>>
>>
>>
>>
>>

>> NULL_SETS

by amy szczepanski + evan meaney

Let us, for a moment, consider the evolving differences between 'data' and 'information.' As I write this, at the CERN facility near Geneva, data is being born—data that may, itself, never become informative. The now-famous Large Hadron Collider (LHC) is an experimental tool of such scope and magnitude that we have yet to fully grasp the potential of its result, and proof of this can be found in its component parts. Just one of its data recorders, the Compact Muon Solenoid (CMS) detector, runs as a functional black box to the system's ongoing investigations, and over the intended 10–20 year experiment, the CMS will record more bytes than every word ever spoken by humankind since the dawn of time. In an era where 'big data' has become a grant-writing catch-phrase, this may, perhaps, be the biggest data of them all. But of what use is it to us? How can we hope to make something so large our own?

There is no human way to examine the CMS results. To do so would mean undertaking a Borgesian project of increasing madness and infinite symbols. Instead, engineers are working to write programs which will comb through the recorder's raw data, approximating and assigning relative value to sections which will then have an audience before human eyes. This is an informatic system so large as to necessitate an initial distance from human analysis; and thus illuminates the divide between data and information—whereas the former produces the latter, but not alone. What we know as 'information' is actually a function of evaluated raw data, positioned for human perception, ready for consumption.

Null_Sets (*N_S*) is a body of 2D work which probes and expresses the gap between the data-centric and informatic emphases of communication. Specifically, it consists of a number of archival prints produced through a data-translation process, whereby our team intentionally encoded text files as images, anticipating that it will not be possible for human users to traditionally decode them.

Using this process, *N_S* takes on the project of mediating ordered sets of human-language text (data) into visual information by rendering those sets as JPEG images. The computer-to-human translation, which normally expects computer-oriented code (binary/hexadecimal/ASCII) to synthesize human-oriented visuals (a picture), explodes at the introduction of unexpected language. In other words, we use the text of *Hamlet* as byte input—exporting a JPEG image which visualizes all of the data in Shakespeare's play, but none of the intended information. The resulting picture appears to be chaotic, when in actuality, it is rigorously ordered under JPEG standards, though not optimized for human eyes or expectations.

Once we have been taught to read, we cannot help but oblige. So good are we at deriving information from data, that we almost never need to think about it. The input symbols S,T,O,P prompt our foot towards the brake instinctually; while this automation is an amazing part of the human cognitive experience, it does reveal a blind gap between recognition and response. Considering that distance and exploring its potential is what *Null_Sets* is all about.

>>
>>
>>
>>
>>
>>
>>
>>
>>
>> Definitions

data
noun [treated as sing. or pl.]

Facts and statistics collected together for reference or analysis.

>> Computing // The quantities, characters, or symbols on which operations are performed by a computer, being stored and transmitted in the form of electrical signals and recorded on magnetic, optical, or mechanical recording media.

>> Philosophy // Things known or assumed as facts, making the basis of reasoning or calculation.

ORIGIN mid 17th cent. (as a term in philosophy)
From Latin, plural of datum.

information
noun

What is conveyed or represented by a particular arrangement or sequence of things.

>> Computing // Data processed, stored, transmitted, and/or interpreted.

>> Information Theory // A mathematical quantity expressing the probability of occurrence of a particular sequence of symbols, impulses, etc., as contrasted with that of alternative sequences.

ORIGIN late Middle English (also in the sense [formation of the mind, teaching]), via Old French from Latin informatio(n-), from the verb informare.

>>
>>
>>
>>
>>
>>
>> Technical Notes

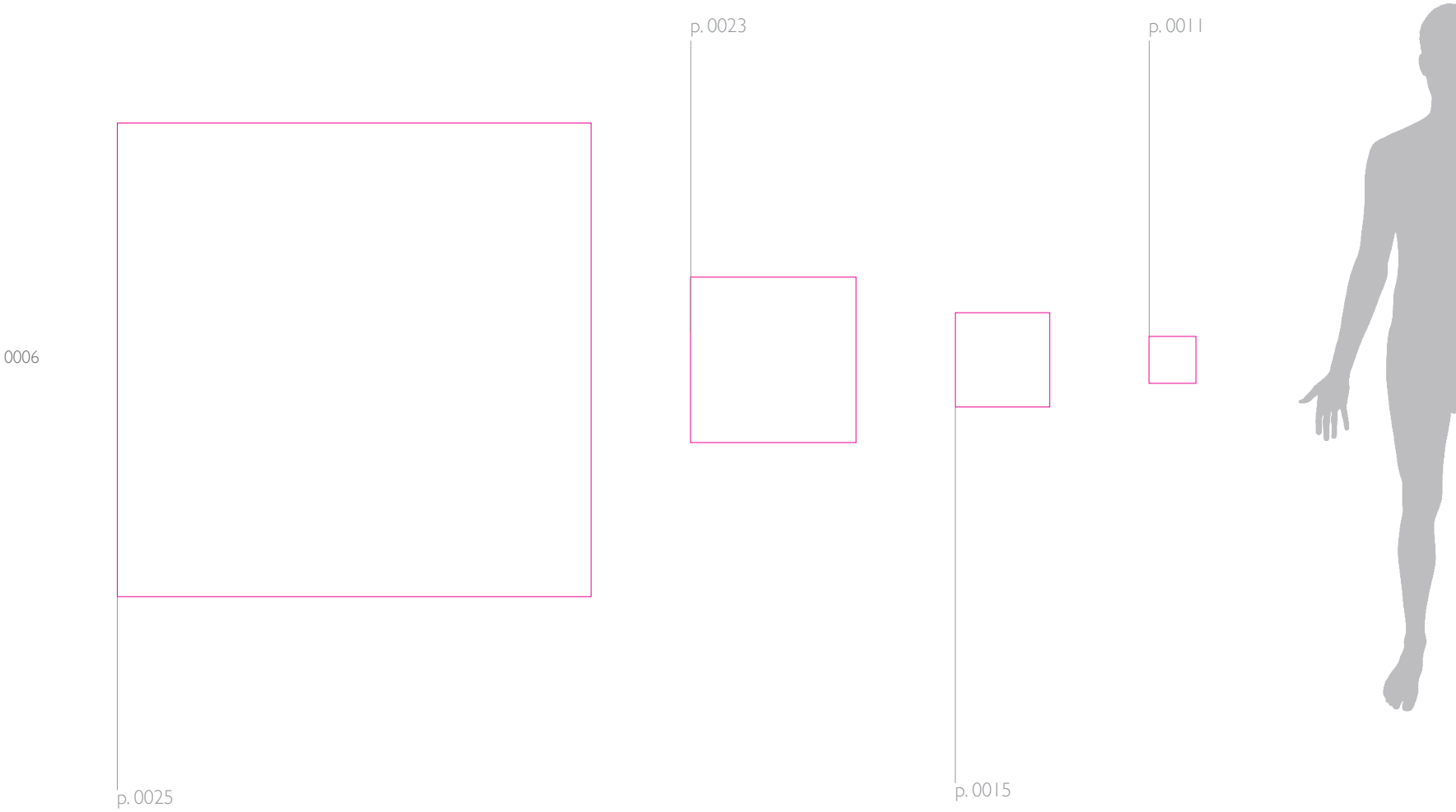
The JPEG standard defines a system for encoding color images with compression. These files can typically be viewed as having three parts: a header, the data that describes the pixels in the image, and a two-byte footer.

The header contains structural data about the image, including the dimensions, the forms of encoding and compression being used, and a series of look-up tables for mediating the compression algorithms. Information in the header is flagged with two-byte markers that begin with the hexadecimal quantity 0xFF and whose second byte identifies the feature being defined. The marker is followed by two bytes that give the length of the table or feature being described, followed by the pertinent information. A typical JPEG image header will have a length of hundreds of bytes, encoding roughly a half-dozen attributes of the method. The footer, on the other hand, is always the two-byte quantity 0xFFD9, which represents the end of the image. For most JPEG-encoded images, the bulk of the content falls between the header and the footer and is encoded as a series of bits. The image data interpreted by the JPEG decoding algorithm includes the average intensity and average color over blocks of 8-by-8 pixels as well as encoded coefficients that describe how the intensity and color vary over each block.

For *Null_Sets*, our team constructed a 607-byte JPEG header for a square pixel image encoded with the discrete cosine transform and non-differential, Huffman coding. We also specified the default look-up tables for image quantization and run-length encoding. The two-byte footer is the quantity 0xFFD9, representing the end of the image. Next, we used the Unix utility 'cat' to concatenate our header file, a text file in rich text format, and our footer. Thus, data that had previously been intended to be decoded as ASCII characters will now be read as if it represents quantities defining the pixels in an image. We gave the resulting file a name with a .jpg file extension and then viewed it with standard image viewing tools. Due to our original texts being encoded in rich text format and consisting of ASCII-encoded text (which never contains 0xFF), we did not need to worry about the byte 0xFF erroneously triggering a marker in the data segment; this would be a real concern if using an arbitrary binary file instead of text.

Once this workflow script had been stabilized, we began to transcode and compare text-bodies from different disciplines within the arts and sciences. These bodies, when translated through the *N_S* / JPEG process, became both the subjects of, and the titles for the individual images. Using the text from Euclid's canonical *Elements* produced our image on page 0015. Appropriating the HTML from a Google search produced the picture on page 0019, and encoding the protein chains for the X chromosome genome led to the image on page 0021. Using this process, any text may be both preserved (as byte data) and estranged (as an image) in a JPEG translation; creating a new means of comparison and aesthetic exploration.

>>
>>
>>
>>
>>
>>
>> Sample Physical Print Size Comparison

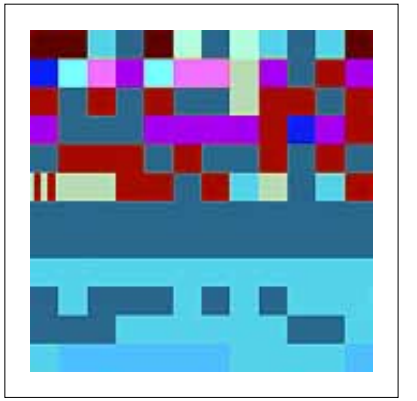


p. 0021

>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>>
>> Sample Prints

Lipsum. (2012). 10,000 Words of Generated Ipsum Lorem Text
Retrieved January 6, 2012, from http://www.lipsum.com/

>> 5 inches^2
>> 30 MB
evan meaney + amy szczepanski
digital image / usa / 2012

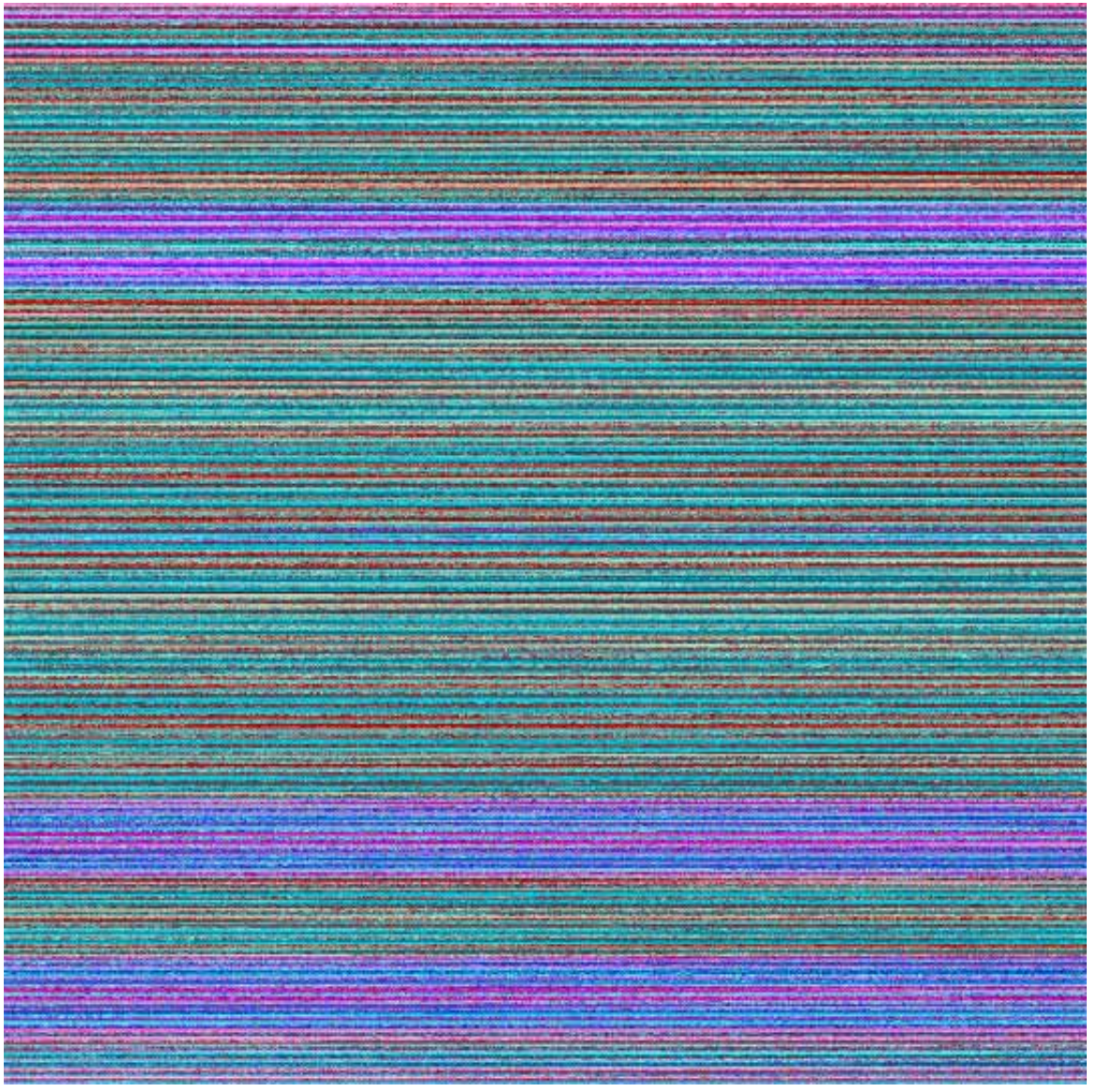


detail

Blender Foundation. (2012). Blender (Version 2.62) [Source Code]:
Blender Foundation. Retrieved from http://www.blender.org/

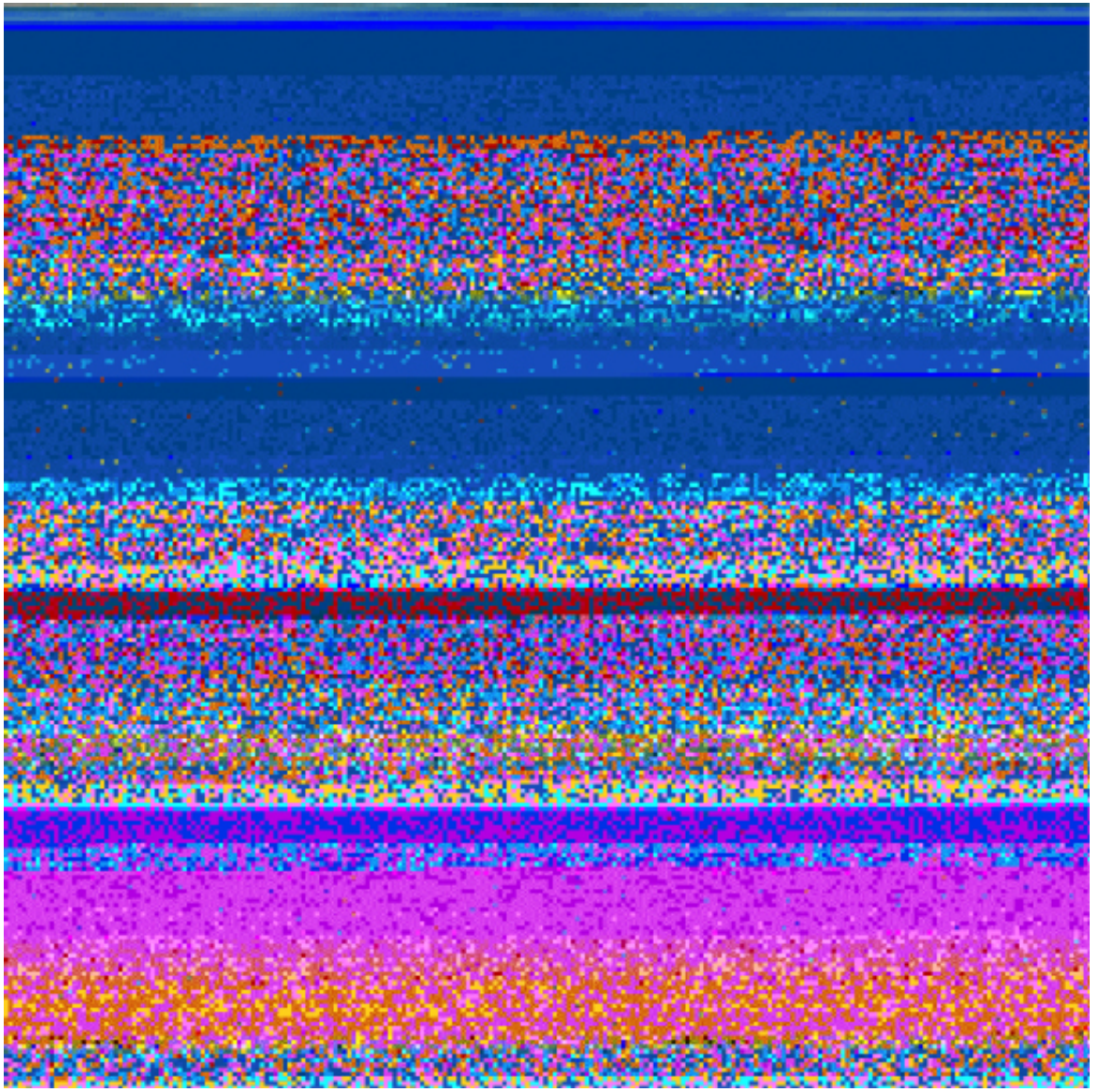
>> 56 inches^2
>> 926.1 MB

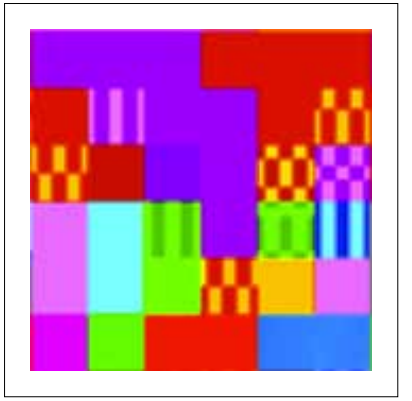
evan meaney + amy szczepanski
digital image / usa / 2012



Euclid. (1885). *The First Six Books of the Elements of Euclid: and Propositions I-XXI of Book XI, and an Appendix on the Cylinder, Sphere, Cone, etc., with Copious Annotations and Numerous Exercises* (J. Casey, Trans.). Dublin: Hodges, Figgis, & Co.

>> 21 inches^2
>> 314.2 MB
evan meaney + amy szczepanski
digital image / usa / 2012

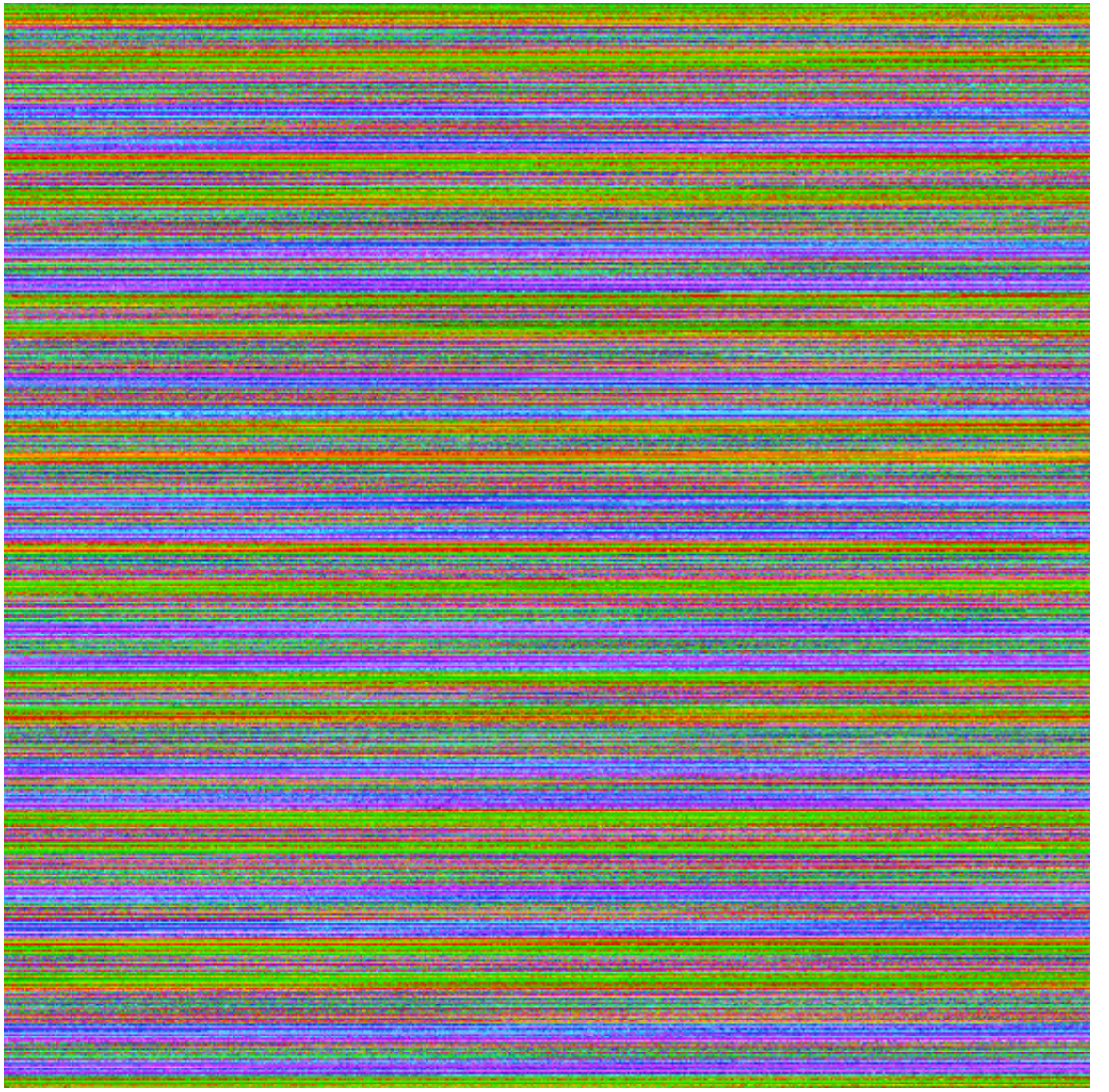




detail

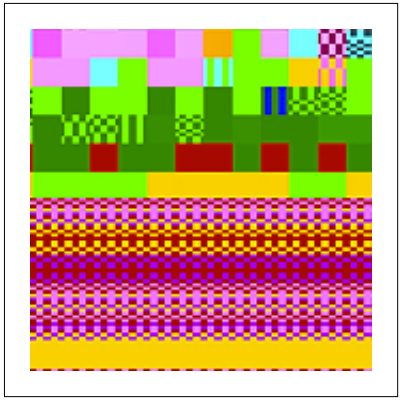
Barthes, R. (1977). Image, Music, Text (S. Heath, Trans.). New York: Hill and Wang.
[Hexadecimal Code for Portable Document Format].

>> 97 inches^2
>> 3.24 GB
evan meaney + amy szczepanski
digital image / usa / 2012



Google. (2012). Google Search Results for 'Google'. Retrieved 1/6/2012, from https://www.google.com/search?aq=f&ix=acb&sourceid=chrome&ie=UTF-8&q=google [Hypertext Markup Language].

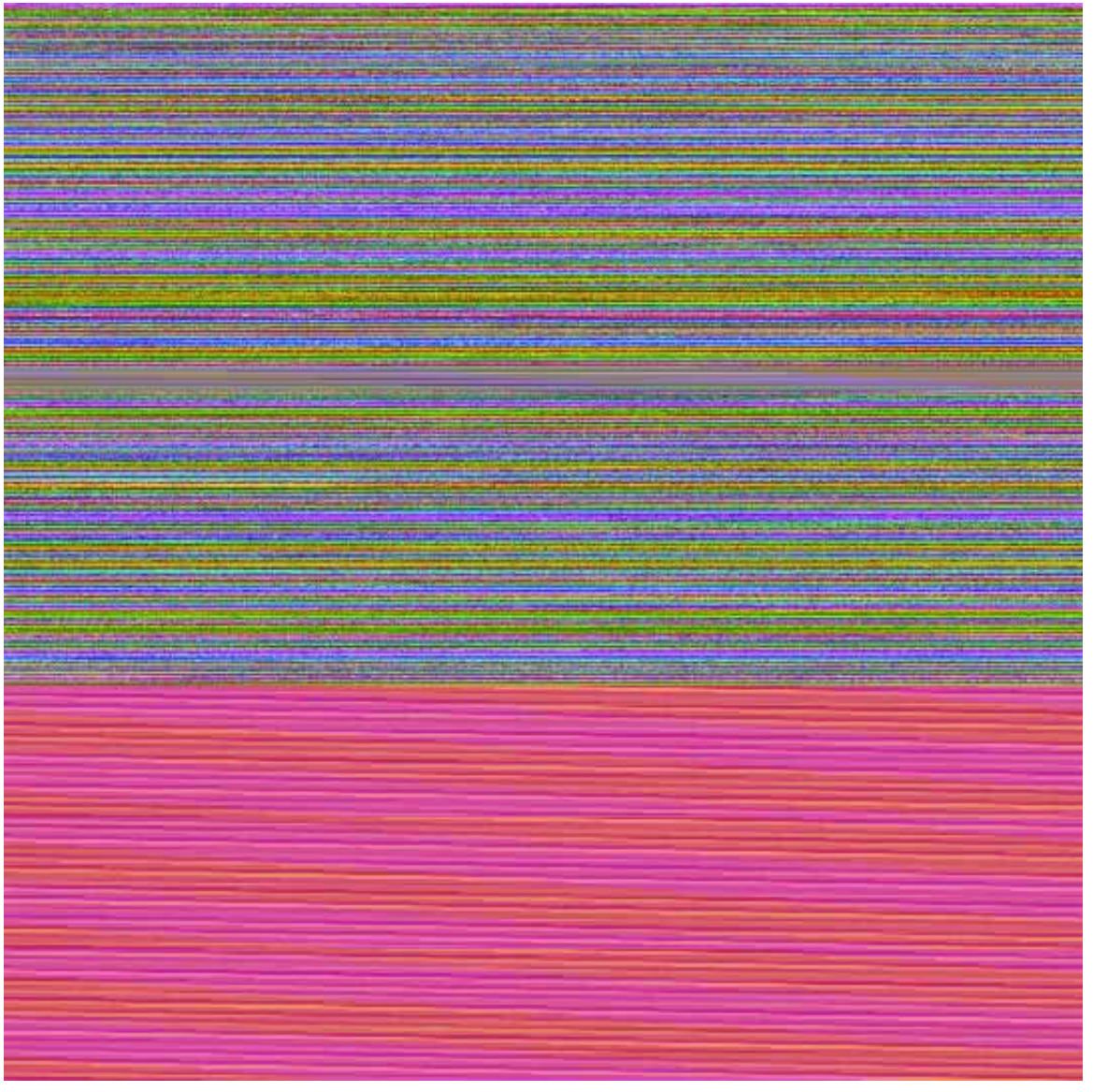
>> 6 inches^2
>> 39.6 MB
evan meaney + amy szczepanski
digital image / usa / 2012



detail

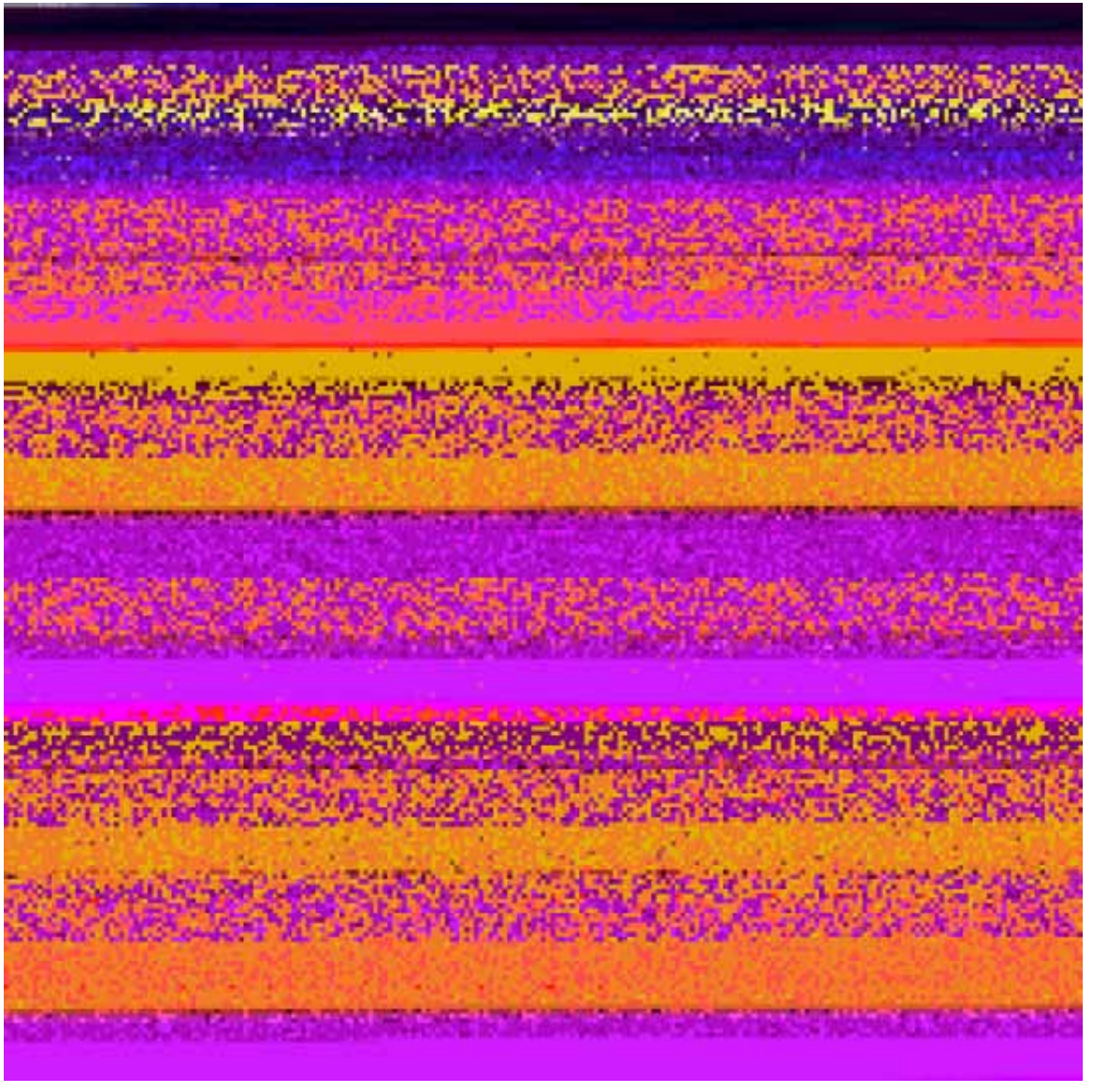
Human Genome Project, & National Center for Biotechnology Information. (2002).
Human Genome Project, X Chromosome.

>> 72 inches^2
>> 1.9 GB
evan meaney + amy szczepanski
digital image / usa / 2012



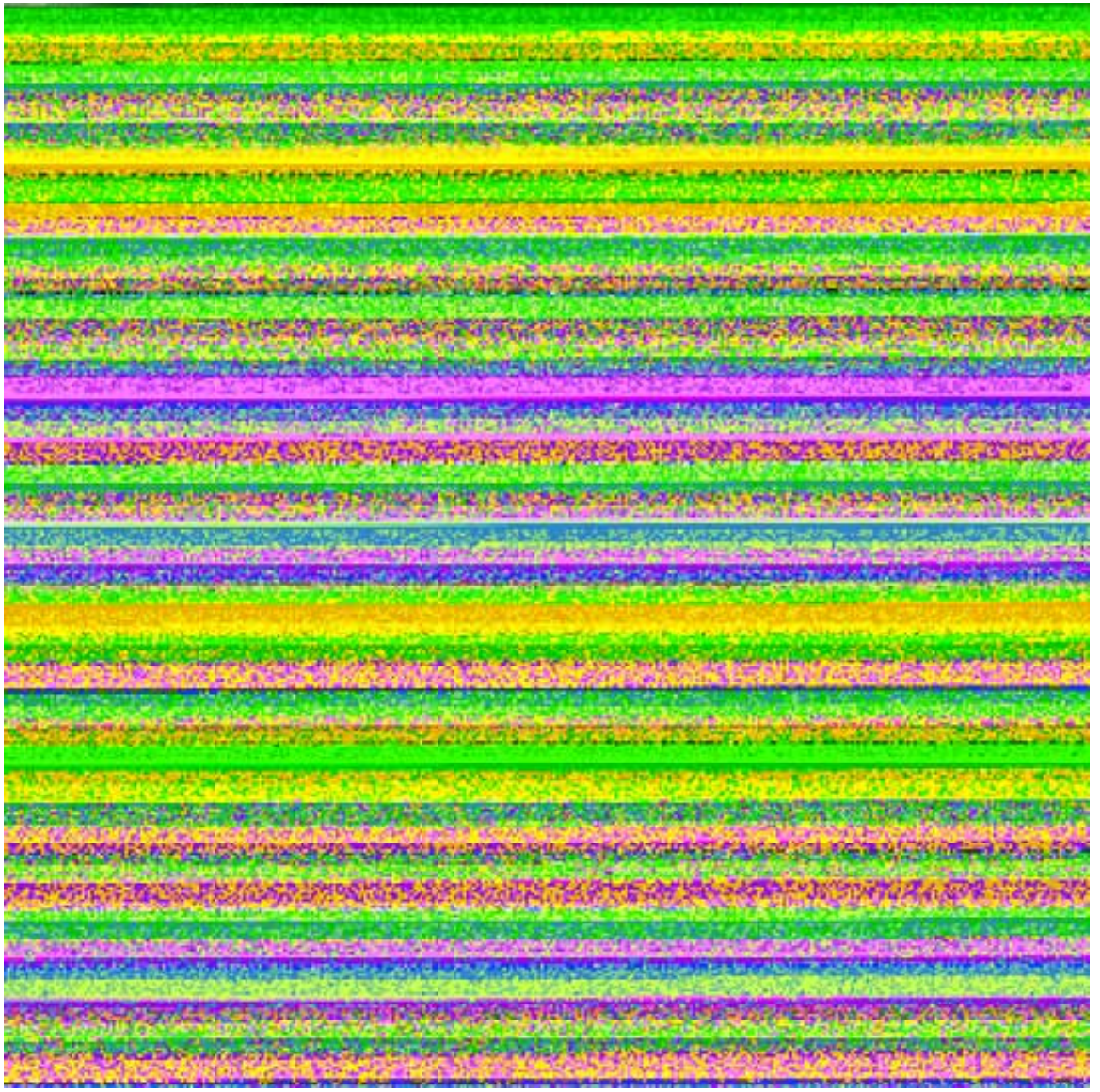
Melville, H. (1851/1929). *Moby Dick*. London:
Oxford University Press.

>> 21 inches^2
>> 285.9 MB
evan meaney + amy szczepanski
digital image / usa / 2012



π Calculated to 1,000,000 Places.

>> 33 inches^2
>> 628.6 MB
evan meaney + amy szczepanski
digital image / usa / 2012

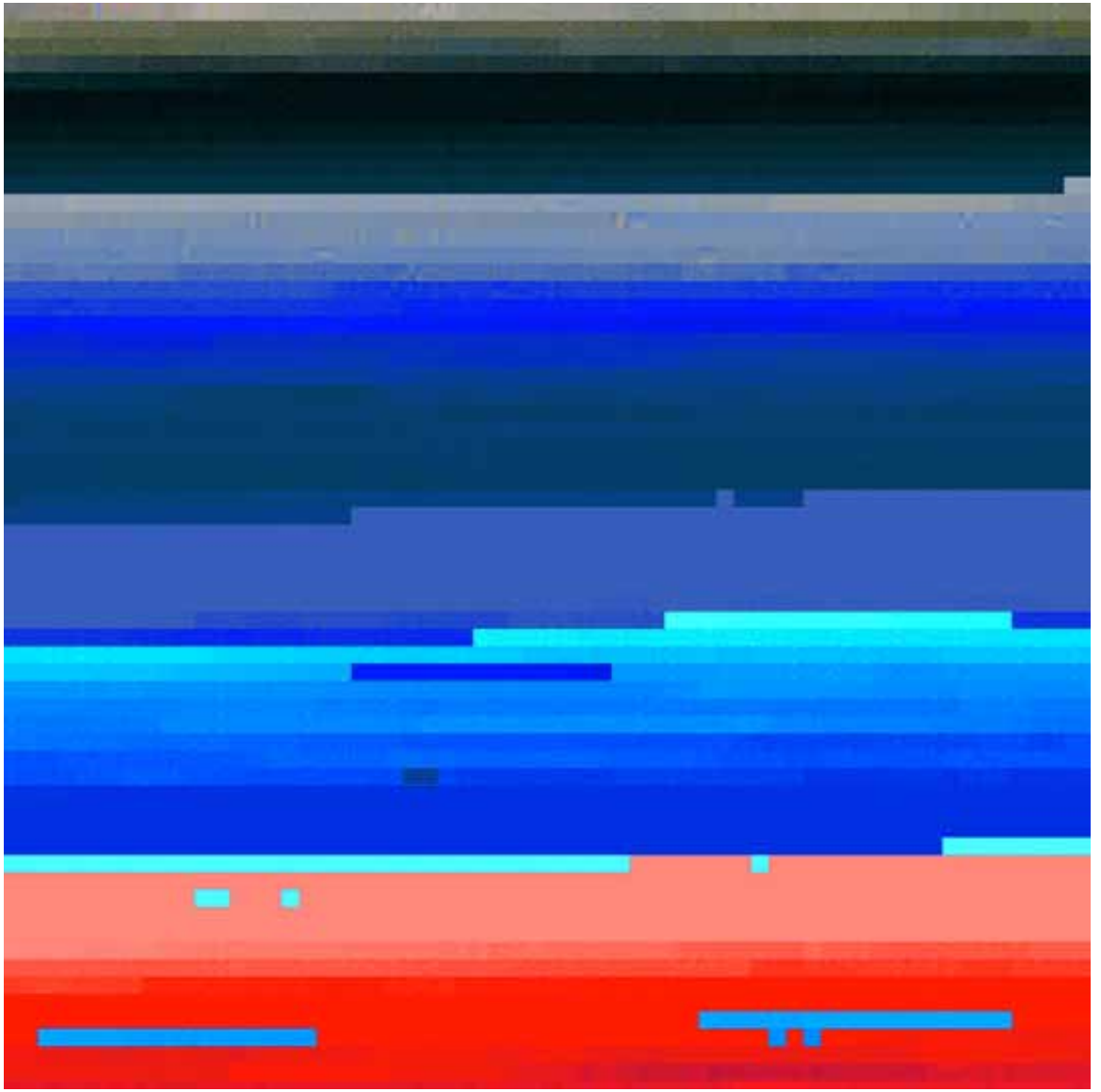


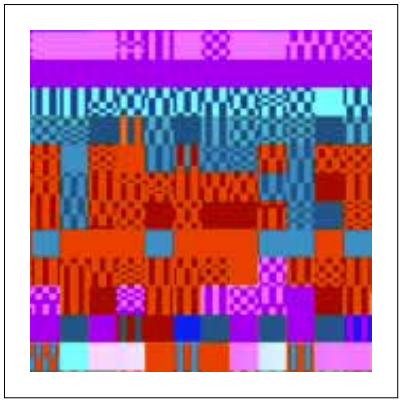
Sato, H. (1980). 殉情詩集. Japan:
Nihon Kindai Bungakukan.

>> 5 inches^2
>> 30 MB
evan meaney + amy szczepanski
digital image / usa / 2012

The Book of Job. (1985) *The Holy Bible, Containing the Old and New Testaments, King James Version*. Nashville: T. Nelson.

>> 6 inches ^{2}
>> 39.7 MB
evan meaney + amy szczepanski
digital image / usa / 2012

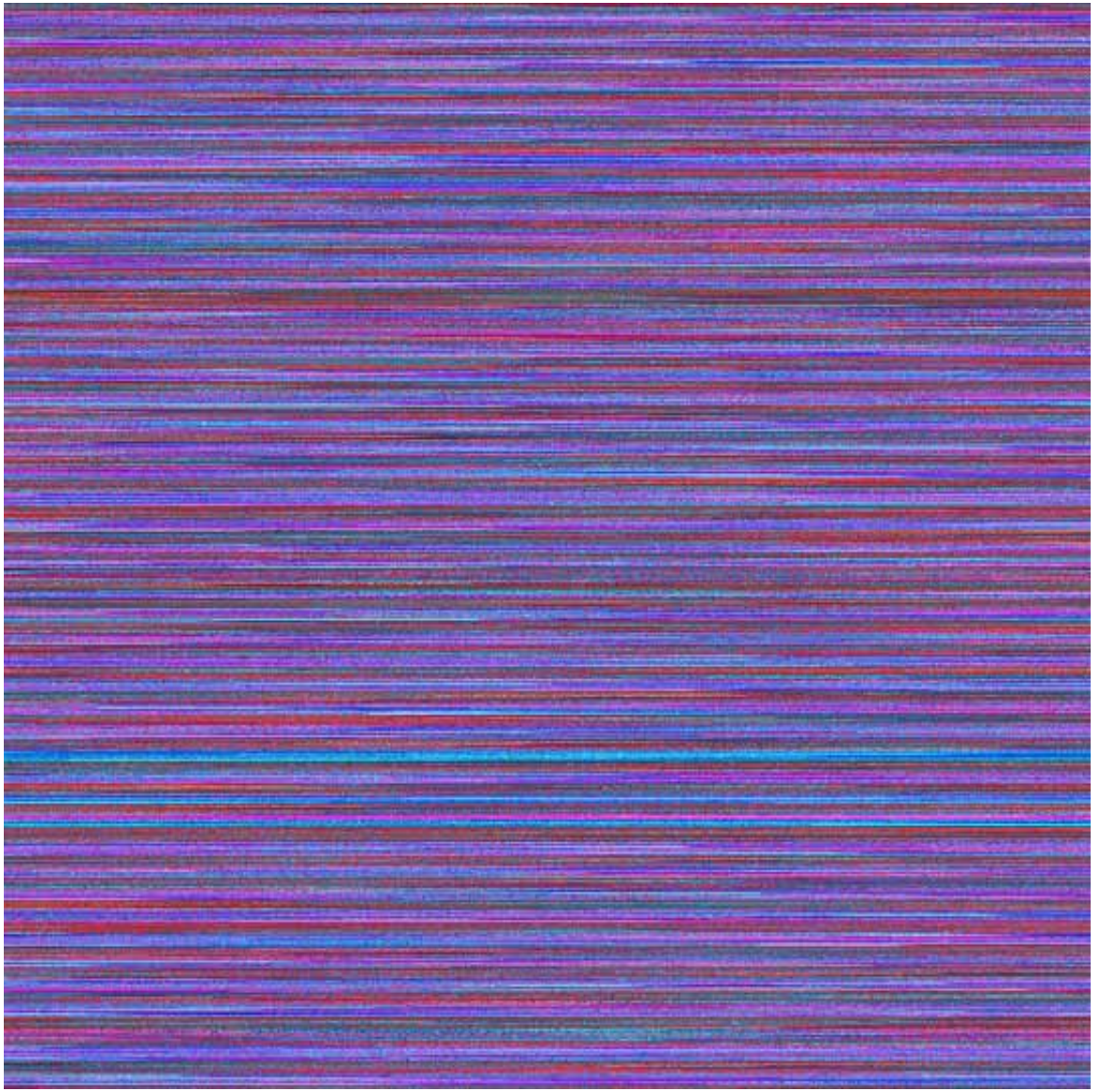


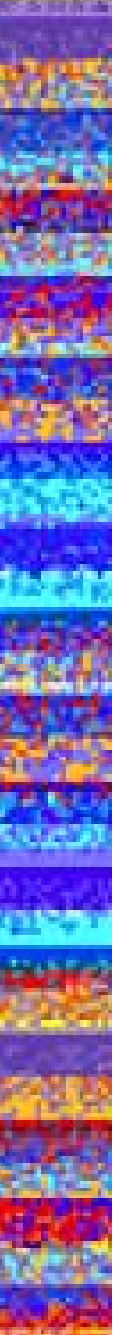


detail

The Oxford English Dictionary. (2000) (2 ed.).
London: Oxford University Press.

>> 62 inches^2
>> 1.1 GB
evan meaney + amy szczepanski
digital image / usa / 2012

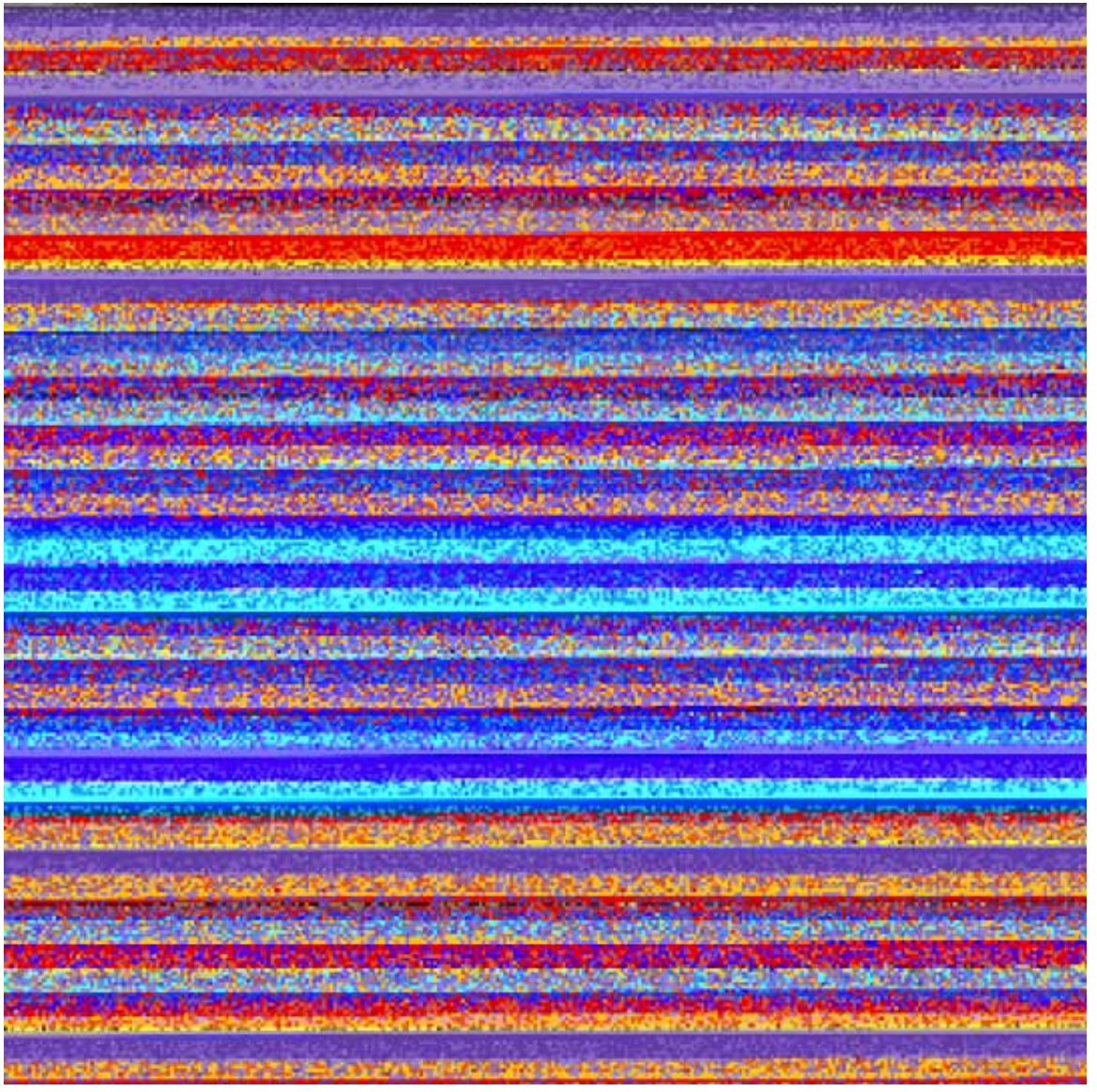


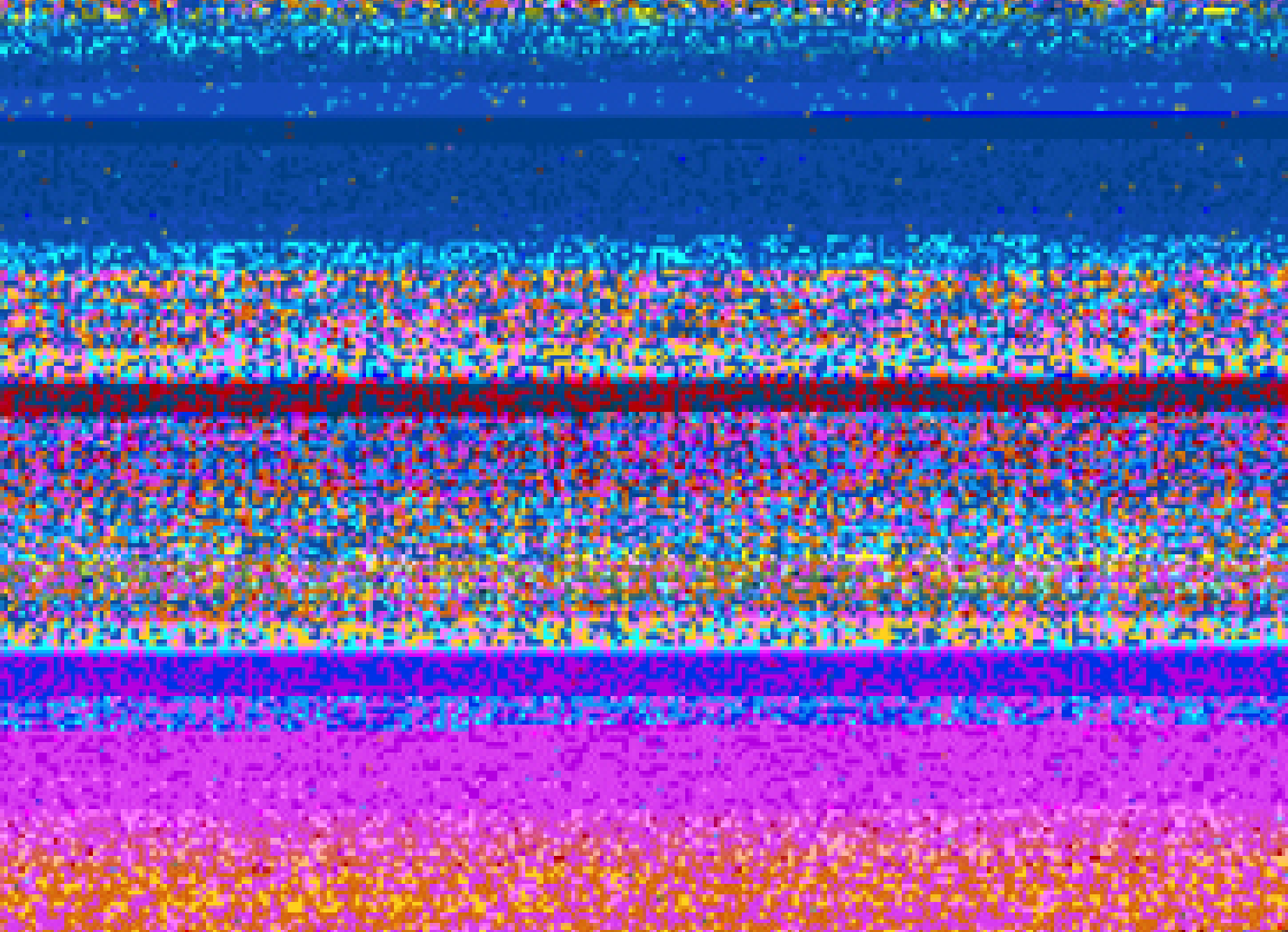


detail

The Square Root of 2.

>> 27 inches^2
>> 511.8 MB
evan meaney + amy szczepanski
digital image / usa / 2012



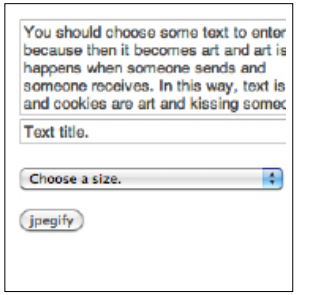


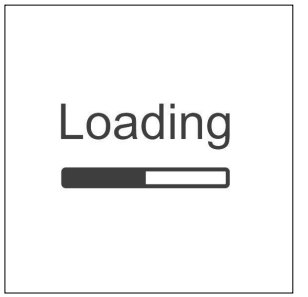
>>
>>
>>
>>
>>
>>
>>
>> Online Workflow Participation

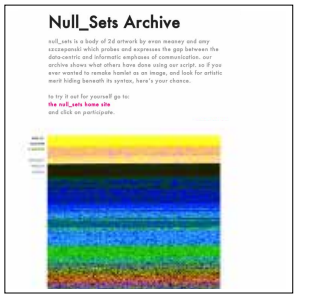
While *Null_Sets* is primarily a pre-selected grouping of images, it is also, itself, a workflow. Using the server-side *Null_Sets script,* anyone can submit text to be converted into a JPEG image and then viewed live on our evolving web-space.

Simply navigate to: http://www.evanmeaney.com/ns, click the participate tab, and follow the on-screen directions. We look forward to seeing what interesting forms of data you translate, and the new outcomes of this ongoing module of the project.

Looking for interesting bodies of text? How about heading over to our friends at Project Gutenberg. Their website, http://www.gutenberg.org has a ton of great text files that can be input into our script.







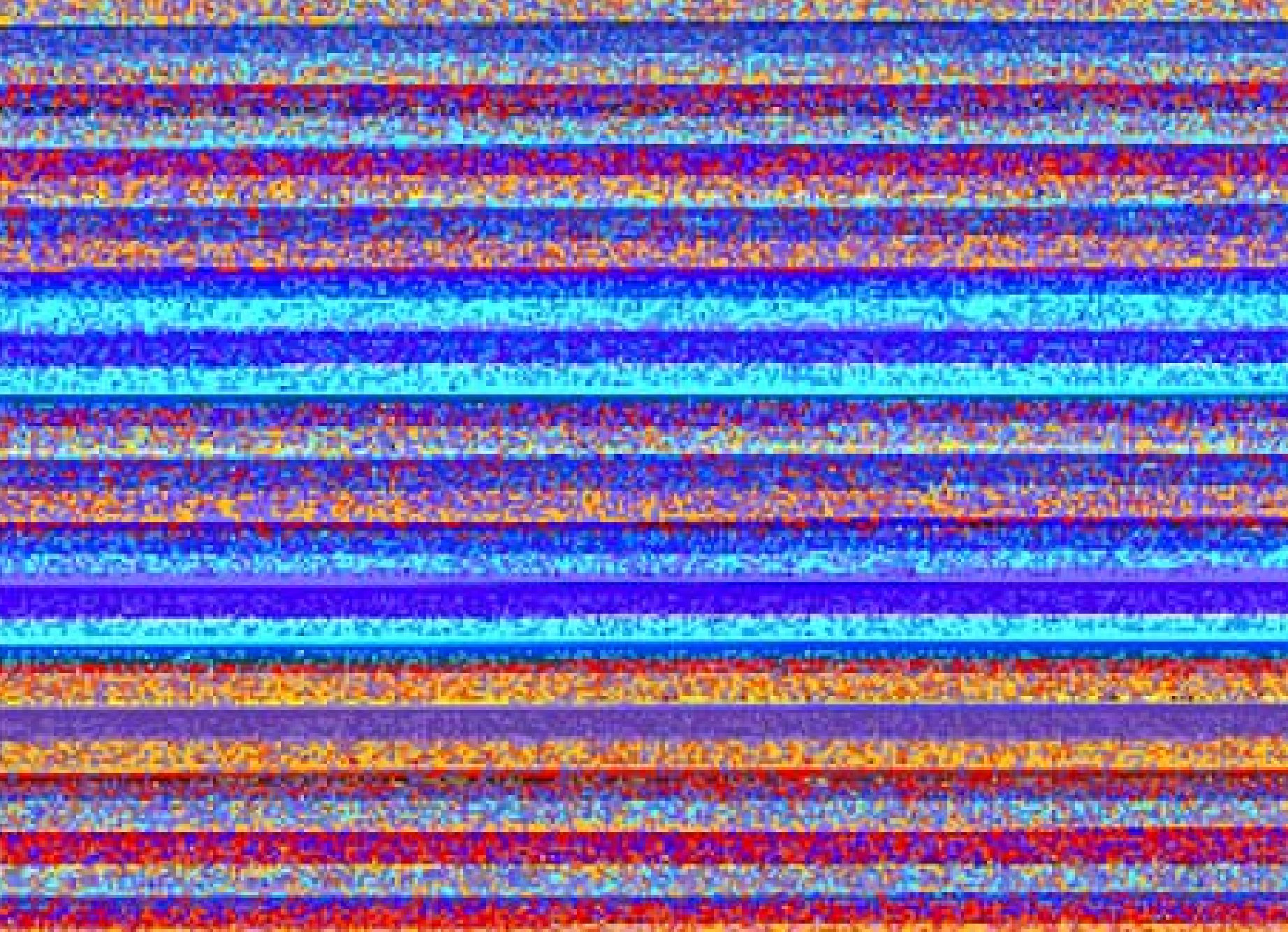
>>
>>
>>
>>
>>
>>
>> Printing Formats

Translating our images into physical prints poses some challenges. Data like this is meant to exist, relativistically, on screens and inside files or directories—bringing it into the real world gets somewhat complicated. When printed, these images run the gamut from .5 inches in length to almost 11 feet, dismissing any hope of uniformity. Our team, however, finds the logistics behind such disparities to be invigorating. To share our work in the most aesthetically interesting, and metrically accurate way, exhibitions of *Null_Sets* will include two modes of presentation. For smaller works, matted and framed archival prints can be shipped in crates, but for the large-scale work, our team will send digital files and/or slides which may be projected onto walls or canvases at a 1:1 scale. To insure correct sizing, easy-to-use measures will be included. In this way we can convey to audiences the concrete side of big-data and use the visual comparison of space and size to drive our point home.

>>
>>
>>
>>
>>
>>
>> Contact Information

Amy Szczepanski
Evan Meaney

The University of Tennessee / ORNL
Knoxville / Tennessee / United States
me@evanmeaney.com
www.evanmeaney.com/ns



>>
>>
>>
>>
>>
>>
>>
>> A Most Sincere Thanks

This work was made possible with generous support from the National Science Foundation (TG-ART110001) and the University of Tennessee's Exhibit, Performance and Publication Expenses Fund. Academic support was also offered by Project Gutenberg and Oak Ridge National Lab.

The eXtreme Science and Engineering Discovery Environment

The Remote Data Analysis and Visualization Center

National Institute for Computational Sciences

The National Science Foundation

The UT Knoxville School of Art

Oak Ridge National Laboratory

The University of Tennessee

Dr. Gregory Newby

Project Gutenberg

Dr. Dorothy Habel

Jeanine Williamson

Stephen Lester

>>
>>
>>
>>
>> Bibliography / Image List

π Calculated to 1,000,000 Places.

The 32nd Mersenne Prime.

Barthes, R. (1977). *Image, Music, Text* (S. Heath, Trans.). New York: Hill and Wang.

Barthes, R. (1977). Image, Music, Text (S. Heath, Trans.). New York: Hill and Wang. [Hexadecimal Code for Portable Document Format].

Blender Foundation. (2012). Blender (Version 2.62) [Source Code]: Blender Foundation. Retrieved from http://www.blender.org/

The Book of Job. (1985) *The Holy Bible, Containing the Old and New Testaments, King James Version.* Nashville: T. Nelson.

Brodziński, K. (2009). Wiesław Retrieved 2/2/2012 from http://www.gutenberg.org/ebooks/27835

Euclid. (1885). *The First Six Books of the Elements of Euclid: and Propositions I-XXI of Book XI, and an Appendix on the Cylinder, Sphere, Cone, etc., with Copious Annotations and Numerous Exercises* (J. Casey, Trans.). Dublin: Hodges, Figgis, & Co.

Google. (2012). Google Search Results for 'Google'. Retrieved 1/6/2012, from https://www.google.com/search?aq=f&ix=acb&sourceid=chrome&ie=UTF-8&q=google [Hypertext Markup Language].

Human Genome Project, & National Center for Biotechnology Information. (2002). *Human Genome Project, X Chromosome.*

Human Genome Project, & National Center for Biotechnology Information. (2004). *Human Genome Project, Build 34, Chromosome Number 19.*

Lipsum. (2012). 10,000 Words of Generated Ipsum Lorem Text Retrieved January 6, 2012, from http://www.lipsum.com/

Melkanoff, M. A. (1961). *A Fortran Program for Elastic Scattering Analyses With the Nuclear Optical Model.* Berkeley: The University of California Press.

Melville, H. (1851/1929). *Moby Dick*. London: Oxford University Press.

Mendeleev, D. (1869). Über die Beziehungen der Eigenschaften zu den Atomge wichten der Elemente. *Zeitschrift für Chemie*, 405-406.

Newton, I. (1687). *Philosophiae Naturalis Principia Mathematica* (Vol. 5). London: Jussu Societatis Regiae ac Typis Josephi Streater. Prostant apud... Bibliopolas.

The Oxford English Dictionary. (2000) (2 ed.). London: Oxford University Press.

Plato. (370 BC/1903). The Phaedrus. In J. Burnet (Ed.), Platonis Opera. London: Oxford University Press. Retrieved 2/2/2012 from http://www.perseus.tufts.edu/

Post, D and P. (1959). Sixteen Reasons (Why I Love You) [Recorded by Connie Frances]. On *Cricket in "Hawaiian Eye"* [MP3]. USA: Warner Bros.

Post, D and P. (1959). Sixteen Reasons (Why I Love You) [Recorded by Connie Frances]. On Cricket in "Hawaiian Eye" [MP3]. USA: Warner Bros. [Hexadecimal Code for MP3 at 48kbps].

Sato, H. (1980). 殉情詩集. Japan: Nihon Kindai Bungakukan.

Section-by-Section Analysis of H.R. 2281 as Passed by the United States House of Representatives on August 4, 1998, The United States House of Representitives (1998).

Shakespeare, W. (1940). *The Complete Works of William Shakespeare: The Cambridge Edition Text, As Edited by William Aldis Wright, Including the Temple Notes*. New York: Doubleday.

Shelley, M. W., & Butler, M. (1818/1998). *Frankenstein, Or, The Modern Prometheus: The 1818 Text*. London: Oxford University Press.

The Square Root of 2.

Stelwagon, H. W. (1890). *Essentials of Diseases of the Skin: Including the Syphilodermata; Arranged in the Form of Questions and Answers.* Philadelphia: W. B. Saunders.

Wikipedia. (2012). JPEG - Wikipedia. Retrieved 1/6/2012, from http://en.wikipedia.org/wiki/JPEG [Hypertext Markup Language].