6th Conference on Computation, Communication, Aesthetics & X
Madrid, Spain

Marco Heleno
mjheleno@gmail.com

Faculdade de Ciências e Tecnologia
Universidade Nova de Lisboa
Lisboa, Portugal

Visualization for Artificial Intelligence Systems

Keywords: Data; System; Visualization; Aesthetics; Artificial Intelligence; Black Box.

Abstract:

This research project aims to study how visualizations can help create forms of literacy through an aesthetic interpretation of Artificial Intelligence (AI) systems. AI is constantly evolving. This is to the extent that not even software engineers can fully understand their intricacy. They are becoming *black boxes*. The study focuses on using Media Archaeology, iterative cycles of design, implementation and validation, experimenting the medium through the development of code-based prototypes, public exhibitions, and experts validation. The outcomes are visualizations of the data flows and narratives of the architectures of these systems.

1. PURPOSE

In the constant evolution of Artificial Intelligence (AI), we are seeing evidence of a growing complexity in these systems to a point that not even their software engineers can understand their intricacy. It is proving that this invisible force, that shapes our everyday lives, is becoming more and more like *black boxes*. There is a need to be able to explain what decisions are being performed between the layers of these systems. What transformations are being done to an input to result into a certain output. It is crucial for building user confidence, further exploitation and other explorations (Olah et al. 2018; Gunning 2016; Mordvintsev et al. 2015).

For example, Artificial Neural Networks (ANNs), that are loosely based on the human brain's neural structure, consist on a series of interconnected processing nodes that perform calculations on data that it receives as input. One node performs calculations on the received data, then passes the results to another layer of neurons that also perform calculations (Baraniuk 2018), and so on. Different layers may perform different kinds of transformations on their inputs. As ANNs can typically consist on a stack of 10 to 30 layers of artificial neurons (Mordvintsev et al. 2015), it is understandable that at some point it gets hard to grasp what calculations are be done between the initial input and the final output of ANNs.

Even though the science community has been designing technical illustrations of the inner workings of these complex AI systems. We find that these can require some previous knowledge and believe that they lack a connection with the non-scientific community.

2. EXPECTED CONTRIBUTIONS

> We need art to surprise us in order to blow up the world, to create fissures out of which the new can emerge. (in Warburton 2017)

We believe that art and aesthetics can have a big role in creating new forms of literacy through an aesthetic interpretation of these systems (Warburton 2017). We aim to study and develop algorithmical data-driven aesthetic visualizations of the data flows and narratives of the architectures of these AI systems. "Explainable AI" visualizations that enable human users to understand the inner workings of these systems (Gunning 2016).

3. APPROACH

The research and development process will focus on experimentation and validation. Using the "theory and methodology of digital media culture" called Media Archaeology, we will study the past and current state of these systems. Dive into their inner workings (Parrika 2012) experiencing these systems "as a medium" (McMullan 2017). Experimentation through a repeated refinement of a design, implementation and validation approach of code-based prototypes. This will enable us to experience the medium, collect knowledge, try out new ideas and / or techniques, change forms and / or appearances, iterate (Grau 2003), using the same language as the medium — code (Maeda 1999). Besides experimentation and exploration, we also think that writing code opens space for the creation of new models of reflection and problem solving (Grau 2003; Reas et al. 2010; Mateas 2008). This also opens the possibility of developing new visual languages with "new and unique powers of expression" (Carpenter, McLuhan 1960).

For validating our work we intend to hold public exhibitions and search for experts validation. We see this public and expert inquire necessary for avoiding bias, guiding

us on developing new aesthetic interpretations, visualizations of AI systems, that fulfill our research goal.

4. RELATED WORKS

The visualizations from an ongoing project, entitled "System Aesthetics" (Field 2017), from the London based creative studio Field, illustrates well how aesthetic visualizations can make a better connection with the non-scientific public. As Field (2017) refer, this project is a group of "studies of form, structure and behaviour of Artificial Intelligence algorithms". They find this exercise mandatory to "understand and discuss these fundamental forces that change our society". They stress that "we need new visual metaphors to help us decide how much influence we want to give to these intangible systems".

The interactive visualization developed by Daniel Smilkov, Fernanda Viégas, Martin Wattenberg, and the Big Picture team at Google, entitled "Visualizing High-Dimensional Space" (Smilkov et al. 2016), aims to help people understand more about Machine Learning models. In Smilkov et al. (2016) words, this project gives "a peek into how machine learning works, by visualizing high-dimensional data".

"The Building Blocks of Interpretability" (Olah, et al. 2018) shows interfaces that "explore the space of interpretability (...) that show what the network detects and explain how it develops its understanding, while keeping the amount of information human-scale." They "treat existing interpretability methods as fundamental and composable building blocks for rich user interfaces."

We believe that these projects and testimonials show the need, importance and urgency of these visualizations.

5. PROGRESS

There is a long-established interest and some experience with previous developed artefacts and / or visualizations working with data as a medium. This project is in a very early stage, where progress has begun, but is in a reading and knowledge collecting fase.

References:

Baraniuk, Chris. *For AI to Get Creative, It Must Learn the Rules—Then How to Break 'Em*. In Scientific American, 2018. Read on the 04/02/2018, available in www.scientificamerican.com/article/for-ai-to-get-creative-it-must-learn-the-rules-mdash-then-how-to-break-lsquo-em/?utm_source=twitter&utm_medium=social&utm_campaign=-sa-editorial-social&utm_content=&utm_term=tech_news_&sf180234982=1.

Carpenter, Edmund, and Marshall McLuhan. *Explorations in Communication: An Anthology.* Boston, USA: Beacon Press, 1960. Read on the 11/08/2015, available in https://archive.org/details/explorationsinco00carp.

Field. *System Aesthetics: Studies of form, structure and behaviour of Artificial Intelligence algorithms.* In Field.io, 2017. Read on the 11/03/2018, available in www.field.io/project/system-aesthetics/.

Grau, Oliver. *Virtual Art - From illusion to immersion.* Cambridge, London: The MIT Press, 2003.

Gunning, David. *Explainable Artificial Intelligence (XAI).* In DARPA, 2016. Read on the 29/03/2018, available in www.darpa.mil/program/explainable-artificial-intelligence

Maeda, John. *Design By Numbers.* Cambridge, MA: The MIT Press, 1999.

Mateas, Michael. *Procedural literacy: educating the new media practitioner.* Pittsburgh, PA, USA: ETC Press, 2008.

McMullan, Thomas. *The Work of Art in the Age of Algorithmic Reproduction.* In Medium, 2017. Read on the 02/02/2018, available in https://medium.com/s/living-in-the-machine/the-work-of-art-in-the-age-of-algorithmic-reproduction-bd3bd9b4e236.

Mordvintsev, Alexander, Christopher Olah, and Mike Tyka. *Inceptionism: Going Deeper into Neural Networks.* In Google Research Blog, 2015. Read on the 10/02/2018, available in https://research.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html.

Olah, Chris, Arvind Satyanarayan, Ian Johnson, Shan Carter, Ludwig Schubert, Katherine Ye and Alexander Mordvintsev. *The Building Blocks of Interpretability.* In Distill, 2018. Read on the 10/03/2018, available in https://distill.pub/2018/building-blocks/.

Parrika, Jussi. *What is Media Archaeology?.* Cambridge, UK: Polity Press, 2012.

Reas, Casey, Chandler McWilliams, and LUST. *Form+Code in Design, Art, and Architecture.* New York, USA: Princeton Architectural Press, 2010.

Smilkov Daniel, Fernanda Viégas, Martin Wattenberg and the Big Picture team at Google. *Visualizing High-Dimensional Space.* In Google AI Experiments, 2016. Read on the 21/03/2018, available in https://experiments.withgoogle.com/ai/visualizing-high-dimensional-space.

Warburton, Nigel. *Art by algorithm.* In Aeon, 2017. Read on the 04/02/2018, available in https://aeon.co/essays/how-algorithms-are-transforming-artistic-creativity.