw data" is that which is captured directly in a way that it can later be measured and compared such as sensors, surveys, forms, lab and field experiments or any other method that produces measurements. But the term should be treated with caution during any interpretation of it as it is always limited by the potential

failures of the capturing method. A clear example of this was encountered in the interpretation of several datasets used during this investigation. For example, the seismic data produced by the Colombian institution *Ingeominas* has collected measurements in their databases dating back to 1993. A first glance, the amount of seismic activity shows a dramatic increase in the year 2002, and upon further inspection of the data tables, a difference in the number of stations was noted. In their reporting²⁶, *Ingeominas* notes that they updated their measurement systems and increased the number of stations throughout the country in 2002. Which explains the abnormal increase in the data. Even in its raw state, the data may be misleading if the interpreter is not wary of the methods used for collecting the data, which can lead to potentially distorted assumptions. A more critical case was found in the databases from the “United Nations Office for the Coordination of Humanitarian Affairs (OCHA)” also known as *Monitor*. This is one of the most exhaustive datasets found that report the general categories of violence in Colombia. Being independent of the Colombian government, it further provides a sense of confidence that the data collected is unbiased of any specific political agenda. Yet, the cases of sexual abuse are barely present. A total of 10 cases were reported in 2008, 12 in 2009 and 2010, 9 in 2011 and the highest was 49 in 2016²⁷. In this particular case, the suspicious absence of data is not a limitation of the methods and sources used to collect the data, but reflective of a more complex social issue that prevents these cases to be reported. The numbers in themselves cannot be

²⁶ The report is published online by Ingeominas as “Administración de equipos operacionales de apoyo de medición” Retrieved July 31, 2018 from <https://www2.sgc.gov.co/Publicaciones/Paginas/estandares-geocientificos.aspx>

²⁷ These values can be consulted in the open data distributed by Monitor. Retrieved August 12, 2018 from <https://monitor.umaic.org/>

treated as reflective of the phenomenon, but the absence of numbers is indicative of the type of questions that can be posted in this section of data. This latter case is indicative of some of the issues pertaining to the indexicality of data. As it has been discussed so far, by virtue of the process, data is able to remain in suspense of its potential interpretation, but its indexical relation to the phenomenon is not an unbiased exposure of it due to the methods and context that should be taken into consideration as to how the data was captured. The idea of the index in data is slightly problematic in these terms and demands some clarification.

The theoretical discussion of the index is rooted in the comprehensive studies of Charles Sanders Peirce and the signifying elements in the icon, symbol and the index (Peirce, 1932). These signifiers are in the icon a “likeness” to the object it denotes, it possesses in its form a reference to its object by similarity, the happy face is an icon as it resembles a human face by the abstracted eyes and mouth. In the symbol, we rely on general conventions that require interpretation, a word in a particular language is a symbol as it needs the person to know the language to interpret it, even if the same word in a different language bears no meaning to that same person. Lastly, the index has a tacit connection to the object and this direct relation points to its existence. Peirce uses “instantaneous photographs” as an illustrative example of the differences between the icon and the index. For Peirce, a photograph may seem closer to the icon as it is “like” the subject being captured, but explains that in its nature, the photograph has just been made to correspond exactly to what was in front of the camera, therefore the image is produced by a physical connection and refers back to the existence of that being reproduced, making it an index. Rosalind Krauss contextualizes Peirce’s ideas of the index in relation to 1970’s abstract art in her

essays “Notes on the index” 1 and 2 (Krauss, 1986). Her central argument is that the influence of photography, and particularly the semiotic definition of the index proposed by Peirce, can be identified in what is otherwise a very disperse set of styles and forms of artistic production that proliferated during the '70s.

If we could say of several generations of painters in the late 19th and early 20th centuries that the conscious aspiration for their work was that it attain to the condition of music, we have now to deal with an utterly different claim. As paradoxical as it might seem, photography has increasingly become the operative model for abstraction. (Krauss, 1986, p. 58)

She begins the discussion with the analysis of artworks that paved the way to thinking about the index in art with a lengthy analysis of Marcel Duchamp's work. For example, in the painting “Tu M” (1918) a series of distorted shadows of his ready-mades are painted into the canvas as referring to the physical objects which are not present in the painting, the shadow as a trace left by the object is descriptive of the index. At the same time, a pointing hand at the center of the canvas is indicative of a second interpretation of the index as that which points to the existence of something else. The elements portrayed in the painting are referring to something not present in it and this strategy is highlighted by Krauss as an important precedent of an artist being influenced and establishing connections between painting, the index, and photography. In the second part of her essay, Krauss articulates the influence of the index in abstract art of the '70s with a series of case studies in which the artworks serve as traces that indicate a particular phenomenon, breaking in the process the normal codes that govern their mediums. She begins with the description of Deborah Hay's piece in which she stands in front of the audience and instead of the expected dance performance, she stays still and describes in an hour-long monologue how the movements of the cells in her body

are the ones performing the dance in that instant. Later continues with a series of works that were part of the 1976 exhibition “Rooms” that took place at a public-school building in “P. S. 1” Long Island City neighborhood in New York. “Doors, Floors, Doors” by Gordon Matta-Clark, “P.S.1 Paint” by Lucio Pozzi are notably transforming the space in which the pieces are located to refer back to the history of the building in which they are installed. Matta-Clarks piece cuts a hole in the floor of three consecutive stories for which the size and form are similar to the door adjacent to it as if it was the shadow of the door. This hole allows us to see the structure of the building and room underneath. Pozzi’s paintings are integrated into the walls of the school, she finds specific walls that are painted in two colors and places a series of plates in the middle of where the colors border each other. She then paints these plates in the same colors of the wall and carefully aligns them with the geometric pattern that already existed in them. Pozzi’s paintings and Matta-Clarks intervention are blending into the building while at the same time pointing back to it.

Krauss is carefully stating that not only the influence of photography and the index are present in these works, but this influence forces the normal codes that govern painting, sculpture, and dance to be transformed. The index is operative in them because they are all pointing to the existence of something outside them but having a physical correspondence to the object they refer to. In Hay’s she is refereeing to the molecules in her body and contextualizing them as dance, Matta-Clarks and Pozzi are pointing to the building in which the objects are physically located and contextualizing them as art. Each of these pieces is reshaping themselves to be able to correspond physically to the phenomenon they are referring to, and this is a seminal precedent for data art practices and what is proposed here as data driven drawings.

Within the conventions that have been created for representing data²⁸, and particularly visual articulations of it, the intent of what kind of information they should produce is critical to the formulation of these conventions. For a scientific community, the consensus is central to the set objectives of building knowledge as a collaborative exercise where ideas are tested by replicating the results of one and other. So, the data used in an experiment not only follows statistical models that have proven useful but also the visual representations of the data that serves as evidence of the results often is created under agreed procedures that allow for comparison between them. This leads to more homogenous forms of data representations that correspond to the community's objectives. On the other hand, in artistic practices, the use of data is not necessarily a given, with accessibility to open data and digital technologies it is now easier to consider it, but the mediums have not developed in consideration that data could be represented in them. Drawing on paper is a natural response to how we normally engage with it which is to register with lines while seeing or imagining something. But when data is invoked into the process, the formal characteristics as to how that process is going to physically respond to the measurements of data forces the drawing on the paper process to be revised and transformed. This is what Krauss was pointing out as happening for abstract art of the '70s and in today's proliferation of data art, it is happening again as old and "new" forms of artistic creation figure out ways to allow data as an index to enter the production.

²⁸ Further examination of this is explored in Chapters 2 and 3 in relation to the ideas of Jaques Bertin and Edward Tufte.

Artistic data representation

The following discussion focuses on the artistic representation of data as a way of re-presenting the “world out there” through metaphors and analogies that indicate a personal view of the information abstracted in the data. More precisely, artistic data representation proposes a contentious relationship with the object contained in the data to signify not the object itself, but establish symbolic associations with that object. The goal is not to make evident the phenomena that the data signifies, but rather to generate new associations between events of the world we live in and our human condition. Nor is it intended to answer a particular question, but rather either to new issues or remain in a state of ambivalence towards the subject matter.

This position towards representing data is not new within artistic practices (Husbands, 2013; Lupi et al., 2016; Manovich, 2010; Viégas & Wattenberg, 2007). The artistic representation of data as cultural artifacts will often be presented in venues such as galleries and museums that define the type of relationship between the piece and those who experience it. Usually, the audience is invited to create a subjective dialogue with the piece, construct associations between the metaphors presented in the work, our past experiences and acquired knowledge. This relationship differs from the one often associated with data representation as a functional artifact in which the goal is to aid in the process of acquiring concrete knowledge of the phenomena embedded in the data. This functional approach serves very specific purposes when the goal is to understand the phenomena unambiguously in order to communicate its behavior to the user. Despite this, the differences between data representation as a cultural or functional artifact are persistently vague since the boundaries that separate them are

easily crossed. A data piece exhibited in a museum does not necessarily remove its functional capacities to inform us clearly about a phenomenon, nor a data representation produced outside of the art world impedes its revision as a cultural artifact. This is the case because data in itself, as argued by Daniel Rosenberg, operates in the representation as a rhetorical constituent to develop an argument, therefore, the representation of data does not impose a strategy on how this index is presented to the audience by itself.

Having said this, it is only recently that curators of art exhibitions have drawn attention to the potential use of data outside its traditional fields, contributing in this way to an important shift and awareness of what can be considered data representation. The outset of art exhibitions around data practices is conceptually critical because the art space, gallery or museum, is in these cases an important context that forces the audience to revise the relationship with data representation. In this particular setting, the rhetorical nature of data turns “visible” as we are invited to examine these works in similar ways that we normally do with other cultural objects that inhabit these spaces. We are intrinsically told that there was an artist behind the creation of the piece and that his or her intentions are there to be examined in relation to what we experience in the piece. Utilitarian functions of these works are less important in these spaces since the audience is not expected to be expert in decoding the codes under which the phenomena represented can be understood functionally. In this sense, the art space equilibrates form and function for us to establish the complex process of understanding through aesthetic and logical routes.

A good example of this was the major itinerary exhibition called “Big Bang Data”²⁹ (2015) curated by Olga Subirós and José Luis de Vicente for the *Somerset House Trust* and the *Centre de Cultura Contemporània de Barcelona (CCCB)*. The exhibition fulfilled two important roles in the process of widening the understanding of data representation for general audiences. It contextualized early examples of data visualization as milestones that anticipated contemporary data art practices, and the carefully selected works that range from those in which the authors are more concerned with clearly communicating the information found in the data³⁰ and those who test the boundaries of abstraction where the information in the data is purposely obscured in order to give the audience access to other ways of experiencing and understanding the data³¹.

One of the most interesting curatorial decisions made by Subirós and Vicente was the inclusion of 3 historical milestones in the exhibition. Not only because they highlight the fact that data representation is a practice with a long history, but more importantly, I would argue, the pieces they selected were testing the

²⁹ The exhibition has a permanent website where the catalogue can be consulted and also includes ongoing events that expand from the exhibition in different venues around the world: <http://bigbangdata.somersethouse.org.uk/>

³⁰ See for example “Selfiecity London” (2015) by Lev Manovich and Moritz Stefaner which is included in the exhibition. The piece is a good example of functional data visualization applied within an overarching goal of an artistic project, a piece that sits in the middle of both worlds without any major compromise. The style of the visualization is classical and within traditional functional information graphics, but the way it is exhibited as an installation and web project explores modes of engagement with the audience that surpass the just informative nature of the visualizations. It can be seen online at: <http://selfiecity.net/london/>

³¹ See “data.tron” (2007) by Japanese artists Ryoji Ikeda who embraces the exhibition space both physically and conceptually to investigate the vastness and complexity of data in the articulation of immersive experiences. His is concerned more with the experience of data in our daily life rather than what the data means or is able to inform. His large projections wash over the spectators that is forced to surrender to the vastness and intense velocity of large amounts of data. Excellent documentation of this piece can be seen in the artist website: http://www.ryojiikeda.com/project/datamatics/#datatron_wuxga_version

boundaries of an emerging practice which is precisely the position in which data art stands to move forward. These historical pieces where the William Alfred's "Stowage of the British slave ship Brookes under the regulated slave trade act of 1788" (1788), John Snow's map of the cholera epidemic in London (1854)³² and the diagrams included in Florence Nightingale's publication "Notes on Matters Affecting the Health, Efficiency and Hospital Administration of British Army"³³ (1858). The inclusion of these pieces in parallel to more contemporary artists and locating them inside the museum is an invitation to displace their original intentions as functional images, in order to revise them as images and cultural objects in their own right.

The case of the Brookes slave ship diagram is a clear example of an image denouncing the conditions in which slaves were transported in the ships. The speculative decisions made by the author such as drawing crawling and hunchbacked figures is as important in the rendering as the careful diagrammatic elements that respect scales and density of the subject matter in relation to the data used to produce the representation. Outside the art space, this image might seem problematic for its bias and author's intrusiveness in what the data has to say on its own, yet the image inside the museum and art

³² See in the Introduction, section "Data Drawings" for a discussion of these two early examples of data visualization.

³³ Florence Nightingale's diagrams are discussed in more detail at the end of the current chapter, yet her work is argued here as one of the first approaches to artistic data visualization due to her decisions to compromise clarity and readability in exchange of emotional impact through aesthetics. Even if she never intended for her diagrams to be exhibited as art, revising her work in light of contemporary data art practices reveals her progressive views on data representation. She understood that part of the knowledge that a diagram carries is contained in its all-encompassing aesthetic construct, not only in its ability to inform directly the information. Therefore, her approach to data representation anticipates the procedures that data artist repeat today which is to negotiate between what the piece is able to communicate aesthetically and what should be exposed unequivocally and without obstructing what the data has to say.

context accepts the authors intent as a given. If Picasso's painting of "Guernica" (1937) was challenged for its unrealistic depiction of those who suffered the bombing³⁴, any discussion on the importance of the painting as an anti-war posture would be obstructed. With the caution of invoking one of the most iconic and widely written about artworks, the point of comparison is to illustrate the distance between what is expected from an image that as a cultural artifact takes a political stance and an image such as the Brookes diagram that normally exists outside this particular context and is created to inform rather than evoke. The procedures under which Picasso refers to the bombing are exposed clearly, the title of the painting is the name of the town where the event took place, the painting was produced two months after the bombing and what is depicted inside is a series of creatures in angst and suffering. While the contextual information required to infer meaning from the painting is clearly revealed, the historical event is only alluded to and not described inside the image, the author expects its audience to bring this information with them and his concern turns entirely to the modes of representation.

In the Brookes diagram, the procedure is different because the data needs to be exposed first and foremost, not only because the image is produced outside the art world, but due to the exposure that audiences have to the underlying information. Once the data is taken out of the reports and analyzed as is, the terrible tragedy is revealed, the ship was not big enough to humanely transport so many people. The function of the image is then to expose these findings, to articulate in an image the unbiased data analysis. All of

³⁴ Guernica was the name of the town in the Basque Country of Spain that was bombed April 26, 1937 during the Spanish civil war by the fascist national government of Francisco Franco with the help of his allies Nazi Germany and Mussolini's Italian air force.

the diagrammatic decisions in the image serve this purpose: the scale of the ship and median height of the slaves, they place each body proportionally in the available space, then include descriptive side notes to clarify the diagram and the image serves its purpose. Why did the author include what seems like an unnecessary rendering of characters moving around the ship in uncomfortable positions? Why is the author suggesting symbolically the suffering of the slaves? Was the author making a calculated decision or is this a naïve gesture due to the lack of codes and norms for diagrams in this point in time³⁵? These questions cannot be answered since there is no record of the author explaining his decisions, this particular intrusion is not only intriguing but arguably anticipates the type of procedures and negotiations that authors are forced to make today when the underlying data is so powerful on its own. If the procedure of exposing data for what it is has the ability to provoke an emotional experience, why not just allow the representations to communicate the data transparently? We can begin to answer one of the central concerns of this thesis which is to identify the authors intent. In this image, we can infer two intentions: the dominating one is to inform the findings produced by the data analysis in an official document. The image at first glance has an authoritative context, the source of the data is reliable and clearly stated³⁶, the diagrams rigorously constructed and presented to the public as such. Jessica A. Glickman who dedicates her thesis to the analysis of the slave vessels argues the following about Alfred's image:

³⁵ In 1788 information graphics was in its infancy so any approach to diagrams and data visualization was full of speculations. It will be much later with investigations such as Jaques Bertin's "Semiology of Graphics" (1967) that any attention was given to a more homogenous approach to information graphics.

³⁶ The data was recorded in the Dolban Act, also known as the Slave trade act of 1788. For more information about this document, read: "The Allotment of Space for Slaves aboard EighteenthCentury British Slave Ships" (Charles & Herbert, 1985)

The original diagram was printed along with twelve hundred words about the conditions on board. Viewed as an eyewitness testament, no moral or special pleading included, it was a unique approach. The purpose was to present a forensic case, which had no reference to religion and presented the facts in a non-emotional way. (Glickman, 2015, pp. 15-16)

The approach to the report further clarifies Alfred's original intention to distance his personal opinion from the rendering. But the aforementioned crouching characters cannot be omitted under close inspection, they are siding notes, easily omitted and overpowered by the evidence it holds, yet William Alfred is planting his opinion inside the image without compromising his dominant goal. For the context in which Alfred is presenting his work, the calculus is clear, facts should be the priority otherwise the object is compromised. But when the context is the art exhibition space, this negotiation is not so clear. This is not to say that subjectivity is a must in art, but the traditions of these spaces expand the ability to test the different calibrations between presenting facts, opinions, or as discussed in the following section, evoke and position the audience in an experience that relates back to the natural complexity of what data originally registered.

Evocative data representation

In order to make a distinction between what data representation is in this research and what it is not, two general categories are proposed here where the main criteria of distinction are defined by the intent of the author, these proposed categories are the evocative and the functional summarized in Table 2:

Table 2

Summary of the Definition and Differences Between Evocative and Functional Data Representation.

Evocative	Functional
Those in which the subjective experience of the author prevails in the process of representing the data:	Where the author's subjective view is undesirable in comparison with a less biased presentation of the information:
Embraces ambivalence.	Seeks evidence.
Poses questions.	Expects affirmation.
Transforms phenomena through metaphors.	Simulates phenomena through mimesis.
Results in contingencies.	Results in concrete knowledge.

To clarify these differences, it is important to note that either approach is intrinsically based on the interpretation that the author has of the data. Inferring any meaning just by looking at a spreadsheet full of records will often present a monumental challenge even for experts. Therefore, any step of moving away from this stage of the data implies a series of human decisions as to how it will be articulated and presented to others. How this is ultimately done responds to the goals described in the categories above.

In the functional category, the primary goal is to move from the obscurity of the raw data to the clarity of concrete knowledge, the intent is to provide the means for reason to operate by providing a clear representation. In the evocative, the path to knowledge is not rejected entirely, but the intent is to forestall reason in accordance with specific research objectives. The resulting representation in the latter may be convoluted and challenging for the audience, but this

should not be equated with the authors inability to articulate what the data signifies, for data artists that seek to provide an evocative experience in their audience will also require some mathematical and statistical knowledge to indicate the phenomenon at hand, yet the established goal would strategically seek to give prominence to the emotional and intellectual operations that emerge before we can attain the phenomenon in a singular and concrete understanding. This discussion is further developed later in this chapter in relation to the theories of the beautiful and the sublime.

A second clarification before moving forwards has to do with the idea of the subjective experience used in the definitions of each category. They both refer to how the author of a representation personally relates to the phenomenon signified in the data. This central issue to the definition of the proposed categories informs a critical aspect of the relationship between practice and theory at the center of this research. In the practical work that tests the ideas articulated in the theory, the author establishes a general premise for identifying the datasets to be used in artworks which is that these datasets should somehow contextually relate to a traumatic personal event. Instead of forcing a complete detachment from the information that the datasets provide, the author seeks to be so closely affected by the events they denote that removing oneself is deliberately complicated. From this obstacle, the author interrogates the role and potential for subjectivity in the task of artistically representing to the audience the information contained in the data that has been otherwise collected and published in the impersonal form of “raw data” regarding the violence in Colombia. This intersection between impersonal datasets with its analysis and exposition informed by subjective experience forces a shift in the forms of enunciation. The practical work does not present itself as the factual account of the violence in Colombia, but as an

interpretation of the facts that the data denotes. In this sense, the data is rigorously analyzed and present as an index³⁷ in the artworks, yet the departure from conventional visual articulations of data, such as diagrams and charts we are accustomed to and replaced with hand-drawn animation, intend to expediently inform the audience that these objects are personal interpretations of the underlying datasets. Within this approach, both functional and evocative procedures operate in the creation of Data Driven Drawings. The analysis of the data and its presence in the artwork implies a functional approach in its preliminary stages of representation³⁸, while the final articulation of the artworks emphasize from its formal enunciation and underlying goal that a subjective sentiment is driving the representation.

Therefore, the proposition of the two categories, functional and evocative, is not intended as a contentious division, but rather as a clarification in order to narrow the scope of this research. At the same time, it does not intend to fundamentally separate science, technology, and the arts, but rather to clarify how these disciplines intersect when the final goal is an artistic gesture.

It is important to note that even if the goal is an artistic piece, working with data will always encompass the intellectual and technical knowledge advanced in other disciplines. The methods used in the creation of Data Driven Drawings follow the same steps

³⁷ In all the Data Driven Drawing projects, the datasets are informing each instance of the visual composition of the animation. The algorithms that compose each piece are going through the dataset, line by line, and the program decides based on the variables of each instance how to draw the images on the screen. This approach allows for the data to be present as an index of the phenomenon.

³⁸ Chapter 2 develops this argument in more detail, the use of data visualization as graphic notation is where functional data representation becomes important in the development of Data Driven Drawings that seek to be evocative in their final result.

as those proposed in scientific schemes: We identify a phenomenon, then abstract it into data (or use the data collected by others); we then filter the raw information to isolate the specific areas we are interested in, and ultimately represent the dataset. Benjamin Fry clearly summarizes this process as follows:

Figure 8. Structure of steps in data representation, from “Computational Information Design”³⁹, by Fry, B., 2004.

Fry also claims that even though working with data is often a multidisciplinary effort, it is detrimental to the end result to completely separate the concerns of each discipline at each stage of the process:

The initial format of the data (how it is acquired and parsed) will often drive how it is structured to be considered for filtering and statistics or data mining. The statistical method used to glean useful information from the data might drive how the data is initially presented—the representation is of the results of the statistical method, rather than a response to the initial question. (Fry, 2004, p. 14)

This is a critical idea for Data Driven Drawings. The goal is never to represent a dataset as a response to questions posed by others, but rather to impose a new view onto the data collected which then

³⁹ Fry describes on his thesis two further steps which he calls “refine” and “interact”. The former refers to the process of iterating over the representation in order to produce the desired result, this process is deliberately omitted here because in my opinion, this refers to the actual process of representation and not something aside from it. The last step he calls interact is only optional, the representation may be interactive in its presentation or not and equally implies the kind of decisions made in the representation since representation refers to the entire form in which the piece is presented.

informs its consequent filter, analysis, and final representation. Otherwise, the initial premise of creating an evocative experience will be compromised before it is represented. To clarify this point, the process for developing a Data Driven Drawing project will be described in relation to the aforementioned stages of producing for data representation in more general terms as described by Fry. This first example of a Data Driven Drawing was initially conceived as a technical exercise in order to understand the process of creating a data driven animation. It was not intended to become anything more than a simple exercise since it was developed at an early stage of research, at this point only a few public databases from Colombia had been consulted and for the purposes of the exercise it seemed of secondary importance what data to use, the goal was to learn the technical process. For these reasons, the dataset that was chosen pertains to seismic data from Colombia which is very well recorded by the Colombian government and seemed large enough for the technical challenge. Not having any specific question or intention with the data beforehand is not being described here apologetically, but precisely because the result turned out to elucidate an important theoretical idea. This significant theory is what John W. Tukey describes as *Exploratory Data Analysis* (1977)⁴⁰ which is an invitation to only ask questions to the data once it is being visualized instead of defining a consolidated hypothesis before entering it. What Tukey argues for is that the impulse to force the data to answer, or be in service of, an already formulated premise restricts our own ability to explore it for what it is. He argues that by restricting the analysis to a hypothesis the data is then limited from potentially interesting and unconceived questions that can be

⁴⁰ Tukey's proposition of "exploratory data analysis" is further developed in Chapter 3 but briefly introduced here for the purpose of clarifying the argument.

posed to the dataset. Tukey makes a case for establishing a more consequential research approach of going back and forth between our own intellectual impulses and an openness to “explore” the data to guide and inform possible paths to what we can do with it. Since the author’s expertise is not of technical or scientific understanding of seismic activity, the dataset was being explored from personal experience with earthquakes. The inclination to draw connections between the natural phenomena and personal experience in this open exploratory approach, allowed for imagining the following piece by combining two non-related phenomena: seismic activity and terrorist attacks, not to claim they have to do with one another physically or socially, but because their combination emerged to construct a particular metaphor based on personal history.

Example 1: “Rite of Passage”

“Rite of Passage”⁴¹ is an animated flock of birds performing a highly coordinated murmuration, in which the flight pattern created is controlled by a program that is reading annual data from seismic activity and terrorist attacks in Colombia. At first, the birds are located randomly on the screen, and from each of their positions, they fly towards the geolocation of the most recent seismic event. The constant seismic activity makes them gravitate to different parts of the screen, creating a dance which is personally associated with something natural and “beautiful”. Whenever there is a terrorist

⁴¹ Three version of this project were created to test different technologies and approaches to animating the data driven flock. The first version made in Processing uses single dots to show the birds (see at: <https://vimeo.com/70717163>). The second version using hand-drawn cycles of birds was coded in JavaScript (see at: <http://www.dddrawings.com/lab/flock-a>). The third version integrated WebGL to the JavaScript code to simulate depth (see at: <http://www.dddrawings.com/lab/flock-b>).

attack that matches the date of the seismic event, the flock turns disparate; they fly rapidly in opposite directions, deconstructing their coordinated natural flow. Depending on the intensity of the attack, the birds move further away from each other, which results in the flock to taking longer to recover the initial course of synchronized action. With these rules set in place, the viewer can interact with the different years in the dataset to view the flocks' behavior which changes every year.

These two seemingly unrelated events do not have a statistical correlation in the sense that seismic activity does not produce terrorist violence or vice versa, or at least this is not the claim. Hence, attempting to represent these two sets in the context of a functional data project is incongruent. But in the author's case, these two events are connected in childhood memories and therefore the link between them becomes a potential metaphor to be explored. The following personal account serves as a backdrop for understanding the evocative intentions of the piece.

While growing up in Colombia, a country that is part of what is known as the "Seismic belt"⁴², it was not uncommon to experience seismic activity. As a child, experiencing these events had a certain excitement to them; they did not produce fear but an inexplicable thrill. At the same time, during the late '80s and early '90s, Colombia was experiencing one of the most violent eras with a convoluted war between the state, drug lords and insurgent groups. One of the particularities of this situation was the choice of

⁴² The seismic belt is a particular geographical area of the Earth where most seismic activity is felt in the surface. It covers a big part of the Pacific side of the Americas and stretches through to the Himalayas and South East Asia. Chile, Japan and California are part of the seismic belt and historically report massive earthquakes.

battleground. It was a common terrorist practice to detonate bombs in rural areas occupied by regular civilians. Since my family and I often traveled to Medellin, a hub for violence due to the Medellin cartel, sometimes we happened to be near those explosions. We were never at risk, but close enough to feel the consequent tremors of the ground, which were similar to an earthquake. Even as a child, this connection did not produce joy, but an intense fear. While I understood the potential risks of an earthquake, this natural occurrence seemed tolerable and natural, while the tremors of an exploding bomb produced a completely opposite impression. I used this jarring connection as the premise to examine the data. Having established a personal connection with specific events that have publically accessible data that defines them, the project followed the aforementioned stages of production described by Fry as “Acquire / Parse”, “Filter / Mine” and “Represent / Refine”.

Acquire / Parse

The seismic data was obtained from the official government institution called *Ingeominas*. On their website, they freely share the entire database of seismic activity in Colombia since 1993 and is updated on a daily basis. The violence dataset was extracted from the publically available “Atentados Terroristas 1988-2012” published by *Centro Nacional de la Memoria Histórica* (National Centre of Historical Memory). In this database, they had already filtered out, from multiple sources, the events that fit their criteria as to what constitutes a terrorist attack: “Understood as any indiscriminate attack perpetrated with explosives against civil objectives in public places and with a high potential for devastation

and lethality.”⁴³(Histórica, 2012) This particular filter was the perfect fit for this project since it restricted the events to those with explosives which are specifically the type of events that relate to the personal experience. Regardless of the distant nature of the two datasets, they both presented two unifying variables in them which are the geographical location and a timestamp. These variables allowed for the datasets to be combined into a unified representation by comparing the date of each event and the location in Colombia where it took place. At this stage of the process, the data needs to be “parsed” in order to normalize the data type that describes each variable.

While the seismic data already contained numerical values for the latitude and longitude describing the locations, the violence data described the place in a text with names of state, city, municipality, and place. In this case, the numerical value was preferred since it could be mapped visually into screen coordinates in a potential representation. To translate the violence textual location to longitudinal coordinates, custom software⁴⁴ was developed during the research to help with the parsing and cleaning of this and any other dataset. For the timestamps, the original datasets used

⁴³ Translated by the author directly from: “*Se entiende como todo ataque indiscriminado perpetrado con explosivos contra objetivos civiles en lugares públicos con un alto potencial de devastación o letalidad.*” (Histórica, 2012)

⁴⁴ The Software is currently in beta version and is intended to be released as open source in the future. The motivation for developing it stems from a growing frustration during the process of this research to find a good application for parsing excel tables and translate them to the JSON format that is preferred in web applications. The software is tentatively named DDD-Parser and is capable of finding anomalies in the input data, combine variables, normalize variables to data types that define place, time, text, number. An important feature of the Software is that it is able to translate place names into longitudinal values by searching in public API’s for the corresponding values of a place name. The user is then able to select and accept the translation for cleaning and parsing the original dataset. The resulting structure can then be saved as a minified JSON file.

different formats to describe the date and time of the events, while the seismic data used the UNIX timestamp⁴⁵, the violence data contained local dates for Colombian time zone. Again the preferred format was taken from the seismic dataset in order to use the datasets freely in the future in case they are combined with events in other places of the world. In this stage, these normalizations and cleanup are intended specifically for the type of application they will be used on, in this case, a programming language that articulates the data visually, and do not anticipate the forms of representation. While some variables are transformed into their data types, the information they represent remains untouched. Once this technical process is finished, the data can be explored visually in order to have a clearer understanding of what it is able to describe.

Filter / Mine

Filtering and mining the datasets is the stage in the process of selecting the variables that one would like to understand by establishing the mathematical logic that describes the data to us, humans. In this stage, we are able to observe the patterns that emerge from the data and begin to infer meaning and identify potential ways in which we can describe these datasets to others. Tukey's (1977) proposal of *Exploratory Data Analysis* suggests that visualizing the data as early as possible and fluidly move back and forth between parsing, filtering, and this preliminary representation, allows the researcher to investigate the complex patterns that the data is capable of producing. From them, new questions previously

⁴⁵ UNIX timestamp refers to the number of seconds since January 1, 1970 in UTC-0. The seismic data was captured as UTC-0 and the violence data as UTC-5 which is the time zone of Colombia.

not considered may arise as was the case in this project. As described before, the starting point for this project was guided by a personal intuition that seismic activity and terrorist attacks had a unifying relationship rooted in a personal experience, but during the process of filtering and visualizing the contrasting datasets, a new intersection to a personal event become more relevant, shifting the attention from events in the author's early childhood (late 90's), to a more recent incident. From the preliminary visualizations, the most prominent pattern based on the violence data showed an increase in terrorist attacks in 2002. This coincides with the election of former Colombian president Alvaro Uribe Velez who ran a campaign promising strong military against insurgent groups. He won the elections and his presidency provoked violent reactions from these groups who started attacking civilians in urban locations with more frequency.⁴⁶ The notably highest spike in the violence data pointed to the event on February 7, 2003, which was an attack on the social club *El Nogal* in the capital city of Bogotá. This particular event had personal significance because in it a close friend of mine lost his life. As the highest spike of violence pointed to a specific event of personal significance, the project shifted its point of origin from the broader association of tremors from explosions as a child and earthquakes, to the particular connection the death of a friend. This unexpected but emotionally charged encounter provided a suitable case for testing the more general premise of Data Driven Drawings as an approach in which subjectivity deeply permeates the representation. The specificity of having a personal connection with a singular datum elucidates a critical problem in data representation, which is the difference

⁴⁶ See online the interactive data visualization to compare each year at: <http://www.dddrawings.com/lab/spikes>

between denoting the microcosms (that is, a singular event in the dataset) as opposed to the macrocosms (that is, the phenomenon in itself). Here is how Tukey describes this “phenomena vs numbers” problem:

Much of what we want to know about the world is naturally expressed as phenomena, as potentially interesting things that can be described in *non* numerical words. That an economic growth rate has been declining steadily throughout President X's administration, for example, is a phenomenon, while the fact that the GNP has a given value is a number. With exceptions like "I owe him 27 dollars!" numbers are, when we look deeply enough, mainly of interest because they can be assembled, often only through analysis, to describe phenomena. (Tukey, 1990, p. 330)

The general phenomena that this project describes is terrorist attacks in Colombia which the original dataset is capable of representing on its own by assembling the values into an understandable form. Either visually in a diagram or by summarizing the numbers into for example: in the terrorist attack of *El Nopal* in 2003, 200 people were injured and 36 died. What Tukey is describing above refers back to Daniel Rosenberg's ideas discussed earlier in this chapter as data being strictly rhetorical (Gitelman, 2013). Tukey suggests that numbers are of interest once assembled, after being analyzed, to describe the phenomena. As such, filtering and mining the data while being predominantly in the domain of mathematics, statistics and logic (Fry, 2004) it is the first construct of the numbers within an argument. It is moving from the raw unrepresented data, into a rhetorical form that begins to describe the phenomenon in a particular way.

Representation

While the previous to stages of the process defined the particular way in which the datasets were being interpreted, the actual

articulation of these findings takes place in the final form of presenting the argument to others. In this case through the animated image of flocking birds that move in synchrony to the events in the datasets as described before. In this stage, the main concern in Data Driven Drawings is to negotiate between the subjective connection to the data, the personal story, and the phenomenon at large. In this project, there was an individual datum within the data that was personal, from the 36 fatalities in the singular event of a terrorist attack in 2003, 1 is someone the author knew. The database describes him within the number 36, no allusion to his name, his life story or anything else, a number within a number that motivates the creation of an artistic project, but in the general scheme of the dataset is so buried inside and personal that is informative only to the author. Considering these contrasting scales, the subjective experience of the data is of importance to how rhetorically the data is to be presented and informs the aesthetic articulation of the piece, yet the potential of data representation encompasses the larger scales in which the phenomenon is abstracted and contained in the datasets.

The general consensus in the extremes of the ideological spectrum (Manovich, 2010; Tufte, 2006) agrees that one of the main characteristics of the medium at hand is that we take a large set of numbers that indicate something in reality and reduce all that complexity to a perceivable scale. This grants us access to something that otherwise exceeds our own scales of perception and rational capacities.

Regardless of whether these are people, their social relations, stock prices, income of nations, unemployment statistics, or anything else. By employing graphical primitives (or, to use the language of contemporary digital media, vector graphics), InfoVis is able to reveal patterns and structures in the data objects that these primitives represent. However, the price being paid for this

power is extreme schematization. We throw away 99 percent of what is specific about each object to represent only 1 percent – in the hope of revealing patterns across this 1 percent of objects' characteristics. (Manovich, 2010, p. 38)

What Manovich refers to as “extreme schematization” relates to an earlier idea he proposed as the “anti-sublime” (Manovich, 2002), in which he argues that this compromise of reducing the information into a perceivable scale, not only grants us rational access but fundamentally subverts a previous inability in the arts to represent certain phenomena that were too large to be reduced into a single object. While Manovich makes an important connection between data representation and the philosophical tradition of the sublime, his proposition of the “anti-sublime” needs to be contested. For, the key issue Data Driven Drawings is what the representation is able to achieve or evoke in the viewer rather than a formal ability to reduce a phenomenon into a perceivable object.

The Sublime

The functional and evocative modes of data representation proposed here are (not by coincidence) related to the ideas of beauty and the sublime respectively. While functional representation and beauty ought to manifest through simplicity, order and concrete knowledge, the evocative and the sublime oughts' to upset these premises and remain suspended in a state where knowledge through reason cannot be fully apprehended. The sublime does not reject a path to knowledge or rational inference, but functions within the premise that our innate resolve to understand, forces new types of knowledge that are not exclusively rational, what Jean-François Lyotard calls gratification of effort in his essay “Presenting the unrepresentable: the sublime”:

The sublime is not simple gratification but the gratification of effort. It is impossible to represent the absolute, which is ungratifying; but one knows that one has to, that the faculty of feeling and imagining is called upon to make the perceptible represent the ineffable – and even if this fails, and even if that causes suffering, a pure gratification will emerge from this tension. (Morley, 2010, p. 134)

Lyotard argues that feelings and imagination are invoked in the operation of interpreting the sublime, these subjective experiences allow for reaching an understanding and not just obscure knowledge when confronted with the ineffable. This operation is a complete experience in “tension” between reason and other forms of knowledge that have no consensus. Yet we arrive, resist to surrender reason entirely, and subjectively construct an understanding over the unbound.

The relationship between data representation and the sublime are presented here through the analysis of authors that have defined our understanding of the sublime (Pseudo-Longinus, Burke, Kant, Deleuze), and those who had already made the connection between the term and issues of data representation (De Landa, 2000; Heinrich, 2015; Manovich, 2002; McCosker & Wilken, 2014). In the developing field of data representation, the understanding of the sublime requires a revision in light of the affordances of digital technologies. For phenomena that were though as impossible to be contained in traditional artistic mediums, can now be bound and represented by delegating to the computer what was before a monumental task. For example, calculating and mapping all known stars of our galaxy by hand in a painting or drawing is an overwhelming human undertaking, but a trivial operation for computers. Therefore, the types of phenomena that were traditionally associated with the subliminal experience of the world, can now be abstracted to the scales of human reasoning using digital technologies. Even with these capabilities, the argument

presented here focuses on the sublime, not as something that emerges exclusively from vast phenomena in the world, but the experience they produce in us once confronted with them. In other words, being able to “understand” and “reason” over the vast and complex manifestations of the world through data representation has only shifted the mechanisms that allude and provoke a subliminal experience. For example, in the practical projects of Data Driven Drawings, the underlying datasets are able to describe, to a certain extent, vast and complex phenomena such as the historical violence in Colombia, yet the artistic goal is to investigate the ever-shifting flux of life as it relates back to these events. Being able to describe with clarity the events bounded in the data is only a new formal capacity for artistic practices, but the ways in which we as artists relate back this information to the audience preserves the potential of evoking as proposed earlier in this chapter.

The study of data representation, at least in the dominating field dedicated to its visual aspect, gives prominence to the association with beauty: “Beautiful Evidence” (Tufte, 2006), “Beautiful Data: a history of vision and reason since 1945” (Halpern, 2014), “Beautiful Data: The stories behind elegant data solutions” (Segaran & Hammerbacher, 2009), “Complex Beauty” (Franceschet, 2014), “Beautiful Visualizations” (Steele & Iliinsky, 2010), among many others. This weight on beauty can be understood as a fulfillment of a dream, being able to take on phenomena whose experience is associated with the sublime, those which are too vast to apprehend, and finally reduce them into an object that allows for reason to fully operate. But for this to be true, one has to study the subject also from its counterpart, the sublime. This perspective has received little attention but it is presented here as critical to the subject.

The first author that draws a connection between the sublime and data representation practices is Lev Manovich in his essay “Data Visualization as New Abstraction and as Anti-Sublime” in which he argues:

If Romantic artists thought of certain phenomena and effects as un-representable, as going beyond the limits of human senses and reason, data visualization artists aim at precisely the opposite: to map such phenomena into a representation whose scale is comparable to the scales of human perception and cognition. (Manovich, 2002, p. 8)

Manovich’s proposition presents a fundamental problem as it fails to consider the effects a representation has on us, in other words, the process of decoding the information presented in a visualization may yield concrete and discernable knowledge, but this is not always the case. Borrowing from Kant’s theory of the mathematical and dynamic sublime (Kant, 1987), the logical comprehension *comprehensio logica* attainable through the numerical views of the data, do not equate the aesthetic comprehension *comprehensio aesthetica* that the representation may incite in the reader. Kant also argues that the “object” that provokes the sublime is not necessarily “un-representable” as Manovich suggests, but it can be contained in a “formless object”, an object that suggests or evokes in us an idea of the totality that is the phenomenon in its natural state:

The beautiful in nature concerns the form of the object, which consists in [the object's] being bounded. But the sublime can also be found in a formless object, insofar as we present *unboundedness*, either [as] in the object or because the object prompts us to present it, while yet we add to this unboundedness the thought of its totality. So it seems that we regard the beautiful as the exhibition of an indeterminate concept of the understanding, and the sublime as the exhibition of an indeterminate concept of reason. (Kant, 1987, p. 98)

The idea of data representation that Manovich offers is comparable to the functional goals of providing evidence and clarity, therefore the “anti-sublime” he proposes is not exactly a new concept negating the sublime, but argued here that it closely refers back to its opposite, the beauty. The main contradiction in Manovich’s proposition of the “anti-sublime” has to do with the motives an artist has for scaling down the phenomena, he claims that data visualization artists seek to map the phenomena into scales of “human perception and cognition” (2002, p. 8), but fails to consider the emotional underpinnings that are intrinsically connected to the sublime. If the artist scales down the phenomena to allow for all of our cognitive abilities to reason over it, then the object is bounded, it regains the form that makes it beautiful. Manovich is conflating the artist ability to bound a vast and complex phenomenon into something perceptible with the potential of representing it. Referring back to Lyotard, the sublime is the effort “to make the perceptible represent the ineffable” (Morley, 2010, p. 134), therefore a data representation, even when it is able to bound phenomena in nature that was before considered unrepresentable in artistic domains, even when these complexities are turned perceptible, its representation can still provoke a deeply emotional experience such as the sublime.

A more comprehensive account of the relationship between the sublime and current practices in data representation can be found in (Heinrich, 2015; McCosker & Wilken, 2014) which share the common premise of studying the sublime through the effects that an object has over those who experience it (aesthetics), concentrating on Burke, Kant and the latter proposal of the diagram in Deleuze as proposed by De Landa (2000).

It is contested here that idea of diagrammatic thinking in Deleuze is analogous to the proposal being put forth here as functional and evocative representations. De Landa describes the diagram in relation to the visual representation of data in science and technology as a practice that makes an effort to produce concrete knowledge, this in contrast with the diagram in Deleuze which he argues operates more like a “virtual” state of potential associations. A good example of this is what Deleuze and Guattari refer to as “rhizome” in the introduction of “A Thousand Plateaus” (Deleuze & Guattari, 1988, pp. 3-25). The rhizome rejects the hierarchical, cause and effect systems and celebrates multiplicity and ambivalence as a way of thinking that is always in an undefined state:

A rhizome has no beginning or end; it is always in the middle, between things, interbeing, *intermezzo*. The tree is filiation, but the rhizome is alliance, uniquely alliance. The tree imposes the verb “to be,” but the fabric of the rhizome is the conjunction, “and. . . and. . . and. . .”

De Landa argues that the diagram in Deleuze has “no intrinsic connection with visual representation” (De Landa, 2000, p. 33) because it is not contained in the form, but operating in our own minds creating a short circuit with unobstructed rational thought, “always in the middle”. What De Landa is doing here is a semantical exercise of expanding the definition of the word “diagram”, from (a) its common use of an image that produces concrete knowledge, to (b) Deleuze’s proposition of it as a state of the mind before knowledge. I argue that in a similar way (a) is equivalent to what is proposed here as *functional* and (b) as the *evocative*. But with the distinction that the evocative is considered here as a problematic representation, one in which the intention is to summon an aesthetic experience that surpasses logic. This state of the mind before knowledge is the space where Lyotard’s “gratification of

effort” takes place in the form of an interpretation of the object that is informed both by arresting emotions and rationality.

To further clarify this idea, we can use Deleuze’s explanation of the diagrammatic in Bacon. He claims that the violent strokes, wipes, rubs and other means to suggest the figure in the painting, infuses the image with a sense of what he calls “catastrophe”.

For these marks, these traits, are irrational, involuntary, accidental, free, random. They are nonrepresentative, nonillustrative, nonnarrative. They are no longer either significant or signifiers: they are asignifying traits. They are traits of sensation, but of confused sensations. (Deleuze, 2003, p. 100)

They are “nonrepresentative” of the concrete form, or to draw a parallel with the matter at hand, of the what the data concretely signifies, but these formal qualities are there to evoke in the viewer an uncontrollable sensation. In these terms, the image “results in contingency” (evocative) and not in “concrete knowledge” (functional) as defined at the beginning of this chapter. These formal qualities that Deleuze refers to, are there not to represent unambiguously, but to forestall and overwhelm reason so new forms of knowledge are evoked in the combination of the “confused sensation” and our drive to reason over them. Or what is discussed here as the sublime.

Edmund Burke and Rudolf Otto argue that the sublime operates under a collision of opposing psychological effects. Otto proposes a concept analogous to the sublime called “numinous” that incites a sense of “terror” (*Tremendum*) and in parallel a fascination (*Fascinans*) for it. This idea was initially proposed by Burke in which he argues that a manifestation of beauty produces “delight” - “I make use of the word *Delight* to express the sensation which accompanies the removal of pain or danger.” (Burke, 1757, p. 54),

and the sublime on the other hand produces “horror” and delight at the same time, a tolerable pain or delightful horror.

The passions which belong to self-preservation, turn on pain and danger; they are simply painful when their causes immediately affect us; they are delightful when we have an idea of pain and danger, without being actually in such circumstances; this delight I have not called pleasure, because it turns on pain, and because it is different enough from any idea of positive pleasure. Whatever excites this delight, I call *sublime*. (Burke, 1757, pp. 84-85)

Kant differs with Burke as he does not consider horror or fear as defining qualities of an aesthetical experience and therefore related to the sublime “Just as we cannot pass judgment on the beautiful if we are seized by inclination and appetite, so we cannot pass judgment at all on the sublime in nature if we are afraid.” (Kant, 1987, p. 120) But, they concur when Kant discusses beauty in terms of something finite and measurable in contrast to the boundless characteristics of the sublime. Kant expands this idea through the proposition of the “mathematical” and the “dynamic” sublime. The mathematical sublime refers to something in terms of its size or the impossibility to actually measure it by means of our own intuition, and the dynamic in terms of “might” or the overpowering sensation that suspends our ability to reason over it.

Borrowing from these ideas, let’s consider the proposed categories of functional and evocative representations to argue that the sublime helps us to better understand the underlying goals of the evocative. As defined earlier in this chapter, the evocative is intended as a representation that “embraces ambivalence” and “results in contingencies”, which is equated here with the suspension of reason in the sublime. In other words, the sublime is an experience a priori to our ability to reason, hence it can remain ambivalent in its inability to be fully apprehended by the mind. Also, it remains as a permanent contingency since it is not universally

defined, but allows for the subjective experience to discern over it without overcoming the mathematical threshold of measuring its magnitude.

Also, it is worth clarifying neither the sublime nor the evocative, should be equated exclusively with artistic objects, but as Kant proposes, it depends on the object in itself, where it be “bounded” to measurable scales (beauty) or boundless (sublime). This is to say that the way I, in my own subjectivity, experience a representation, defines when it is beautiful or sublime, even if the original intention is to clarify.

There is one particular scientific simulation⁴⁷, that in the author’s opinion, is able to provoke a subliminal experience despite its own intent. The “NASA/Godard Space Flight Center” published in 2012 a computer simulation of the galaxy evolution tracing its development shortly after the big bang to present times. Its objective, Reddy explains, is to “help astronomers better understand the new findings in galaxy evolution” (Reddy, 2012). Therefore, it is clearly conceived as a functional representation of this particular phenomenon. This is a good example of the ways in which the most advanced technologies today are capable of processing massive datasets, they can turn perceptible something that could only be imagined before, and still what it represents is ineffable. Reddy explains that “The simulation ran on the Pleiades supercomputer at NASA's Ames Research Center in Moffett Field, Calif., and required about 1 million CPU hours”. The seemingly simple animation of particles swirling around and developing its spiral form fascinates

⁴⁷ The video can be seen here:
<https://svs.gsfc.nasa.gov/vis/a010000/a011000/a011087/>

as it is beautiful and elegant. But as the animation progresses, I personally can't help to be overwhelmed by it, or in Burke's words, astonished:

The passion caused by the great and sublime in nature, when those causes operate most powerfully, is astonishment; and astonishment is that state of the soul, in which all its motions are suspended, with some degree of horror. In this case the mind is so entirely filled with its object, that it cannot entertain any other, nor by consequence reason on that object which employs it. Hence arises the great power of the sublime, that, far from being produced by them, it anticipates our reasoning, and hurries us on by an irresistible force. Astonishment, as I have said, is the effect of the sublime in its highest degree; the inferior effects are admiration, reverence, and respect. (Burke, 1757, pp. 95 - 96)

He argues it "anticipates our reasoning", or in this case my own reasoning. Not because I am unable to even begin to understand the intricacies of the scientific ideas that inform the intended audience (astronomers), that is the intellect, which in this case I do not dominate, but my own reason, my ability to understand this particular representation of time. I understand 13.5 billion years as a number and that the galaxy forms from chaos to its current form. But seeing time compressed in this video astonishes me uncontrollably. This problem elucidates what Kant referred to as the mathematical sublime, we can't fully measure the immensity of the phenomenon and this incites a deep emotional feeling that can only describe as sublime. So what kind of knowledge is emerging from this experience? Certainly, we are learning something, it is not simply exposing our ignorance, but other rational thoughts are operating here. The nearly 200 thousand years of humankind is a concrete piece of information that if it is calculated into the 13.5 billion years compressed in a 2:16 minutes video, it does not amount to even a second of the visualization. This means also that our own existence or even our long family history is so infinitesimal within this grand scale of time, that all concerns and traumas are

reduced to nothing in this context. This is both pleasing and frustrating. The rational process of trying to calculate our own life into the video is precisely the path to encountering that it is impossible to understand how it makes any sense in this macro view of the universe, what Lyotard lucidly describes as “gratification of effort”. This new knowledge attained is ambiguous but can be articulated nonetheless.

The last idea discussed in this chapter pertains to the first account we have of the sublime. All the previously mentioned authors start their discussions with Burke and follow it under the premise that the ideas of the sublime pertain to issues of aesthetics in representational and performative arts. But, as described by Maria del Rosario Acosta:

The category of the sublime, specifically as an aesthetic category, really emerges during the XVIII century. As mentioned before ... it existed in ancient debates on rhetoric, and, in this sense, it is not only a modern category.⁴⁸ (Acosta, 2012, pp. 93-94)

Acosta is indicating here an unknown author which is referred to now as Pseudo-Longinus. His text *On the Sublime* dates back to the first century and was only rediscovered after Nicolas Boileau translated it into French in 1674. Pseudo-Longinus defines the sublime as:

That which is admirable ever confounds our judgment, and eclipses that which is merely reasonable or agreeable. To believe or not is usually in our own power; but the Sublime, acting with an

⁴⁸ Translated into English by myself, here the original text in Spanish: “*La categoría de lo sublime, como categoría propiamente estética, surge sólo realmente hasta el siglo XVIII. Como ha sido mencionado ... existía ya en las reflexiones antiguas sobre la retórica, y, en este sentido, no es sólo una categoría moderna.*”

imperious and irresistible force, sways every reader whether he will or no. (Pseudo-Longinus, 1890, p. 2)

Acosta's observation points out that Pseudo-Longinus introduces the sublime in a juncture between aesthetics and rhetoric, experience and language, which anticipated the proposals by Burke and Kant of the sublime as an aesthetic experience in which an object (Kant) denotes what is otherwise obscure and unattainable.

Extending on Acosta's observation, let's examine the case of Florence Nightingale who is an important precursor of artistic data visualization. Not in terms of the context and objectives of her work (her work was not intended as a cultural artifact to be shown in artistic venues), but because she is one of the pioneers in using rhetorical strategies to persuade her audience into the issues she portrays. Nightingale anticipates artistic intents when she introduces unconventional metaphors into her data visualizations as a strategic device to provoke precisely an emotional impact first, and once the reader is swayed in, present an important issue with factual provenance.

Florence Nightingale was a statistician who worked as a nurse during the Crimean War (1853 - 1856). She noticed that the main cause of mortality in the British army was due to the poor sanitary conditions in the battleground nursery. There were more deaths caused by infections and other curable diseases, which meant that they could save a lot of lives just by improving hygiene and reinforcing medical supplies. To make this plea, Nightingale produced a series of arguably ostentatious graphics to report this problem back to the authorities in London (Nightingale, 1858). The one she is most recognized for is the "Diagram of the Causes of Mortality in the Army in the East" (Nightingale, 1859, p. 19), but more casually known as the "Rose Diagram" or the "Coxcomb".

These are polar diagrams representing time in a circular progression and the number of deaths extends from the center⁴⁹. The blue “wedges” as she describes them, represent all curable diseases, while the black and red ones indicate all other causes.

Nightingale was noticeably versed in the “best” practices of information graphics, yet she decided to avoid the more conventional strategies such as lists or column charts for displaying transparently a time-based dataset. Her “Rose diagram” distorts the underlying data, yet the originality of the diagram at the time made it noticeable and this invited her readers into it. Furthermore, her descriptions in the report use a “colorful” language, she described the resulting form in one of the first iterations as “the great black bat’s wing” (Nightingale, 1858, p. 5), or the “wedges” extending from the center. What is noticeable here is that Nightingale understood her audience and appealed to their *aesthetical comprehension* (to use Kant’s terms) to initially attract them to her cause. Once the viewer was compelled by the conversational language and oddly captivating graphics, she then was able to communicate the facts, and therefore address the *logical comprehension* of the reader.

The conceptual range between the opposing ideological categories of function and evocation presented here have been this far given more weight to developing a defense for subjectivity and evocation in artistic data representations. This focus is aiming at the objectives of the end results in Data Driven Drawings. But for data to be given meaning through representation, even if its form of enunciation is convoluted and challenging, it has to be somehow

⁴⁹ I drew direct inspiration from Nightingale “Rose Diagram” for the graphic notations presented earlier in the chapter as “Spikes”.

understood by the author. Otherwise, the procedure of presenting it to others removes from the data its ability to be an index. As argued by Floridi (2011) earlier in this chapter, data is a transient object that can be given multiple meaning to produce information with it, this also indicates that in order to build an argument with it, one has to investigate its possible meanings. This is where functional data visualizations turn useful for data artists, to “visualize” and draw meaning from the data before its potential representation. The next chapter examines this issue in relation to the traditions of graphic notations in music and animation as objects developed during the preproduction of artistic pieces.

Chapter 2: Data visualization as graphic notations

In this chapter, the goal is to reinforce the idea that the functional and evocative modes of data representation do not have to be mutually exclusive in the production workflow of an artistic piece. Even if the intended representation seeks to be evocative, being able to understand the data often requires making it “visible” for the one who is studying it. In this sense, the early stages of a data driven animation project demand a preliminary engagement with the data that is exploratory and a priori to defining the intended representation. In these terms, the approach to data visualization as graphic notation is understood as an internal working document for the authors to explore and understand the primary source of data to work with. These visualizations are not meant to be published, printed or shared with the public and serve as a functional, yet very important, document for the production team. Its importance and reason to include it in the present discussion lies in its point of departure from functional data visualizations as described in the previous chapter, which is that instead of serving as documents that communicate information to others, these images are the initial place for the authors to speculate, recognize potential associations between the data and their artistic intentions, identify the structures and patterns in the data that may transfer to the final production, and as such, is the first space for the author to think and engage in the creative process of producing data driven animation.

Data visualization as graphic notation is in simple terms, any expedient image that allows us to move from the complex nature of data to an abstraction that facilitates the authors an understanding of the sample at hand, furthermore, the expediency of these visualizations allows for multiple iterations over how to “see” the

data by refining a particular visualization or testing different ways of structuring it visually. Each of these iterations is useful for the initial exploration of the data and to understand what aspects of it can be enhanced or discarded later on during production. For example, as explained in Chapter 1 “Example 1: Rite of Passage”, a data visualization was developed during the filtering and mining process and from its exploratory analysis a pattern showed that the highest “spikes” of violence pointed to an event that occurred in 2003 with personal significance. Multiple versions of this visualization were made during this process and only in the version discussed earlier, the spikes clarified a path to pursue working with the combination of these datasets and identify the artistic intention moving forward. If this preliminary visualization was not developed and, in turn, strictly define how the data was going to be represented, this information would have been easily missed. In this case, that violence data of terrorist attacks was the one creating in the visualization the distinctive sporadic spikes, while the seismic data created more of a mass that had minor variations but was constant and repeated frequently throughout the years. This visual difference and pattern suggested then the proposed interpretation that the flock of birds would move “naturally” to the flow of the seismic activity, while the terrorist attacks interrupt and disperse the flock sporadically. Since the earthquakes in these visualizations appeared constant, this pattern was speculatively associated with what is constant in the flock which is them flying in an organized way, while the terrorist attacks only created the dispersion at a frequency that is not constant or predictable, which in turn made the animation of each year unique even if it follows the same logic developed in the algorithm.

Even though these visualizations are used in most cases for internal study, the process of abstract visual analysis to aid preliminary

ideas is not new or exclusive to data representations. There is a long history of creative practices that use graphic notations and anchoring this approach to that tradition is an important association that is useful to understand how data visualization is not only functional for an artist to recognize the data at hand but serves as a creative space for thinking and articulating the intentions to be pursued in the final artwork.

To contextualize this tradition and its applicability to data driven animation, particular attention is given to a subject presented in this thesis that, in the author's opinion, deserves wider attention in animation studies which is the use of graphic notations in preparation for an animated film production. As seen in the small sample of case studies presented in this chapter, the use of graphic notations has been used throughout the history of animation in both commercial and independent films, yet at this point there has not been any comprehensive study of this production process that can be used as a starting point to compare these practices with the ideas presented in this thesis. Therefore, this chapter serves two purposes within this research: first, it proposes an initial method of analyzing graphic notations in animation, to then have a base to build a connection between the tradition of graphic notations and the use of data visualization as such. Considering the scope of the present research, the first contextualization is brief and the author suggests that more research should be developed in this area to provide future researchers with a more comprehensive academic ground, but in the absence of other investigations in this area and to keep the attention on data driven drawings, this study is presented here as an introduction to these ideas and then use this overview as the base for understanding the approach of data visualization as graphic notations.

The way this argument is structured throughout this chapter is as follow: first, explain what graphic notations are and describe the tradition from which they emerge in animation practices. Then introduce a methodology for identifying the principal visual signifiers in graphic notations and apply it in the discussion of a series of three case studies of prominent authors who have used them. Lastly, an example of a personal film made by the author in the past is discussed for its use of graphic notations and from it draw thematic and conceptual connections with Data Driven Drawings.

Graphic Notations

Based on the small sample of case studies presented here, we can identify the use of graphic notations as a strategy for preparing animated films of very different nature, cases for both commercial and independent films were found. Yet the scholarship and practice of animation have privileged two other methods: the literary script and the storyboard. Graphic notations do not necessarily replace them, but it is contested that they provide a powerful scheme for concentrating on the orchestration of the different elements of the film and encourage creative exchanges between the ideas or plans that anticipate the film and the end result. Additionally, in the development of the practical works of Data Driven Drawings, data visualizations have been used as the device for anticipating the final work. This approach stands on the tradition of graphic notations explained here.

One of the primary challenges encountered during this research is the lack of a singular term that defines what is described here as graphic notation, almost every author assigns a different name to it. Norman McLaren simply calls it “chart”, the Fleischer Studio coined

the term “Story Mood Chart” in reference to the particular interest in mapping the emotional trajectory they wanted their audience to have. In the case of Polish filmmaker Zbigniew Rybczinski’s, the name changes depending on what the graphic notation is intended for: In the film “Tango” (1980) he refers to it as “Planning the action’s progress in time”. Or in the case of the film “New Book” (1975), it is called “The trajectory of the story”, which in this particular case, aptly indicates the unconventional network-like structure of the film. For clarity, hereafter all of them will be referred to as graphic notations.

Two use cases of graphic notations in animated films have been identified here. The first one is intended to provide clear and rigorous instructions to the team members. The second is meant as a suggestion of the general structure and elements that compose it, encouraging team members to use a more intuitive approach to the ways in which the visual cues are translated into the final animation.

We can think of these two in comparison to the history of music scores in the Western tradition. The first approach aligns with the more traditional model of the score. Here, the composer defines with absolute precision how the piece should be performed by means of the formalized syntax in the music sheet. The second aligns with avant-garde practices of the graphic scores (Boutwell, 2012; Cage & Knowles, 1969; Evarts, 1968; Sauer, 2009). This latter started to emerge in the 20th century as a challenge to the relationships that the traditional score establishes between the composer, musicians and in consequence, the audience. Not as a fashionable rejection of the status quo, but as a philosophical and aesthetic exploration of what music could be when intuition and improvisation are provided a structure to operate under open

guidelines established by the composer. Theresa Sauer explains this way of thinking as:

Their philosophies encompass a desire to improve communication amongst composers, performers, and audiences, to develop a wholly different language, to encourage creative improvisation, and to challenge the way we understand music and sound. Some seek to create from the viewpoint of function, and others from the viewpoint of aesthetics. (Sauer, 2009, p. 10)

A parallel can be established here between the formalized syntax of music scores and the pursuit of functional data visualization to be an effective path to understand the information being presented. As Sauer points out, music scores are functional objects so the interpreters are able to execute with precision the task they are given, within these constraints, any interpreter of the score can reach a similar outcome, the music produced is a recognizable composition with subtle variations based on the orchestration and the musicians own abilities. A functional data visualization tends to operate similarly as we are all invited to interpret the same information that the visualization provides, subtle variations may take place based on our own expertise, but an effective visualization in these cases should be able to narrow down the margin of interpretation to the minimum. Sauer aptly argues that graphic scores depart from this to position the premise “from the viewpoint of aesthetics” (2009, p. 10), compromising the possibility of narrow interpretation and functionality, but precisely this challenge opens new paths for musical explorations. Data driven animation pursues this same type of challenge not only to animation practices but to the way in which we relate to data. By distancing itself from singular interpretations, our perception of data is ambivalent and as such suggests new ways in which we can relate back to the phenomena and the ways in which we communicate to it others aesthetically.

In the body of work by the influential American composer John Cage, we can find a good example where this shift is effectively articulated. For example, his piece “4’33’” composed in 1952, famously instructed the musicians not to perform any music, but to bring the instruments to the stage and wait. It was composed for any instrument since the performative act was to establish a codified setting in which music is normally played and in turn let the sounds in the surroundings become the music. Cage was exploring ways to remove his classical training as a musician to give way to his experience of the world as an artist that happened to make music. By rejecting the music score, or even the articulation of music with instruments as in 4’33’, Cage is establishing a creative process in which the flux of life, and the environment, are contained and articulated in an aesthetic experience for us to reflect on. 4’33’ is not simply the sounds of the environment as they are, but the transformation of them into the outlining of music as art to be experienced. He distances himself from conventional modes of making music but doesn’t remove himself from pointing us to what he considers evocative. In his essay “No Ear for Music: The Scary Purity of John Cage” (2008, pp. 261-279) Richard Taruskin clearly explains this affordance of the composition when he writes the following about 4’33’:

A work that is touted as a liberation from esthetics in fact brings an alert philosopher to a fuller awareness of all the constraints that the category of “the esthetic” imposes. Sounds that were noise on one side of an arbitrary framing gesture are suddenly music, a “work of art,” on the other side; the esthetic comes into being by sheer fiat, at the drop of a piano lid. The audience is invited—no, commanded—to listen to ambient or natural sounds with the same attitude of reverent contemplation they would assume if they were listening to Beethoven’s Ninth. This is an attitude that is born not of nature but of Beethoven. By the act of triggering it, art is not brought down to earth; “life” is brought up for the duration into the empyrean. (Taruskin, 2008, p. 275)

Taruskin is pointing out a key idea in the radicalism of Cage, he does not abandon music entirely but supports his work in the conventions that surround the event of listening to music. The ceremonial event of a concert, the revered place in which it takes place, our own disposition towards music as an experience, is left intact for us to become aware of its artificial construct once the event itself of music being played is subverted. This same idea was argued earlier in Chapter 1 in relation to the displacement of data representation to the art gallery or museum. The audience in these spaces establish a very different relationship to the objects presented here, we do not demand from art evidence of life and the world, but a path to experience it. The pursue of ambivalence, pose questions and result in contingencies that data driven drawings stands on, requires precisely this disposition towards the data representation. Therefore, these representations may reject formalized syntax and conventions, yet framing them within art is a fundamentally necessary context for the audience to be, paraphrasing Taruskin, “invited—no, commanded—” to “see” the data representation within the same predisposition as other art objects like a painting or sculpture that normally inhabit this space.

In his research attitude towards music, Cage focused his attention to graphic scores⁵⁰, for which he famously developed a procedure of writing music based on the ancient Chinese manual *Book of Changes* or *I Ching*. The manual offers a complex system of symbols under which chance events can be ordered, Cage would ask

⁵⁰ In 1969 he published an anthology of graphic scores in the book “Notations” which was the source of inspiration for the previously mentioned publication “Notations 21” by Sauer in 2009. Sauer’s book started as an inquiry on the influence of graphic scores in composers from the 21th century and the publication reveals that Cage’s and his colleagues approach to music scores remain valid and continue to develop in contemporary music.

questions about his composition to the manual and systematically proceed based on the answers. Marc G. Jensen examines Cage's approach to the *I Ching* in his article "John Cage, chance operations, and the chaos game: Cage and the I Ching" (2009) and explains how Cage was delegating and seeking information in a structured way to the *I Ching* for which he clarifies that:

An important element of the I Ching is the nature of its role as an oracle, and what it means to consult it. Rather than giving information about the future, the book is intended to produce a detailed examination of the present moment. Based on a system of holistic philosophy, the production of 'random' events is intended to elucidate subtle currents and tendencies in the universe. (Jensen, 2009, p. 98)

This procedure is closely related to data collection as a form of containing in a structured way the events produced by a phenomenon. The events registered in a dataset are in a similar way chance operations of how the world out there is operating in the present. Also relevant is the distinction that this procedure is not meant to produce information about the future, in other words, it does not produce an opinion in the same way we expect raw data to be. But the ability to access these chance operations in a structured form is what attracts Cage to work with it. Taruskin explains how this method became a clear articulation of Cage's philosophy towards art as:

Self-schooled in "a spirit of acceptance, rather than a spirit of control," he gladly acknowledged the incomprehensibility of his results—and of theirs as well. Catching at the ancient aura of the sublime and so being truer than his opponents to the impulse that brought forth Schoenbergian atonality, Cage maintained "that the division is between understanding and experiencing, and many people think that art has to do with understanding, but it doesn't." (Taruskin, 2008, p. 265)

Expanding on Cage's idea, data art is not only attached to a common drive to demand clear understanding from a data

representation as argued in the previous chapter, but even Cage proposes that art itself may impose the same restriction. His approach to musical composition and the use of graphic scores based on chance operations allowed him to liberate this constraint in the pursuit of new modes in which music, and in Cage's case life itself, is experienced. Data Driven Drawings are located within the same pursue of "acknowledging the incomprehensible" (Taruskin, 2008, p. 265) as the resulting place, and the graphic score tradition developed by Cage and his colleagues are seen here as a coherent method to develop procedures for developing these types of work.

Distinct from the history of music, in the case of graphic notations for animated films, we cannot clearly divide the two approaches to different genres or movements in the field. This is due to the fact that animation studies and practices do not depend so strictly on a tradition of using graphic notations in the same way that scores were developed for the practice of Western music. Therefore, we can only analyze them in the context of the particular strategic goals to which they respond: to explain with precision, or to provide a structure to activate intuitive interpretations.

The prime example of graphic notations used for providing strict guidelines in animation is the "exposure sheet". This method became popular in commercial animation studios from around the world as a way of formalizing into the production pipeline a clear form of communication between directors (or lead animators) and their secondary animators or camera operators. The exposure sheet details how the animation is constructed frame by frame by abstracting the directions into graphical cues. For instance, the camera operator can be given the animation assets to be shot and the exposure sheet will have labels for all the camera movements and the order of the sequences. In this sense, the exposure sheet

was not to be “interpreted” by the operator, but precise instructions to follow that were provided by it.

On the other hand, one of the earliest uses of a notation system in animation in which the goal is to establish a more open interpretation is, ironically, with the use of classical music scores as guidelines. Here, the score is a structural guideline for animation artist to interpret the musical language into a visual one. This method gained prominence with the rise of abstract animation and visual music in Europe with the early experiments in music and painting correspondences of Leopold Survage in 1912-1914, and later in the pioneering work of Walter Ruttmann, Viking Eggeling, Hans Richter, Oskar Fischinger and Len Lye during the 20’s and early 30’s (Brougher & Mattis, 2005; Guldemonnd & Keefer, 2012; William Moritz, 2004; Rogers & Barham, 2017; Russett & Starr, 1988).

The type of relationship that these authors established between music and image provide a clear premise of what the graphic notations are able to afford for animation artists. In the case of visual music, the goal has never been to “illustrate” music, but to explore creative translations of musical constituents into visual ones, and to find aesthetic correspondences between the musical rhythm, pitch, and orchestration, with visual elements like shape, color, and composition.

A visual music artist does not read the music score as a set of instructions to follow but as an indication of how the musical elements are orchestrated in a composition in order to explore potential analogies in the visual realm. The scores for a visual music artist are the space for analysis in order to pose the right questions to be investigated in the representation.

Holly Rogers and Jeremy Barham report that we can trace this strategy back to Walter Ruttmann's film *Lichtspiel Opus I* (Lightplay Opus I) from 1921 which "is thought to be the first screening of an abstract animated film seen by a general audience anywhere in the world." (Russett & Starr, 1988, p. 40). Ruttmann collaborated with the composer Max Buttin who created an original composition before the film was made. In the music score presented in Rogers & Barham's book *The Music and Sound of Experimental Film* (2017), we can see that Ruttmann drew exploratory sketches on top of the score, indicating that he was carefully studying it in anticipation of his animation.

Not much is known about the collaboration between Buttin and Ruttmann, but the surviving score with a little drawing by Ruttmann suggests a close correspondence of sound and image, as does the overall rhythmic structure of both (Rogers & Barham, 2017, p. 35)

This methodology was later adopted by animators working both with abstract and figurative⁵¹ images and is still being used today in visual music practices at large. But the type of graphic notations to be examined in more detail here, are the ones that are created originally within the needs of an animated piece. Expanding on the traditions of visual music and graphic scores, the authors examined in the following case studies find the need to develop their own graphic notations with a unique visual grammar. The result is an incredible array of styles and particular nomenclatures that may be

⁵¹ The German animator Lotte Reiniger, who was close friends with Ruttmann, borrowed this approach in her works based on classical compositions during the 30's. In a short documentary by John Isaacs titled "The Art of Lotte Riniger" (1970), she reconstructs her workspace of the piece "Papageno" based on the opera "The Magic Flute" by Mozart. Here we can see that she places the music score prominently in the side of her animation table, indicating that she was following the cues on the score to interpret them in the animation of her silhouettes. The rhythms and musical flow in which these characters' act and move, show a close correspondence to the musical piece they are paired with.

often obscure and difficult to decode. But this lack of unity does not mean we cannot find methods for studying them.

The proposed methodology for studying graphic notations

This research proposes four types of signifiers that have been identified in all graphic notations for animation which are: time, resemblance, order, and proportion. The last three are borrowed from Jaques Bertin's theories proposed in his seminal book *Semiology of Graphics* (1967)⁵². But since time is not always a property represented in data visualizations⁵³, Bertin does not prioritize it in his theory. In the case of animation, time is inescapably a property of the object and therefore added here as a fundamental signifier.

- **Time:** The visual representation of the time structure, here we find examples of linear structures by horizontally or vertically disposing of the graphics. But also others such as circular, elliptical or network like structures that explore time as a nonlinear phenomenon.
- **Resemblance:** These are often abstract shapes that represent an element of the film. For instance, a character is represented by a particular shape or color and we see where in the time structure it appears. By resemblance, we can see

⁵² See section on Bertin in Chapter 3.

⁵³ For example, a simple bar chart displaying an ordinal dataset of how much A compares to B, does not have time as a fundamental property.

where in time this item is placed and how it relates to the other elements of the film.

- **Order:** The order denotes the particular orchestration of the elements in the film. It exposes the rhythms, durations, and/or spatial compositions.
- **Proportion:** Represented by variations in size of the different elements, the proportion denotes a particular intensity, often signifying emotional fluxes, or variations of a character's particular action.

These four specific signifiers are proposed here after the empirical study of a series of examples for which both in the graphic notations and the final films were accessible. Since each of them has a unique visual nomenclature and structure and since no other academic study was found that proposed a methodology for studying them, this research proposes the above-mentioned signifiers to unify their study. Following is a description of the empirical process developed to study the notations with their corresponding film and how these steps led to the proposed signifiers as the fundamentally unifying elements that they all share. As a foundational methodology for studying graphic notations, the main focus is on identifying what properties they share over those that make them unique. The particular styles, tools used for producing them, who specifically was it intended for, how and when was it used during production and the effectiveness of the notation, are things discussed in the findings discussed later in the chapter, but initially left aside in the formulation of the signifiers. For these signifiers need to describe the internal mechanics of any notation and not what they mean, they should describe the form so any content can be applied to them in the future.

The central procedure established to identify the signifiers was to translate the notations to data. This process of abstraction was useful to discriminate the elements that pertain to the internal mechanics of the notation because they can be organized in a database as variables. The resulting structure of the dataset should, in turn, describe the original object logically.

For example, any systematic use of a graphical element which describes the pattern of events that the author is trying to register is defined in the dataset as “event”. For example:

Figure 9. Detail section of the graphic notation used in the film “Tango”, from “Tango – 1980 Lodz, Planning the Action’s Progress in Time”, by Rybczyński, Z., 1980.

Figure 10. Detail section of the graphic notation used in the film “A Chairy Tale”, from “Visual Outline of A Chairy Tale, prepared by McLaren for Ravi Shankar”⁵⁴, by McLaren, N., 1957.

⁵⁴ McLaren describes the need for the notations as an effort to provide the composer and musician Ravi Shankar an extra tool to understand the already finished film. The musicians were given both the chart and the film to develop the music in this particular case. (McLaren & McWilliams, 1991, pp. 61 - 64)

In Figure 8, each horizontal line describes a character and the vertical line when they enter and leave the frame. In this case, an event in the database is created for each vertical line. While in Figure 9, the author is not describing when a character enters and leaves, but the type of action they are performing, his system for describing when a particular action begins and ends is discernable by the change of line or horizontal marks. How they draw it stylistically is not described in the dataset because they are all unique and as seen in the examples above, are there to inform when an event begins and ends. Yet, all of the notations studied offered a very easy to identify cue for creating blocks of events in the database. These cues are traceable because the notation system establishes a logical progression of a particular abstract shape, the indication that a shape repeats and progresses somehow in the system is what is referred to as “resemblance”. Without the ability to recognize where a particular element is placed before and after, by resemblance, no logical flow of the events can be extrapolated.

Within these blocks of events, we can describe the position inside the plane as spatial coordinates of “x”, “y” and/or “z” in numerical values refereeing to pixels in the digital image. With the separation of blocks of events and their spatial location, the structure of the database reveals the patterns implicit in the orchestration of the film that allow the study of the notation. Take for instance the following sample corresponding to Figure 9:

Table 3
Sample Data for the Graphic Notations of “A Chairy Tale”

Title	Y (in pixels)
Timid Touching	890
Comic Attempts to Sit	1000

Tries sit - chair moves to right	1020
Tries sit - chair moves right again	1040
Tries sit - but chair moves left	1055
Tries sit - (confused movement of head)	1095

The very simple structure has enough clues for us to infer how the notation was made. Considering that each event is contained in each row of Table 3, the most noticeable element is the name of the coordinates variable. In this case “y” which means that the notation is visually describing time vertically. In the case of Figure 8, the name of the variable repeating throughout is instead “x” which means that time flows horizontally in the image. In more complex structures⁵⁵, the event could contain “x”, “y”, “z” and “r” where “r” stands for rotation. In theory, and so far applicable to the notations found throughout this research, all notations describe time by systematically placing events in the plane. We can see in the data that the values of the “y” variable increment linearly, which means that time is described visually vertically and from top to bottom. The time signifier can always be deduced from the logical flow of the coordinate variables. Then by comparing the distance from one event to the other, the data can also indicate the third signifier of order. In the sample above the distance from the first event to the second is of 110 pixels, while the second and third event are separated by 20 pixels. This tells us that the event titled “timid touching” is much longer than the following “Comic Attempts to Sit”. What the titles suggest can be interpreted and performed in widely different ways, but the distance that separates them tells how the author decided to give “order” to such events.

⁵⁵ See Figure 11 in case study D for an example of a circular time structure. The resulting database included a rotation variable.

Lastly, during the translation of the notations to data, an attempt was made to give a value that could describe the discrete changes in the proportion of each figure by including the variables of width and height, but this variable was not consistently showing up in all of the notations. Figure 8, for example, does not display any shift in proportion, they are all waving lines of the same height and thickness, or in Figure 9 the width and height were not necessarily descriptive of what the author is suggesting with the graphical changes. But, during the analysis presented in the findings of the case studies, it became evident that drawing attention to the proportional changes of the shapes was also critical for understanding what the author is intending to represent. In some notations, it was the main signifier as is the case of Figure 9 in which the height of the line is describing the intensity of the actions throughout the film.

Figure 11. Detail section of the graphic notation used in the film “Mr. Bug Goes to Town”, from “Story Mood Chart”, by Fleischer Studio, 1941.

Recognizing that the sample under which this methodology has been tested is small, only 4 notations this far, the author expects these signifiers to be revised as more notations are tested to it.

Custom software for studying graphic notations

During the first examinations of these artworks, looking separately at the graphic notation and the film made it difficult to infer if they both had strict correspondences. To aid this limitation, a series of interactive web applications were developed to match the films to the corresponding elements in the graphic notations. Using the

databases described above, the application is designed to match the time in the video playback to the structure of the database. The basic logic of the algorithm is to compare the length of the video to the spatial length of the notations. Once the video starts playing, the application is able to search inside the database for the elements that correspond to the particular time of the video and draw a red line superimposed over the graphic notation. In this way, the viewer can clearly compare both elements side by side with great precision.

After manually creating all of the notation databases⁵⁶ with the same conventions and logical structure, a JavaScript library was written so the application is able to build the notations automatically if provided the three required elements: a video, the database, and the digital 2D image. Very often upon the first review, the application would display some anomalies that usually correspond to human errors in the original notations. For example, it was common to find in this sample that the authors would miss or skip a step in the grid and therefore time and order cannot be matched by the algorithm anymore. The red line would suddenly fall behind or skip ahead of the video, but this can be easily corrected by manually compensating for the error back in the data.

These applications using the custom JavaScript library are presented on the website: <http://www.dddrawings.com/notations/> that was made complementary to the present document. This particular approach of studying graphic notations through an interactive application is an original contribution to animation studies developed here. The author expects to continue developing the

⁵⁶ The final databases used in the applications can be consulted in the following repository: <https://github.com/lcgonza/dddrawings/tree/gh-pages/data/notations>

application as more case studies are implemented into the workflow and analyzed under the proposed signifiers described above. In this particular methodology, the practical and technical processes are intricately related to the theoretical articulation, therefore, both aspects of the methodology should continue to be examined and reviewed in parallel until a wider sample is tested. Due to the amount of manual labor that is required to expand the sample and considering it is not the central pursue of the present research, this introductory examination is sufficient for the argument presented here. The author considers that the ideas presented in this chapter deserve a wide encompassing research endeavor of its own.

The study was developed in Singapore, at the Ph.D. laboratory of the Arts, Design and Media program of the Nanyang Technological University in Singapore during a period of 8 months in 2013. The software was written in JavaScript coding language using a Windows workstation.

Results of the case studies

Case Study A: Fleischer Studio – “Mr. Bug Goes to Town” (1941)⁵⁷

The second and last feature-length film from the Fleischer studio was released in 1941 at the end of a very successful life of the commercial studio. Their main competitors where the Disney studio and one of the most significant differences between the two is that the Fleischer’s relied heavily on gags and wonderfully extravagant

⁵⁷ See interactive application at: <http://www.dddrawings.com/notations/mr-bug-goes-to-town/>

character animation, while Disney was more focused on the overall story development of the characters (Bendazzi, 2016, pp. 115-118, Vol I.). While the gag after gag approach provided the Fleischer studio great success in short formats, this was not as successful in their feature-length films. They noticed how Disney feature films received praise for their emotional impact on the audiences and in an effort to parallel this, they developed a graphic notational system they called the “story mood chart”.

The underlying strategy in the story mood chart was to map the emotional trajectory they want their audience to have throughout the film. To do this, they created a histogram of two corresponding variables: “degree of emotional pace” and time. In the left of the chart, we can see that they encode the “degree of emotional pace” into numbers from 0 to 100 and from the corresponding notes at the bottom of the chart, we can infer what they meant by it. For example, the lowest lines of the histogram in the “fourth sequence” match with the note: “Hoppties deepest sorrow!!!”. In contrast, the highest lines in the “fifth sequence” say: “Fastest most exciting tempo of picture selling danger distress of earthquake ... Total reversal pictures, deepest sadness WHAM SUCCESS, highest happiness, great contrast!!!”. The resulting graphic notation shows a particular interest for modulating these different emotions throughout the story by spacing highs and lows evenly, placing the lowest towards the beginning of the second half of the film and providing the climax with the most emotional intensity before the ending.

The four proposed graphical signifiers are represented in this graphic notation as follows:

The “time” structure of the film follows a linear storyline, time, in this case, is represented linearly in the histogram by starting on the

left and ending on the right. The horizontal placement of the graphical elements denotes the linear progression over this timeline.

“Resemblance” in this case is the graphical cues that describe each instance of the emotional pace which is represented by straight vertical lines. The repetition of the same graphical element shows clearly that they denote the same variable (emotional pace). Since the intention of this graphic score is to exclusively map the emotional trajectory, the use of a singular graphical element makes the overview of the emotional pace transparently displayed.

The “order” of each individual line is strictly placed one after the other in the horizontal plane, which reinforces the visibility of the linear time structure. At the same time, the distribution of the different high and low blocks shows an intended orchestration of the emotional journey. There is prominence to higher emotional paces throughout the story and the low sections often anticipate an emotionally heightened story moment. This awareness of the potential physiological effects on the audience shows how they were trying to subvert their previous use of gag full stories and in change, modulate the different emotions strategically.

The “proportion” of each vertical line expresses the intensity of the particular emotional pace within the scale of 0 to 100. The variations of size for each line indicate, with clarity, were in the general scale that particular moment should be considered. For example, towards the end of the “second sequence,” we can see a deep change from lines near 75, plummeting down below the scale of 25. In the film, we see this in the change from a comedic scene in which the burglars describe to their boss how the bugs community were having a party, to a darker scene where the boss starts to conceive a plan to sabotage the bugs’ happy life.

By following these four visual signifiers in the graphic notation, we can see how it could have been used by the directors, animators, and composers to create their work based on this general map of the emotional journey. Following the movie in relation to the graphic score shows that was translated with great precision, particularly in the music score and performance of the characters.

Lastly, this early example of an original graphic score is particularly meaningful because of the context in which it was devised. The fact that it was incorporated into the production pipeline of a commercial feature-length animation is important because it expands the affordances of graphic notations. It would be a mistake to associate the use of graphic notations exclusively for abstract or experimental work, because, as shown here, the Fleischer studios found it to be a useful scheme for a narrative, character-based and commercial product.

Case Study B: Zbigniew Rybczyński - “Tango” (1980)⁵⁸

In this widely praised⁵⁹ film by Polish filmmaker Zbigniew Rybczyński, we have an example of a profound reflection on the representation of time and space. “Tango” is a short film in which a series of characters enter and leave a room, do not seem to be aware of the others’ presence and as soon as they leave the room, enter it again immediately. The progression of these repeating actions starts to reveal that what we are seeing is the amalgamation

⁵⁸ See interactive application at: <http://www.dddrawings.com/notations/tango/>

⁵⁹ “Tango” is his best-known work which received the “Academy Award for Best Animated Short” in 1982.

of multiple time spaces at once. As the title suggests, the film is more a choreography than the articulation of a story or anecdote. In this sense, it becomes clear from the graphic notation that Rybczyński was orchestrating the characters in a musical sense, rather than in a linear story fashion.

I argue that it is in these types of narrative explorations where graphic notations prove to be particularly useful. He also developed story-boards (Rybczyński, 2009, p. 156) which are in themselves stunning drawings, but very difficult to follow, and a potential literary script may only be a series of nonsensical series of descriptions of people entering and leaving a room. The graphic notation on the other hand clearly exhibits the conceived musical orchestration of the film and is clearly understood.

This is how Rybczyński articulates the four graphical signifiers in his graphic notation which is in terms of data visualization a “time-series diagram”:

The “resemblance” signifier is represented by a series of waving lines distributed horizontally, and each of these lines denotes a character of the film which is listed vertically. He also uses small vertical lines that interrupt each of the waving lines to show where the action starts and ends. Different from the lines we saw in the previous example, the waving lines do not define any characteristic of the action or intensity of the mood, all characters are treated equally, and what they do on screen is of secondary standing to give importance to their temporal and spatial relations. In this sense, the use of “proportion” is in the duration of these lines across the plane, and indicate when a character appears for the first time and when it leaves.

“Time” is operating in two ways for this film, one that adheres to the film format itself which is inevitably linear (the film starts in second 0 and ends 8 minutes later). This linear progression of time is represented horizontally from left to right, and he uses graph paper in which 5 squares represent 1 second of film. The second manifestation of time, which is the conceptual exploration of multiple times collapsing into one, is represented by the “order” in which the different lines are displaced vertically. We gain clarity of this complex relationship by visually separating these two parallel manifestations of time, one in the horizontal plane and the second in the vertical one. This visual solution to an intricate problem is where this particular graphic notation proves its powerful affordances, as it disambiguates an otherwise complex issue that may be on the directors’ mind, but with the graphic notation, it is clearly represented and therefore easier to follow as a guide for a production team.

This is only one of the many different graphics notations that Rybczyński has created throughout his career. A fascinating collection of these is contained in his book “Traktat O Obrazie / A Treatise on the Visual Image” published in 2009. Here we can find further explorations of unconventional representations of time and space and his ability to draw parallels between conventions used in data visualization and their potential to clarify complex narrative structures. Most of the graphic notations are variations of “time-series diagrams”, except for the one used in the film “New Book” (1975) which is a map. The film shows a series of parallel scenes in a city, presented all at the same time on a grid of 3x3 split screen. The graphic notation stems from a cartographic approach that shows, simultaneously, the network of relations that each event has in the city. The book is full of seemingly complicated diagrams,

which at a closer look, reveal to be very clear graphic notations that elucidate the abstract thinking process of this outstanding artist.

Case Study C: Norman McLaren – “A Chairy Tale” (1957)⁶⁰

From all the examples of graphic notation that I have been able to find so far, I consider the ones made by Norman McLaren to be the most sophisticated in the use of “resemblance” and “order” signifiers. Throughout his body of works, McLaren displays a great mastery in the use of abstract images, visual music and character animation alike. This eloquence for communicating with a purely visual language is demonstrated in the graphic notations he created for “A Chairy Tale” (1957), “Vertical Lines” (1960), “Canon” (1964) and “Mosaic” (1965)⁶¹.

In the case of “A Chairy Tale”, the “resemblance” signifiers are a series of interplays between two morphing shapes that indicate each of the characters of the film, the man is represented in black color and the chair in white.

The “ordering” of these shapes throughout the graphic score shows the relationship that these characters established in the story: the man is trying to sit on the chair but it does not let him do so. For the first half of the film, the man is making different attempts to sit

⁶⁰ See interactive application at: <http://www.dddrawings.com/notations/a-chairy-tale/>

⁶¹ Unfortunately, these graphic notations are not easily accessible. The one for “Vertical Lines” is featured in the book *Art in Motion: Animation Aesthetics* (2007) by Maureen Furniss. The others were featured in the National Film Board of Canada’s website, which had a special section dedicated to McLaren’s archive. But this archive has since been removed and no longer publically accessible, which means that for accessing them it is necessary to visit their archives in Montreal.

on the chair, making him our focus of attention as we follow his thought process. In the graphic notation, we see the dark black lines representing each of his failed attempts. He then grows frustrated and decides to sit on the floor instead, opens a book and disregards the chair. At this point, we see in the graphic notation that his line turns from black to a lighter gray and the white line of the chair goes to the center. This is shown in the film by the change of dynamics between them, where the chair is now trying to call for his attention. From this point forward we can see that for the moments in which the man is trying to sit on the chair, the black shapes are in the center, and when the chair seeks attention, the white ones turn central. When the man finally discovers that what he needs to do is let the chair sit on him, we see in the graphic notation a newly formed abstract shape that combines both colors into one. Represented this way, they had finally understood each other.

The film presents a linear storyline and the “time” structure is implied vertically starting at the top and ending at the bottom. McLaren uses graph paper in which every square represents one second of the animation.

Lastly, the “proportion” of each shape indicates the particular intensity of the action that each character is performing. For example, every time the man makes an attempt to sit on the chair, the width of the shape tells us how far down he reaches into the chair, very wide for moments in which he gets really close, and smaller ones for when he barely attempts the action. Also, there is a moment in which the man invites the chair to dance the tango and the shapes turn wavy and loose to denote the levity of the moment. The overall clarity that one can perceive by comparing the abstract shapes with the final film, shows us McLaren’s ability to translate

from one language to the other, and demonstrates the enormous potential for displaying in a graphic notation a wide range of interacting elements of a film.

One important clarification about this particular graphic notation is that the one we see here was made after the film was shot and edited. The one produced before the film⁶² is a much cruder version of this one, without all the intricate interplays of shapes and color distinctions. Nonetheless, this one was produced to provide the composer Ravi Shankar with a visual score for him to compose the music. McLaren describes this interaction as:

The duration of sequences, episodes, actions and gestures was indicated precisely on this chart, by the use of colours, diagrammatic marks, names and numbers. I then spent an afternoon screening the film many times for Shankar and his percussionist. Between every screening we would identify each sequence on the chart. After about a dozen screenings, they were both thoroughly acquainted with the film and the chart as it related to the film. They required three weeks to evolve the music based on the chart. (McLaren, 2006, p. 17).

This shows us that McLaren believed he was able to communicate something unique through the graphic notation. The film was done and Shankar could have created the music based on it. Also, they could just speak to each other to come up with the ideas for the music. But McLaren finds the need to express what he is thinking through the abstract graphics in this image. He is teaching us that there are a series of complex thought processes that graphic notations afford to communicate that other languages might not be able to do.

⁶² The one developed in pre-production is available at the "Collection Cinematheque Quebecoise" in Canada.

Case Study D: Juan Camilo González – “sisisisisisisisisisisi” (2011)⁶³

For this case study, the author uses a short animated film of his own creation which was developed during his MFA studies at the University of Southern California. During its preproduction, a graphic notation was developed in an intuitive exploration of visually preparing the film. In light of the present research, this notation has been revised in the context of the proposed methodology. The work developed during this research on top of this previously made film was exactly the same as the other case studies which was to translate the notation into data and implement it into the custom web application that tests its correspondence to the final film. Describing the personal motivation for developing the graphic notation of this film, should help to clarify the connection between the premises established in a traditionally animated film and the current Data Driven Drawings proposal.

In this short film, the goal was to examine a particular dynamic of the relationship I had with my father. I was always very polite towards him, he would never tolerate that I challenged his opinions and I always avoided doing so. After his passing, I had the feeling that our relationship could have been better if I had the courage to be more open about my opinions, even if that meant disagreeing with him from time to time. Furthermore, I believe this behavior continues to be an issue for me when I relate to others because I always make an irrational effort to be very polite. The result of this dynamic is an always growing tension inside me: I want to express

⁶³ See interactive application at:
<http://www.dddrawings.com/notations/sisisisisisisisisisi/>

This satisfied the goals of the film because it showed a clear map of how the growing tension would be orchestrated in time, it also described where in the paper characters should be drawn and indicated the overall circular structure of the film. This guide carefully followed throughout the year of work and at the end, graphic notation had been translated into the final film. In Figure 12 taken from the last frame of the film, we can see that the overall shape of the graphic notation is the same underlying one in the resulting image.

Figure 13. Graphic notation and final frame superposed, from “SiSiSiSiSiSiSiSiSiSiSi”, by González, J, J. C., 2011.

The proposed graphical signifiers also apply to this graphic notation in the following way:

The repeated graphical unit of “resemblance” was a series of cubes or cages. These were associated with every instance of a character

performing a polite gesture which is like being in a cage and the accumulation of cages entraps them.

The “time” structure of the film, as described before, was circular as a way of showing the overall repetition of this behavior in being portrayed. This circular representation of time is shown in the distribution of the cubes starting in the center-left side of the screen, expanding outwards towards the right and lastly returning to the beginning.

The “order” in which the cubes are displayed represent the way the characters will move in the space of the paper. Following the circular structure, the goal was to have small deviations from a perfectly circular trajectory and a sense that the characters were always getting closer to the viewer. Also, there is a short moment in which the cubes start to get very small until they become a waving line before they grow again. This anticipated the climax of the film with a short moment of serenity, in order to give the viewer space to breathe before the most intense moment is presented.

Lastly, “proportion” is denoted by the size of the cubes which were there to indicate the intensity of the gesture that had to be animated at that moment of the film. This particular element proved particularly useful in a long production like this one. The production took around one year to finish, and in this process, considering that no script or story-board have been produced, it would have been easy to forget what was supposed happen within the larger context of the film.

From graphic notations to data visualization

The study of graphic notations in animation deserves further investigation beyond what has been discussed here. This introduction to the issue presented itself as a sub-argument of the larger investigation of Data Driven Drawings, and therefore, it can only be discussed within a narrow scope. To initiate the discussion, a unifying term is proposed here: “graphic notations” to encompass the different approaches similar to the ones presented and discussed above. At the same time, four graphical signifiers have been proposed aid similar studies that seek to have a starting ground to study the structures of such objects. We observe that the conceptual and aesthetic traditions of abstract animation, visual music, and graphic scores, are transferred to the creative use of graphic notations for animated films. And through the analysis of the different case studies we identify the potentials of graphic notations: the ability to translate complex thought processes to clear visual presentation, the affordance to articulate and orchestrate a wide variety of properties of a film that may escape traditional methods like the script or story-board, and the possibility to contain the idea and structure of the film and yet open a space for exploration and poetic associations during production.

For Data Driven Drawings, the graphic notations developed to follow the same signifiers but with the difference that the visual articulation of the elements is not drawn systematically and intuitively by the author, but emerge from the structure of the data. Within the terms presented in Chapter 1, these are functional data visualizations, they seek to allow for clear reasoning over them and are seen in the creative process as a utility. In relation to the ideas presented in this chapter, they are not merely descriptive objects

but graphic notations that allow for creative interpretations on how the information in them is ultimately presented to the audience.

In relation to the overarching question 2 of this research, what is at stake for the animation film form is that it draws attention to the existing practice of preparing a film with graphic notations, such tradition forces us to observe that index of reality through an object that in itself is representing. One could claim that a sound recording used in an animated documentary is also a representation, but in these cases the index is tangible can be presented back to the audience. In data representation, we have to construct everything because the data in its raw state is imperceptible and amorphous. If this index of reality is to be presented somehow to the audience, it requires a process of an embodiment for it to be perceptible.

In the next chapter, the discussion returns to the main objective of the thesis which is to build a defense for subjectivity in data art. In the logical progression of Data Driven Drawing project, the ideas presented next arise precisely after the exploratory analysis through data visualizations has been established. At this point, the author is acquainted with the data at hand, has inferred meaning from it and potentially strategized ways in which it could be represented. What arises is a question of how the findings at this stage are either clearly communicated to others or transformed in its final articulation.

Chapter 3: The Artist Dilemma

The ideas presented in this chapter assume that an artist has already established some level of understanding of the data to be used in his/her artwork. In the previous chapter, an approach for investigating the data was proposed by considering functional data visualizations as graphic notations. Going through those exploratory analysis allows for then concentrating on what is presented here as the general dilemma confronted by an artist in the process of conceptually defining “how” to represent reality through data. This issue relates to finding where within the conceptual spectrum of intentions, proposed earlier as either functional or evocative data representation, one wants to be in.

The goal of the practical work developed here is to use data as an index of reality as a way to interrogate issues that are rooted in personal trauma, and in this sense, what they seek to communicate is not the mimetic representation of the author’s biographical life, but the subjective experience that emerged from the events that took place in his life. Bringing forth a personal view not as a way of focusing the viewers’ attention towards the author, but towards the experience in itself, one that can be personal to others as well, or expose the ways that violent realities affect our human condition. This presents itself more in the form of a question rather than an affirmation. What is at stake in the rejection of a clear exposition of reality and allow for the subjective experience to mediate this exposition to others?

This question is confronted and problematized in relation to the ideas proposed by authors that defend two opposing views: on one hand, we have the dominating views on data visualization academics, who see the representation of reality as a functional issue that should reveal the complex phenomena of the “world out

there” for reason to be invoked and attain concrete knowledge, in other words, the beautiful representation of reality. The opposing view is discussed here in light of animated documentary theories that propose a mediated representation of reality through the subjective experience, in an effort to validate this approach in the wide epistemological project of documenting reality. The original artworks that were developed for this research affiliate with the latter point of view and are discussed in more detail in the following chapter.

Functional Data Visualization, the case against subjectivity

The academic research on data visualization as a functional art has a strong history. For an account of the historical trajectory, Friendly and Denis made an outstanding research titled “The Milestone Project”⁶⁴ (2001) and related publications (Akerman & Karrow, 2007, pp. 207-253; Friendly, 2008; Friendly & Denis, 2001) tracing the long history of the medium dating back to early cartography such as the Komya town map (6200 BC), all the way to present times where computation and representing data permeates a wide variety of disciplines. Perhaps missing from this exhaustive collection would be the tally sticks from central Africa during the Upper Paleolithic, which reveal that humans have been counting and physically recording data far before the development of mathematics.

With the advent of computation, we have an enormous expansion of the possibilities to collect and process complex data, which today

⁶⁴ The Milestone Project is presented as an interactive archive that can be freely consulted in: <http://datavis.ca/milestones/>

gives access to a wide variety of disciplines to embark on this task from each of their perspectives. This increased interest and use have led to the necessity of a more conceptual understanding of the medium. The most influential figures in this intellectual development are the cartographer and graphic designer Jaques Bertin, mathematician John W. Tukey and statistician Edward Tufte. I will now describe their contribution in relation to how the visual representation of data is commonly produced today, and at the same time describe how their ideas either transfer to the approach proposed here or need to be challenged.

Jaques Bertin – Formalization of a visual grammar

Bertin's major contribution is contained in the book *Semiology of Graphics* originally published in 1967⁶⁵, in which he formalizes a theory coupling perception to visualization. Bertin rigorously catalogs the different types of variables we find in data and the potential graphic encodings that are best suited for them. He argues that there are three types of visual cues that are fundamental to the process of signifying in a graphic what is operating in the data:

Resemblance, order and proportion are the three signifieds in graphics. These signifieds are transcribed by visual variables having the same signifying properties. (Bertin, 2011, p. 177)

The “visual variables” Bertin refers to are a set of conventions that we now normally see in data visualizations such as “primitives” (Manovich, 2010), dots, lines, geometric shapes. Also, the use of

⁶⁵ This edition is in French, so the edition consulted here was (Bertin, 2011)

textures and color to distinctively recognize differences in the patterns being presented. In Bertin's theory, he creates a taxonomy of visual variables and argues which are "better" to represent specific types of data. Listing these conventions and their applicability to this research extends beyond the scope of the present discussion, but the purpose of bringing forth Bertin's theory here is to say that his ideas and methods have provided a solid ground for the current conventions we use today in data visualization. Even if they restrain the type of visual signifiers to simple shapes, their applicability has proven to be robust and abiding by the field.

Let's take for instance one the graphic notations⁶⁶ made for the project "Rite of Passage" in which Bertin's ideas are applied:

⁶⁶ See image graphic notations at: <http://www.dddrawings.com/lab/spikes/>

Figure 14. Graphic notation from data visualization of 2003 data of seismic activity and terrorist attacks, from “Rite of Passage”, by González J., J. C., 2013.

Each event in the graphic is represented by a triangle of a particular color (resemblance) and by repeating the same shape for each of the events, one can visually discern that they represent something that is similar in nature as the others of the same shape and color. Another variable in the data is time, which here is represented by the distribution of the triangles around a circle (order). Appealing to the visual association between circles and time - clocks, sun, moon, the path of shadows through the day -, this particular ordering is intended to signify the time of the year in which the event happened. Lastly, the size of the triangle denotes the intensity of the event (proportion). The resulting image is transparently applying

Bertin's core concepts in order for it to denote the information of the dataset.

Bertin's work has inspired the further standardization of a graphical language to represent data (Mackinlay, 1986; William S. Cleveland & Robert McGill, 1984) and modern data visualization software and tools such as D3, Gephi, RStudio or even the graphical possibilities of the widely accessible Microsoft Excel, are heavily influenced by these standards to the point that it is possible to automate, to a certain extent, the process of visualization. For example, let's say we have a clean dataset, organized in rows and columns in Microsoft Excel. This software allows us to select these values and create different types of charts: pies, columns, scatter plots, etc. It also offers the "best" possible chart for the data at hand through the "recommended charts" option. This will recognize the type of variables in the data and match them with the corresponding type of chart that better fits it. This automation is an extension of Bertin's rigorous proposals, and in many cases, the resulting image is good enough to recognize the patterns to fulfill the functional task.

The consequence of Bertin's ideas and its articulation into modern visualizing software is the possibility of having a common grammar and the homogenization of this visual language. This is not a problem in certain disciplines in which the goal is to produce functional visualizations. For instance, scientific journals are filled with data visualizations under common standards allowing specialized readers to understand the information presented without having to develop new abilities for decoding the graphics. In artistic data visualizations, these standardized graphics may not always conform to the research goals of the project. Rejecting the prevailing forms of a particular visualization technique is ascribed to the conceptual and aesthetical needs of the idea pursued, or

what is intended to communicate. As described in Chapter 1, Florence Nightingale was a pioneer of this disruption with the “Rose Diagrams” in which she compromises the “best” type of chart and invents her own. Not to obscure the information, but to create an emotional engagement with the viewer before appealing to their reasoning.

Andres Gaviria introduces this issue into the specific field of artistic data visualization with a persuasive argument in which he proposes a distinction between “Genre Art” and “Research Art” (Gaviria R., 2008, p. 481). To draw a parallel with Bertin’s contribution, visualizations in “Genre Art” share formal and aesthetic qualities that with time become standard of a particular movement. On the other hand, visualizations in “Research Art” may not always exhibit recognizable visual signifiers that respond to the formal characteristics of a movement. The singularity of the pieces in this category also respond to a need for experimentation and develop new forms that contribute to the general knowledge of the field, even if they are challenging to the viewer. Gaviria concludes that visualizations in the latter category may not comply with a homogenous form, but respond to the research needs of the artist:

Artistic information visualizations operate through visual metaphors that need not be easily decipherable nor aesthetically pleasing so long as they are reflectively interesting and in compliance with research’s dictum that they be innovative. Knowledge, in this case, is reached obliquely, through associations that may be elaborate and complex and not at first distinctly understood, but whose theoretical implications are, in some sense, significant to the field or related topics. It is, therefore, clear that artistic information visualization projects are not generic but specific in form and intent (Gaviria R., 2008, p. 482)

The research goals established in this thesis led to the use of non-generic graphics. In most cases, using drawings made with analog techniques such as ink on paper as an experiment to bridge the

personal intuition and subjective view, with the schematization of data. The themes of each piece explored the association of contrasting ideas: death and lightness of the body in “Dibujo Defunciones”, earth’s tremors of seismic activity and explosions with a flock of birds murmuration in “Rite of Passage”, types of violence in Colombia and rivers in “N. N.” (Chapter 4). These premises in which violence and “beauty” collide, established a space for exploring in the forms of data representation, the use of personal drawings. Even if Bertin’s core ideas of “resemblance”, “order” and “proportion”, persist in Data Driven Drawings, their goal is to evoke through the metaphorical association of opposing signifiers, rather than faithfulness to simulate the phenomena that the data signifies.

John W. Tukey – Exploratory vs confirmatory data analysis

Even though Tukey is specifically debating with the scientific community, his ideas on exploratory data analysis are critical observations that validate some of the methodologies that operate in artistic works with data. As the name suggests, the exploratory approach is an invitation to delay any assumption of what the data might be able to reveal. In contrast with the confirmatory, in which he argues that giving too much emphasis to statistical schemes and the effort to prove an already formulated hypothesis, leads to obfuscating potential issues worth investigating.

‘Exploratory data analysis’ is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there. (Tukey, 1977, p. 806)

This “attitude” is not an invitation to delay the visualizations, but pleas for the use of visualization to investigate what may be hidden in the data. What Tukey proposes is that a visualization is not only a tool for explaining to others but also the means for understanding our own questions about the phenomena at hand. This intrinsic potential for a visualization to pose questions is what the evocative approach proposed here intends to exploit. Furthermore, the evocative representation aims to forestall the confirmation and invites the audience to explore the significant problems of the phenomena by means of their own intuition.

The exploratory methodology proposed by Tukey has been very influential on the use of computers, development of interactive software and automation of the visualization process⁶⁷. Going back to Benjamin Fry’s description of the different stages of a visualization described in Chapter 1, we can think of the exploratory method as an iterative process, moving back and forth between representation and the earlier steps of capturing and mining the data. In this context, the use of computers, and especially automation through specialized software, enhance our ability to efficiently iterate as many times as needed and therefore support the exploratory premise.

Additionally, Tukey’s methodology or “attitude” as he describes it, help understand the use of what is proposed here as “graphic notations” developed with more detail in Chapter 2. The goal of a graphic notation is to precisely explore the data, understand the

⁶⁷ Throughout his career, Tukey himself expanded on Bertin’s work to develop further graphical schemes such as the “box plot” or the “stem and leaf display”, all which responded to his need for exploratory analysis. Furthermore, he was also involved in developing one of the earliest computer programs for interactive visualization called PRIM-9 in 1972.

phenomena, recognize meaningful information from the emerging patterns through a visual display, help us pose meaningful questions. In this sense, the images that are created as graphic notations are not conceived as the final representation, but the point of access for us to define what aspects of the data will be ultimately used in the final representation which is here a Data Driven Drawings piece.

Edward Tufte – Graphical Excellence and Moral Responsibility

Edward Tufte's impressive body of work presented in four books (Tufte, 1990, 1997, 2001, 2006) remains today one of the most influential contributions to the understanding of data visualization as functional art. In his works, we find three recurrent strategies for bringing forth his ideas:

- The historical study of use cases in which he uncovers the potential of data visualizations to both experts and non-specialized audiences.
- Proposes methodologies for creating a visualization that provides accessibility and abides by the "truth".
- Suggest that representing "reality" has intellectual and moral implications.

One of Tufte's major contribution is an elucidation that data visualization pertains to a wide range of human inquiries, his historical revision of the medium is passionate and reveals the overwhelming democratization of data visualization beyond specialized fields. Tufte argues that the goal is not the presentation

of a scientific method, or showcasing a sophisticated graphical technique, but the revelation of an aspect of the world that is hidden and complex.

In Tufte's view, and evident in the fascinating stories he showcases throughout the different case studies he analyses, the data in itself carries with it non-fictional descriptions about natural and social events around us, these are in their own right as interesting and revealing as any story we may come up with on top of them. These stories are buried inside the raw data and our job is to extract them and present them to others as they are. In this sense, Tufte sees as unnecessary the imposition of our own subjective perspective that may distort or obscure this process of extracting the truthful report.

While Tukey makes the case for separating the imposition of statistical methods from the exploratory analysis, Tufte makes an equally important plea for designers to have a stronger understanding of "quantitative skills". In other words, Tukey is trying to invoke our intuition into the normally rigorous process of analysis, and Tufte advocates for rigorous thinking in a normally more intuitive process like the design⁶⁸.

lack of substantive and quantitative skills of the illustrators, dislike of quantitative evidence, and contempt for the intelligence of the audience-guarantee graphic mediocrity. These conditions engender graphics that (1) lie; (2) employ only the simplest designs, often unstandardized time-series based on a small handful of data points; and (3) miss the real news actually in the data. (Tufte, 2001, p. 87)

His argument against subjectivity is positioned within the premise that any personal imposition over the stories in the data poses a

⁶⁸ While Tufte refers to the field as "design" or the practitioners as "designers", this applies to art as well.

threat to the revelation of truth. This position aligns specifically with functional representations but fails to consider the possibilities of the evocative ones described in chapter 2. Mainly because in the evocative representation, what is being presented to the audience is not evidence of the truth, the evocative intends to forestall reason and therefore, the idea of truth as something that can be attained and reasoned over is inaccessible by design. In this sense, the threat that Tufte sees in subjectivity pertains to the goals of functional representations where the “world out there”, for which data is the index of, does not need subjective points of view that may end up in patronizing distortions of the “real news”. But, as discussed in the second part of this chapter, subjectivity also empowers the creation of new relationships with reality as seen in the case of an animated documentary. Particularly with the idea that reality is more than what happens “out there”, but also includes internal, emotional and subjective experiences as well.

Here it is important to clarify what is meant by subjectivity. For Tufte, it pertains to the personal opinions and bias in the author of a representation. In these terms, it is understandable how it can become a risk for a functional representation. He considers this to be the central moral dilemma in data representation and therefore critical that subjectivity is avoided:

Making an evidence presentation is a moral act as well as an intellectual activity. To maintain standards of quality, relevance, and integrity for evidence, consumers of presentations should insist that presenters be held intellectually and ethically responsible for what they show and tell. Thus, consuming a presentation is also an intellectual and a moral activity. (Tufte, 2006, p. 9)

For an animated documentary, on the other hand, it pertains to a set of personal experiences that are part of reality. A relative one, one that does not necessarily apply to everyone and is often

internal, “the world inside”. In relation to the ideas of beauty and the sublime presented earlier, Tufte is focusing on the possibility of showing “the world out there” for reason to be activated over it, as the title of his book *Beautiful Evidence* (2006) suggests. In the sublime on the other hand, an internal experience is activated by a particular manifestation of external phenomena, and here, reason is not able to fully operate. Which means that the external phenomena are only the way to reach an internal experience. Animated documentary uses this same strategy for activating an internal experience in the audience, it makes reference to something that exists in reality, not to present it in itself, but to stimulate an emotional response. This approach may not be where reason stands, but is part of reality, one that is subjective because it is internal and is provoked by the equally subjective point of view of someone that “animates” it. In contrast to Tufte’s proposition to set aside subjectivity from the epistemological project of documenting reality, Michael Renov argues that documentary works that do embrace it, do not necessarily fail our moral responsibility as authors or viewers:

...documentative work that invites radical doubt, ambivalence, the embrace of contingency rather than certain knowledge ought not be viewed as simply fashionable or facile in its skepticism. For this value exists both as challenge and affirmation: provocative in its refusal of individualist truth, profoundly moral in its call for and reliance on individual moral responsibility. (Renov, 2004, p. 147)

Lastly, it is contested here that Tufte’s position is even too narrow for functional representations. Going back to the case of Florence Nightingale, her work meets his plea for rigorous statistical knowledge on the part of the designer. She was a statistician first and clearly versed in both the quantitative skills and graphic design. But she allows her passion and concern for the people dying in the camps to inform another aspect of the representation that aligns

with her own subjective experience. Her decision to create not the “best” diagram, but the most impactful one is strategic. She postpones the revelation of the truth by appealing to an aesthetic experience first, incite an emotional response in her audience, and then, only if the reader has been provoked, the truth can be deciphered from her diagrams. The compromise she makes is very small, partially distort the visual presentation in order to engage with her audience in an emotionally meaningful way.

Tufte does not mention her in his books, only a brief reference can be found on his website. Here, he focuses only on the fact that her proposed diagram is more difficult to read and a preferred method would have been to simply present a table.

The diagrams improve on pie charts in exchange for a greater difficulty of reading. As is the case for pie charts, the inherent problem is the difficulty of making good comparisons across the wedges, as Sally Bigwood points out. In general, for such small data sets, tables (or for maps, numbers distributed on the surface perhaps sized as names of cities are sometimes sized in proportion to their population) will outperform graphics. (Tufte, 2017)

As has been argued this far, locating the potential strategies for creating evocative data representations do not always align with the ideas presented in the literature of data visualization. To aid this problem, we can borrow from the theories of animated documentary instead. It is in this field where the subjective experience is not only allowed but explored as a subject matter in itself. This will serve as the last model that help enclose the theoretical context from where Data Driven Drawings are proposed here. Data Driven Drawings is an effort to create evocative data representations, up to this point we have identified the “why” by contextualizing the potential of data to grant access to phenomena that reaches beyond normal comprehension. With this ability, we have two options: behold it to

show it as beautiful or expand it to provoke the sublime. Data Driven Drawings align with the operating function of the sublime as it looks to expand our understanding of the phenomena by inciting an emotional experience of it. “How” to do this has not been completely uncovered using the existing theories of data representation, hence the need to look somewhere else.

Evocative representation of reality in Animated Documentary

Data Driven Drawings as an approach that uses animation and data representation as its formal characteristics coincide with animated documentary in two ways: firstly, the general premise of both sees the representation of reality as a creative act. Secondly, an animation is articulated by conceiving creative relationships between distinct entities, one frame is created to dialogue with the next frame and the next. What we see in an animation is an impression of movement that signifies something not because of each individual frame, but because of their association. In a data representation, we perceive the phenomena that the data indicates because each individual datum is exposed in relation to the rest. The creative act is not in the exposition of these individual entities as they are on their own, but in the proposed relationships between them. Norman McLaren famously said that in animation:

What happens between each frame is much more important than what exists on each frame. Animation is, therefore, the art of

manipulating the invisible interstices that lie between frames.
(Hébert, 2014, p. 64)⁶⁹.

This same suggestive definition can be applied to data representation where datum and frame are analogous. In this context, Data Driven Drawings find in animated documentary its closest theoretical correspondence.

The theories on animated documentary are inclusive of the artists' point of view, celebrate subjectivity as a fundamental entity from which artists present their view of the world, and study it through the possibility of representing internal experiences that may not be possible to directly portray in other documentary forms. It was already in 1933 when John Grierson proposed that documentary is "The creative treatment of actuality" (Grierson, 1933, p. 8), a definition that has remained valid throughout the years for being broad enough to allow documentary filmmaking to manifest in different forms, even in animated documentary which seems at first like an oxymoron:

Animation and documentary make an odd couple. Theirs is a marriage of opposites, made complicated by different ways of seeing the world. The former conjures up thoughts of comedy, children's entertainment and folkloric fantasies; the latter carries with it the assumptions of seriousness, rhetoric and evidence.
(Honess Roe, 2013, p. 1)

⁶⁹ Norman McLaren proposed this famous definition for animation which has been cited extensively, Pierre Hébert clarifies its provenance and cites it in his article "A Matter of Affiliation" (2014) (Hébert, 2014, p. 64): "It was reproduced for the first time by the film magazine *Séquence* in Montreal (I've just discovered this) and later by André Martin in the magazine *Cinéma 57*. It was a handwritten sheet on which McLaren had written, apparently for himself, this now famous text: *Animation is not the art of drawings that move but the art of movements that are drawn. What happens between each frame is much more important than what exists on each frame. Animation is therefore the art of manipulating the invisible interstices that lie between frames.*"

Honess Roe, who thus far has published the only comprehensive book on *Animated Documentary* (2013)⁷⁰, proposes three functional categories for animated documentary: “mimetic substitution”, “non-mimetic substitution”, and the one that particularly aligns with this research, “evocation”⁷¹. Each will now be examined articulated in animated documentary theory and compare them to contemporary practices of artistic data representation.

Mimetic substitution

This category refers to the ability for documenting events or phenomena where no other forms of recording are possible. Such is the case of an early⁷² example of animated documentary: “The Sinking of the Lusitania” (1918) by Winsor McCay. The event in which a German U-boat torpedoed and sank the ocean liner known

⁷⁰ Despite the fact that there is only one book solely dedicated to this subject, the scholarly research on animated documentary is abundant in journals, conference proceedings, festival programs and specialized symposiums.

⁷¹ Her categories are conscious of the precedents set by earlier theorists that had proposed alternative taxonomies. Eric Patrick and Paul Wells propose each a set typological modes (Patrick, 2004; Wells, 1997). Paul Wells developed his modes in an effort to find a distinction between live-action and animated documentary. His modes: “imitative”, “subjective”, “fantastic” and “post-modern” proposed a framework specifically for animation that was different to the influential modes proposed by Bill Nichols for live-action (1991). Patrick proposed a different set of modes he called “structures”, which are based on the filmmakers’ approach to signify reality within the animated language. He called them: “illustrative”, “narrated”, “sound-based” and “extended” structures. Both Wells’ and Patrick’s efforts were important because they provided the first steps towards the development of animated documentary theory.

⁷² Honess Roe argues that McCay’s film is the first “commercially released animated documentary” (Honess Roe, 2013, p. 8) preceded by cases in which animation was used to illustrate sections of other documentary films (Percy Smith, Max Fleischer). In the chapter “The Archeology of Memory: The Explorations of Animated Documentary” the author Chunning Guo reports on an open debate about the claim that McCay’s was the first, contested by Marcin Gizycki who argues that an earlier animation of this same event was made by John Bull in 1915 and a copy of it is archived at the British Film Institute (Bobrowska & Bobrowski 2016, p. 53).

as “The Lusitania”, was never recorded on camera and McCay was commissioned to recreate it through animated drawings. McCay’s approach was to craft drawings that seek realism in comparison to his earlier cartoons, in order to set a tone of verisimilitude with reality. Yet the film is not completely detached from a point of view, it is highly emotional and portrays the events from the perspective of a wounded society that sees tragedy in the event. The slow sinking of the ship with bodies falling from the sides is paralleled with photographic portraits and short biographical texts of American civilians who died in it. At the time, the film was received by the audiences, not as an entertainment piece, but the opportunity “to witness the whole tragedy” (Honess Roe, 2011, p. 219).

In relation to the previously discussed ideas from Edward Tufte, his proposal that the subjective point of view threatens the “presentation of evidence” because it can be “patronizing” or “over-decorated”⁷³, McCay’s film is clearly problematic. One of the intertitles complicates the relationship established with the audience when it says: “From here on you are looking at the first record of the sinking of the Lusitania”, asserting to itself a sense of veracity that what is shown is truthful. This proclamation may obscure other operating characteristics such as the fact that is a reconstruction, and that it is made from a patriotic stand that shows the events from a particular side of the conflict. Even if one agrees with the underlying cause as something just, the subjective point of view, in this case, does threaten what Tufte contested as the “moral responsibility” with the facts. Honess Roes proceeds with other

⁷³ Tufte discusses the potential for patronizing the audience and overly decorating the presentation at length in the chapters: “Graphical Excellence” and “Sources of Graphical Integrity and Sophistication” (Tufte, 2001)

examples of mimetic-substitution⁷⁴ where she clarifies that animated documentaries under the mimetic-substitution category are not inherently problematic for the presentation of facts, this complication arises in the discursive strategies of the author and not in what this type of animated documentary is able to achieve.

This same argument can be applied to artistic data visualization that operates within the characteristic of mimetic-substitution and abides by the ability to represent something that was impossible to record by other means, it reports faithfully on what the event was, and shows coherence between these operating characteristics and the proposed discourse. Such is the case of Nicholas Feltron's series "Annual Report" (2005-2014). His process for creating this project was to record obsessively an enormous amount of data related to seemingly mundane events of his life: how many coffees he drank, every single person he meets, the trajectories he embarks every day in the city, the amount of time he spends doing different activities, music he listens to, etc. He quantifies all these events and at the end of the year, he translates the raw data into a series of infographics that clearly display these otherwise ephemeral events of his life. In terms of data visualization, the charts and diagrams are virtuous examples of mimetic-substitution, they clearly substitute the transient event for a representation that beholds them. Each diagram on their own is a functional data representation as discussed here, they disambiguate and attain complex phenomena. But they also display the type of affordances of mimetic-substitution as a way of reconstructing reality from a personal point of view.

⁷⁴ Other examples of mimetic-substitution appear with more frequency as sections inside live-action documentaries where there is no other record of the event. As for fully animated films, Honess Roes provides the examples of "Chicago 10" (2007) by Brett Morgen and "Walking with Dinosaurs" (1999) TV series by the BBC.

Feltron is able to infer personally meaningful insights from the visualizations, such as how much time he spends with his girlfriend in comparison to his mother (2012). But the meaningful aspects of the piece are not in what each individual visualization is able to report on, but the general impression of what this exercise means. They reveal a sort of personal catharsis in the compulsive exercise to look at every detail of his intimate life, discourses with timely issues of surveillance in an era where our digital footprints are being tracked by governments and corporations⁷⁵, invites on self-reflection by comparing his behavioral patterns with our own. His strategy is inverse to what Florence Nightingale did (Chapter 1) in her “Rose Diagram”. She appealed to an emotional response first by presenting estranged but beautiful diagrams in order to attract the audience before presenting the facts. Feltron, on the other hand, presents “beautiful” but clearly readable facts (mimetic-substitution) to then activate in us an intellectual and emotional response. In Feltron’s work, we can see that mimetic-substitution affords the presentation of facts, personally subjective intents and equally subjective readings in the audience.

Non-mimetic Substitution

The second category proposed by Honess Roe is the “non-mimetic substitution” in which the imaginary reconstructions of the real events is set aside. The artist in these cases develops metaphorical

⁷⁵ Feltron’s works shows this by showing how much meaning can be inferred from seemingly mundane data. Our digital footprints are recorded in social media interactions, search engine behavior, places visited via mobile phones geolocation, and alike. An excellent interactive documentary presenting these issues was made by Brett Gaylor “Do Not Track” (2015) and produced by the National Film board of Canada: <https://donottrack-doc.com/>

associations between the real events and their personal interpretation of them. A common approach for these types of animated documentaries is the use of sound recorded interviews, in which the audio is maintained as the indexical relationship to reality and then interpreted visually in the animation. Sheila Sofian's body of work is an exceptional case for this category. In "A Conversation with Harris" (2001) we listen to the voice of Harry, a young boy who describes his point of view on the Bosnian War after having to immigrate to the US with his family during the conflict. In "Survivors" (1997) she presents the voices of multiple women that have been mentally and physically abused by their spouses. The visuals that Sofian proposes to accompany these interviews are highly metaphorical, they are not literal illustrations of what the interviewee says, but the artist owns interpretation of the content. Sofian suggests that both live-action and animated documentary are the artist's point of view on the subject, but since the animated image is unambiguously presented as artificial, the audience has the opportunity to instantly be aware that the artist personal voice is present. She explains:

Animation is more transparent in its construction. The audience understands that the image is created entirely from the artist's hand. Unlike live-action, there is no pretense to represent a 'true' replica of events onscreen. (Sofian, 2005, pp. 9-10)

Sofian's non-mimetic approach to documentary challenges Tufte's proposition that the "artist's hand" (as Sofian puts it) or the subjective point of view, puts into question the moral responsibility with the audience. Sofian contests that the viewer is not tricked into thinking that the representation is the "truth", but in the formal presentation of reality through metaphors and synthetic imagery, the audience becomes aware that what is presented is subjective and therefore open to being contested.

Example 2: “Eterna”

To provide an example of non-mimetic substitution in the context of data representation I will now explain the second example of a practical project I developed during my research. In 2015, I was asked to develop an artistic piece to be presented in the symposium “Art of the Networked Practice”⁷⁶ co-chaired by NTU professors Randall Packer and Vibeke Sorensen. Given complete creative freedom, I took the opportunity for experimenting with real-time processing and representation of social media data⁷⁷. In relation to the defining category presently discussed, I “substituted” textual information on Twitter messages for an animated character that performed an action depending on the emotional decoding of the message, therefore imposing a metaphorical association between an animated character and the stream of messages in social media.

The way the application works is that we can use a custom search engine to query in the entire Twitter live activity for messages that use the word one searches for. In return, the Twitter server sends us, in real-time, any message that uses that word. The result varies depending on the query, to test it I would often use common words like “happy” or “sad” and the application would receive hundreds of messages per minute. If I used less common words like “network” or a particular hashtag, the application received messages at slower rates. My first impression of the resulting accumulation of disparate

⁷⁶ In this case, real-time refers to the ability of capturing the data from social media as it is happening, different to other projects presented here where the data was already collected. Even though the other projects process internally the data in real-time from pre-recorded data (the code reads line by line the database and translates this information to what we see on the screen), here the collection of data is also happening in the moment the viewer is experiencing the piece, giving it a true sense of “real-time”.

⁷⁷ This is inspired by the work of Vibeke Sorensen, discussed in Chapter 4.

messages was that there was a mysterious coherence, an aggregated voice given by different peoples' singular gestures. This reminded me of the Argentinian writer Macedonio Fernández and his novel *The Museum of Eterna's Novel* published posthumously in 1967.

Fernández speaks to his reader throughout the entire novel, in which he suggests that the future of prose is in the creation of the “good” novel, one which requires a new type of reader that he refers to as *lector salteado* “skipping reader”. He professes that he has made the last “bad novel”⁷⁸, one in which there is a linear story we can follow, with classical forms of introducing the universe of the characters, developing conflict and ending in a resolution. This type of novel expects, and requires, a “linear reader”. The future “good” novel that he is proposing, anticipates the hypertext as one in which the reader can escape the current text at any point to “skip” into a new one, constructing at will the totality of the text. Also, it involves the reader as an active agent of the novel, in *Museo de la Novela de la Eterna*, Fernández speaks to the reader constantly by breaking away from the narration and making us aware that we are also part of the text, inviting us to “interact” with the novel and suggesting that we can “skip” at any given time.

In this sense, the text that emerged from the disparate sources in the Twitter stream seemed to me like the resulting construct made by a virtual “skipping reader”: dislocated, seemingly chaotic and yet singular and poetic in its own way. In homage to Fernández, I named this piece “Eterna”⁷⁹, which in the novel is the central

⁷⁸ Referring to his previous novel “Adriana Bueno Aires” (1922) as the “bad novel”.

⁷⁹ *Eterna* has the same meaning in Spanish as the word eternal in English, yet in Spanish is a female name.

characters that serve as the linkage between the different sections. Inspired by this, I created an animated character that embodies and unifies the disparate messages of the stream. Each message is processed in real-time by an open-source library called “Sentiment”⁸⁰, this library takes in the text and returns a single number that represents the “sentiment” score: below 0 when the text has “negative” connotations and above 0 for “positive” ones⁸¹. The result is a stream of “skipping” numbers that I then “substituted” with animated sequences of the character “Eterna” performing a short action that related to the sentiment score.

For example, the following are frames from the animation where the score is encoded. The left image is on the negative side of the scores spectrum, the central is a neutral score and the right is a positive one:

Score = -7

Score = 0

Score = 7

⁸⁰ The sentiment library uses the AFINN-165 wordlist to process the text and return a sentiment score. The AFINN method was developed by Finn Årup Nielsen and a detailed explanation of his method can be found in his paper “A new ANEW: Evaluation of a word list for sentiment analysis in microblogs” (Nielsen, 2011). The library used for this project can be found at: <https://www.npmjs.com/package/sentiment>

⁸¹ In the library tutorial, they give the following example to demonstrate: the phrase “Cats are stupid” returns a score of -2, and the phrase “Cats are totally amazing!” returns a score of 4.

Figure 15. Frames corresponding to the sentimental analysis score, from “Eterna”, by González J., J. C., 2015.

All the sequences ranging from -7 to 7 in the sentiment score were placed in a single image so the application could select the corresponding sequence from the score that each individual tweet yielded.

Figure 16. Animation sequences contained in a sprite sheet, from “Eterna”, by González J., J. C., 2015.

The music was created by NTU professor Ross Williams and he followed the same strategy for his composition. Given the scale of scores from -7 to 7, he created a series of short musical sequences performed with a flute, ranging from 1 to 4 seconds long. In the final application, the program selected the musical file that corresponded with the score and this is then played with the matching animated sequence. The resulting data driven animation was the floating character of Eterna performing an undetermined dance, delegating the order and frequency of the visual and musical construction to the behavior of the social media activity. The actual data that produced the animation was also displayed on the screen to allow the viewer to understand what was activating the animation.

The data was shown as the raw text of the original message with its corresponding sentiment score and metadata of the time, place and author of the tweet and a set of diagrams that displayed how the data was behaving in time. The diagram in the bottom shows a series of triangles for which the size and color showed the intensity of the score throughout time: blue for positive, red for negative and the size of the triangle represented the numerical value of the score. This purposeful exposition of the data to clearly be read in the functional term served as a disclosure of the index that match the presented interpretation, or the transparent disclosure of the non-mimetic substitution of the data.

In the “non-mimetic substitution” category of animated documentary, we find the affordances of animation to propose metaphorical associations with a given index of reality. This index can be sound interviews, archival footage, or data, as shown in the previous example. The reference to reality is there for the artist to establish creative associations and present to their audience a

transformed version of the world. This proposed perspective on reality is overtly constructed and makes no claims of offering the audience a direct path to knowledge about the original phenomena. In this sense, we can go back to the proposed definitions of evocative and functional data representations in Chapter 1. As proposed by Kant, these types of works appeal to the “aesthetic comprehension” rather than the “logical one”, they begin to embrace the ambivalence of how phenomena may affect our lives rather than seek for evidence of the phenomena itself. This is done so by the transformation of the phenomena through metaphors and sets aside the need to simulate phenomena through mimesis.

Evocation

The third function proposed by Honess Roe, “evocation”, refers to those documentaries that address the particular world of emotions, states of minds and in general phenomena that eludes the observations of the physical world. These complex ideas are represented by the artist from his/her own point of view and the audience is invited to observe these realities through the artist constructed viewpoint. In these cases, the subjective point of view is part of the subject matter, a type of vessel that intermediates our access to the “world out there” through their eyes and thoughts. Complicating the faithful exposition of reality in exchange for investigating how that world affects our experience of it. Honess Roe argues:

“By visualizing these invisible aspects of life, often in an abstract or symbolic style, animation that functions in this evocative way allows us to imagine the world from someone else’s perspective” (Honess Roe, 2011, p. 227).

As Honess Roe suggests, these approaches often appear in “abstract or symbolic styles”, with intricate narratives that avoid a clarifying or explanatory discourse about reality. In evocative films, audience participation is as important as the artist personal view. William Moritz suggests that these types of films, where reality is often obscured and abstracted in their formal presentation, are “‘interactive’ systems”. Not necessarily within the understanding of interactive in new media and digital technologies, but in the sense that if the audience does not participate as critical thinkers, the film has no meaning and the reality they address remains concealed. Moritz explains this idea after the content analysis of four⁸² iconic films produced in Eastern European countries during the communist era:

Home, Tale of Tales and Déjeuner sur l'herbe all posit an ‘interactive’ system, in which the purposefully convoluted narrative structure must be unraveled by the viewer. While these strategies may have been devised to circumvent censors, they result in a rich experience that rewards repeated viewings, since the films are composed freshly in the viewers' minds, with new connections, new perceptions, and new feelings each time. And the intricate artistry of all four filmmakers does appeal as strongly to the emotions as to the reason of the viewer, which enhances our sympathy for the characters and makes the films transcend the narrower political issues that they originally protested. (William Moritz, 1997, p. 47)

Moritz gives importance to the “convoluted narrative structure” as a strategy to enhance the observations of the social conditions that these films referred to. By deliberately complicating the presentation of reality, these films seek to evoke rather than explain, allow the audience to draw their own conclusions rather than patronize or impose a biased view. In the same way that Sofian

⁸² *Home* (1959) by Walerian Borowczyk and Jan Lenica in Poland, *The Hand* (1954-5) by Jiri Trnka in then Czechoslovakia, *Tale of Tales* (1970) by Yuri Norstein in Russia and *Picnic on the Grass* (1986) by Priit Pärn in Estonia.

contested that animated documentary is able to be “transparent” with its audience by revealing the “hand of the artist” in the visual presentation, Moritz proposal is that the same can be done in the narrative forms. The complex narratives are the immediate revelation that an artist is making very conscious decisions about what the audience perceives, inviting the viewer to challenge these views and interact with them emotionally and critically.

Pointing back to research question 1 of this thesis, these convoluted narrative structures are particularly relevant precedents to identify how an artist intent that aligns with evocation, may consider this strategy for alluding back the index that the data refers to. Even when this index is not clearly discernible from the complications presented aesthetically, the audiences of the films that Moritz examines were clearly able to find a path to those real events. But the veil of the subjective articulation in the animations affects emotionally how we observe back that reality.

In the next and final chapter, the discussion points to the particular strategies used to develop the main Data Driven Drawings projects that have informed the theoretical framework presented in this thesis.

Chapter 4: Production of Data Driven Drawings

In this final chapter, the goal is to bridge the theoretical ideas discussed so far with a series of three personal projects that thematically explore the death of my father. The approach I have developed here for representing this momentous event in my life, is motivated by a turn of attention onto what I ought to represent: in previous animation projects, the focus was on internal states of mind that emerged from my relationship with my father, while in this instance, I confront specifically the context of his death and how I experience it now. In the former, I had explored the possibility of embodying these internal states of mind through animated short films. In the latter, I investigate the possibility of addressing concrete data that signify the context in which my father was murdered, therefore motivating the present investigation named Data Driven Drawings.

In the context of the “evocation” category proposed by Honess Roe and discussed in Chapters 1 & 3, my previous work was partially abiding by the definition of this category which intends to represent internal states of mind based on real experiences. But they do not refer to the particular reality from which they came from, rather they were more the manifestation of an imagined state of mind, even if inspired by real events. For example, in *Mecanismo Olvidador* “Forgetter Mechanism” referred to in Chapter 1, the internal state of mind was a struggle between the “oscillating” feelings towards my father: love and hate, a clashing effort to remember and forget him. In the case of “SiSiSiSiSiSiSiSiSiSi” discussed in Chapter 2, the state of mind is prompted by an irrational need to always be polite towards my father and confusing this with an inability to express myself openly, provoking an ever-

expanding frustration. The former was represented through a man trapped inside a pendulum swinging back and forth between these two states, and the latter through the incessant repetition of animated characters performing polite gestures. These projects anticipated the themes that I am interested in now and fundamentally inform my personal aesthetics in animation, but the central point of departure from which I developed my new work in Data Driven Drawings is in the introduction of data as an index of reality.

While this research started with the intention of adding data as a signifier of reality to the type of films being produced before. Understanding this initially as a transition into animated documentary within the inclusive categories proposed by Honess Roe as “mimetic substitution”, “non-mimetic substitution” and “evocative”. But the present research revealed that the transition was more complex than what it was initially anticipated and the expectation for seamlessly expanding the previous work into animated documentaries was therefore conceptually and formally defied. This problem will now be discussed in relation to the research questions of the thesis and discuss how the theoretical research provides a conceptual ground for resolving these issues, concluding with the description of a series of Data Driven Drawings in which the proposed theoretical framework is probed.

The central problem of using data to denote reality is that its formal characteristics require a deep level of abstraction. The reason it can contain aspects of reality that are complex and inaccessible to our regular senses is that it can reduce these phenomena to a form that is very different from how we experience the particularly signified phenomenon. For example, let's say we have a sound recording of thunder; the most immediate revision of the raw sound as it was

captured would allow us to identify the event by recognizing it from other sounds of thunder we have experienced before. But if we capture this natural event as data, the most immediate access provided in the raw information would not manifest as what we normally recognize as a thunder. The way it is recorded conceals in the numbers both what we may be able to recognize as the physical manifestation of the thunder and other invisible interstices of the phenomenon. In this sense, data would always require us to process this information before we start to recognize what it is. We could translate the data to the sound of the thunder, or visually simulate it, or codify it graphically into a functional data visualization. In all of these cases, the goal is to disambiguate what the data signifies. But as discussed throughout this thesis and particularly in the context of the evocative data representations and artistic practices, other forms of embodying the data start to emerge and purposely complicate how we denote the original phenomena in alignment with specific research and artistic goals.

Here we can begin to answer the first research question: What is the artist's intent when data is the index of reality?

This answer should not be absolute, but instead, be considered a conceptual spectrum of possible intentions to be within the functional and evocative categories. This thesis makes no assumption that these categories belong to a specific discipline, but relate to how the author seeks to inform others of the meanings found in the data. An eloquent and emotional argument could be presented through functional data representations such as the cases of the Brookes slave ship, Nightingale's Coxcomb or the Nasa galaxy simulation. Particularly in the case of the Nasa simulation, the subjective experience that overwhelms and evokes can be

imposed by the spectator even if its original intent was completely functional.

Since an artistic intent does not only have to align with the objectives of the evocative category but can also investigate within the affordance of the functional to bound and clarify the complex. An artistic goal could be to align with the philosophical tradition of the beautiful and not the sublime, in search of an aesthetic experience that is discernable. Further exploration of the functional within artistic intents was discussed in Chapter 2 with the proposition of using data visualization as graphic notations. In these cases, the representations have functional purposes of clarifying the orchestration of a film to be. They also provide a space for reasoning over the metaphorical association between the graphical elements and how they end up in the films and offer a structure for articulating a potential evocative representation.

Not ascribing the artistic intent to a specific extreme of the conceptual spectrum was then contextualized in chapter 3 in which it pointed how data representation and animated documentary intersect as two methods that seek to represent reality while validating subjectivity in their process and conception. The current academic studies in animated documentary are discussed as a starting point to reason over the possible manifestations of artistic data representation, in other words, while Honess Roe proposes three categories for animated documentary: “Mimetic substitution”, “non-mimetic substitution” and “evocative”, this thesis proposes that we see data representation from within these modes of representing reality. But there are some formal distinctions that need to be clarified before I can finish formulating this argument.

Coming back to the example of the thunder recorded as sound and data, consider a potential visual representation of these two modes

of capturing the event in the context of functional and evocative representations and animated documentary. The functional and mimetic substitution categories operate under the same premise: the visual articulation of the signified phenomenon contained in the sound recording or dataset, visualizes as closely as possible the thunder as it happened, removing any ambiguity and helping us to reason over it from its visual representation, or better yet, its visual simulation.

In a non-mimetic substitution approach to animated documentary, the sound recording may then serve as a point of departure to impose metaphorical associations, the visual presentation is not there to allow the thunder to express its magnitude and power on its own but used instead within the rhetorical strategies of an author with poetic intentions. An animator may then allow the sound recording to signify the source in the sound space, and articulate a comparison with something else in the visual space to create a metaphor. Thus, animation and cinema as audiovisual media are able to afford this direct comparison between the two senses and audiences are able to discern on the proposed comparison by connecting the two in the mind. This potential is clearly explained by Russian filmmaker and theorist Lev Kuleshov in what is known as the “Kuleshov effect”:

Kuleshov articulated the possibilities of association between images in film montage in a widely accessible didactic film⁸³ that explains the Kuleshov effect: the face of a man and then a soup platter creates a different set of associations in our mind in comparison to the same image of the man edited against a scene of a girl in a

⁸³ The film experiment is cited extensively in film theory literature and it is available online from multiple sources:
https://www.youtube.com/watch?v=_gGl3LJ7vHc

coffin or a woman laying on a divan. How these associations unfold in our mind show that we impose meaning on what the man may be thinking or feeling. In the first example, he might be hungry and wanting to eat the soup, in the second one he may be mourning and in the third, we could infer he has feelings of desire. The fact that the sequence of the man's face is exactly the same in all three of them and that his facial expression, taken out of any of these contexts, is very neutral, has provided a powerful example of how we decode meanings in film language and how metaphors and associations manifest in the medium. Sergei Eisenstein expands on this in his theory of montage⁸⁴, but describing the theories of film montage in depth extends beyond the scope of the present research. For purposes of simplicity and to maintain a connection with data representations, we remain now in the domain of a singular set of associations between two signifiers that allow us to create a metaphor.

In the non-mimetic substitution category of animated documentary, this collision of signifiers is further explored by the possibility of using imagined imagery as a counterpoint of a recognizable signifier. Different from the Kuleshov effect, now we have one element that we recognize from reality and a new one that is proposed and created from imagination. A good example of the possibilities that arise with this can be seen in the film "Survivors" (1997) by Sheila Sofian, where she has sound interviews of multiple women who suffered from abuse with their partners. One of the

⁸⁴ Sergei Eisenstein (1898 - 1948) was a Russian filmmaker and film theorist that formalized a wide encompassing theory that derives from the Kuleshov effect described above. He developed a series of "methods of montage" known as: metric, rhythmic, tonal, overtonal and intellectual. These theories have endured over time and remain valid in today's film theory and practice. (Eisenstein, 1975, 1977)

strategies she uses when there is a momentous description of a violent event is that she contrasts the audio with abstract imagery instead of the realistic visual representation of the event. The result stands on the tradition of the Kuleshov effect but necessitates acute participation of the audience to build the metaphorical connection. In Sofian's work, the abstract images are an approximation of concrete emotional and mental realities that do not have visual similes, but the abstract images provide a path for us to build the connection in the sense that what is shown is a representation of an internal state.

In the case of data representation, the comparative signifiers often manifest in abstract or imagined constituents. Since data in its raw state is complex, we have to decide whether to bring it forth to the audience as something clearly recognizable or denote its existence in other ways. Considering the evocation category in animated documentary as the one that explores those realities that escape our observable world, the reality signifier is conveyed to the audience after it has been interpreted by the author and is represented "often in an abstract or symbolic style" (Honess Roe, 2011, p. 227). I argue that here we are presented with a particularly challenging problem that evocative animated documentaries and evocative data representations are equally confronted with: How do we convey reality if the source has to be processed first by the author? And in this process, is there a compromise for the potential of building metaphors that depend on reality being signified?

A good example of this can be seen in Vibeke Sorensen's series: "Mood of the Planet" (2015), "Mood of New York" (2016) and "Mood of Singapore" (2016). In them, the reality signifier comes from the

“sentiment analysis”⁸⁵ of social media activity and this is then transformed into an abstract animation that follows the tradition of visual music (chapter 3). If we compare this to the Kuleshov effect, we no longer have such clear elements to compare, on one hand, we have the unbound and complex phenomenon of social media activity and its counterpoint is a type of representation that seeks to evoke in the audience an emotional experience by not simplifying the representation of the data, but expanding its association with an abstract audiovisual catalyst. But, it is different from what we are able to see in the Kuleshov effect with the recognizable images of a man and the counter scenes, or the case of Sofian’s work where the sound interview denotes the source of the imagined representation.

In Sorensen’s artworks, the only way we can begin to engage with the poetic associations is due to the fact that the data is not left in its raw state, but undergoes the process of sentiment analysis to turn it concrete. It may not be as close to the way we experience the world through external perception, but it is intellectually concrete as it bounds the complex phenomena to an accessible concept that denotes our human emotions. As seen in Sorensen’s work, data needs to undergo some level of processing until we reach a point where the data is no longer completely inaccessible but is able to dialogue with the proposed modes of representation to provide the audience space for drawing associations. This also indicates that the raw data in the state that is closest to reality, is inescapably interpreted by humans (authors of a representation) so it can be

⁸⁵ Vibeke Sorensen has been working in collaboration with NTU professor Erik Cambria from the School of Computer Science and Engineering to incorporate into her artistic work the ideas of an emerging an active research area concerning Big Data and Sentiment Analysis on social media activity. The work that they do together is an active academic, scientific and artistic research for which the aforementioned series serves as a space for probing their findings in these areas.

associated with a second element (mimetic or non-mimetic representation) for the viewer to experience it (functionally to be reasoned over or evocative to embraces ambivalence, pose questions, transform phenomena through metaphors and result in contingencies).

A topic that Sorensen's work also elucidates is the aesthetic possibilities of representing the experience of the world as denoted by Big Data. She does this in the articulation of her work as an immersive installation in which a vastly large and complex phenomenon that emerges from social media interactions is represented as an overpowering sensorial experience. The magnitude of how "Big" the data is, demands a technical shift from traditional computational and analysis models to be able to process it at a reasonable speed that permits its intended use, and this magnitude is then exposed by Sorensen as an active agent of the aesthetic proposal. She does this by immersing the viewer in the abstract representation of the unbound data, the data is not contained in the abstraction, but the abstraction and immersive presentation jointly signify how "Big" the data is, allowing us to experience the phenomenon for its vastness and complexity. The types of phenomena that Big Data affords and investigates are manifesting a version of the world that expands on the issues of the sublime.

Big Data as an area of academic and artistic research is one of the most promising subject matters for future investigations on evocative data representation. Throughout this thesis, the term Big Data is not used or discussed at length because topics being investigated in here in the practical work do not pertain to Big Data directly but to a series of issues that precede the problems that Big Data poses. The datasets being used to develop the Data Driven

Drawings pieces are technically not “Big”, in the sense that they do not surpass the traditional computational capacities and analysis models that Big Data challenges. The social and natural data used here may be complex and denote phenomena of great magnitude in our experience of the world (violence in Colombia or seismic activity), but the phenomena that are contained in Big Data extends far past the magnitudes of these.

Before answering the second research question, the practical work will be discussed in more detail:

The three Data Driven Drawings projects titled “Vista al Mar”, “Dibujo Defunciones” and “N.N.” are conceived as a series in which I probe three modes for representing the experience of my father’s death. They are, in a sense, a series of steps in abstracting the history of my personal biography, and are all infused with my subjective view of the data in an effort to evoke in the audience a personal reading of each piece. “Vista al Mar” is purposefully the most autobiographical one, in which I refer to my father directly and the context that led to his death. In “Dibujo Defunciones” I step away from his specific story and represent his death in the context of other homicides in Colombia as they are reported on the national statistics of this particular type of violence. Lastly, “N.N.” uses a wider encompassing dataset of multiple categories of violence in Colombia to represent a more general view of the phenomena. This strategy of moving from the microscopic view of the personal event to the macroscopic view of the general context in which that initial event took place, allowed me to probe different types of representations that Data Driven Drawings can afford and can be better understood in dialogue with the theoretical ideas discussed earlier. Furthermore, it is in the creation of these pieces and investigating through creative practice that I am able to better

understand the potential future applications of the research presented in this thesis.

Project A: “Vista al Mar”

“Vista al Mar” is an interactive web application in which I invite the viewer to visit the place where my father died and explore through my own experience of it. The user is first presented with a map view of Cartagena city in Colombia. I knew he was in Cartagena on that day but only learned that he was killed in a public street in a neighborhood called *El Gallo* from the death certificate⁸⁶ that the police filled out when they attended the case. It mostly has information that I already knew such as his name, gender, age, ID number and place of residence. But the information that was news to me was contained in the box “Place of Death” which stated that the event took place in a public street, and the name of the neighborhood: *El Gallo* in Cartagena, Colombia.

⁸⁶ In Colombia, death certificate forms are standardized and provided by *DANE* which stands for National Administrative Department of Statistics.

Figure 17. Data Driven Drawing introductory interface delineating the neighborhood *El Gallo*, from “Vista al Mar”, by González J., J. C., 2018.

In the piece, I delineate the geographical limits of this neighborhood by superposing a drawing made with ink on paper over the *Google Maps* interface. The viewer is then able to select a particular starting point inside the limits of *El Gallo* and we then zoom in to the *Google Street View* area that was selected. The user is then able to navigate through the streets using Google’s interface, but the picture pulsates between the street view and personal drawings. This pulsation is set to the rhythm of my own heart rate which is taken from data I recorded myself when reading a document that he wrote. At the same time, there is a minimalist soundtrack of a sound recording of the ocean in a loop. The sound changes volume depending on how far the user is from the actual ocean by calculating in real-time the distance from where the spectator is in the street view and the coast which is located a few kilometers north.

In this project, I seek to represent what happens in my mind when I think of the day that my father died, or better yet, to exorcise it out of myself to make peace with it. This internal experience is recurrent, and to use John Berger's description of drawings made from memory: "haunting" and "in its own way unbearable" (Berger, 1987, p. 60). The way this memory manifests is two folded: I have different pieces of information that I use to partially reconstruct what might have happened, and for what is left uninformed I seem to use my own imagination to fill the blanks. In this piece, I show this by inviting the viewer to virtually visit the actual neighborhood in which my father died through *Google Street View*, and superpose images that I have drawn and associate with memories of my father. The intention is to have the viewer visit this space and share my point of view, one that is informed by a series of perplexing emotions I feel when I remember or imagine what happened the day that my father died.

Figure 18. Data Driven Drawing final visual composition of *Google Street View* and superposed drawing, from "Vista al Mar", by González J., J. C., 2018.

The production of this piece did not start with the concrete idea of representing this state of mind but emerged from a series of metaphorical associations conceived and produced with data I had collected. In the same way that Tukey advocates for using data visualization for “exploratory data analysis” (Chapter 3) before formulating the hypothesis in functional data representations, I argue that this idea can be extended to artistic productions by using graphic notations of the data to explore the potential associations we can establish with the phenomenon at hand. In the case of “Vista al Mar” I was only able to formulate what the piece was about and start its production, after engaging in an exploratory analysis of data from my own heartbeats. To better explain this, I will describe in more detail the process to elucidated how the idea for the piece emerged.

During my first year of research and 10 years after my father’s passing, I decided to record my heartbeats while reading the only document I have access to and was written by him. I knew it existed and was archived in my mother’s house, but I never had the courage to read it before. The contents of it are a series of letters and newspaper clippings that my father used to explain to my family and me, the case that led to his imprisonment. It is 206 pages long and explains in great detail his case and how it was portrayed in the local media⁸⁷. I wanted to incorporate my own subjective experience as an intermediary between this particular story and how it may be ultimately represented, but not knowing in advance how much of it I would finally reveal to others. The heartbeat recording while reading the document served as the first level of abstraction to study my own reaction to the concrete story, a reaction I had no control over

⁸⁷ The newspaper clippings are primarily from “La Patria” which is the main newspaper in the city of Manizales, where I am from.

and could empirically provide a sense of how it affects me internally.

For this recording, I used a pulse sensor⁸⁸ that was connected to an Arduino⁸⁹ which then saved the data to a computer. In this way, the data provided me access to something that would be otherwise invisible to perceive and contain it in a concrete object which is the dataset. But in the early stage of capturing data, this was intended as a non-mimetic substitution of an unbound idea or unanswered series of questions: How do I react to my father's biographical story? Does it still affect me when confronted by it? Is there anything, in particular, that has a greater impact on me? In this sense, the function of capturing the data was not to analyze my own physical health as intended by the technology of the pulse sensors in medical contexts, but already in the capturing stage, I was exploring a metaphorical association between a subjective point of view and a biographical story.

This exercise was inspired after examining the work of two contemporary artists and their approach to data representation: The Mexican Rafael Lozano-Hemmer and the French Christian Boltanski. Lozano-Hemmer has developed a series of interactive installation pieces called "Pulse Room" (2006), "Pulse Front" (2007), "Pulse Tank" (2008), "Pulse Spiral" (2008), "Pulse Park" (2008), "Pulse Index" (2010), "Pulse Drip" (2013) and "Pulse Corniche" (2015). In these, Lozano-Hemmer incorporates pulse sensors to capture in real-time

⁸⁸ I used the "Pulse Sensor Amped" by *World Famous Electronics llc*. This sensor was created to be used with Arduino and similar electronic devices, making it an affordable and efficient heart monitor for this project. More information on the sensor is available in their website: <https://pulsesensor.com/>

⁸⁹ Arduino is an open-source electronic hardware. In this case, it allowed me to connect a pulse sensor to it and transfer the data to my personal computer where it could be stored in a document. See <https://www.arduino.cc/>

the audiences' heart rates, and in turn, use it to activate the systems that are presented to them in the installation. For example, the first one of the series is titled *Almacén de Corazonadas* "Pulse Room"⁹⁰ and is installed in a large room with hanging light bulbs that range from one to three hundred depending on the gallery space where it is presented. A pulse sensor is located near the entrance of the room and the audience is invited to record their own heartbeat in it. After this recording is captured, the lightbulb hanging nearest to the station modulates the light it emits to the rhythm of the person's heart. All the light bulbs are connected in series and each represents a person that was there at some point. Every time a new person records their heart, the last one in the series is taken away, all others move one step closer to the end and in this way, allow the new one to join. The result is an ever-changing pattern of modulating lights that fill the darkroom in a poetic and rhythmic depiction of a life cycle. Lozano-Hemmer explains the sources of inspiration as:

This work was inspired by *Macario*, directed by Roberto Gavaldón in 1960, a film where the protagonist suffers a hunger-induced hallucination in which every person is represented by a lit candle in a cave. Other references for this work include minimalist, machinic and serialist patterns in music (for example in scores by composers Conlon Nancarrow, Steve Reich and Glenn Branca) and the postulation of the theory of Cybernetics at the National Institute of Cardiology in Mexico City to explain the process of self-regulation of the heart. (Lozano-Hemmer, 2006)

The type of representation that is afforded by the data is particularly powerful in this piece, the presence of human life is abstracted in the lightbulb and we are able to understand this by association. It draws attention to our own understanding of the heart as a physical

⁹⁰ The documentation for this piece is available in Rafael Lozano-Hemmer's personal website: http://www.lozano-hemmer.com/pulse_room.php

indication that someone is alive when it pulsates and removes all other indicators such as biography or physiognomy. For my own work, I was particularly interested in the possibility of using the heartbeat data as a non-mimetic substitution of myself.

Furthermore, the contrast between my own heartbeats as a physical signifier of me being alive also highlights the contrast with the fact that what I am reading is something that was written by someone that is dead, emphasizing in this way his absence.

The idea of presence and absence is especially relevant in the piece *Les Archives du Coeur* “The Heart Archives” by Christian Boltanski. It is located in the inhabited Teshima Island in Japan in a museum built specifically to contain this piece. Boltanski has been collecting people’s heartbeats since 2005 in an ideally endless repository. The only metaphorical inhabitants of the island are signified through audio representations of the heartbeats collected from diverse countries around the world. Similar to Lozano-Hemmer’s work, the abstraction of human life into the heartbeats, forces us to focus our attention towards a trace left by human life, activating in this way an evocative experience where imagination begins to draw associations and potentially asks self-reflecting questions. For Boltanski, what gets highlighted in his own work, is a sense of absence rather than presence. The trace of a human heartbeat does not immortalize the source from where it was collected but emphasizes the latent possibility that this person is, or will be, dead. Boltanski explains it in an interview with Sarah Rosenbaum-Kranson for the “Museo Magazine”:

In fact, the first idea was a little bit like a joke—I made a piece, I think, five years ago with my own heartbeats, and I thought this time it would be good not to have a photo album, but a heartbeats album, that it would be possible to go home and say, oh, I’m going to hear my grandmother tonight. And after that, I decided to collect heartbeats. But, in fact, people are not going to survive with

this, it's going to show absence. If you go to this island in Japan [Boltanski's heartbeat recordings are archived on Teshima Island in Japan's Inland Sea], and if you hear the heartbeats of your grandmother, your grandmother's going to be very dead, and she's going to be more dead because you hear these heartbeats, and it is going to show you more that she's absent. In fact, it's like when you see a photo of somebody, you feel more that this person is dead. (Rosenbaum-Kranson & Boltanski, 2010)

The relationship I had with the document my father wrote coincided with Boltanski's description of the heartbeats, it was a constant reminder of his absence and as time passed by, it only intensified the fact that he was dead. From this association and inspired by the work of Lozano-Hemmer and Boltanski, I decided to record my own heartbeats while reading it in an effort to contain in a dataset the feeling of absence. Once I had the recording, I developed a series of interactive visualizations of the dataset so I could investigate any potential representations of it.

The first one called "Pulse Analysis"⁹¹ was intended as a strictly functional visualization since I wanted to clearly understand the raw data that was recorded by the pulse sensor. It visually displays the pulse as a constant waving line and highlights where the beats actually happen. It also shows me the frequency and places where the heart rate is going up or calming down. I can match these peaks with the time it took me to read and the data shows me that my pulse was mainly going up in pages that had my father's handwritten signature and the newspaper clippings which would often have pictures of him. After this functional analysis, I started to make less descriptive renderings of the data, trying to draw with it using a variation of the underlying code in a similar way that I engage with my sketchbook, as a free space for exploring with the line and allow

⁹¹ See "Pulse Analysis" in the code sketchbook website:
<http://www.dddrawings.com/lab/pulse-analysis/>

for unexpected associations to emerge in the drawings. From this exercise, I began to draw the pulse as an ocean⁹² in what seemed at first like a loose association between the form of the pulse and the ocean waves, but this led to making a connection between the document and the day that my father died.

The day my father died, he was on a business trip to the city of Cartagena, Colombia for the first time in his life. This was an exciting thing for him because he had never been in a city with an ocean and in preparation for his trip this was all we talked about. He took the earliest flight and as soon as I awoke I started to telephone him to learn his reaction to the ocean. His phone was off and we never spoke again. From what we know, he was killed soon after he arrived in the city and I never knew if he had a chance to go to the beach and enjoy the ocean as he had wanted to do for so long. This personal detail has deep significance to me and is what I remember the most when I think of his death. It fills me with sorrow and from thereafter, the beach has been a place I associate with grief and mourning. The piece “Vista al Mar” which translates from Spanish to “Ocean View” is a cathartic ritual in which I virtually visit the place where my father died and fill it with my own memories.

The treatment of the sound encloses the unfulfilled dream to know if he visited the beach or not, the closest the viewer is to the nearest beach in relation to its current geolocation, the louder we hear the ocean waves. It is both an invitation to “look” for the beach and a deliberate complication to reach it. But the viewer is not obliged to perform this task, for the piece is open-ended as to what trajectories the viewer embarks on and how long he or she spends in it. Using the *Google Street View* interface, it takes a long time to

⁹² Ibid: <http://www.dddrawings.com/lab/pulsescape-waves/>

actually travel from “El Gallo” to the nearest beach that is located at around 9.7 kilometers north-west from the center of the neighborhood. There is no visual indicator to help the viewer know where the ocean is located or if the viewer is traveling in the right direction. Taking advantage of the interface provided by Google, we are only able to travel through the streets that the car that took the pictures went through, in consequence, many of the streets are virtually inaccessible in the interface, making it easy to get lost and reach dead ends. The only indication of how near the ocean is from the current location is the very discreet variations of the volume to the soundtrack, the nearest we are, the louder we hear the ocean waves. Even for me that I have familiarized myself with the trajectory and trying my best to reach the beach as fast as possible using *Google Street View*’s interface, it often takes me hundreds of clicks and between 18 to 20 minutes to reach it. This deliberate complication to fulfill this task serves well my open question as to whether my father was able to experience the ocean. The ocean is relatively near as suggested by the sound and the viewer may or may not be able to find it, in the same way, that I don’t know if my father found it.

The goal is for this piece to be a personal repository of memories of my father, which can grow for as long as I add new drawings and animated sequences to it. It will mutate and expand as I keep adding memories, in the same way, that I feel his memory operates inside me. When I remember this day, it is never a fixed image and this memory dialogues with the current state of my life. As I grow and experience new events in my life, this memory changes and interacts with my evolving life.

I invite the viewers to establish this journey from my own perspective, and the mutating and ever-changing characteristics of

the piece allows me to only partially attain this internal state of mind. I argue that in my previous work, even though I was representing the internal states of mind as well, the formal characteristic of a short film establishes a different relation to memory. A film is contained in a singular state that is the finished film in celluloid or digital file and every view repeats the fixed version of the memory. On the other hand, here I began experimenting with the critical use of digital technologies, an evolving database of personal drawings and animations and taking advantage of the fact that Google keeps updating the pictures in their own database of street views. This allows for a closer representation of memories that are not fixed and evolve with time. I am conscious that every time we see something it may not suggest the same ideas in the viewers' mind, especially in evocative works where the goal is to result in contingencies that reinforce the possibility of a piece to be "interactive" in the sense that Moritz describes Eastern European animations with "convoluted narrative structure" (Chapter 3). Yet the possibility of creating art objects that mutate themselves and are interactive and open-ended as afforded by digital technologies is something I foresee as a formal and conceptual premise for my future artistic work as started to investigate in this research.

Project B: "Dibujo Defunciones"

In this second project, I am exploring an internal relationship I build with Colombian statistics of homicides as reported throughout the years. In a way, I remember my father every time I am aware of one of these crimes, I draw an emotional connection with the possible stories of the victims and their families and in turn, I relate these to my own personal story. This information is a constant reminder of

my own father and how his death stills affect me personally. Figuratively speaking, I see my father every time I become aware of a homicide.

To represent this in the piece, I have taken the pictures from the newspaper clippings in the document mentioned earlier and cropped the face of my father from them. There are four of them: 3 that show his face from the front and one in profile. I then created a loop of a constant cross-dissolves between each of these pictures in sequence, slowly transitioning between each of the pictures that appeared in the newspaper to create a morphing portrait of my father. This sequence is playing under a white layer I place on top of it and then reveal parts of it using a technique in which another drawing is used to crop the white layer to partially show what is underneath. The following diagram shows a simplified version of this technique. For purposes of clarification, the drawings layered on top of the portrait in the final piece are not as the one shown in this diagram, but to clearly show how it works, I used a solid abstract shape that simplifies the reading:

Figure 19. Diagram of simplified cropping technique, from “Dibujo Defunciones”, by González J., J. C., 2018.

The actual drawings that crop and reveal the portraits are a series of abstract animation sequences of line drawings for which I have used

as a reference to my father's hand-written signature. As described in the earlier project "Vista al Mar", I had noted that my heart rate was significantly higher when I was looking at both the pictures of my father and his signature. Based on this started to reflect on his signature and associated it with a vital drawing. The lines that compose it remind me of his personality as someone who was strong and intense. For example, the way I remember my father when he was signing something is that his gestures were so vigorous that he would hold the pen, lift his hand and move it in the air in a circular motion as in preparation for something that required a concentrated anticipation, then forcefully launch his arm towards the paper and write his name with such intensity that usually the surface that was holding the paper would shake from side to side. When I look at the signature now, I can see why he would anticipate it with those circular motions in the air, the big "C" at the beginning is a semi-circle and the rest of the lines maintain that circular motion:

Figure 20. My father's signature, from "Dibujo Defunciones", by González J., J. C., 2018.

He was a lawyer by profession and never particularly inclined towards drawing, or at least I have no recollection of him drawing, but his signature manifest to me now as such. Using his signature/drawing as a point of departure, I then took parts of the lines in it and used them as frames in a series of an abstract animated sequence. I anticipate each of these lines with a series of frames that construct its shape and then continue animating them in

different motions until they are reduced to small shapes and disappear into the paper. The sound that is synchronized to these sequences is recorded from me drawing on paper, these sounds emerge from the friction between a pen and the paper, translating in this way my memory of him vigorously writing his signature. This treatment of the sound allows me to embody and denote his performative gesture in the air when anticipating his signature and the memory of the surface shaking while he was “drawing” these lines on the paper. Not by realistically portraying this memory visually, but abstracting it into a sound that accompanies the abstract animations.

Once all these elements were created, I moved on to the association with the data of homicides in Colombia. This information was extracted from the “United Nations Office for the Coordination of Humanitarian Affairs (OCHA)” that has a project called “Monitor”⁹³ in which they publish regularly updated datasets of different types of violence in Colombia on their website⁹⁴. One of the types of violence indicated in these datasets refer specifically to homicides and each event has information on the location, date, time, victims count and gender. From the location indicated by the place name, I used the application *ddd-Parser* developed during this research to translate the places into geolocation data that I could map on the screen. Then for each event indicated in the data, the application that runs the piece randomly selects one of the animated sequences created before and places it on the screen in relation to the geolocation of the event.

⁹³ See the “Monitor” website where they map the data and allow for downloading it as spreadsheets: <http://monitor.umaic.org/>

⁹⁴ See source and description of the “Monitor” datasets in: <http://www.dddrawings.com/datasets/>

The selected animation sequence then starts playing from that particular location on the screen and reveal parts of the morphing portraits that are playing on a loop underneath.

Figure 21. Data Driven Drawing state of final visual composition, from “Dibujo Defunciones”, by González J., J. C., 2018.

The internal logic of the application is processing line by line the dataset that is ordered by the time of the year, in this way it progressively reveals the frequency with which these events occur. What this means is that when there is a high concentration of homicides or if they are occurring very frequently one after the other, there is a higher accumulation of line animation filling the screen and in turn revealing more of the underlying portrait. In association with my own relation to the data, the more homicides denoted in it, more of the image of my father is shown and in contrast if the homicides number decrease, the screen would remain whiter and the portraits invisible, but this never happens based on the dataset currently used.

The result is a foggy image slowly revealing and hiding parts of my father's face in a synchronized but unpredictable order. I argue that the sense of order and synchrony is given by the schematization of this violent phenomenon into data, which I can then use to build a logical set of rules that translate the data into representational elements. For instance: geolocation to screen coordinates, number of victims to abstract animation sequence and time of the event for when these elements are activated and shown. On the other hand, the unpredictability is provided by the complexity of the phenomena being represented. The frequency and patterns that emerge are given by the phenomenon itself and not ordered by myself, if the data changes, the resulting representation does as well. To reinforce this, the viewer has access to a menu placed on top of the screen with the different years that are available from the "Monitor" project. Currently, they provide data from 2008 to 2017 and each of these years will construct the resulting animation uniquely because the phenomenon never manifests, in the same way, every year. The goal is to keep adding data to the application as it becomes available. The "Monitor" project currently updates their datasets on a daily basis and as long as they keep publishing them, I would like to keep updating the data in my piece, and if they stop, I would find a new source to collect this information. Unfortunately, this type of violence is recurrent and resists to decrease in Colombia as reported in these datasets⁹⁵. Which means that every time I study the data I am confronted again and again with the memory of my father as one of the victims of this violent context. The way I see this is that the piece demands from me that I keep updating it and I

⁹⁵ For example, there were 3818 victims in the year 2008, 4672 in 2009, 4268 in 2010,

embrace this as a cathartic ritual for which I can meditate over my father's memory.

Project C: "N.N."

At the beginning of this series we started in the microscopic representation of a particular place in Colombia with personal significance, the second one steps further away to establish a point of view of all the homicides contained in the data, but denoting a singular person in the partially revealing portrait of my father. In this last one, I propose a view of violence in Colombia as denoted in the data from "Monitor", from far by looking at all of the indicators interplaying at the same time. I seek to probe what Burke suggested as the "delectable terror", not to suggest that there is a schematic path to evoking the sublime by pairing these two types of experiences, but it indicates a potential strategy I decide to investigate. The personal memory from which I make the aesthetic decisions for this representation is one that I consider to be a "delectable" memory: We went on a short excursion to one of the highest mountains in the center of Bogotá called *Cerro de Monserrate* which is at 3,152 meters above the ocean level. It was not crowded despite it being a popular tourist attraction for the view and the iconic church. Once there, we found a place with a good view of the city and stood there in silence for a very long time. This was rare on him, he liked to talk a lot, was hyperactive and rarely stood still. He once told me that staying still reminded him of his time in prison, and sometimes he would find himself aimlessly walking around the city just so he could feel that he was free. So, having him in silence and sitting for a long period of time made a strong impression on me. I don't remember it as an uncomfortable silence, I was happy to share this with him and he seemed calm and

happy as well. At some point, the silence was broken and we pondered over the sound of the city, from where we were, it sounded like a calm river despite the chaos and cacophony of sounds that compose it. This memory manifests to me like an explanation of data representation, from the complexity of sounds in the city, an abstraction emerges that describes it metaphorically: the city sounds like a river from afar, I do not know what it means, but at that moment with my father, it seemed evocative to us.

In this piece, I reconstruct this memory by abstracting each victim of violence in Colombia as denoted in the dataset by “Monitor” as a single dot on the screen: one victim is represented by 1 pixel. Depending on the type of violence this dot gets assigned a particular color. The geolocation is translated to screen coordinates from where the dot is first drawn and from this position it starts to move in a waving motion⁹⁶ as if it was a river line seen from afar. The progressive accumulation of lines denotes the amount of violence throughout the year, revealing a terrifying reality of the vast number of victims. The application goes through one year of data in about 3 minutes, thousands of dots rapidly fill the screen, yet this macroscopic view and this particular form of representation seem like a calm and organic generative system. The resulting image is also a distorted version of the Colombian map, showing that the violence and victims are from all around the country.

⁹⁶ The type of motion is defined programmatically using a Perlin Noise algorithm, originally created by Ken Perlin in 1983.

Figure 22. – Data Driven Drawing resulting state of the year 2008, from “N. N.”, by González J., J. C., 2018.

No sound is used in this project for the following reasons: the particular “delight” of that moment with my father was the comfortable silence we shared and since the rivers are represented visually, I decided to allow for silent contemplation of the animation.

In relation to the categories of data representation, this animation is able to fulfill, to a certain extent, a series of functional objectives. It reveals without demanding too much effort that the violence is distributed throughout the entire country, and by looking at the evolution of the image, usually the highest concentration of violence emerges from the center. This information is generic and ambiguous in comparison to what it could be if it was fully

functional. The graphic notation⁹⁷ I developed to study this data fills me more directly with “terror”, it shows the different types of violence as a list with concrete numbers that count the total number of victims per year and violence category, and through a simple histogram the frequency of the events is shown as vertical lines across the screen.

Figure 23. Graphic notation of violence data, from “Monitor Info”⁹⁸, by González J., J. C., 2017.

The graphic notation does not attempt to obscure any aspect of the data and aesthetically does not fulfill any particular tasks other than to clearly inform, but the revelation of the phenomena is impactful in the way that Tufte (Chapter 3) argued that functional visualizations are able to perform on their own when subjectivity is

⁹⁷ See graphic notation “Monitor Infor” at: <http://www.dddrawings.com/lab/monitor-info/>

⁹⁸ It is recommended to see it online at: <http://www.dddrawings.com/lab/monitor-info/>

removed. In this piece, I am betraying the moral responsibility that Tufte defended, not with the data, but with the impactful information it denotes. Yet, having represented this dataset as both purely functional and with the intent to evoke, I argue that the evocative provides a space for reflection that the functional removes. In the theoretical discussion of the sublime, Kant disputed Burke's proposition that the sublime is associated with terror. Kant argued that "Just as we cannot pass judgment on the beautiful if we are seized by inclination and appetite, so we cannot pass judgment at all on the sublime in nature if we are afraid." (Kant, 1987, p. 120) I concur with Kant in the sense that while the apprehension of the concrete knowledge causes terror on me, it does not suspend my ability to reason and understand the phenomena. On the other hand, when my father and I were impressed by the abstraction of the city sound as the river, it was because reason was forestalled, as mentioned earlier: "I do not know what it means", but find it poetic and evocative. If anything, this is the type of experiences I seek to represent in my work, ones that embrace ambivalence, pose questions, transform phenomena through metaphors and results in contingencies.

To answer the second question posed in the introduction: What is at stake for the animated film form when data is introduced to the production workflow? We should consider that the Data Driven Drawings examined above do not necessarily look like animated films in the traditional sense, but as examined in the theory, they stem from its tradition. As animated documentaries, they use data as the index of reality and therefore this source is intangible. Alluding back to it is a challenge of representation on its own. In "Vista al Mar", for instance, the audience may be able to infer that the data comes from heartbeats presentation is similar to other manifestations of it such as the sound or images produced in a

heart monitor in a hospital. But in “Dibujo Defunciones” and “N.N.”, the source of the data is not presented in any way to the audience. These abstractions seek to investigate in the image how the subjective experience relates to the data.

As discussed in the Introduction, using hand-drawn animations as is the cases of “N. N.” and “Rite of Passage” (Chapter 1), we borrow from video games the possibility to produce modular animations that blindly anticipate the complex structure of the datasets. While a traditional hand-drawn animation takes a lot of time to produce, these modular sprite sheets, at least the ones used for these projects, required very few drawings in comparison. But, even when pre-rendered drawings are made, the emerging compositions and sequences that emerge are highly complex. This allows the artists to iterate and investigate the representation as one has to negotiate between what we think might happen and what the data and algorithm produce the end. This means that different from traditional animated films, the artist loses control of how images are composed, and often how they move around the screen if this is implemented in the code. This forces us to re-examine the types of images and narrative structures that are possible to create in traditional films and move towards a creative approach to what we instruct the application to do with our images.

The type of negotiations we are forced to by animating with data are aligned with the overarching theoretical questions of artistic intent to represent functionally or evocatively. While functional representation requires precision and control to inform a singular argument, in evocation we do not necessarily conform the information for clear interpretation but open these for audiences to construct. This is fundamentally an attitude that allows for a rich exploration of the potential arguments we can articulate in

negotiation with the data, the realities they behold and the audiences that are capable of decoding our forms of communicating information in ways that are and should always be in a permanent state of flux.

Conclusions

In conclusion, this thesis should serve as introductory research to the issues concerning artistic data representation in general and more specifically to the ways in which animation and data can be combined in new forms of expression as proposed in the Data Driven Drawings projects. The combination of data and animation has proven to transform the regular conventions of what one might think animation is, at least the traditional format of film or digital file as its support was set aside to give way to interactive applications that are able to import the data and articulate it based on algorithms that interpret and give form to its representation. The preparation for these new kinds of animation is informed by the tradition of graphic notations as developed in chapter 2 and the results discussed in chapter 4 have successfully given this author the space to explore the premise set at the beginning: to articulate states of the mind in a constant state of flux and build a defense for subjectivity in data representation.

In the growing field of data representation within digital technologies, a lot of tools are readily available allowing for authors with different skillsets to develop data visualization. Despite this, the homogenization of codified standards to visualize datasets is concerning. The repetition of these codes is creating a sense that these are the forms in which data is able to inform us. Data is in itself formless (Floridi, 2011, 2016) and very importantly, it is in support of an argument (Gitelman, 2013), how we give form to our arguments with data should be assumed with an attitude of permanent exploration (Tukey, 1962, 1977, 1990) and research (Gaviria R., 2008). This research insists on this attitude towards data not for fashionable purposes of creating unconventional visualizations, but precisely in the challenges presented to the

audience, there is a clear intention to re-present, evoke and pose questions of the realities contained in the data. From the traditions of avant-garde art (Krauss, 1986) and experimental animation (Willaim Moritz, 1997) we know that audiences are capable of decoding the most unconventional accounts of reality through the lens of artists that will not sub estimate their audience and excite us through bold experimentation. The goal of data artists should not be to tame the data, but to investigate and develop eloquent articulations of it, even if they are difficult to “understand”.

Future Research

The next steps for this research have several objectives. Firstly, all of the practical projects and experiments are published online and are constantly referred to in the thesis, in this sense, the preferred form of presenting this to a wider audience would be to use a digital publication and also translate it to the authors native language (Spanish) since the contents relate closely to Colombian history.

The section dedicated to graphic notations did not have enough space to develop within the overall scope of this thesis. It deserves a research endeavor on its own to allow for expanding the sample of case studies. Since the source material for these objects is rarely published, they need to be extracted from both personal and public archives. The theoretical framework developed for studying the sample presented here should be tested and refined once there are more cases to examine. A project like this should also be presented online as the applications developed for studying them have not only been useful during the process of this research but have become a tool for teaching about alternative forms of preparing a film or animation.

An important and unique potential of working with data in an online platform is the possibility of working with the idea of real-time. In the past, this has only been explored in the telematic art as discussed by authors such as Roy Ascott, and pirate radio projects. While telematics was initially empowered by live T.V. systems, its ideas could be renewed with real-time communication systems that are common now on the internet. Yet the connection between telematic art and emerging real-time data applications has been barely explored artistically or theoretically. This approach is also of interest as it provides new methods for exploring the underlying theme of life in an ever-changing state of flux that motivated the practical projects here.

Also, the lack of examples of data driven animations seems to be rooted in the gap of knowledge and academic discussion in animation and art scholarship, which this thesis expects to address in part. But also in the fact that it requires both technical and artistic aptitudes to produce. The interdisciplinary nature of producing data art projects makes it difficult to adapt to conventional art, film or animation academic programs. For the ideas presented here to develop further, it is imperative to educate others in the theory and practice of Data Driven Drawings.

Lastly, and related to the previous point, another difficulty for developing these projects was the lack of Software needed for them. Even though each piece is developed as a web application using JavaScript programming language, the author felt the need to create custom Software for parsing and cleaning the data, another for creating sprite sheets from animated sequences and a JavaScript library for data visualization that does not impose codified forms visualizations and still is able to aid the developers in the common tasks of animating and visualizing data. In an effort to help others

find an easier way of producing Data Driven Drawings, these applications will be released in the future as open-source projects.

Bibliography

- Acosta, M. d. R. (2012). Desde el Umbral de las Palabras : Sobre lo Sublime a Partir de Pseudo-Longino. *Revista de Estudios Sociales, Iss 44, Pp 91-101* (2012)(44), 91.
- Akerman, J. R., & Karrow, R. W. (2007). *Maps : Finding Our Place in the World*: Chicago : University of Chicago Press, 2007.
- Barsam, R. M. (1992). *Nonfiction Film : A Critical History*: Bloomington : Indiana University Press, c1992. Rev. and expanded.
- Bazin, A. (2005). *What is Cinema?* : Berkeley : University of California Press, c2005.
- Bendazzi, G. (2016). *Animation : A World History*: Boca Raton, FL : CRC Press, Taylor & Francis Group, [2016].
- Berger, J. (1987, September). To Take Paper, to Draw. A World Through Lines. *Harper's Magazine*, 57-60.
- Bertin, J. (2011). *Semiology of Graphics : Diagrams, Networks, Maps* (W. J. Berg, Trans.): Redlands, Calif. : ESRI Press : Distributed by Ingram Publisher Services, 2011. 1st ed.
- Bihanic, D. (2015). *New Challenges for Data Design*: London : Springer, [2015].
- Bobrowska, O., & Bobrowski, M. (2016). *Obsession, Perversion, Rebellion: Twisted Dreams of Central European Animation*: Galeria Bielska BWA.
- Borges, J. L. (1964). *Other Inquisitions, 1937-1952. Translated by Ruth L.C. Simms. Introd. by James E. Irby* (R. L. C. Simms, Trans.): Austin : University of Texas Press, 1964.
- Boutwell, B. (2012). Morton Feldman's Graphic Notation: Projections and Trajectories. *Journal of the Society for American Music*, 6(4), 457-482. doi:10.1017/S1752196312000363
- Bret, V. (2012, 16 May). *Stop Drawing Dead Fish*. Paper presented at the SIGGRAPH, San Francisco.
- Brougher, K., & Mattis, O. (2005). *Visual Music : Synaesthesia in Art and Music Since 1900*: [New York] : Thames & Hudson ; Washington, D.C. : Hirshhorn Museum ; Los Angeles : Museum of Contemporary Art, c2005. 1st ed.

- Brown, R. H. (2012). The Spirit inside Each Object: John Cage, Oskar Fischinger, and "The Future of Music". *Journal of the Society for American Music*, 6(1), 83-113. doi:10.1017/S1752196311000411
- Burke, E. (1757). *A Philosophical Enquiry Into the Origin of Our Ideas of the Sublime and Beautiful*. London: London : printed for R. and J. Dodsley.
- Burke, E. (1796). *The Sublime and Beautiful*: Oxford : printed and sold at the Universities of Oxford, Cambridge, Edinburgh, Glasgow, St. Andrews, Aberdeen, and Dublin, 1796.
- Cage, J., & Knowles, A. (1969). *Notations*: [New York : Something Else Press, 1969].
- Cairo, A. (2013). *The Functional Art : An Introduction to Information Graphics and Visualization*: Berkeley, CA : New Riders, [2013].
- Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in Information Visualization : Using Vision to Think*: San Francisco, Calif. : Morgan Kaufmann Publishers, c1999.
- Charles, G., & Herbert, S. K. (1985). The Allotment of Space for Slaves Aboard Eighteenth-Century British Slave Ships. *The William and Mary Quarterly*(2), 238. doi:10.2307/1920430
- Clarkson, T. (1839). *History of the Rise, Progress, and Accomplishment of the Abolition of the African Slave Trade*: New York : Published by John S. Taylor.
- Damasio, A. R. (2000). *The Feeling of What Happens : Body and Emotion in the Making of Consciousness*: New York : Harcourt Inc., 2000, c1999. 1st Harvest ed.
- De Landa, M. (2000). Deleuze, Diagrams, and the Genesis of Form. *Amerikastudien / American Studies*(1), 33.
- Deleuze, G. (2003). *Francis Bacon : The Logic of Sensation* (D. W. Smith, Trans.): London : Continuum International Publishing Group, 2003.
- Deleuze, G., & Guattari, F. (1988). *A Thousand Plateaus : Capitalism and Schizophrenia*: London : Athlone Press, 1988, c1987.
- Eisenstein, S. (1975). *The Film Sense* (J. Leyda Ed.): New York : Harcourt Brace Jovanovich, c1975. [Rev. ed.].
- Eisenstein, S. (1977). *Film Form : Essays in Film Theory* (J. Leyda Ed.): New York : Harcourt Brace Jovanovich, 1977.

- Evarts, J. (1968). The New Musical Notation - A Graphic Art? *Leonardo: Journal of the International Society for the Arts, Sciences and Technology*, 1(4), 405-412.
- Eyler, J. M. (1979). *Victorian Social Medicine: The Ideas and Methods of William Farr* (1st ed.): The Johns Hopkins University Press.
- Eyler, J. M. (2001). The Changing Assessments of John Snow's and William Farr's Cholera Studies. *Sozial-und Präventiv Medizin*, 46, 225-232.
- Feltron, N. (2012). The Annual Report of Feltron 2012. Retrieved from <http://feltron.com/FAR12.html>
- Floridi, L. (2011). *The Philosophy of Information*: Oxford University Press, 2011.
- Floridi, L. (2016). *The Routledge handbook of philosophy of information. [electronic resource]*: New York : Routledge, 2016.
- Franceschet, M. (2014). Complex Beauty.
- Friendly, M. (2008). The Golden Age of Statistical Graphics. *Statistical Science*, 23(4), 502.
- Friendly, M., & Denis, D. J. (2001). Milestones in the history of thematic cartography, statistical graphics, and data visualization. Retrieved from <http://www.datavis.ca/milestones/>
- Fry, B. (2004). *Computational Information Design*. (PhD), Massachusetts Institute of Technology, Retrieved from <http://benfry.com/phd/dissertation-110323c.pdf>
- Fry, B. (2008). *Visualizing Data*: Sebastopol, CA : O'Reilly Media, Inc., 2008.
- Furniss, M. (2007). *Art in motion : animation aesthetics / Maureen Furniss*: Eastleigh, UK : John Libbey ; Bloomington, IN : Distributed in North America by Indiana University Press, c2007. Rev. ed.
- Gaviria R., A. (2008). When Is Information Visualization Art? Determining the Critical Criteria. *Leonardo*, 41(5), 479-482.
- Gitelman, L. (2013). *"Raw data" is an oxymoron*: Cambridge, Massachusetts : The MIT Press, [2013].
- Glickman, J. A. (2015). *A War at the Heart of Man: The Structure and Construction of Ships Bound for Africa*. (Masters of Arts in History), University of Rhode Island,
- Grierson, J. (1933). The Documentary Producer. *Cinema Quarterly*, 2, 7-9.

- Guldemon, J., & Keefer, C. (2012). *Oskar Fischinger 1900-1967 : experiments in cinematic abstraction / edited by Cindy Keefer and Jaap Guldemon*: Amsterdam : EYE Filmmuseum ; Los Angeles : Center for Visual Music, 2012.
- Halpern, O. (2014). *Beautiful data : a history of vision and reason since 1945*: Durham : Duke University Press, 2014.
- Hébert, P. (2014). A Matter of Affiliation. *Animac Magazine*.
- Heinrich, F. (2015). (Big) Data, Diagram Aesthetics and the Question of Beauty. *MedieKultur: Journal of Media & Communication Research*, 31(59), 73-94.
- Hempel, S. (2007). *The Strange Case of the Broad Street Pump: John Snow and the Mystery of Cholera* (First Edition ed.): University of California Press; First edition.
- Histórica, C. N. d. M. (2012). Atentados Terroristas 1988-2012. Retrieved from <http://www.centrodememoriahistorica.gov.co/micrositios/informeGeneral/basesDatos.html>
- Honess Roe, A. (2011). Absence, Excess and Epistemological Expansion: Towards a Framework for the Study of Animated Documentary. *Animation* (17468477), 6(3), 215.
- Honess Roe, A. (2013). *Animated documentary / Annabelle Honess Roe*: Houndmills, Basingstoke, Hampshire ; New York : Palgrave Macmillan, 2013.
- Husbands, L. (2013). The meta-physics of data: Philosophical science in Semiconductor's animated videos. *Moving Image Review & Art Journal*, 2(2), 198.
- Iversen, M. (2012). Index, Diagram, Graphic Trace. *Tate Papers*, No. 18(Autumn 2012).
- Jensen, G. M. (2009). John Cage, Chance Operations, and the Chaos Game: Cage and the "I Ching". *The Musical Times*(1907), 97.
- Kant, I. (1987). *Critique of judgment* (W. S. Pluhar, Trans.): Indianapolis, Ind. : Hackett Pub. Co., c1987.
- Kant, I., & Goldthwait, J. T. (2003). *Observations on the feeling of the beautiful and the sublime / Immanuel Kant ; translated by John T. Goldthwait*: Berkeley, Calif. ; London : University of California Press, 2003. 2nd paperback ed.

- Kitchin, R. (2014). *The data revolution : big data, open data, data infrastructures & their consequences*: London : Sage Publications Ltd, 2014.
- Koch, T. (2004). The Map as Intent: Variations on the Theme of John Snow. *Cartographica*, 39(4), 1-14.
- Krauss, R. E. (1986). *The originality of the avant-garde and other modernist myths*: Cambridge, Mass. : MIT Press, 1986, c1985.
- First MIT Press paperback ed.
- Lozano-Hemmer, R. (2006). Almacén de Corazonadas | Pulse Room. Retrieved from http://www.lozano-hemmer.com/pulse_room.php
- Lupi, G., Popova, M., & Posavec, S. (2016). *Dear data*: New York : Princeton Architectural Press, First edition.
- Mackinlay, J. (1986). Automating the design of graphical presentations of relational information. *ACM Transactions on Graphics*, 5(2), 110.
- Manovich, L. (2001). *The language of new media / Lev Manovich*: Cambridge, Mass. : MIT Press, c2001.
- Manovich, L. (2002). Data Visualization as New Abstraction and as Anti-Sublime. In *Small Tech [electronic resource] : The Culture of Digital Tools* (pp. 3-9): Minneapolis : University of Minnesota Press, 2008.
- Manovich, L. (2010). What is visualisation? *Visual Studies*, 26(1), 36-49. doi:10.1080/1472586X.2011.548488
- McCosker, A., & Wilken, R. (2014). Rethinking 'big data' as visual knowledge: the sublime and the diagrammatic in data visualisation. *Visual Studies*, 29(2), 155-164. doi:10.1080/1472586X.2014.887268
- McLaren, N. (2006). *Technical Notes by Norman McLaren (1933-1984)*: National Film Board of Canada.
- McLaren, N., & McWilliams, D. (1991). *Norman McLaren on the Creative Process* (D. McWilliams Ed.): National Film Board of Canada.
- Moritz, W. (1997). Narrative strategies for resistance and protest in Eastern European animation. In J. Pilling (Ed.), *A reader in animation studies / edited by Jayne Pilling*: London : J. Libbey, c1997.
- Moritz, W. (2004). *Optical poetry : the life and work of Oskar Fischinger*: Eastleigh : John Libbey, 2004.

- Morley, S. (2010). *The Sublime*: London : Whitechapel Gallery ; Cambridge, Mass. : MIT Press, 2010.
- Neurath, O. (1936). *International Picture Language*: Department of Typography & Graphic Communication, University of Reading.
- Nichols, B. (1991). *Representing Reality : Issues and Concepts in Documentary*: Bloomington : Indiana University Press, c1991.
- Nielsen, F. Å. (2011). A New ANEW: Evaluation of a Word List for Sentiment Analysis in Microblogs.
- Nightingale, F. (1858). Mortality of the British Army : At Home and Abroad, and During the Russian War, As Compared with the Mortality of the Civil Population in England. Retrieved from <https://archive.org/details/mortalityofbriti00lond>
- Nightingale, F. (1859). *A Contribution to the Sanitary History of the British Army During the Late War with Russia*: Harrison.
- Nørretranders, T. (1999). *The User Illusion : Cutting Consciousness Down to Size*: New York, N.Y. : Penguin Books, 1999.
- Patrick, E. (2004). Representing Reality: Structural / Conceptual Design in Non-Fiction Animation. *Animac Magazine no. 3*, 39-47.
- Peirce, C. S. (1932). *The Collected Papers of Charles Sanders Peirce, Vol. II: Elements of Logic* (C. W. Hartshorne, Paul Ed. Vol. II). Cambridge: Harvard University Press.
- Pseudo-Longinus. (1890). *On the sublime* (H. L. Harvell, Trans.): London and New York, Macmillan and co.
- Reddy, F. (2012). Astronomers Uncover a Surprising Trend in Galaxy Evolution. Retrieved from <https://sys.gsfc.nasa.gov/vis/a010000/a011000/a011087/>
- Renov, M. (2004). *The Subject of Documentary*: Minneapolis : University of Minnesota Press, c2004.
- Rogers, H., & Barham, J. (2017). *The Music and Sound of Experimental Film*: Oxford University Press.
- Rosenbaum-Kranson, S., & Boltanski, C. (2010). Christian Boltanski Interview. Retrieved from <http://www.museomagazine.com/802500/CHRISTIAN-BOLTANSKI>
- Russett, R., & Starr, C. (1988). *Experimental Animation : Origins of a New Art*: New York, N.Y. : Da Capo Press, 1988, c1976. Rev. ed.

- Rybczyński, Z. (2009). *Traktat o Obrazie / A Treatise on the Visual Image*: Art Stations Foundation.
- Sauer, T. (2009). *Notations 21*: Mark Batty Publisher.
- Schopenhauer, A. (1969). *The World as Will and Representation*: New York : Dover Publications, [c1969-.
- Segaran, T., & Hammerbacher, J. (2009). *Beautiful Data : The Stories Behind Elegant Data Solutions*: Beijing ; Sebastopol, CA : O'Reilly, c2009.
- Sofian, S. (2005). The Truth in Pictures. *Frames per Second Magazine of Animation, Vol. II*, 7-10.
- Steele, J., & Iliinsky, N. P. N. (2010). *Beautiful Visualization : Looking at Data Through the Eyes of Experts*: Beijing ; Sebastopol, CA : O'Reilly, 2010. 1st ed.
- Taruskin, R. (2008). *The Danger of Music and Other Anti-Utopian Essays*: University of California Press.
- Tufte, E. R. (1990). *Envisioning Information*: Cheshire, Conn. : Graphics Press, c1990.
- Tufte, E. R. (1997). *Visual Explanations : Images and Quantities, Evidence and Narrative*: Cheshire, CT. : Graphics Press, c1997. Second printing, with revisions, May 1997.
- Tufte, E. R. (2001). *The Visual Display of Quantitative Information*: Cheshire, Conn. : Graphics Press, c2001. 2nd ed.
- Tufte, E. R. (2006). *Beautiful Evidence*: Cheshire, Conn. : Graphics Press, c2006.
- Tufte, E. R. (2017). Florence Nightingale's statistical graphics. Retrieved from https://www.edwardtufte.com/bboard/q-and-a-fetch-msg?msg_id=00007o&topic_id=1
- Tukey, J. W. (1962). The Future of Data Analysis. *The Annals of Mathematical Statistics*(1), 1.
- Tukey, J. W. (1977). *Exploratory Data Analysis*: Reading, Mass. : Addison-Wesley Pub. Co.
- Tukey, J. W. (1990). Data-Based Graphics: Visual Display in the Decades to Come. *Statistical Science*(3), 327.
- Vesna B., V. (2007). *Database Aesthetics : Art in the Age of Information Overflow*: Minneapolis : University of Minnesota Press, c2007.

- Viégas, F., & Wattenberg, M. (2007). Artistic Data Visualization: Beyond Visual Analytics. In D. Schuler (Ed.), *Online Communities and Social Computing* (Vol. 4564, pp. 182-191): Springer Berlin Heidelberg.
- Ward, P. (2005). *Documentary : The Margins of Reality*: London : Wallflower, 2005.
- Wells, P. (1997). The Beautiful Village and the True Village: A Consideration of Animation and the Documentary Aesthetic. *Art & Design Magazine, Profile no. 53*.
- William S. Cleveland, a., & Robert McGill, a. (1984). Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*(387), 531. doi:10.2307/2288400
- Yau, N. (2011). *Visualize This : The FlowingData Guide to Design, Visualization, and Statistics*: Indianapolis, Ind. : Wiley Pub., c2011.
- Youngblood, G. (1970). *Expanded Cinema*: New York : Dutton, 1970. [1st ed.].