



# EXPLORING TEXTUAL DATA: TRANSMUTABILITY AS A CREATIVE CONCEPT AND PRACTICE

## CATARINA LEE

Independent researcher,  
Lisbon, Portugal  
catarinaleecastanheira  
@gmail.com

## LUÍSA RIBAS

ID+/ Faculty of Fine Arts,  
University of Lisbon,  
Portugal  
l.ribas@belasartes.ulisboa.pt

### Keywords

Digital data  
Text  
Visualization  
Sonification  
Transmutability

This paper explores the creative potential of the transmutability of digital data, from a theoretical discussion of the concept to its contemporary manifestations in creative practices. It begins by addressing the creative possibilities associated to the topic and then provides an overview of artifacts that imply or express transmutability as an artistic concept and method, while focusing on data sources of a textual nature.

To this end, we resort to a framework for the description and analysis of these artifacts, focusing on their conceptual dimension, on their mechanics and on the elements of their experience. In particular, we address the concepts they approach through the use of data in textual formats as source information or content, we consider the processes for its manipulation, and describe the resulting sensory manifestations while emphasizing their dynamics and variability.

In this manner, this study seeks to highlight how transmutability becomes relevant as an artistic argument, by proposing aesthetic experiences that explore the ubiquity of data in our contemporary world.

2016.  
**xCoAx**  
.org

Computation  
Communication  
Aesthetics  
& X  
Bergamo, Italy

## 1 INTRODUCTION

In order to understand the creative potential of the mutability of digital data, we can begin by considering that, within the computer, “all media objects are composed of digital code; they are numerical representations” (Manovich 2001, 27). Thus, digital data, when regarded as raw material, can be translated into any tangible form through algorithmic manipulation. This creative potential is explored through practices that rely on software as their medium and involve articulations between sounds, images, or other physical or sensory realms.

According to Golan Levin (2010), the underlying principle motivating the development of such artworks is the transmutability of digital data. This notion becomes the conceptual and technical starting point of this study, which seeks to provide an understanding of the concept and examine creative practices that not only explore an analytical view of data, but also develop expressive audiovisual languages that provide new perceptions or aesthetic experiences of data.

Following this idea, the focus of this work is on the exploration of textual material, given that “a lot of the richest information we have” is available in text formats (Viégas *qtd. in* Heer 2010, 7) and “grows on a daily basis” (Nualart-Vilaplana *et al.* 2014, 224), while considering the “advances being made in text analysis research” and computational manipulation (Nualart-Vilaplana *et al.* 2014, 221). This represents a transformative potential worthy of development and exploration, which implies a reflection on data that entails the creative potential of the “text processing algorithms” (Kucher and Kerren 2015, 117) applied to its manipulation and transformation.

## 2 TRANSMUTABILITY AS A CONCEPT

The principle of transmutability relies on the mapping of any input data stream into sounds and images; as Golan Levin emphasizes, the “premise that any information can be algorithmically sonified or visualized” can be the “starting point for a conceptual transformation and/or aesthetic experience” or a means of “enabling some data stream of interest to be understood, experienced, or made perceptible in a new way” (2010, 273-4).

In this sense, the notion appears associated to other concepts that similarly express the inherent mutability of digital data or the potential of mapping any (digitized) physical or sensory phenomenon into new tangible forms. It is related to the *transcoding* of digital data, as a direct consequence of describing information

numerically (Reas *et al.* 2010, 79) and evokes *transmediality* as a “translatability across media” (Hayles 2006, 194). This notion can also be associated to *transmateriality* as a term that sees “digital media and computation as material flows (...) transducing anything to anything else” by “sourcing new inputs and/or manifesting new outputs” (Whitelaw 2009).

Therefore, we can say that artifacts that explore this inherent mutability of digital data creatively question the “nature of our now ubiquitous data systems” by making data “explicit” and tangible, while probing its “potential, and significance” (Whitelaw 2008). In this process, different approaches and methods for re-configuring data may be involved, following mainly analytical or aesthetic purposes. This means that the aim of the project, “from poetic to functionalist” (Whitelaw 2009), can be to provide a “new reading or understanding of information” or, in turn, to explore data in order to “create expressive languages or sensory experiences” (Lee *et al.* 2014, 420).

### 3 APPROACHES TO TEXTUAL DATA

In accordance with these analytical or expressive approaches, when considering data in textual format, we can identify different conceptual purposes and aesthetic intents.

Some projects assume textual data *per se* as the subject matter of the work, that is, they consider text as raw material (“the text as it is”) or extract and consider “a representative part of that text” (Nualart-Vilaplana *et al.* 2014, 224) as the result of “text mining algorithms” (Kucher and Kerren 2015, 117). The focus of these works is on the exploration of the formal specificities of text as source material, considering that a text can have “multiple internal structures”, a specific morphology (paragraphs, sentences, words), diverse data types or formats (txt, html, etc.) and different patterns, as well as “a subjective component and an abstract structure that is not readily analysed by a computer” (Nualart-Vilaplana *et al.* 2014, 223-224).

Other projects consider textual data as *content* that conveys some kind of meaning, or represents a given subject matter. In these cases, the focus is on semantics rather than form, and the aim is to propose a new “understanding, perception or experience” of that content (Levin 2010, 274) or to “portray not merely data, but the personal, emotional reality that the dataset refers to” (Whitelaw 2008).

Finally, text can be considered as an *abstraction*, in the sense that what is conceptually emphasized is the translation process itself (Levin 2010), or the possibility of mapping any kind of data

into a new tangible representation. In such cases, and “depending on how the text is treated and processed”, it can be detached from its semantics, being that the textual source or origin “is not always relevant” (Nualart-Vilaplana *et al.* 2014, 228). The dataset is treated as “an abstract set of potentials”, since “the process doesn’t care what the dataset is, or was”, and treats it as “just input” (Whitelaw 2008).

These different strategies thus expose the potential of translating and revealing inherent, and eventually latent or hidden, dimensions of text into a new expressive manifestation, relating to its *formal* specificities, *semantic* aspects, or *abstraction* through a translation or mapping process.

## 4 TRANSMUTABILITY AS A CREATIVE PRACTICE

In order to provide an overview of the range and scope of creative approaches that are tied to the principle of transmutability, in their potential diversity, we selected a group of artworks corresponding to the following criteria: (1) use software as medium; (2) explicitly work on or explore information in textual format; (3) entail visualization and/or sonification methods; (4) whose result emphasize the significance of data and/or the transformational process involved as subject matter of the work.

1. Dragulescu, Alex. *Spam Architecture*. 2005.
2. DuBois, Luke. *Hard Data*. 2010
3. Fry, Ben. *On the Origin of Species*. 2009
4. Harris, Jonathan and Sepandar Kamvar. *We Feel Fine*. 2006
5. Harrison, Chris. *Bible Cross-References*. 2008
6. Jevbratt, Lisa. *1:1*. 1999–2002
7. Luining, Peter. *ZNC Browser 2.0*. 2003
8. Maigret, Nicolas. *Pure Data Read as Pure Data*. 2010
9. Rubin, Ben and Mark Hansen. *Listening Post*. 2001
10. Rubin, Ben and Mark Hansen. *Shakespeare Machine*. 2012
11. Viégas, Fernanda and Martin Wattenberg. *History Flow*. 2003
12. Whitelaw, Mitchel. *Weather Bracelet*. 2009

In order to analyze these works we resorted to the frameworks proposed by Wardrip-Fruin (2006) and Hunicke, LeBlanc and Zubek (2004) for understanding aesthetic artifacts that are digital computational systems, or works that are *driven by processes*, as *dynamic systems*. These frameworks highlight that, when examining these artifacts, we must consider not only their sensory results or modes of expression but also their procedural modes of expression and dynamics (Ribas 2014, 53).

In this sense, the model proposed by Wardrip-Fruin addresses the interplay between data, processes, surface, interaction, author and audience (2006, 9-11). It also considers the “forms and

roles” of computation that distinguish the ways in which the work operates, according to its computational variability, interaction and source of interaction (2006, 398). In addition, the MDA framework provides different but interrelated perspectives focused on their mechanics, dynamics and aesthetics (Hunicke *et al.* 2004).

Drawing on these frameworks, our analysis highlights the alignment between the works’ themes and concepts, as implemented through specific data and processes, while considering the elements of their experience, namely the surface elements and dynamic behavior that define the works’ experience (Lee *et al.* 2014, 423), according to the following dimensions:

**Conceptual dimension** (theme and content) – considering the subject matter of the work (relating to its content, such as its approaches to text), while addressing the significance and relevance of transmutability as an artistic argument;

**Mechanics dimension** (data and processes) – regarding the implementation of concepts with specific data and processes as constituent elements of the system (data collection, values and input method, as well as mapping processes and their possible articulations);

**Experience dimension** (surface and dynamics) – contemplating the sensory outcomes (output format, modes of expression) and the observable behavior of the work (output nature, system behavior), as aspects pertaining to the nature of the work as a technological and aesthetic artifact, and relating to the variability and determinability of its behavior.

By considering such views we seek to describe the salient traits of these projects, while tackling into the questions that their conceptualization, enactment and experience may raise.

## 5 ANALYSIS

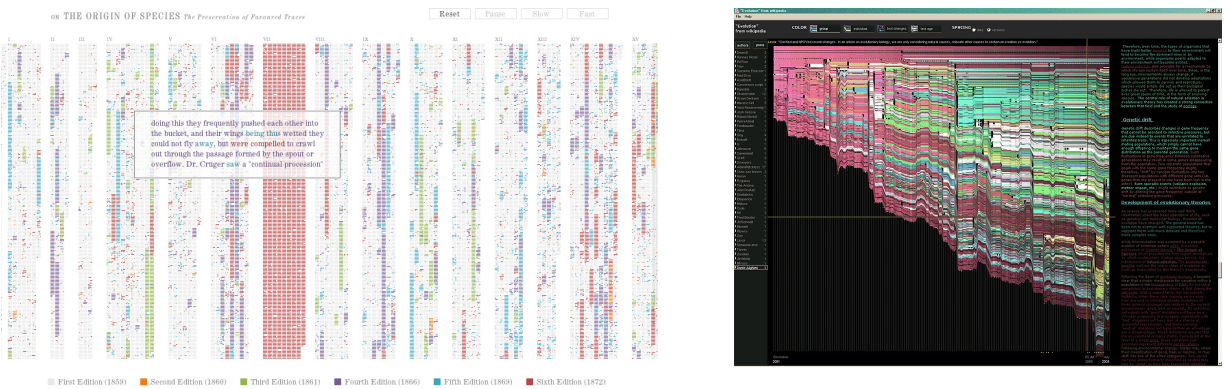
### 5.1 CONCEPTS: THEMES AND APPROACHES

According to the previously mentioned approaches to textual data, we can identify diverse creative and aesthetic intents, as well as relationships to text as the main referent or subject matter.

We distinguish projects that tend to explore the *formal* and *material* qualities of text (its format or internal logic), for example, manifesting a particular interest in literary works as “a field that, apart from being characterized by complex combinations

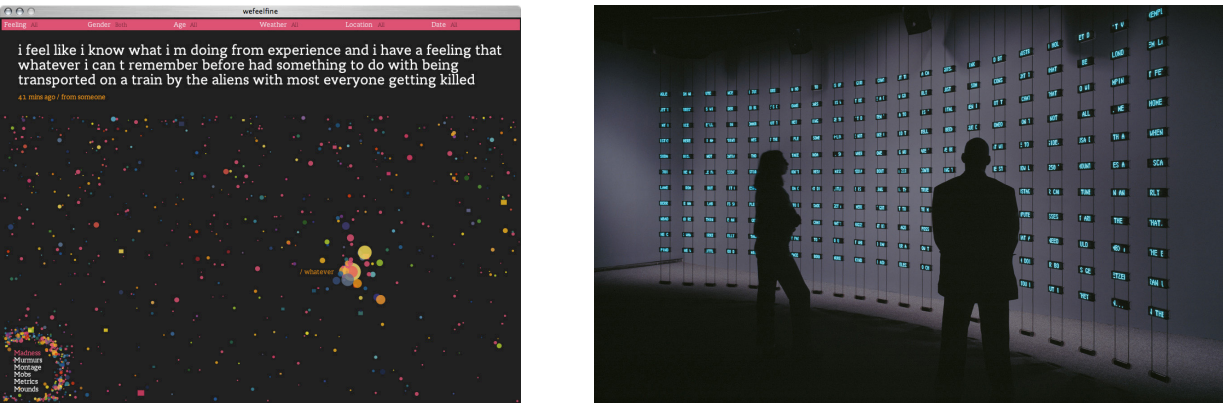
of words, can present high levels of human abstraction and freedom of structure and experimentation” (Nualart-Vilaplana *et al.* 2014, 234). Works such as Ben Fry’s *On the Origin of Species* (2009) give us a perception of the evolution of scientific ideas and the gradual refinement of Darwin’s discourse over several editions of the book. Another example is *History Flow* (Viégas and Wattenberg 2003) that visualizes and reveals patterns emerging from the editing history of Wikipedia articles.

**Fig. 1.** *On the Origin of Species* (Fry 2009) (left) and *History Flow* (Viégas and Wattenberg 2003) (right).



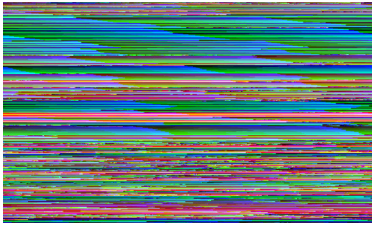
Other projects, in turn, focus on content, using text as a means to explore a given subject matter. Rather than focusing on the text format, these projects focus on the meaning that the text conveys, seeking to express or portray the reality that the textual data refers to, as an “index of reality” (Whitelaw 2008). For example, *We Feel Fine* (Harris and Kamvar 2006) is defined as an “exploration of human emotion”, by gathering “emotional data” on a global scale, through the search of blog entries with occurrences of the phrases “I feel” and “I am feeling”. Another example is *Listening Post* (Rubin and Hansen 2001) that provides an audiovisual reading of online conversations in real-time, by collecting data from unrestricted blogs and forums, as a reflection on the “immediacy of virtual communication”.

**Fig. 2.** *We Feel Fine* (Harris and Kamvar 2006) and *Listening Post* (Rubin and Hansen 2001).







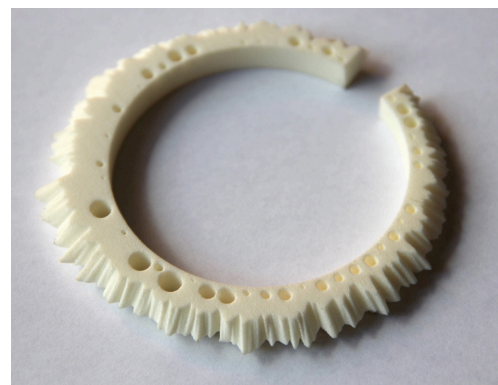
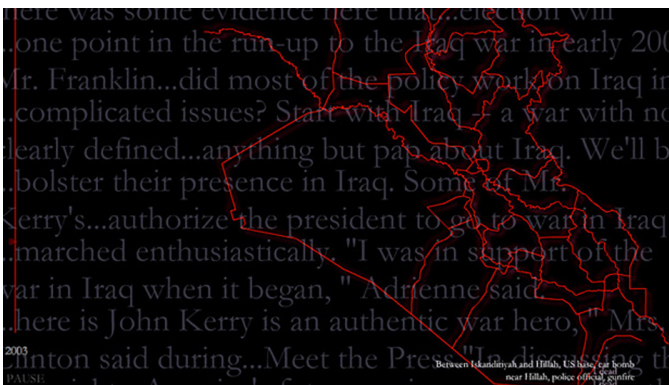


**Fig. 5.** 1:1 (Jevbratt 1999-2002).

However, some projects, use a continuous data stream, whose values are changing in real time, or even chunks of information that gradually update the values. These streams or chunks are usually captured through computational processes and inserted into the system automatically, as in *1:1* (Jevbratt 1999-2002) that uses web crawlers to search for IP addresses, which are then stored in databases that are visualized through different interfaces.

### 5.3 SURFACE: SENSORY RESULTS AND EXPRESSION

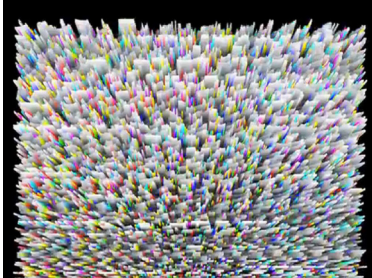
The diversity of sensory modes of expression and formal aspects of representation that we see in these projects are tied to their different aesthetic intents and approaches to textual data as subject matter. We observed that many of these works use visualization methods, proposing a purely visual expression of data, while only a few examples use sound or sonification in addition to, or as a complement, to the visualization procedures. That is the case with *Hard Data* (DuBois 2009), in which the author seeks to re-contextualize “formal stochastic music in the context of real-world statistics”, while creating abstract audio-visual experiences based on data from the American military actions in Iraq. Among the projects analyzed, we also included one example of a physical rendering of data, that is, *Weather Bracelet* (Whitelaw 2009) in which the author creates a “wearable data-object” generated from daily weather data sourced from the Bureau of Meteorology.



**Fig. 6.** *Hard Data* (DuBois 2009) (left) and *Weather Bracelet* (Whitelaw 2009) (right).

In terms of formal aspects of representation and expression, it is not always evident what aspects or parts of the text are actually represented in the output, through the visualization or sonification process, being that many projects don't even present textual information as output, and only a few present parts of the





**Fig. 7.** *Pure Data Read as Pure Data* (Maigret 2010).

source text. In *Shakespeare Machine* (Rubin and Hansen 2012), fragments of speech “appear, dissolve, and move like a choreographic dance”, according to an algorithm that sets rules for the combinations of words. In turn, *On the Origin of Species* (Fry 2009) presents the whole source text in the output.

Nevertheless, most of the projects analyzed approach structural aspects of the text, such as grammatical or morphological attributes as parameters that are mapped into graphic or audio features. Many of these projects resort to abstract elementary figures and sounds that, when combined, can reveal unexpected patterns or rhythms, or even complex configurations emerging from the data. For example, *Pure Data Read as Pure Data* (Maigret 2010) translates the source code of the application Pure Data into sounds and colored pixels, in order to promote a “physical experience of the digital data”.

#### 5.4 BEHAVIOUR: DYNAMICS AND VARIABILITY

Adding to the mentioned formal aspects of expression, the source of data also influences the nature of the output and the dynamic behavior of the work, depending on whether the work is open or not to interaction with external input.

The use of a fixed dataset usually corresponds to a system that is closed to external input. As such, the output is an instance that the system generates each time it runs, resulting in either a static or a transient (non-variable) output that promotes a contemplative experience based on the formal or semantic qualities of the source data.

In this case, the output can be a static image resulting from a process of ‘filtering’, such as a selective “snapshot” of the final state of the work or of “accretions” of processes over time (Dorin et al. 2012, 247). For example, the project *Bible Cross-References* (Harrison 2008) presents a global view of the “textual cross references found in the Bible” through diagrams that “honor and reveal the complexity of the data”.

When the output is transient, as a time-based or animated sequence (usually in response to a time dependent dataset), the work privileges a perception of patterns emerging from the text or a way to “understand or follow its evolution over time” (Nulart-Vilaplana et al. 2014, 230). An example of that is *On the Origin of Species* (Fry 2009), whose animated visualization demonstrates the changes and additions of text over the successive editions of the book.

Conversely, a continuous data stream can be used to gradually determine output variations, providing an immediate perception of input fluctuations coming from external data sources or

processes. That is the case in *Listening Post* (Rubin and Hansen 2001) that culls information from online sources in real time. Also, in *We Feel Fine* (Harris and Kamvar 2006) we can see that the interface grows and changes as new updates in the blog entries are found.

In addition, when the user can explore or navigate different views, the experience of the output becomes varied, even if the system is not necessarily producing variable results while acting on the same input. For example, in *1:1* (Jevbratt 1999-2002) the user is allowed to navigate through the interface, being able to “query the (visualization) system and obtain a unique representation for each search” (Nualart-Vilaplana et al. 2014, 230).

## 6 DISCUSSION

According to these observations we can highlight what these projects share, as a creative exploration of textual data, and how they diverge, regarding their conceptual approaches to the source data, as well as the different aesthetic intents and kinds of experience they propose.

Based on our selection, we observed that projects that use literary works put an emphasis on form since these texts present a “high level of abstraction and little formal structure” (Nualart-Vilaplana et al. 2014, 228). The potential lack of regularity in terms of vocabulary or length of the texts and their subjective discourse structure, result in more creative freedom and expressive possibilities, since there are no given conventions or rules for representation.

The fixed nature of these texts is usually associated with a sequential analysis of the whole, that is, the visualization often follows the texts’ sequence or order. One exception is *Shakespeare Machine* (Rubin and Hansen 2012) in which parts of the text are selected according to different rules, and the reference to the original text sequence is then discarded.

The examples analyzed, however, seldom explore aspects inherent to the literary text by means of sonification processes. And this is something we consider worthy of further examination and exploration, given the mentioned high level of abstraction and formal structure of the text and their openness to subjective interpretation.

When the focus is on meaning, a wider scope of themes emerges, ranging from human social dynamics (e.g. virtual online communication, identity, or different kinds of statistics), to natural phenomena (e.g. meteorological data), or even to the density and complexity of the web structure. These projects tend to either

work with a fixed dataset or sequential updates of that data, presenting an indexical narrative of a reality, thus putting to the fore its latent, or even hidden, patterns.

Finally, the exploration of text as an abstraction is mostly related to an analysis of data as raw material, pertaining, for example, to web content or digital data that can be readily analyzed by computational means; it can be considered as it is and subjected to any kind of arbitrary mapping, hence emphasizing its “malleability” and “susceptibility to transformation” (Whitelaw 2008).

The dataset is thus detached from any given meaning and treated according to a subjective process or conceptual approach, being that the output does not necessarily point to a direct relationship with the source data. Since the nature of the source data does not determine or condition the mapping process, this kind of approach is more prone to involve sonification and audiovisual results.

## 7 CONCLUSION AND FUTURE WORK

The previous discussion also suggests that the aesthetic experience of these works is not merely focused on their sensory results, but on the understanding of the processes leading to the observable results. Accordingly, we can consider that “what we experience, even as static displays”, are the results of “software performances”, which give us not objects but instances or occasions for experience (Manovich 2013, 33). So we interpret these outputs as the products of processes. In this sense, these projects entail a process of “procedural interpretation” or an understanding of the work that often involves “mental simulations of the processes behind the surface” (Carvalhais and Cardoso 2015, 143-144).

According to this idea, we acknowledge the potential for a deeper examination of these forms of procedural interpretation, namely, through a refinement of the framework concerning the distinctive features of the experience of these artworks. This implies considering their dynamics and variability, and therefore, further discussion of what we consider to be the aesthetic artifact in question; the system and the outcomes it presents to the audience as instances or events. Consequently, when examining the dynamics of the work, it is important to consider both the variability of the system and the variability of the outputs, given that the aesthetic artifact can be considered both the software and its outcomes, as in *1:1* (Jevbratt 1999-2002).

Furthermore, an examination of a broader scope of systems that are open to interaction with external data or processes can

be of interest, in particular, considering human input, or audience interactive work as well as the possibilities that are given to the audience for accessing, influencing or determining variable outcomes.

Acknowledging the multiplicity of transmutability as a creative concept and practice, this study sought a deeper understanding of artistic approaches to textual data, highlighting their focus on form, content or abstraction. To this end, it described a set of aesthetic artifacts according to a framework focused on their themes and subject matter (concepts), their data and processes (mechanics) and their surface and dynamics (as the elements of their experience).

With this approach, this study sought to reveal the creative and expressive potential of transmutability and to emphasize its relevance as an artistic argument that comments on the growing amount of digital data that permeates our contemporary world.

## REFERENCES

- Carvalhois, Miguel, and Pedro Cardoso.** "Beyond Vicarious Interactions: From Theory of Mind to Theories of Systems in Ergodic Artefacts". *xCoAx 2015: Proceedings of the 3rd Conference on Computation, Communication, Aesthetics and X*. Glasgow: 139-150. 2015.
- Dorin, Alan; Jonathan McCabe, Jon McCormack, Gordon Monro, and Mitchell Whitelaw.** "A Framework for Understanding Generative Art". *Digital Creativity* 23, n. 3-4, 239-259. DOI: 10.1080/14626268.2012.709940. 2012.
- Hayles, Katherine.** "The time of digital poetry: from object to event." *New Media Poetics: Contexts, Technologies, and Theories*. Morris, Adalaide and Thomas Swiss (eds). Cambridge, Massachusetts: MIT Press, 181-209. 2006.
- Heer, Jeff.** "A conversation with Jeff Heer, Martin Wattenberg, and Fernanda Viégas". *Queue*, March, v. 8, n. 3. <<http://doi.acm.org/10.1145/1737923.1744741>> [visited: 06-10-2015]. 2010.
- Hunicke, Robin; Marc LeBlanc, and Robert Zubek.** "MDA: A Formal Approach to Game Design and Game Research." *Proceedings of the Challenges in Games AI Workshop*. Nineteenth National Conference of Artificial Intelligence, San Jose, California. 2004.
- Kucher, Kostiantyn, and Andreas Kerren.** "Text Visualization Techniques: Taxonomy, Visual Survey, and Community Insights". *Proceedings of the 8th IEEE Pacific Visualization Symposium (PacificVis'15)*. IEEE Computer Society Press, Hangzhou, China: 117-121. 2015.
- Lee, Catarina; Luísa Ribas, and Miguel Cardoso.** "On the notion of transmutability: from concept to practice". *xCoAx 2014: Proceedings of the 2nd Conference on Computation, Communication, Aesthetics and X*. Porto: 416-429. 2014.
- Levin, Golan.** "Audiovisual Software Art". *Audiovisuality: Compendium*. Daniels, Dieter and Sandra Naumann (eds). Cologne: Verlag der Buchhandlung Walther König, 271-283. 2010.
- Manovich, Lev.** *Software takes command*. Open access ed. New York, London: Bloomsbury Academic. 2013.
- . *The Language of New Media*. Cambridge, MA: MIT Press. 2001.
- Nualart-Vilaplana, Jaume; Mario Pérez-Montoro, and Mitchell Whitelaw.** "How we draw texts: a review of approaches to text visualization and exploration". *El profesional de la información*, May-June, v. 23, n. 3, 221-235. <<http://dx.doi.org/10.3145/epi.2014.may.02>> [visited: 06-10-2015]. 2014.
- Reas, Casey, Chandler McWilliams, and LUST.** *Form+Code in Design, Art, and Architecture*. New York: Princeton Architectural Press. 2010.
- Ribas, Luísa.** "Digital Computational Systems as Aesthetic Artifacts: Practices, Systems, Processes and Perspectives". *xCoAx 2014: Proceedings of the 2nd Conference on Computation, Communication, Aesthetics and X*. Porto: 52-63. 2014.
- Wardrip-Fruin, Noah.** "Expressive Processing: On Process-Oriented Literature and Digital Media". *PhD Thesis*. Rhode Island: Brown University. 2006.
- Whitelaw, Mitchell.** "Art Against Information: Case Studies in Data Practice". *The Fibreculture Journal*, 11. <<http://eleven.fibreculturejournal.org/fcj-067-art-against-information-case-studies-in-data-practice/>> [visited: 03-10-2015]. 2008.
- . "Transduction, Transmateriality, and Expanded Computing." (*the teeming void*). <<http://teemingvoid.blogspot.com/2009/01/transduction-transmateriality-and.html>> [visited: 23-10-2015]. 2009.

## CITED WORKS

---

**Dragulescu, Alex.** *Spam Architecture*. 2005.

<https://sq.ro/spam-architecture.htm> [visited: 20-01-2016]

**DuBois, Luke.** *Hard Data*. 2010

<http://archive.turbulence.org/Works/harddata/> [visited: 20-01-2016]

**Fry, Ben.** *On the Origin of Species*. 2009  
<http://fathom.info/traces> [visited: 20-01-2016]

**Harris, Jonathan and Sepandar**

**Kamvar.** *We Feel Fine*. 2006

<http://wefeelfine.org> [visited: 20-01-2016]

**Harrison, Chris.** *Bible Cross-References*. 2008

<http://www.chrisharrison.net/index.php/Visualizations/BibleViz> [visited: 20-01-2016]

**Jevbratt, Lisa.** 1:1. 1999–2002

[http://128.111.69.4/%7Eje-vbratt/1\\_to\\_1/index\\_ng.html](http://128.111.69.4/%7Eje-vbratt/1_to_1/index_ng.html) [visited: 20-01-2016]

**Luining, Peter.** *ZNC Browser 2.0*. 2003

<http://runme.org/feature/read/+ZNC/+11/> [visited: 20-01-2016]

**Maigret, Nicolas.** *Pure Data Read as Pure Data*. 2010

<http://peripheriques.free.fr/blog/index.php?/past/2010-pure-data-read-as-pure-data/> [visited: 20-01-2016]

**Rubin, Ben and Mark Hansen.** *Listening Post*. 2001

<https://vimeo.com/3885443> [visited: 11-04-2016]

**Rubin, Ben and Mark Hansen.**

*Shakespeare Machine*. 2012

<http://ocr.nyc/public-space-interventions/2012/10/01/shakespeare-machine/> [visited: 11-04-2016]

**Viégas, Fernanda and Martin**

**Wattenberg.** *History Flow*. 2003

<http://fernandaviegas.com/wikipedia.html> [visited: 20-01-2016]

**Whitelaw, Mitchel.** *Weather Bracelet*. 2009

<http://mtchl.net/weather-bracelet/> [visited: 20-01-2016]