Dataremix: Aesthetic Experiences of Big Data and Data Abstraction

West, R.

https://doi.org/10.34737/vq9zx

Dataremix: Aesthetic Experiences of Big Data and Data Abstraction

Ruth West

A thesis submitted in partial fulfilment of the requirements of The University of Westminster for the degree of Doctor of Philosophy by Published Work

December 2021

Abstract

This PhD by published work expands on the contribution to knowledge in two recent large-scale transdisciplinary artistic research projects: *ATLAS in silico* and *INSTRUMENT | One Antarctic Night* and their exhibited and published outputs. The thesis reflects upon this practice-based artistic research that interrogates data abstraction: the digitization, datafication and abstraction of culture and nature, as vast and abstract digital data. The research is situated in digital arts practices that engage a combination of big (scientific) data as artistic material, embodied interaction in virtual environments, and poetic recombination.

A transdisciplinary and collaborative artistic practice, *x-resonance,* provides a framework for the hybrid processes, outcomes, and contributions to knowledge from the research. These are purposefully and productively situated at the *objective | subjective* interface, have potential to convey multiple meanings simultaneously to a variety of audiences and resist disciplinary definition. In the course of the research, a novel methodology emerges, *dataremix,* which is employed and iteratively evolved through artistic practice to address the research questions: 1) *How can a visceral and poetic experience of data abstraction be created? and 2) How would one go about generating an artistically-informed (scientific) discovery?*

Several interconnected contributions to knowledge arise through the first research question: creation of representational elements for artistic visualization of big (scientific) data that includes four new forms (genomic calligraphy, algorithmic objects as natural specimens, scalable auditory data signatures, and signal objects); an aesthetic of slowness that contributes an extension to the operative forces in Jevbratt's inverted sublime of looking down and in to also include looking fast and slow; an extension of Corby's objective and subjective image consisting of "informational and aesthetic components" to novel virtual environments created from big

(scientific) data that extend Davies' poetic virtual spatiality to poetic *objective | subjective* generative virtual spaces; and an extension of Seaman's embodied interactive recombinant poetics through embodied interaction in virtual environments as a recapitulation of scientific (objective) and algorithmic processes through aesthetic (subjective) physical gestures. These contributions holistically combine in the artworks *ATLAS in silico* and *INSTRUMENT | One Antarctic Night* to create visceral poetic experiences of big data abstraction.

Contributions to knowledge from the first research question develop artworks that are visceral and poetic experiences of data abstraction, and which manifest the *objective | subjective* through art. Contributions to knowledge from the second research question occur through the process of the artworks functioning as experimental systems in which experiments using analytical tools from the scientific domain are enacted within the process of creation of the artwork. The results are "returned" into the artwork. These contributions are: elucidating differences in DNA helix bending and curvature along regions of gene sequences specified as either introns or exons, revealing nuanced differences in BLAST results in relation to genomics sequence metadata, and cross-correlation of astronomical data to identify putative variable signals from astronomical objects for further scientific evaluation.

Contents

List of Figures

List of Tables

List of Artworks and Publications in Volume 2

Artworks

ATLAS in silico

INSTRUMENT | One Antarctic Night

Publications related to ATLAS in silico

West, R.G. et al. (2014). Scalable metadata environments (MDE): artistically impelled immersive environments for large-scale data exploration. The Engineering Reality of Virtual Reality 2014 / IS&T/SPIE Electronic Imaging. 28 February 2014. San Francisco, California, USA, 901205.

West, R. et al. (2009). Sensate abstraction: hybrid strategies for multi-dimensional data in expressive virtual reality contexts. IS&T/SPIE Electronic Imaging. 2009. International Society for Optics and Photonics, 72380I-72380I.

West, R. et al. (2009). Algorithmic Object As Natural Specimen Meta Shape Grammar Objects From Atlas In Silico. Leonardo Electronic Almanac, 16 (6–7), 1–5.

Gossmann, J. et al. (2008). Scalable Auditory Data Signatures for Discovery Oriented Browsing in an Expressive Context. Proceedings of the 14th International Conference on Auditory Display (ICAD 2008). 2008. Paris: International Community for Auditory Display, ICAD08 1-6.

Publications related to INSTRUMENT | One Antarctic Night

West, R. et al. (2020). Designing a VR Arena: Integrating Virtual Environments and Physical Spaces for Social Sensorial Data-Driven Experiences. Electronic Imaging, 2020 (13), 360-1-360–9.

West, R. et al. (2020). Immersive Rendering and Sonification of Large Scale Antarctic Astronomy Data in Virtual Reality: The Making of INSTRUMENT | One Antarctic Night. Scientific Committee on Antarctic Research (SCAR): Astronomy and Geo-Space Observations from Antarctica Session. 3 August 2020. Hobart, Tasmania, Australia: International Science Council: Scientific Committee on Antarctic Research, 30.

West, R. et al. (2018). Experiencing a Slice of the Sky: Immersive Rendering and Sonification of Antarctic Astronomy Data. Electronic Imaging, 2018 (3), 4491–44910.

Publications related to Dataremix and Data driven immersive experiences

West, R. et al. (2015). DataRemix: Designing the Datamade. Leonardo, 48 (5), 466–467.

West, R. et al. (2013). Dataremix: Designing the datamade through artscience collaboration. Proceedings of the IEEE VIS Arts Program (VISAP).

West, R.G. et al. (2013). Art, science, and immersion: data-driven experiences. The Engineering Reality of Virtual Reality 2013. 4 March 2013. International Society for Optics and Photonics, 86490H.

Contextual Materials in Appendix

See **Appendix 1** for materials provided for context yet not within the scope of this thesis commentary related to the artwork *Ecce Homology* and associated publications

Acknowledgements

I would like to express my deepest appreciation to my supervisor Neal White for his wisdom and support throughout this journey, and to Nicola Triscott for encouraging me to undertake this PhD and for her guidance and support as my initial supervisor.

I would also like to extend my deepest gratitude to my colleague JP Lewis, for his generosity of spirit, insight and support throughout the writing of this PhD. I would like to extend a special thanks to Bill Seaman, for his encouragement, inspiration and support in the writing of this PhD and throughout my path in becoming a transdisciplinary artist. I'm extremely grateful to my colleagues and mentors, Cindy Grimm, Andrea Polli, Eitan Mendelowitz, Hepi Wachter, Roger Eastman, Tom Grubb, Ross Sowell, Bill Smart, Margaret Dolinsky, Andres Burbano, Victoria Vesna, N. Katherine Hayles, Donna Cox, Roger Malina, and Christine Malina-Maxwell, without whose inspiration, support, and understanding throughout my journey, the completion of this thesis would not have been possible.

I cannot begin to express my thanks to all the artists and scientists with whom I have collaborated on the projects discussed in this PhD, Jeff Burke, Eitan Mendelowitz, Tom Holton, Cheryl Kerfeld, JP Lewis, Ethan Drucker, Todd Margolis, Iman Mostafavi, Joachim Gossmann, Trevor Henthorn, Rajvikram Singh, Javier Girardo, Andrew Prudhomme, Chris Poovey, Luke Hillard, Ian Afflerbach, Hannah Helgesen,Violet Johnson, I-Chen Yen, Zach Thomas, Lars Berg, Alejandro Borsani, Roger Malina, Lifan Wang, Weihong Yan, Steve Lamont, Loren McQuade, Ben Hackbarth, Alex Horn, Jurgen Schulze, and Weizhong Li.

I am also grateful to the students and researchers of the xREZ Art + Science lab, Kathryn Hays, Luke Hillard, Zach Thomas, Ian Afflerbach, Jess McCurry, Chris Poovey, and Hannah Helgesen for their willingness to jump into the unknown territory in between disciplines daily, and for their support of my undertaking this endeavor.

Lastly, a heartfelt thanks to my friends and soul family who have encouraged and supported me throughout this adventure in life, this PhD and in the arts, Emilia Jimenez, Cindy Grimm, Margaret Dolinsky, Thomas Britt, Patricia Amrhein, Eitan Mendelowitz, Claire Austin, JP Lewis, Varuni Kulasekera, Brian Ritchie, Christopher Lueg, Mary White, Craig Penner, Lauren Zemmelman-Schneider, Sascha Schneider, Leong Tan, Meg McLaughlin, and the Teacher.

Dedicated to my parents and family who overcame the unimaginable with grace.

Author's Declaration

I declare that all the material contained in this thesis is my own work.

Sections

1.0 Introduction

This PhD by published work expands on the contribution to knowledge in two recent large-scale transdisciplinary artistic research projects: *ATLAS in silico*, and *INSTRUMENT | One Antarctic Night* [IOAN] and their exhibited and published outputs. Digital documentation of the exhibited works and their associated peer-reviewed publications to accompany this commentary are provided in the digital *Volume 2*. The overarching aim of this practice-based artistic research is to explore the paradox of big data as a proxy for '*reality confronting its representation*[1] '. The research interrogates data abstraction: the ever increasing digitisation, datafication[2], and abstraction of the individual, culture, and nature as vast and abstract digital data at scales outside that of the human body. At is core, the research asks: " How do we see and know beyond what the framing narratives of our instruments, algorithms, representational schemas and training guide us to see and know? How do we look for what we don't know we're looking for?" In response, the works and publications in *Volume 2* explore the artistic possibilities for destabilizing the framing narratives of data creation, analysis and representation to create visceral and poetic experiences of data abstraction and potentials for generating artistically-informed (scientific) discovery.

Data abstraction, the production of digital data at unprecedented velocity and scale, occurred in both science and society in the early 2000's and has since accelerated to become a defining characteristic of our contemporary moment. This includes data from sources as diverse as automated genomic sequencing, terrestrial observatories, image or video sharing applications (e.g. flickr.com launched in 2004 or YouTube in 2005), social media, high-resolution sub-cellular

---

1 (Hughes and Brecht, 1979, p36–37)

2 Datafication is the "unearthing of data from material no one thought held any value."(Mayer-Schönberger and Cukier, 2013, p77) It translates what would have previously been considered invisible or opaque processes or events into data that can subsequently be tracked, aggregated, analyzed and used in limitless ways.

volumetric imaging, surveillance, financial transactions, space born telescopes, Internet-of-things devices, health trackers and the personal data streams of the "*quantified self*[3]" movement. By 2007 the amount of digital data produced or replicated exceeded global available storage space for the first time (Gantz et al., 2008). It was predicted that with ten-fold growth occurring over five years, by 2011 over half of the "digital universe" could not be stored, would not have permanent storage and exists in transient states. Data abstraction fuels our "digital shadow," information about individuals (surveillance imaging, financial records, phone calls, medical data, sensors, Internet of Things, social media etc.) as compared to that created directly by individuals (texts, emails, documents, videos, photos or audio media sharing etc.) that now comprises over fifty percent of all digital data (Gantz et al., 2008). This digital shadow underlies the mechanisms of surveillance capitalism (Zuboff, 2015) that mine its contents in pursuit of profit and behavioral/societal control, and the relentless scrutiny of a data analytics industry that subjects us to an inescapable "data gaze"(Beer, 2019). In its unrelenting acceleration data abstraction fuels the ongoing expansion of a "digital universe" predicted to reach 175 zettabytes of new data created by 2025 (Seagate and International Data Corporation, 2020).

An early driver of data abstraction and significant contributor to expansion of this digital universe is the scientific quantification and abstraction of nature as high-velocity, large scale, digital data (Mayer-Schönberger and Cukier, 2013, p82). This abstraction of nature into digital data underlies the ongoing transformation from hypothesis-based to discovery oriented science initiated in the early to mid-2000s. Sequencing the human genome, a fifteen year endeavor completed in 2000-2003, relied upon high-throughput, high-velocity robotic, and factory-like, data acquisition (Lander et al., 2001; International Human Genome Sequencing Consortium, 2004). Shifts in approaches to measurement, such as those brought about by the techniques of

[3] (Wolf, 2011)

the genome project, have historically lead to substantial change in the conduct of science (Cukier 2010 cited in (Kitchin, 2014a, p128)). Genomics ushered in a shift from hypothesis-driven experiments on "isolated parts" of nature with "data closely linked to observation" driven by individual investigators, to non-hypothesis based discovery oriented research based on high-throughput large scale data (Brent, 2000). This shift also lead to the prominence of consortia-based "big science"(Frazier et al., 2003) powered by ubiquitous use of black box algorithms, such as BLAST (Basic Local Alignment Search Tool) (Altschul et al., 1990; Altschul, 1997) running on networked servers with massive data repositories[4]. Gray and Szalay view data abstraction, and the application of concepts and techniques developed through the genome initiatives to other domains[5] resulting in widespread digital datafication, as a turning point for science in which theory, experiment, and simulation are unified by a data-intensive scientific method applicable across all domains. They consider the emergence of this data-intensive method as the start of a "*fourth paradigm* for scientific exploration"(Hey, Tansley and Tolle, 2009, pxix).

The exhibited and published works in *Volume 2* (and earlier works that inform them) evolved at, and from, this pivotal juncture – the transition from hypothesis-driven to discovery-oriented science grounded in high-velocity big data ushered in by the sequencing of the human genome. The works are located at the intersection of artistic practices the utilize (scientific) data as a material, virtual reality, digital abstraction and poetic recombination. The novel, very large

[4] Worldwide BLAST servers supported by genome sequencing consortia include the US National Center for Biotechnology Information (NCBI) https://www.ncbi.nlm.nih.gov/, GenBank, Ensembl http://www.ensembl.org/ at the European Molecular Biology Laboratory European Bioinformatics Institute in the UK, and DNA Databank /GenomeNet in Japan https://www.genome.jp/. (Stevens, 2018) For a sense of the impact and ubiquity of the BLAST algorithm in generating knowledge in science in 2003, the year the human genome sequence was completed, the NCBI server alone supplied 100,000 unique BLAST runs from 70,000 unique IP addresses every 24 hours. That is ~ 36.5 million analyses from ~ 25.5M IP addresses during 2003 from just the one BLAST server in the US. This does not account for the myriad private proprietary implementations of BLAST or any of the other public online genome consortia web servers. To put this volume in context, ecommerce sites were just gaining popularity, and the stock price of Amazon in the NYSE averaged ~$50 USD in 2003, versus average $3500 USD per share in 2020. BLAST is one of the most widely used data-mining tools in history. And yet, due to the black box nature of the algorithm, and the privileging of velocity and high throughput analysis, the majority of users in 2003 (along with contemporary users) accept the default heuristic scoring matrix settings that ultimately determine the ranked results from the analyses. (NCBI usage data provided by W. Matten, National Center for Biotechnology Information (NCBI), User Services, personal communication, March 2003.)

[5] Informatics concepts from the genome initiatives have been applied in a broad context including, for example bioimaging data and Earth terrestrial or oceanic sensor network data.

scientific datasets used to create the works are high-dimensional, multi-scale, and were not publicly accessible at the time of their incorporation in to the works. Produced through high-velocity robotic data capture in both metagenomics[6] and astronomy they span multiple scales from the immaterial informatic to molecular to sociodemographic, environmental and astronomical. The contributions to knowledge from the exhibited works and publications share a common goal of interrogating data abstraction, specifically through problematizing the generation of meaning and knowledge in science arising from the abstraction of nature into big (scientific) data. The problematizing occurs through a set of two research questions (section 1.2 below) that ask how visceral and poetic experiences of data abstraction can be created, and how one would go about generating an artistically-informed (scientific) discovery.

1.1 Context and Rationale

"[The artist] must reduce all of the subjective and objective with the end of informing human sensuality. He tries to give human beings direct contact with eternal verities through reduction of those verities to the realm of sensuality, which is the basic language for the human experience of things...Sensuality stands outside of both the objective and subjective. It is the ultimate instrument to which we must first refer all our notions, whether they be abstract, the result of direct experience or of some circuitous reference to such experience. Sensuality is our index to reality."

The Artist's Reality: Philosophies of Art, Mark Rothko, Christopher Rothko, Yale University Press, 2006.

The artistic research is inspired in part by what Rothko (2004) termed the *"realm of sensuality,"* which in his view resides outside of the *objective* and *subjective*. The artworks in *Volume 2* utilize whole-body interaction within generative virtual environments constructed from unique large scale scientific datasets as a way of referencing data abstraction back to our sensuality. This approach brings embodied experience in relation to the ineffable scales of digital data,

[6] Metagenomics is the study of a collection of genomes from a mixed community of organisms sampled in their natural habitats. It comprises a set of techniques and defines a field of research. Direct sampling of organisms in natural habitats overcomes both the inability to culture the majority of microorganisms in the laboratory as well as the challenge posed by microbial genomic diversity. Metagenomics is made possible through the integration of high-velocity DNA sequencing, high-performance computing (compute clusters), and bioinformatics algorithms. (*The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet*, 2007)

irrespective of its velocity, size or the scales of phenomena it represents in search of a sensate abstraction. Also reflected in the works and publications is bringing together and facilitating extensive transdisciplinary collaborations spanning multiple disciplines. This is an integral part of the artistic practice-as-research that explores possibilities for creating artworks with multiple entry points that can exist concurrently as aesthetic and poetic experiences of vast and abstract data, public engagement, and as investigations of approaches to creating artistically-informed (scientific) discovery. The artistic research develops novel transdisciplinary approaches and artworks that are purposefully located at the interface, along the *subjective | objective* edge.

1.2 Research Questions

Artworks and peer-reviewed published outputs in *Volume 2* are interrelated and build upon each other in response to two simultaneous streams of inquiry.

The first, related to aesthetics and formal attributes (ways of experiencing) asks: 1) *How can a visceral and poetic experience of big data and data abstraction be created?* Can this create a sensual/sensate abstraction, that exists beyond either the objective or the subjective, and existing on the interface, on an edge in between? Can sensuality be not only an "index to reality[7]", but an index to the abstraction of reality as vast digital data? How can poetic alternatives to cinematic zooms be constructed to span vastly differing spatial and temporal scales within data?

The second, related to discovery (ways of seeing and knowing) asks: 2) *How would one go about generating an artistically-informed (scientific) discovery?* How do we look for what we don't know we are

[7] (Rothko and Rothko, 2004)

looking for? How do we detect the unexpected, when we see what we're trained to see, what our framing narratives, methods and instruments enable us to "see"?

2.0 Contextual Review

The exhibited artworks in the digital *Volume 2* and their associated peer-reviewed publications are the outcomes of artistic research that spans 10+ years and is conducted throughout the development of the artistic practice. The work and its aims are multifaceted and as a result this contextual review draws upon multiple selected themes. These include the framing narratives of (big) data abstraction, artistic practices that utilize (scientific) data as a material, virtual environments, digital abstraction and poetic recombination.

2.1 Artistic Practice: Aims and Methodology

This section (2.1) of the contextual review also draws on established definitions of artistic research in discussing the aims and methods of the author's artistic practice-as-research.

2.1.1 Artistic Research

Borgdorff (2010, p118, 2012, p162) argues that artistic research is "thinking in, through, and with art." It articulates the experiential "pre-reflective, nonconceptual content of art," and "…thereby invites 'unfinished thinking'." He provides the following operational definition:

> *"Art practice qualifies as research if its purpose is to expand our knowledge and understanding by conducting an original investigation in and through art objects and creative processes. Art research begins by addressing questions that are pertinent in the research context and in the art world. Researchers employ experimental and hermeneutic methods*

that reveal and articulate the tacit knowledge that is situated and embodied in specific artworks and artistic processes. Research processes and outcomes are documented and disseminated in appropriate manner to the research community and the wider public." *(Borgdorff, 2012, p53)*

The research unfolds through artistic practice and both the research and new knowledge is embodied in its products and are shared publicly. In this way, artistic practice can advance knowledge (Candy, 2011, p55; Scrivener, 2011a, 2011b). In differentiating the varied aspects of artistic research, Borgdorff goes on to describe the ways in which artistic research can be considered "boundary work[8], and "boundary objects[9]." In regards to artistic research, boundary work relates to "negotiations that are required along boundaries" or between academia and the artworld (Borgdorff, 2012, p196). Additionally, boundary objects, are objects that have a given "meaning in a certain research environment and another meaning in another research environment" and can span both academic and non-academic contexts (Borgdorff, 2012, p196). He expands on artistic research as boundary work through the "blurring of art and other life domains" and the ability to articulate things that "bear on who we are, [and] where we stand" as also transdisciplinary research that exemplifies Mode 2[10] knowledge production (2012, p198). In arguing for the ability of art to generate new knowledge, Borgdorff (2012, p207–208) draws from Hans-Jörg Rheinberger's conceptualization of experimental systems that dynamically interrelate

---

[8] Gieryn (1983) examines the characteristics of science and how it is demarcated from non-science to acquire its "intellectual authority" in articulating the concept of boundary-work. It defines the ways in which science, as an institution, constructs social boundaries between itself and non-science domains to protect its interests (authority or resources). Boundary-work functions as the negotiations of a fluid social boundary to demarcate what counts as science or non-science for the purposes of expansion or monopolization of professional authority over scientific domains and acquisition of material resources, such as public funds for research or inclusion in educational curricula, or the protection of autonomy over professional activities or to displace responsibility for undesired impacts of consumption of science by non-science domains. As a fluid ideological style, it is used to describe science as empirical, theoretical, pure or applied, depending on what is most suited to a claim of authority in a given context to protect its interests. Boundary-work is seen as the process by which the "intellectual ecosystem" is "carved up" in to knowledge-production "niches" to differentiate "goals, methods, capabilities and substantive expertise(1983, p783)" and as a way of favoring the authority of the institution of science (or of other domains) in knowledge production. Artistic research is thus boundary-work through its assertion of the territory of knowledge production as extending to artistic practices and negotiations in that territory spanning academia and the art world (Borgdorff, 2012).

[9] Star and Griesemer (1989) conceived of boundary objects in the context of facilitating cross-disciplinary collaboration. "Boundary objects are both adaptable to different viewpoints and robust enough to maintain identity across them." They function as "means of translation" between different worldviews, and can embody, and represent, or concretely convey knowledge, and facilitate the transformation of knowledge(Carlile, 2002).

[10] Mode 2 knowledge as defined by (Nowotny, Scott and Gibbons, 2003) is characterized as transdisciplinary knowledge engaging multiple theoretical and methodological perspectives, generated in varied contexts of application, including social contexts, is a reflexive process that goes beyond the objective, with outcomes having multiple validieis depending on their context. This is in comparison to mode 1 knowledge generated through foundational (basic) research, conducted within a disciplinary context with a defined set of methods and as part of traditional academic discourse.

technical objects, the experimental conditions, and objects through which we know, and corresponding knowledge objects (epistemic things) that they engender. Within these systems epistemic objects may transform to technical objects, (or technical objects transform to epistemic objects) and retain their dynamics and epistemic potentials by becoming embedded within additional layers of experimental systems and continue to "produce unprecedented events in different contexts"(Rheinberger, 1997, p80). Further Borgdorff concludes:

> *"In the context of artistic research, artworks are epistemic things and events that have not yet been 'understood' or 'known' - or, to be sure, that resist any such epistemological grip. Art's knowledge potential lies partly in the tacit knowledge embodied within it and partly in its ability to continuously open new perspectives and unfold new realities[11]. I have elsewhere described this 'knowing' as pre-reflective and non-conceptual. I would now like to characterize it, with Rheinberger, as a productive not-yet-knowing against the backdrop of an ever-receding knowledge horizon" (Borgdorff, 2012, p213).*

In line with Borgdorff's definition, the exhibited artworks and peer-reviewed publications in *Volume 2* are artistic research "where the end product is an artefact – where the thinking is, so to speak, embodied in the artefact (Frayling, 1993). These artefacts are both "the work of art" and "the works of art" that generate new knowledge[12] (Scrivener, 2011a, p61). They are the outcome of a creative engagement with the issue of high-velocity digitization of nature and culture and data abstraction.

The methods and knowledge acquired in the creation of *ATLAS in silico*, are informed by a prior work that is outside the scope of this thesis commentary, *Ecce Homology*. In turn, knowledge gained in development of *ATLAS in silico*, is expanded upon, in the development of *INSTRUMENT | One Antarctic Night*. The outcomes of the works in *Volume 2* were neither pre-

[11] Bordorff (2012, p45)also expands upon artistic practices as both and/or either revealing the world to us or constructing the world, and details the multiplicity of points of view within artistic practice as aesthetic, interpretive, performative, mimetic, expressive, and emotive, stating that artistic research can engage one or all of these points of view simultaneously.

[12] Here Scrivener refers to the ways in which "the work" of a field, the act of labor in the field itself, relates to the outcomes, or "the works of" a field in which knowledge and understanding is embodied. He writes, "artistic research might be said to be artistic because it incorporates the work of and works of art in the generation and exchange of new knowledge and understanding"(2011a, p61).

determined or previsioned, nor did the author first create theoretical and methodological frameworks and apply them to produce the artistic practice or the artworks in order to generate determined or hypothesis-driven outcomes. Instead the artworks emerge from an initial set of questions (see research questions, section 1.2), and a somewhat underdetermined position. They are by definition not yet fully known, and develop/evolve through artistic practice and a process of exploration and discovery such that the resulting artwork embodies creative engagement with the research questions. In light of Borgdorff's definition of artistic research which draws upon Rheinberger's conceptualization of experimental systems, the artworks also in some ways function analogous to experimental systems. One way that this occurs, for example, is when the artworks are used to investigate how one would go about generating an artistically-informed (scientific) discovery (research question 2, section 1.2). The works are the outcome of a transdisciplinary practice. Aspects of the works are valid in some contexts and not in others, making them boundary objects in the context of artistic research as per Borgdorff's definition.

Within this broader context of practice-as-research, the use of resonance as a metaphor in the author's transdisciplinary practice (see section 2.1.2 below) emerged through experimentation in, and through, artistic practice and its products.

2.1.2 Transdisciplinary and Collaborative Artistic Practice: x-resonance

Resonance, the ability of objects to vibrate at one or more of their natural frequencies in relation to each other, serves as a central metaphor in the transdisciplinary and collaborative artistic practice. This emerged over a span of 10+ years in response to the need to constellate and facilitate knowledge, processes, concepts, objects, materials, media etc., funding, and large-scale collaboration within and across a wide range of disciplines in the process of conducting artistic research.

The author's practice is termed *x-resonance*. Structural collapse of the Tacoma Narrows Bridge on November 7, 1940[13] due to resonance evoked the notion of x-resonance and is a material inspiration. Study of the collapse, through a practice of "failure literacy" common in engineering disciplines (Delatte, 2010), led to new knowledge for design of bridges. Disruption of existing knowledge, materialized as the bridge, led to new knowledge and new material realizations despite that there is yet to be consensus on an explanation as to the sudden appearance of the oscillations (Arioli and Gazzola, 2015) that destroyed the bridge. In this regard, it is of interest to note that, Rheinberger (1997, p80) evokes stabilization, destabilization and resonance in order for experimental systems to remain "productive in an epistemic sense " by maintaining the dynamics of the system "at the borderline of its breakdown." This dynamism, or the "bundle of inconsistencies and unknowns" is "held together by the system's "*differance*" its "capacity to produce distinctions " and the continuous production of epistemic things[14] (1997, p101).

In the author's practice, resonance, a phenomenon we most often associate with the interaction of vibrational frequencies, is a metaphor for broadly (and rather loosely) denoting destabilization/(re-)stabilization, disruption, similarities and/or differences in domains, and the processes by which resonances (in both or either) can act upon existing forms or knowledge, ultimately yielding transformation and the generation of new forms and knowledge. This allows for the transdisciplinary practice to engage with Nicolescu's (2014) concept of the "hidden third," in which the bridging of the objective and subjective arrive at new meaning. Resonance also supports development of the novel aesthetic and high-dimensional data mapping strategies employed in the works in *Volume 2*. These strategies result in articulating *dataremi*x[15], an approach

---

13 Video of the bridge collapse is online at: https://www.youtube.com/watch?v=3mclp9QmCGs
14 Ultimately the structuring of experimental systems in and through this production of differences as their driving force enables systems to be differentially reproduced (Rheinberger, 1997, p224).
15 Dataremix will be discussed further in later sections. The term was coined by the author and published in (West, Malina, et al., 2013; West et al., 2015).

that resonates and remixes scientific (objective) elements with artistic and contextual (subjective) elements. In this respect, the artworks in *Volume 2* represent embodied encounters with vast and abstract data at the *objective | subjective* interface and as a form of materializing the "hidden third" through art within a "transdisciplinary space," a space created at the intersection of the *objective | subjective*, embodied, immersive, "multi-sensory and interactive experiences, and engaging varied worldviews … in the process of making" (Steelman et al., 2019).

Although x-resonance employs a physics-based metaphor, the practice was not originally conceptualized in relation to Barad's (2007) agential realism, which is based in quantum physics. The development of dataremix took place in the context of x-resonance. Even a form of remix that proceeds through the metaphor of resonance implies there are pre-existing entities available to represent or mix, in contradiction to the notion of indeterminacy followed by intra-action that leads to the simultaneous co-creation of meaning and matter in Barad's agential realism. Yet recent work by Arlander (2020) suggests that cut-up and remix methods may share commonalities with diffractive methods of dialogically reading texts through each other by proposing a view of mixture as fluidity (via Coccia (Arlander, 2020, p34)) that retains characteristics of "waves, interference and diffraction patterns" and in which entities mix but do not lose their form or substance. Utilizing concepts such as those proposed by Arlander make it possible to conceptualize dataremix as constellating and resonating the non-opposition of subjectivity and objectivity through x-resonance in part as potential diffractive readings of objective elements such as scientific tools, data, and analytical processes pertaining to the scientific datasets utilized in each of the works, through artistic, aesthetic, interactive and subjective gestures that are embodied (materialized), as the experiential artworks in *Volume 2*.

While the author as practitioner exhibits some attributes of paradisciplinarity, the "parallel and symmetric practice of scientific and artistic activities" (Lapointe, 2012) by simultaneously working and publishing in the sciences[16] and exhibiting and publishing in the arts[17], transdisciplinary collaboration is central to the author's methodology.

The author's transdisciplinary practice develops hybrid processes and outcomes which are purposefully and productively situated at the *objective | subjective* interface, have potential to convey multiple meanings simultaneously to a variety of audiences (e.g. non-specialist lay-public or specialist), and resist disciplinary definition. Elements from distinct disciplines are continually reinterpreted and elaborated to achieve integrative outcomes, that in resisting disciplinary definition may remain incomplete in some aspects and/or fall short when viewed from discipline-specific perspectives. It shares some aspects with the types of collaborations in which scientists and artists are co-authors on both arts outputs and scientific publication outputs (Malina, 2010) and some aspects of Donna Cox's "Renaissance Team"(Cox, 1991, 2002, 2008) model for collaboration in which multi-faceted teams of artistic, computer and other scientific or humanities researchers work together across disciplines to achieve shared outcomes. Its working processes include aspects characteristic of transdisciplinarity articulated in (Mobjörk, 2010) such as self-organizing collaborations that arise outside of academic or institutional structures, developing a shared language, novel methods, and integrating multiple world views and value systems in a manner that evolves with the project.

The foregoing transdisciplinary practice and experimentation in development of x-resonance over 10+ years ultimately led to the author founding the xREZ Art + Science Lab, (http://xrezlab.com). It is an x-resonance studio+lab embedded within a matrix of academic,

---

16 Peer-reviewed publications by the author have appeared in journals/proceedings including: Proceedings of the National Academy of Science (PNAS), Nucleic Acids Research, American Journal of Human Genetics, Genomics, American Academy of Neurology, SPIE Proceedings/IST Electronic Imaging Journal, Human Factors Journal, Scientific Committee on Antarctic Research proceedings, Neural Information Processing (NeurIPs), Symposium on Applied Perception

17 Curated/juried exhibitions of artworks by the author have appeared in venues such as UCLA Fowler Museum, Los Angeles Municipal Art Gallery, WIRED NextFest, ACM SIGGRAPH Art Gallery, MONA (Museum of Old and New Art, Tasmania, Australia) and peer-reviewed publications have appeared in journals/proceedings including CAA (College Art Association Conference), Aminima, Leonardo Journal, Leonardo Electronic Almanac, Artnodes

industry and community partnerships. The author conducts her artistic practice and research at the lab, currently located at the University of North Texas, where the author is a professor cross-appointed in the College of Visual Art and Design and the College of Engineering with a courtesy appointment in the College of Science. Graduate and undergraduate students from the disciplines of new media art, music, political science, design, information science, psychology, history, physics, mathematics, computer science, sculpture, journalism, anthropology, philosophy, and geography have participated in transdisciplinary research and creative work in the lab's highly collaborative environment with outcomes including peer-reviewed publications and public exhibitions.

Within the transdisciplinary collaborations for the works in *Volume 2,* members of the team (and their respective disciplinary expertise) are increasingly affected in regards to their respective disciplinary identities and practices by the outcomes being produced as the artistic research evolves. In their ethnographic study of the renaissance teams of Donna Cox and collaborators at the Advanced Visualization laboratory, (Woodward et al., 2015) articulate a theory of teams developing a collective identity driven by Simondon's concept of a "technological object, " an object that during the process of its development drives members of a team to extend beyond their specialist/disciplinary boundaries to engage in "extraspecialist" collaborative work. This tenet informs the author's research, the broader work at the xREZ studio+lab, and each of the works in the portfolio. In parallel with Woodward's theory, the author's transdisciplinary collaborative process "avoids mapping expertise onto contributions" and instead, in recognition of the "active role of the technological object in its own production," (Woodward et al., 2015, p503) credits any and all collaborators that participate in any phase of a given project, no matter the role, scope or nature of the contribution. For example, at the xREZ lab, a student may participate during a summer, helping set up an audio system for an installation, and they are credited in perpetuity on the project. This is in line with scientific research collaborative practices, and with artistic collaborative practices in theater or film, for example. It also reflects the nature of the transdisciplinary collaboration and artistic research for each of the works in *Volume 2* as one in which "an engaged human's thoughts, practices and 'becomings' are immanent to the workings and becomings" of the project (Woodward et al., 2015), and its underlying technology and data.

2.2 Data Abstraction: Framing Narratives and Big (Scientific) Data

> *"…Darwin faced a very basic visual problem: how could natural selection, a concept almost by definition impossible to illustrate directly, be illustrated, especially when the existing visual conventions of the natural sciences were associated in varying degrees with conceptions of species fixity?"*
>
> *Jonathan Smith, Charles Darwin and Victorian Visual Culture, 2006.*

While offering unprecedented access to the inner workings of nature and ourselves, and seemingly functioning analogous to realism in visual representation, at the same time that big data reveals it also obscures. In their analysis of Magritte's work, "The Human Condition," in which a painting of a landscape is placed by a window, and in the painting the exact portion of the landscape that the canvas obscures is represented, and is seemingly continuous with reality, Hughes and Brecht (1979, p36–37) articulate the paradox of "*reality confronting its representation*" and the contradiction of art that in "seeking to reveal has obscured." Their delineation of the paradox as a vicious cycle is grounded in the contradiction that the image displaces reality and there is no means to determine if the work itself contains every aspect of the original landscape or underlying phenomena.

Similarly, this paradox and aura of incompleteness pervades big (scientific) data as it abstracts nature. Yet, big data capture is routinely designed for extensibility and endless reuse. Examples in non-science contexts include the reuse of retail surveillance camera footage intended to track shoplifting that instead is used to track customer flow and eye gaze for analysis of marketing effectiveness, or Google Street View's data capture that exceeded capture of images of roadways and structures for mapping to include GPS data, and WiFi network names and content from open wireless networks which enabled it to use the data for secondary uses, such as its research in autonomous vehicles (Mayer-Schönberger and Cukier, 2013, p109).

In the sciences, this design for extensibility and reuse, despite the inherent incompleteness, underlies the *fourth paradigm's* data-intensive scientific method that incorporates the state of knowledge of any domain. It spans knowledge incorporated within experimental design, instrument (apparatus) design and specimen preparation or setup, resolution, formats, storage, access and governance, and the design of the of database schemas that integrate these multiple layers of domain and practitioner knowledge. Extensibility also facilitates vast and abstract data that is multi-modal, multi-scale and that spans multiple spatial and temporal scales simultaneously. Discussing data-intensive (big data) science, Gray proposes a thought experiment: "You've done some science, and you want to publish it. How do you publish it so that others can read it and reproduce your results in a hundred years' time?" (Hey, Tansley and Tolle, 2009, pxxix). Gray is addressing the endless reuse that is central to the utility of big (scientific) data that is captured for one purpose then analyzed for alternate purposes, by those other than the original creators of the data, at multiple points in time, even 100 years or more from when it was created. For example, the Biomedical Informatics Research Network (BIRN) (established in 2001) aims to enable this kind of scenario through data federation, sharing, governance and workflows across 30+ international institutions (Helmer et al., 2011). An example of such a resource built upon the BIRN framework is the Cell Centered Database (CCDB) that extends the data paradigms originated in genomics and proteomics to high-resolution 3D light and electron microscopic biomedical imaging. The data model of the CCDB reflects experimental design within the context of imaging techniques. It makes available original datasets, as well as analysis products derived from them for reuse in a manner that retains the provenance of the data capture, analysis and representation (Martone et al., 2003, 2008).

Extensibility, reuse and incorporation of provenance allows data from different domains, both scientific and those spanning societal contexts, with data at varied geographic, conceptual, or other physical and natural scales, to be related and linked to data across scales and sources. boyd and Crawford (2012) observe that size is no longer the defining characteristic of big data, it is this cross-referencing and aggregation, the multi-modal, multi-scale, multi-dimensional characteristics of the data abstraction, as well as the capacity to query and search across dimensions that truly allows for data to be "big."

Irrespective of its extensibility, potential for reuse, scope, complexity or dimensionality (all characteristics that confer its value) big (scientific) data abstracted from any phenomena is often incomplete, ephemeral, and "…always a selection from the total sum of all possible data available – what we have chosen to take from all that could potentially be given. As such, data are inherently partial, selective and representative, and the *distinguishing criteria used in their capture has consequence*[18]" (Kitchin, 2014b, p3). Barad also alludes to how distinguishing criteria matter in how we see determining what we can see and know in writing "The apparatus is an inseparable part of the observed phenomenon" (2014, p180). Sampling methods, regardless of how high a resolution, and the limitations of instruments introduce discontinuities at boundaries between spatial and temporal scales, and the resulting digital data remain approximations of their underlying phenomena. This gap persists within the realm of simulations of nature or culture because our models that underlay the simulations, and the simulations themselves, are based on our digital sampling. Data functions as a material resource in science, yet a gap persists between

[18] Emphasis added.

data, its algorithmic reductions, and the meaning ascribed to the data throughout its manipulation and refinement in the data-to-knowledge pipeline. Distinguishing criteria, the framing narratives of data capture, analysis and representation, are both specific to a dataset and universal in that they circumscribe attributes inherent to all (digital) data (e.g. resolution, format, storage medium, governance and access etc.). They encode agreed upon, unspoken, and often unacknowledged assumptions by the creators of the data about what might be found within it, even upon its reuse for unimagined purposes over time that could span, as Gray proposed in his thought experiment, 100 years or more.

Data abstraction brings about a parallel through big data to the displacement of reality by the image that underlies the "paradox of the realist painter, of art that both reveals and obscures[19]," revealed in Magritte's work. In brief: as drivers of ever increasing data abstraction, framing narratives arise from choices such as what and how to sample (digitally datafy), at what resolution and format(s), database schemas, statistics, algorithms, representational schemas, display and interaction technologies and metaphors used in the processing [20] and representation of big (scientific) data. Combined these framing narratives circumscribe how and what we can see and know.

The author's artistic experimentation and research questions are motivated by the following working observation in regards to framing narratives, data abstraction and the paradox of big (scientific) data as a proxy for reality *"confronting its representation."* Akin to Darwin's challenge of representing natural selection in a 19th Century visual culture vested in concepts of species fixity (Smith, 2006), in an era of data abstraction and big (scientific) data of increasing scope and complexity we face a challenge comparable to Darwin's problem: How do we see and know

---

19 (Hughes and Brecht, 1979, p36–37)
20 This processing is the transformation of data from raw to processed through various algorithmic, statistical and other manipulations within the data-to-information-to-knowledge-to-wisdom pipeline(Ackoff, 1989; Rowley, 2007).

beyond what our representational schemas, algorithms, instruments and prevailing (scientific) and representational culture enable us to see and know? How do we look for what we don't know we are looking for?

2.3. Aesthetic Experience: Data Mappings, Immersive Virtual Environments, and Poetic Recombination

Digital art engaging data is a heterogenous area of practice. Data sources employed by artists to address a limitless array of themes range from personal data streams of the quantified-self movement (Frick, 2012) to collective emotions as reflected in social media (Harris and Kamvar, 2005), to air traffic time series (Koblin, 2005), storm and meteorological event data (Miebach, 2019), particulate air pollution (Polli, 2016), to the creation of immersive artificial nature (Wakefield and Ji, 2021) and beyond. The artistic research is situated within this realm of digital art practices that use data as a material. Its focus and contribution is specifically in the use of large scale, high-dimensional, multi-scale big (scientific) data, as artistic material and as an exploration of aesthetic experiences of big data and data abstraction. It takes the form of generative and interactive immersive virtual environments and utilizes poetic recombination in its conceptual and methodological framework. The artistic research is conducted within a transdisciplinary and collaborative practice that engages scientific and artistic collaborators, and also students from multiple disciplines within an integrative studio+lab setting. Therefore, this section contextualizes the work in *Volume 2* in relation to a few selected works that engage big (scientific) data of similar scale and dimensionality to the datasets utilized by the author, and which are realized within similar collaborative practices that bring together artistic and scientific collaborators, and students from varied disciplines, within the context of artistic studio+labs. It is also contextualized in relevant viewpoints in regards to aesthetic experiences of big data, virtual environments and poetic recombination.

ATLAS in silico, *INSTRUMENT | One Antarctic Night* and a work that preceded them, *Ecce Homology*[21], create novel aesthetic experiences of big (scientific) data and of data abstraction. Eliciting an aesthetic or affective sensation in the experiencing of data is a characteristic of artistic practices that utilize data as a material (Corby, 2008a; Li, 2018).

Lisa Jevbratt (2004), writes of an aesthetic sensation, a new inverted sublime, in relation to artistic visualization of big data in response to Lev Manovich's use of her work, *1:1(2) Migration* (Figure 1) as an example of the ideal of the anti-sublime (Manovich, 2002). Jevbratt formulates this inverted sublime as distinguished from the classical sublime in terms of direction and force, and challenges the need for reduction in formulating the data mappings utilized to engage big data. Manovich(2002) regards visualization as a subset of the mapping of one type of representation on to another, which is made possible through computation. He observes that despite being "one of the more common operations in computer culture" and in new media art, the selection of dimensions from within a dataset and how they are mapped on to dimensions of representation is also one of the most important and political (Manovich, 2002).

[21] *Ecce Homology* informed the development of the works in *Volume 2* yet is outside the scope of this commentary.

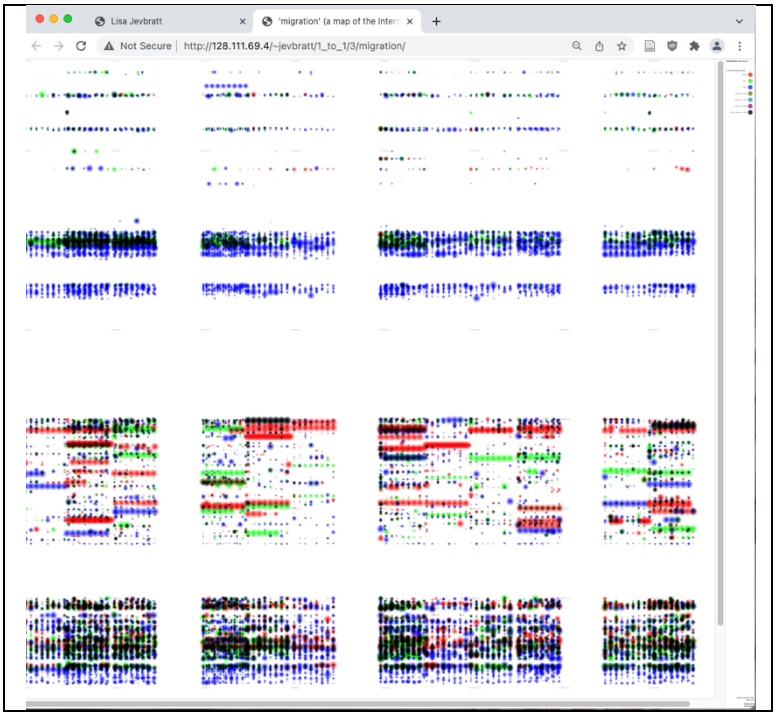


Figure 1: Lisa Jevbratt (1999, 2005) 1:1(2) Migration interface

One of several interfaces for viewing a database of the IP address of every website in the world. Constructed by custom crawlers collecting samples of IP addresses to create a representation that increased in resolution over time. There are 255 IP addresses per pixel. Top left is 0.0.0 and bottom right is 255.255.255. Red is sites from 1999. Green is sites from 2001/2002. Blue is sites from 2005. Size is number of sites on a logarithmic scale. (Source: http://128.111.69.4/~jevbratt/1_to_1/3/migration/ and Jevbratt, 2005, http://128.111.69.4/~jevbratt/1_to_1/index_ng.html)(Jevbratt, 1999)

The pivotal role of data mappings is the foundation for Manovich's critique of visualization and his derivation of the anti-sublime. His argument for visualization as the anti-sublime is, in brief, that data visualization, whether artistic or scientific, proceeds through the mapping of the "un-graspable," that which is at scales "beyond the limits of human senses and reason" to representations at the scale of human perception and cognition, in essence creating the opposite of the sublime (Manovich, 2002). He observes that through data mappings, artistic visualization produces "clear and orderly forms" that reveal patterns and structure from the "informational chaos" of vast data. Data mapping in essence performs a reduction, and data visualization becomes a new form of abstraction. Its dual nature at once abstracts the concrete into minimalist patterns and in turn those same patterns are then capable of generating an endless variety of images(Manovich, 2002, p7). Manovich considers computation's ability to "map anything into anything else" a fundamental problem for data mapping, and artistic visualization in general, in that the choices in mapping appear arbitrary unless a specific motivation is articulated that would relate the content and context of the data itself. As a solution, Manovich proposes the foregrounding of the arbitrary nature of the chosen mappings and a potential for chance operations or irrational operations in the mapping process (Manovich, 2002, p10). He critiques new media art, data mapping projects, for their "rational impulse to make sense of our complex world" and posits that the "typical strategy" of mapping complex and vast data onto representational schemas is a way to ' "read off" underlying social relations" from big data' (Manovich, 2002, p11). Manovich concludes with an appeal for artistic visualization to reveal the "ambiguity always present in our perception and experience" and to address the "challenge [of] how to represent the personal subjective experience of a person living in a data society" a form of "data-subjectivity" in which our human subjectivity includes a new dimension, that of being continuously" immersed in data," in new ways.

To this, Jevbratt counters that we experience an inverted sublime in which we are looking down and in through our creation and use of big data and in the process are being distanced or repelled from that which we want to know, whereas in the classical sublime we look up and out and are attracted towards that which we want to know. She notes, along with Manovich, that reduction, both in terms of the amount of data and options for its mapping to representation, is privileged as a strategy in seeking insight from big data in both scientific and artistic visualization. She states that "the most common mistake in data visualizations, artistic or scientific, is not too much information but too little, their "images" of the data landscape are not high resolution enough for an esthetic decision to be made" (Jevbratt, 2004, p7). Through invoking intuitive esthetic reasoning and a "methodological distancing" she challenges the need for reduction as an approach to dealing with the vastness, or the "un-representable" in big data and advocates that intuitive understandings of big data can be achieved through drawing on sensations of the sublime.

It is useful to note that this debate around the aesthetics of artistic visualization of big data emerged in the early to mid-2000s, at the same time as the completion of the sequencing of the human genome was prompting the shift from hypothesis-oriented science to big data, discovery-oriented science. A wide-ranging discourse ensued pertaining to artistic practices that engage data as a material. It intersects with this dialogue between Manovich and Jevbratt. Several relevant and representative views include: Warren Sack's (2011) response to the anti-sublime which broadens the frame of artistic visualization to move beyond the visual to include an aesthetic of governance to address the data abstraction of the body politic; Andres Ramirez Gaviria's (2008) classification of artistic visualization as conceptually transgressive research art that operates through metaphors that " need not be easily decipherable nor aesthetically pleasing;" Robert Kosara's (2007) classification of pragmatic versus artistic visualization along a continuum of sublimity which emerged as a response to the crisis faced in 2006 in the field of scientific/information visualization due to a lack of a consolidated definition of the science of visualization; Tom Corby's (2008a) aesthetics of a new type of image consisting of "objective and subjective, informational and aesthetic components" capable of producing "critical, aesthetic, and affective experience" and an "embodied knowledge;" and Kim and Park's (2013) classification of artistic visualization along the aesthetic dimensions of viewpoint, interpretation and alternative senses.

In formulating an aesthetics of information visualization, Warren Sack (2011) offers alternatives to the anti-sublime, seeing it as one possibility in a larger set of artistic visualization practices. These include not only an aesthetics of the sublime and the anti-sublime, but also of the uncanny, of administration, and of governance. Establishing a parallel in the genesis of computing as an outgrowth of bureaucracy and artistic information visualization as a critique of the computational "bureaucratic machinery," he frames his argument for an aesthetic of governance in the methods and art historical precedents of conceptual art that critique industrial and administrative (bureaucratic) production. This facilitates the inclusion of social and political dimensions of data within the frame of an aesthetics of information visualization. With this inclusion, Sack proposes a shift in artistic visualization from a visual aesthetics to an aesthetics of governance: a shift from an aesthetics of the body of the individual subject to the body of the collective as a "body politic." Additionally, he proposes an inquiry into the dematerialization of the body and other material objects through digital datafication, which places them further at risk of disappearance beyond dematerialization due to the ephemerality of digital data. In essence, he proposes an exploration of their data abstraction and its relation to the Body Politic, and states that "inventing a visual form to show the Body Politic itself to itself is the outstanding problem of artistic research in information visualization."

This view of artistic visualization as grounded in conceptual art is shared by Andres Ramirez Gaviria(2008). He adopts two categories from Blais and Ippolitos (2006) classification of artistic production, genre art and research art for his argument. Genre art, is artistic production of artefacts that are comprehensible to a broader audience and circulate as part of the established art market's commercial system. Research art, engages experimentation, collaboration and is comprehensible primarily to a specialist audience. Gaviria proposes a classification of artistic visualization as research art, and not solely as aesthetics. He writes that the focus in artistic visualization as research art is on the ability for the work to be "conceptually transgressive" and question or problematize issues in ways not otherwise possible. In regards to data mappings, Gaviria proposes that artistic 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" (Gaviria, 2008, p482).

Robert Kosara (2007) proposes a sublimity scale along which to classify pragmatic, informative and artistic visualization and defines criteria for each according to the degree to which a visualization aims to convey data in ways that are recognizable and readable (Figure 2). He motivates this in a call for development of a "visualization criticism" to create a theory and language for the field of visualization as a whole. At one end of the scale, pragmatic visualization is defined as being utilitarian, based on non-visual data, and as producing a visualization image that is both recognizable and easily readable. A key criterion is visual efficiency in order for charts and graphs and other visualization images used to communicate the data effortlessly with speed and accuracy. Pragmatic visualization seeks generalizable approaches applicable to a wide range of datasets. Informative art is positioned at the mid-point. Examples are ambient visualizations that are not recognizable yet are readable. Artistic visualizations, at the opposite end of the scale from pragmatic visualization, are defined as possessing an aesthetics of the sublime. Neither recognizable nor readable, the objective of artistic visualization is not to convey data but to communicate a concern and while "the underlying problem may not be visible," it is " made visible through the piece"(Kosara, 2007, p3). Instead of visual efficiency, the objective is to enable a viewer to understand the concern. Kosara's proposal follows on from a 2006 panel at the IEEE Visualization conference titled "Is there science in visualization?" (Jankun-Kelly et al., 2006) in which the field acknowledges it is in crisis and lacking a clear definition of visualization as a science.

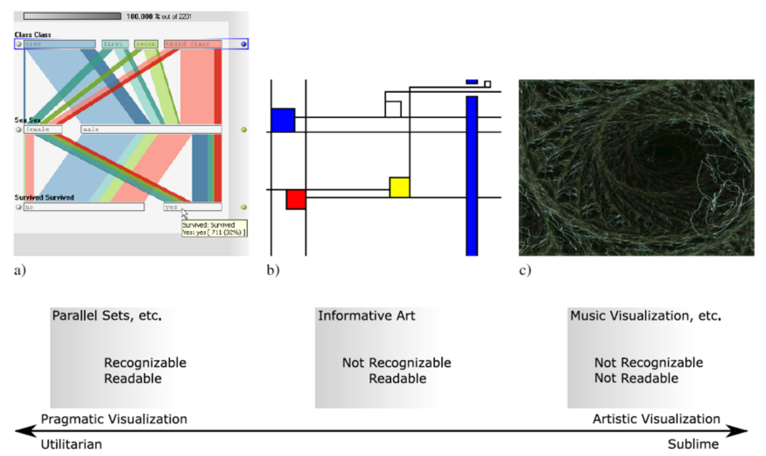


Figure 2: Sublimity Scale developed by Robert Kosara

Sublimity scale with end points as pragmatic versus artistic visualization. Images in the figure: a) pragmatic visualization - parallel sets (data about people on the titanic) readable and recognizable visualization, b) ambient visualization (data about a bus schedule): readable but not recognizable; c) artistic visualization: music visualization (based in data) not readable.

Source: Kosara, R. (2007). Visualization Criticism - The Missing Link Between Information Visualization and Art. *2007 11th International Conference Information Visualization (IV '07).* July 2007. Zurich, Switzerland: IEEE, 631–636. Available from https://doi.org/10.1109/IV.2007.130

Kim and Park (2013)focus their analysis at the artistic visualization end point of Kosara's sublimity scale and amplify it further. They disambiguate three aesthetic dimensions of artistic visualization and propose a classification of works as articulating one or more viewpoints, interpretations, or developing "innovative sensoria[22]." In their schema, different artists working with a data source can have differing motivations and utilize the same data to present differing perspectives on a topic, or different perspectives on the same dataset to establish viewpoints. Artists can also utilize data as a mode of dynamic interpretation of a process and/or objects out in the world or internal to their own artistic practice for which they collect data or to develop narrative views that they surface from within data as interpretations. The third dimension, is the development of new "sensory organs" or enhancing, shifting and/or inventing senses. Works in this category may, for example bring awareness of the limitation of human senses.

[22] (Kim and Park, 2013, p1)

In a parallel to Jevbratt's proposal of esthetic reasoning as a counterpoint to reduction in artistic (information) visualization, Kim and Park remark on the potential value of intuitive analysis through artistic visualization for addressing "problems that cannot be solved with current visualization techniques" (Kim and Park, 2013, p2). Yet they go further to ask "…why do we keep downgrading a multidimensional massive input data into lower dimension for perception? Is it not better to shift or create a new conceptual sensorium which is capable of handling high-dimensional data?" (Kim and Park, 2013, p5). Kim and Park propose the artistic development of alternative senses that acknowledge the limitations of human sense perception and development of sensoria beyond the body as a counterpoint to "low-dimensional quantitative representations" that give rise to the recognizable and readable dimensions of pragmatic visualization (Kim and Park, 2013, p6). In their description of the potentials and challenges of these new senses they identify the author's work *ATLAS in silico* and its multiscale representations created through the novel process of dataremix (described in section 3.0 below) as an example of works that shift and/or invent such new sensoria. In their discussion they point to the inherent complexity, ambiguity and unreadability of these representations, and identify this as a challenge. In section three this characteristic is discussed in terms of representational layers of works intersecting with experiential and novel spatiality. This characteristic is therefore purposeful in the creation of scalable representations for *ATLAS in silico,* and in how they operate within the creation of a mutli-form sublimity and aesthetic experiences of data abstraction achieved through the research.

At the nexus of these views of aesthetic/artistic visualization, Tom Corby (2008a) touches upon the aesthetic of the sublime, the connection to modernist abstraction, and an aesthetic of social practice in concluding that information visualization artworks "develop a new type of image that consist of objective and subjective, informational and aesthetic components"(Corby, 2008a, p467) and produce "critical, aesthetic, and affective experience" and "embodied [and] aesthetically situated knowledge" (Corby, 2008a, p464). He analyzes two works that create a "fictive space that is an amalgam of different data" to produce hybrid images constructed from data captured from social interactions: *Cyclone.soc* (Corby and Baily) and *Mount Fear East London* (Abigail Reynolds). Corby relates the "visual and immersive landscapes produced by these works" to the sublime, articulates the connection of their respective data mapping strategies'[23] to modernist abstraction, and their relationship to social dynamics through their datasets to demonstrate how artistic visualization conveys "complex subjective experiences in novel aesthetic forms" (Corby, 2008a, p464). Functioning both at the limits of how images function in the visual arts and how they function in scientific information visualization, these new types of images produce "meanings that are often ambiguous or multi-leveled and that produce embodied, affective, sensory experiences that elude rational description and measurement" (Corby, 2008a, p467)

In *Cyclone.soc,* (Figure 3) (Corby and Bailey, 2005; 2007) the artists describe their mappings as "Streamed live, newsgroup postings are fitted to the atmospheric topologies of visualizations of cyclonic weather fronts to give the effect of the conversational churn and eddy of newsgroup argument and counter-argument"(Corby and Baily, 2007) and as "thread[-ing] information derived from social interactions into a pre-existing meteorological topology" (Corby, 2008a, p463). Multiple layers of mapping include, text data (letterforms) mapped onto topological contour lines, motion/animation of the letterforms along the contours, projection of the letterform streams/flows into and onto a physical space at architectural scale, and within which the audience can freely move and walk about. Viewers interact with the large-scale immersive projection by walking within the space or using controls to navigate to specific areas of interest.

[23] Corby relates the "looping spiraling texts in *Cyclone.soc*" as reminiscent of Concrete Poetry, and the drawings by Reynolds for the production of *Mount Fear East London* to "late or synthetic cubism and the formal experiments of Russian constructivism" (Corby, 2008a, p464)

Figure 3: Cyclone.soc (2005) Tom Corby, Gavin Baily

Source: Corby, T. and Baily, G. (2005). Cyclone.soc [image]. Available from http://tomcorby.com/cyclone/

As Manovich (2002) observes, data mappings are foundational to artistic or scientific visualization. Mappings are constructed and can operate at multiple levels within artistic visualization, enabling multiple levels of scale and coherence for the work as a whole. For example, works can be viewed through the lens of how the data and mapping strategy employed are reflected at the scale of individual formal elements (e.g. data value mapped to formal attribute such as shape or color), the overall format of the work (e.g. screen-based, multi-projection installation, immersive environment etc.), or mode of interaction (e.g. single or multi-user, tangible interface, tether-less full-body motion capture etc.), and any combinations of these with metaphors, the overall concept, and aims. The multifaceted interrelationships of data mappings, formal elements, materials and media, duration and dimension give rise to our experience of a given work.

These mappings exist in artistic visualization/data artworks that utilize an endless variety of data that span domains as disparate as posts from internet chat rooms and online public forums (Hansen and Rubin, 2002), real time data from a colony of tracked mole-rats (Freeman, 2016), or livestreams from webcams in every time zone on Earth (Thompson and Craighead, 2009). They are also present in artworks that utilize big scientific data and are central to the conceptual and methodological manifestation of the work.

An example of a work utilizing scientific data in astronomy is George Legrady's digital media artwork *We Are Stardust (2008).* Created by Legrady and collaborators within the context of his Experimental Visualization Lab at UC Santa Barbara[24], the work is a two-screen interactive projection that employs data from the orbiting NASA Spitzer infrared telescope and from an infrared imaging camera in the installation space (Legrady, Villegas and Burbano, 2009; Legrady, 2011). On the first screen, 36,064 astronomical locations comprising the 5 year sequence of observations made by the space born telescope is visualized as traces animated and overlaid onto a 2D projection of the celestial sphere showing a star chart of approximately 2,000 of the brightest stars visible from earth with constellations marked and labeled (Figure 4, top panel). This same location data, which is in the form of a time series of equatorial coordinates, is simultaneously used to drive the motion of an infrared camera mounted in the ceiling as it collects infrared imaging of the audience moving about within the installation space (Figure 4 bottom). The infrared camera data presented in the second projection is geometrically altered as the camera changes its orientation to follow the direction of the path of telescope observations.

Visualizing the sequence of locations for the 36,034 observations instead of the imaging data captured by the telescope functions as one of the layers of mapping in this work. The mapping layers employed include: star location in equatorial coordinates (right ascension (RA) and declination (Dec)) mapped onto a shape (a square); category of observation made by the telescope mapped to color applied to the shape; time series of observations mapped to lines connecting two or more shapes; line segments animated in time series amongst shapes; and the

[24] George Legrady Experimental Visualization Lab: https://datascience.ucsb.edu/people/george-legrady and studio http://georgelegrady.com/

difference between RA and Dec of successive observations is mapped to IR camera pan and tilt motions. An additional layer of mapping is duration: the 5 year observational time series is mapped onto 25 hours of operation for the installation. The visual language of spherical star map projections is mapped onto the infrared camera's rectilinear geometry using a sinusoidal projection to achieve geometric distortion. The physical layout of the installation as dual projections on opposite ends of a rectilinear gallery space with IR camera pointing down from above, and projections at floor-to-ceiling scale creates a data and conceptual mapping onto immersive physical space at architectural scale. The audience interaction occurs through walking about in the space and having ones image captured via the IR camera and presented in real-time into the geometrically distorted projection. The interaction functions as a conceptual and metaphorical mapping to the infrared imaging data from the telescope that is absent in the first (astronomical) projection. It is displaced in the work by the infrared imaging of the audience as they walk about in the installation. This displacement operates as a level of mapping described by Manovich (2002) as a strategy employed by Situationist Surrealism, in which the work performs a "spatial "ostranenie" (estrangement): to let the city dweller experience the city in a new way and thus politicize her or his perception of the habitat. One of these methods was to navigate through Paris using a map of London" (Manovich, 2002, p10). In *We Are Stardust* the audience experiences this estrangement by seeing themselves as abstracted through the location data of astronomical observations to scales beyond the human body via the infrared imaging.

Figure 4: We Are Stardust installation (2008) Legrady, Villegas, Burbano

Top: Visualization of Spitzer telescope sequence of observations. Bottom: Overhead infrared imaging of the installation space and audience as a geometrically distorted projection.

Source: Legrady, G. et al. (2008). We Are Stardust [image]. Available from https://www.mat.ucsb.edu/g.legrady/ glWeb/Projects/st/stardust.html

Two additional examples of work that utilize large scale scientific data utilizing mappings that articulate multiple scales are *Southern Ocean Studies* (Corby, Baily and Mackenzie, 2012) and *Carbon Topographies* (Corby et al., 2021) (Figure 5 and Figure 6) Both are part of a larger body of work in which Corby and artistic and scientific collaborators utilize climate modeling data to elucidate interrelationships amongst environmental and societal forces impacting climate change, and engage these through dimensions of climate modeling data. They are created with multiple artistic and scientific collaborators in the context of Corby's transdisciplinary research studio+lab https://www.manifest-data.org/ and https://reconoitre.net.

Two climate models are utilized for *Southern Ocean Studies*. The first is the Global Atmosphere-Ocean General Circulation CM2.X Coupled Climate Model. CM2.X: models the physical, chemical and biological dimensions of the climate system and feedback processes from the Antarctic region going back to 1861. The second OCCAM, the Ocean Simulation Model, models data as a time series that spans 24 years and includes biochemistry, seal level, sea ice, temperature, salinity, and wind velocity.

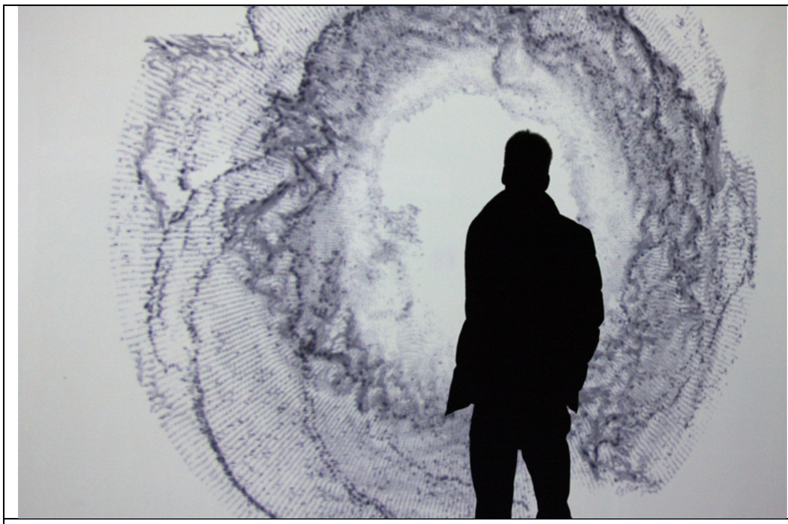
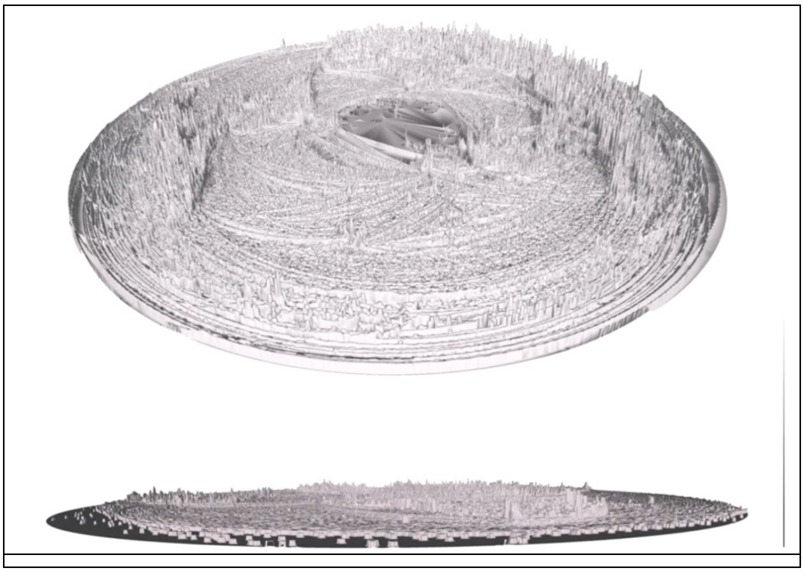


Figure 5: Southern Ocean Studies (2009-2012) Corby, Bailly, Mackenzie

Animation of Antarctic Circumpolar Current climate data including tidal flow, sea ice and wind direction. Large-scale projection installation.

Source: http://tomcorby.com/southern-ocean-studies/

The data utilized for the *Carbon Topologies* is comprised of global emissions of the three primary greenhouse gasses (C02, CH4 and N2O) per emissions sector (e.g. transportation, energy generation, agriculture etc.) per country and is obtained from the Emissions Database for Global Atmospheric Research (EDGAR) maintained by the European Union (https://edgar.jrc.ec.europa.eu/emissions_data_and_maps) (Corby et al., 2021).



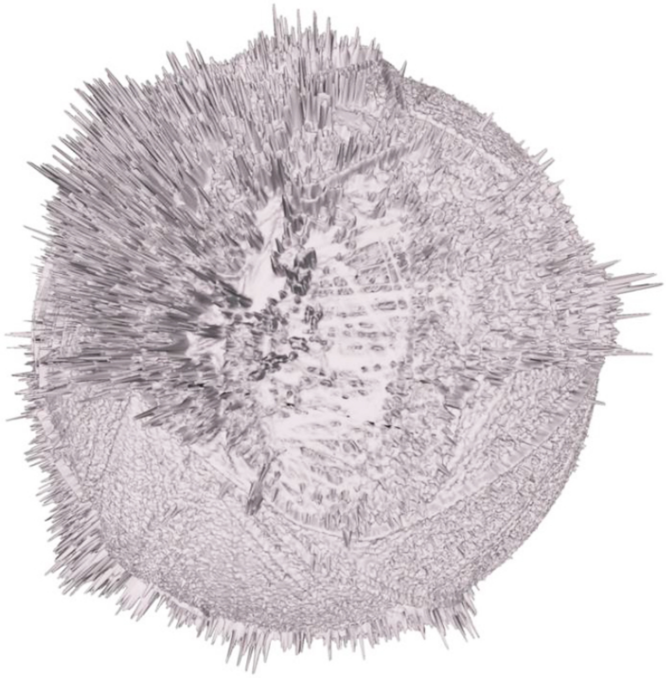


Figure 6: Carbon Topologies (2020) Gavin Baily, Tom Corby, Jonathan Mackenzie, Giles Lane, Erin Dickenson, Louise Sime, George Roussos

Top: Model of the Southern hemisphere of global distribution of CO2 production from energy, industrial, agriculture, flights and shipping 1970 – present. Antarctica is at the center and Europe towards the periphery (1,500 mm diameter). Two views: top and side view.

Bottom: 3D model of global CO_2 from cruise ships and flights 1970 – present (500 mm diameter)

Source: Corby, T. et al. (2021). The Data Imaginaries of Climate Art: The Manifest Data Project. Leonardo, 1–10. Available from https://doi.org/10.1162/leon_a_02136.

In both works, the data are large scale, multi-dimensional and multi-scale. Data mappings for each reflect the underlying phenomena and data sources. For *Southern Ocean Studies*, the dynamism of the climate systems and biogeochemical processes are mapped on to a monochromatic particle system (Figure 5) that is animated in response to climate model data. This mapping enables multiple levels of scale within the data (e.g. wind direction, sea ice, tidal flow and geochemical and atmospheric flux) to be apprehended as dynamic interlaced patterns that cohere into a larger overall pattern. The Antarctic landmass is made evident as a result of the dynamics of the patterns in the particle system of the surrounding ocean data. This creates an aesthetics of interconnectedness in which the landmass, the ocean dynamics and global climate are a system-as-pattern. As in Legrady's work (Figure 4), an aspect of this work is the juxtaposition of presence and absence of different data types to construct visual impact, metaphors, poetics and meaning. The floor-to-ceiling scale of the projection creates visual impact and immersion into the represented environmental system. Interaction is through contemplation or walking in the space in front of the projection. For the *Carbon Topologies*, amounts of greenhouse gas emissions per sector and per country are mapped to geometric and topographical features of a 3D model and materialized as 3D printed sculptures. The 3D prints condense vast scales of landmass, atmosphere and ocean to tactile structures in the near range to the scale of the body (between 1500mm and 500mm). Mapping of scale of emissions amounts per sector and country to height of the geometry as well as density of the geometry for the 3D printed structure creates a tangible aesthetic for an otherwise invisible process of accrual over time of greenhouse gasses in the atmosphere. Differences in types and amount of emissions are mapped on to density of the topographical features of the 3D structures. In this manner levels of scale are mapped to a tactile dimension that is reinforced visually as an irregular yet scaled pattern. In both works, the interrelations amongst societal and environmental phenomena at different spatial and temporal scales, emerge through mappings that generate of multiple types and layering of dynamic patterns.

Legrady's and Corby's works employ scientific data that spans astronomical, planetary, and regional scales in their creation. Collectively, they are situated at the artistic visualization end of Kosara's sublimity scale and employ mappings driven by the characteristics of scientific data. To convey and address relationships of scale amongst the data, the underlying phenomena, and the human body the works use multi-layered mapping strategies at levels of formal features of graphical elements, dynamics, interaction and installation format and spatiality. Scale is mapped onto both the elements of the representations and onto immersion within a projection space, situated within an architectural space, or onto tactile 3D structures that invert massive planetary or regional scales in relation to the body. Scale is also mapped on to interaction through walking within the space adjusting one's proximity to the projections or through approach and position of the body in relation to the 3D printed structures. Through these mappings the spatiality of Legrady's and Corby's installations itself functions as if it were a medium.

While the author's artistic research complements many of the approaches in these works, its original contributions extend them in new directions. The works and publications in *Volume 2* develop a novel framework for aesthetic and artistically-impelled high-dimensional data-driven non-hypothesis-driven mappings that operate on multiple layers within each work: representational elements, embodied interaction and immersive virtual environments whose spatiality emerges through big (scientific) data to create aesthetic experiences data abstraction. The multilayered operation of the data mappings and their functioning in the totality of the artwork, extend Corby's conceptualization of an image that contains both "subjective and objective through informational and aesthetic components" to novel virtual environments created from big (scientific) data that exist on an interface, on the edge, between *objective | subjective*.

Knowledge gained through creation and exhibition of *Ecce Homology*[25] (a work that preceded the works in *Volume 2* and is outside the scope of this thesis commentary) informed the artistic research with immersive virtual spaces and their potential for the creation of aesthetic experiences of big data and data abstraction. *Ecce Homology* is a multi-participant interactive installation at the intersection of comparative genomics and immersive experience. Titled after Friedrich Nietzsche's *Ecce Homo*, it reflects on how one becomes what one is. It is a contemplation of human evolution and the unity of life by examining similarities (homology) between the human genome and the genome of the rice plant. The artwork reveals the operation

[25] http://insilicov1.org

of BLAST, the Basic Local Alignment Search Tool, (Altschul et al., 1990; Altschul, 1997). For information on the significance and impact of BLAST see footnote #4 (above). BLAST is the primary algorithm used world-wide for comparative genomics analyses. It is also a black box whose ubiquity is driven by its speed and heuristics.

To create the artwork, a custom computational brush model was developed, and aesthetic data mappings created to implement a non-phonetic, pictographic calligraphic "alphabet" for nucleic and amino acid sequences of genes and their protein products(West et al., 2005a; West, Lewis, et al., 2013). This is accomplished through a holistic and aesthetic mapping of human and rice gene sequences from Genbank[26] into genomic calligraphy. Inspired by traditions of Chinese and Sanskrit calligraphy and pictographic writing systems, the data mapping creates a parallel between how protein structure is specified through nucleic acid sequences and how that structure reflects a protein's function in an organism with the ways in which form and visual features in pictographs are directly connected to their meaning.

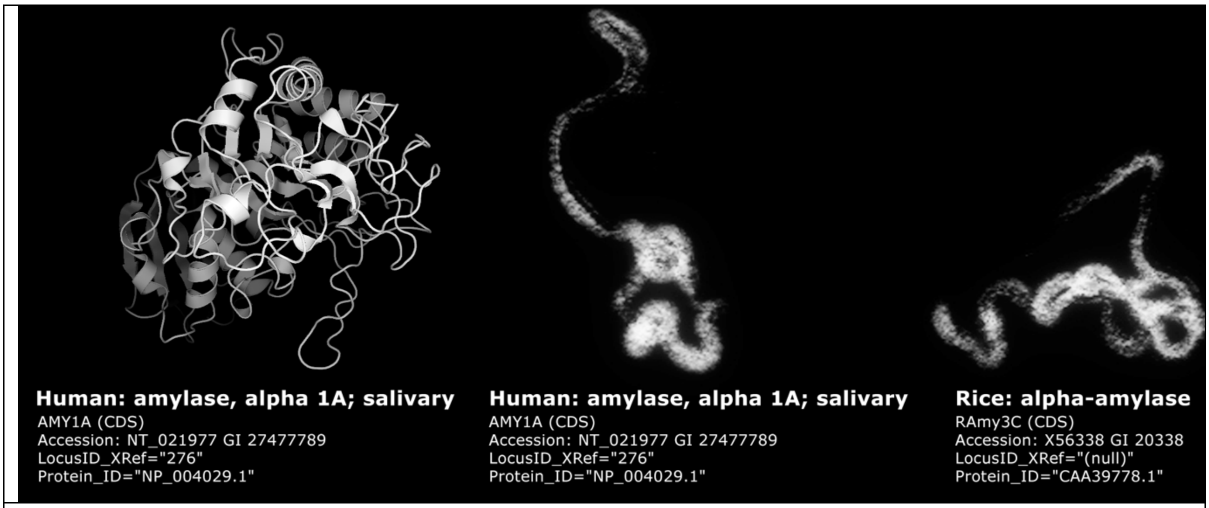


Figure 7: Genomic calligraphy from Ecce Homology for the amylase enzyme.

Middle and Right: Genomic calligraphy, protein strokes for human and rice alpha amylase gene (respectively). Left: 3D model created from x-ray crystallography by a scientific collaborator for the project.

[26] Genbank: https://www.ncbi.nlm.nih.gov/genbank/ is the US National Institute of Health's genetic sequence database containing all publicly available DNA sequences and is part of the International Nucleotide Sequence Database Collection.

A custom computer vision system tracks participants movements and creates light-filled traces into the installation projection from analysis of tracked head and hand positions. This allows visitors to interact with the luminous pictographic projections while they are standing anywhere from a few inches to several feet from the projection surface. Participants initiate BLAST analyses through these light traces that inscribe the pattern of their movement into a space in the projection that is positioned between the human and the rice genome. The comparative genomics analyses explore evolutionary relationships between human and rice genes within the installation space. A series of five tiled projectors present the customarily unseen operation of the BLAST algorithm operating on the strokes and radicals of the genomic pictographic alphabet visualizing human and rice genes across a forty-foot by twelve-foot display wall. The results of the analyses are rendered into the projection.

The installation premiered within the UCLA Fowler Museum of Cultural History and Hammer Museum's exhibition *From the Verandah: Art Buddhism, Presence.* The completion of the human genome sequence in 2003 heightened public interest in genetics and big data. Over the course of the exhibition that ran from November 2003 to January 2004, a wide ranging dialogue with the public occurred about the work, the complex science of genomics, black box algorithms and data abstraction.

Figure 8: Participants interacting with Ecce Homology at the opening evening of its exhibition at the UCLA Fowler Museum of Cultural History

Figure 9: Four views of multi-participant interaction with Ecce Homology as installed at the UCLA Fowler Museum of Cultural History

A comment from an audience participant written in the UCLA Fowler museum exhibition guest book reads:

> *"Dancing with science, with information, with knowledge – with something essentially not human yet wholly human. About humanity. Humanity defined. I've never felt such a connection with a machine, a computer, and I realize the connection, the sense of connection I felt, was with my Self. With myself on a genetic level. That which is similar mapped out a bond."*

In his essay on paradigmatic performance, Stalbaum (2006) discusses *Ecce Homology*. He defines paradigmatic performance as "art practice which allows the data to have co-participatory agency in the outcome of promiscuously varied modes of artistic practice, particularly those that explore emerging relationships between data, considered virtual, and the real"(Stalbaum, 2006, p2). Paradigmatic performance is situated in what Stalbaum terms a "databased" culture, and serves as a way to create a novel mediation of the world and humanity's relationship to it. He cites Hutchins[27] in describing "the wild" as being socially distributed computational and cognitive systems in which exists a shared responsibility for broader societal outcomes. Prior to discussing his work in locative art and GIS data, Stalbaum notes that paradigmatic performance is not limited to locative art practices. It applies to any and all large (big) data along a wide continuum of scales from earth sciences, to astronomy, biology and surveillance. He considers the "scale of performance most applicable to the human body" as "somewhere between the often very small in biology and the incredibly large in astronomy" (Stalbaum, 2006, p3). The remainder of the essay focuses on Stalbaum's work with GIS data, and he states that how datasets such as astronomy or biology might "drive paradigmatic art practices" is an "open question that begs further investigation and discovery" (Stalbaum, 2006, p3). In regards to this open question, Stalbaum writes that the author's work, *Ecce Homology* is "…the only art project that [he] is aware of that explores the aesthetic possibilities of large datasets in biology in a way that mediates how humans (the audience) perform for the system: the database thus producing activity, and that activity potentially playing a role in [a] type of knowledge discovery" (Stalbaum, 2006, p6). Stalbaum's analysis affirms the uniqueness of the contributions made by the works and publications in *Volume 2* that extend the knowledge gained in creating *Ecce Homology* through the research aims of creating aesthetic experiences of big data and data abstraction (research

[27] Via Stalbaum: Edwin Hutchins, *Cognition in the Wild* (Cambridge, Massachuetts: The MIT Press, Feb 1995) ISBN 0262082314, p. 6

question 1, section 1.2) and investigating the potential for artistically informed (scientific) discovery (research question 2, section 1.2).

Reflection upon the audience experience with *Ecce Homology* at the UCLA Fowler exhibition (see quote above and Figure 8 and 9) and at a later exhibition of at ACM SIGGRAPH 2005 (Figure 10) motivated exploration of the creation of interactive and generative immersive virtual spaces from high-dimensional, multi-scale big (scientific) data as the next step in the artistic practice. *Ecce Homology* engaged both spatial and temporal dimensions of scale. The interaction is based in an aesthetic of slowness. The slower and more contemplative participants hand and bodily motions are, the more continuous and persistent the results of their interaction are in initiating BLAST comparative bioinformatics analyses amongst genes from the human genome and those from the rice plant within the artwork in the gallery space.

This aesthetic of slowness calls attention to the black box nature of BLAST that is based on operating at the internal clock speed of processors within networked compute clusters[28]. By privileging an interaction and aesthetics of slowness, it inverts the temporal scale of the black-box algorithm. Spatial scale is mapped to the representational elements (calligraphic pictographs, and algorithm dynamics) and onto a sense of immersion within the large scale projection (forty-feet wide by twelve-feet tall) that is situated in an even larger architectural space. This aesthetic of slowness is extended in *ATLAS in silico* and *INSTRUMENT | One Antarctic Night* through their embodied interaction that recapitulates scientific (objective) processes as aesthetic (subjective) physical gestures within immersive virtual environments created from big (scientific) data. Through developing and employing an aesthetic of slowness to interrogate data abstraction, the research contributes a new set of force dimensions to Jevbratt's conceptualization of the

[28] For a detailed description of *Ecce Homology*, see (West et al., 2005a). For the artwork over 10,000 lines of BLAST code obtained from the US National Center for Biotechnology Information (https://www.ncbi.nlm.nih.gov/) were refactored and deployed and custom code inserted within this implementation to enable the algorithms normally invisible process to be visible to participants in the installation.

inverted sublime. In addition to looking down and in, the "looking fast" forces of black box algorithms are positioned in relation to the forces of "looking slow" developed through the research as an aesthetic of slowness.

Adjusting one's proximity to *Ecce Homology's* projection-based display adjusts viewpoint, which can engage a greater degree of peripheral vision or collective engagement, yet this change in spatiality does not necessarily increase a sense of a visceral or embodied experience of data abstraction, nor enhance the aesthetics or poetics of the interaction. In search of an approach for going beyond the spatiality configured by this relationship between the projection display and the architectural space it is embedded within, the artistic research explores the potentials of immersive virtual reality for creating visceral and poetic experiences of big (scientific) data and data abstraction.

Figure 10: Participants interacting with Ecce Homology's forty-foot wide projection display at ACM SIGGRAPH 2005 Art Gallery/Emerging Technologies Exhibition

The exploration of immersive virtual spaces within the artistic research was informed by the pioneering work of two artists, Char Davies[29] and Bill Seaman[30]. Davies created a new bodily and experiential poetics and aesthetics of immersive virtual space, and also developed novel

[29] Char Davies artist's website: http://www.immersence.com/
[30] Bill Seaman artist's website: https://billseaman.com/

expressive whole body interaction through breath and balance in the iconic works *Osmose*[31] and *Ephemere*[32]. Seaman developed Recombinant Poetics (RP), a theory, method, and genre of techno-poetic construction (Seaman, 1997; B Seaman, 1999) for immersive interactive generative virtual environments. He develops RP through the creation of an extensive body of work that includes: *The World Generator / The Engine of Desire* (1999-ongoing), *Exchange Fields* (2001), *Hybrid Invention Generator* (2002), *The Architecture of Association* [versions 1, 2, and 3] (2008-2009), *A China of Many Senses* (2011-12), *Engine of Many Senses* (2013). He also articulates RP though a series of texts encompassing its derivation within his doctoral dissertation (W C Seaman, 1999) and its subsequent evolution (Seaman, 1998, 2001b, 2003, 2007, 2012b, 2012a, 2013; B Seaman, 2014) and extension to Recombinant Informatics (RI) (B Seaman, 2014; W Seaman, 2014) and Neosentience (Seaman and Rossler, 2008; Seaman, 2015). Similarly, to Legrady and Corby, both Davies and Seaman engage in transdisciplinary collaboration in the creation of their work.

In her artist statement, Davies's writes that her work explores an "alternative VR – approaching the medium as a means of facilitating or 'bringing forth' subjective experiences of being in the natural world that might transcend our culturally-tinted lenses of perception and behavior." She goes on to write that she has "sought to reaffirm the role of the subjectively *lived* body within the virtual realm and deeply engage the participant's sensory imagination" (Davies, 2001). *Osmose* and *Ephemere* engage these and other themes of long term interest to Davies. They include "archetypal aspects of Nature, and the desire to dissolve boundaries between interior and exterior within the context of enveloping luminous space" (Davies, 1998, p2).

Osmose (1995) and *Ephemere* (1998) are virtual reality environments comprised of three-dimensional graphics, spatialized audio, and interaction through a head-mounted display

[31] http://www.immersence.com/osmose/
[32] http://www.immersence.com/ephemere/

combined with a breath and balance sensor vest. *Ephemere* continues the work undertaken in *Osmose* (Davies et al., 1998). The *Osmose* environment contains "nearly a dozen realms, of forest, pond, subterranean earth and so on, all situated around a central clearing" (Davies, 2004, p78) and is structured on a vertical axis such that participants navigate the world by floating up and down. The primary metaphors in *Osmose* and *Ephemere* are nature and landscape. Davies considers participants immersants. Through use of their breath, immersants float throughout the virtual world.

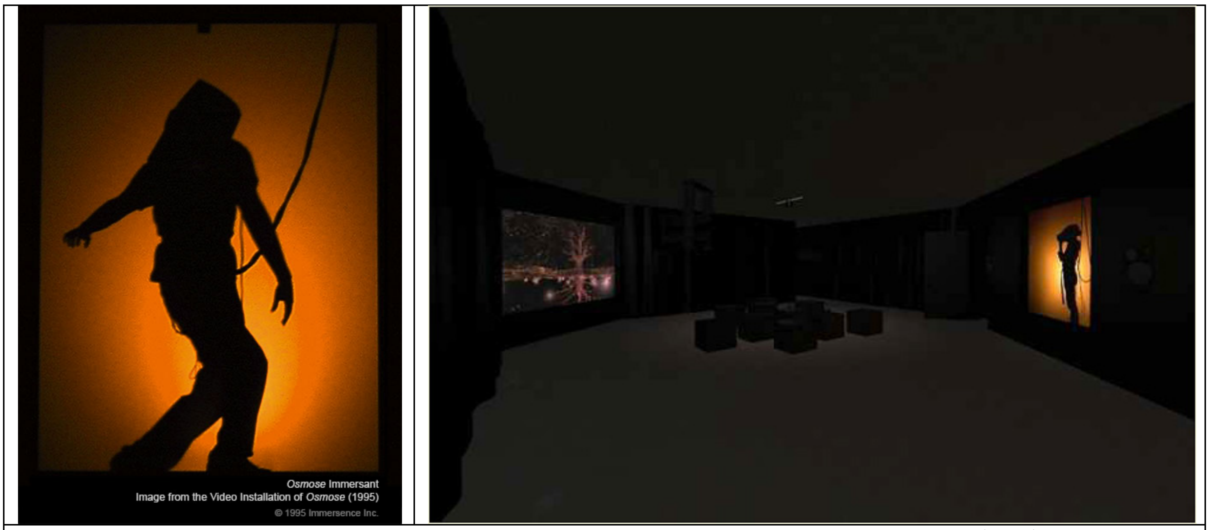
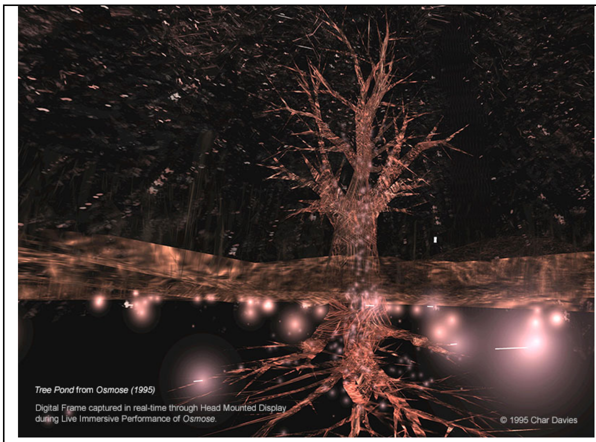


Figure 11: Left: Char Davies. Immersant during performance of immersive virtual environment, Ephémère (1998), at the Australian Centre for the Moving Image, 2003. Right: Char Davies. Installation view of live performance of Ephémère (1998), with immersant seen through shadow-silhouette screen, at the Australian Centre for the Moving Image, 2003. Images used with permission.

Source: http://www.immersence.com/osmose/
Source: http://www.immersence.com/osmose/os_screenshadow.html
Source: http://www.immersence.com/osmose/os_installation.html

As a participant experiences *Osmose's* virtual world their silhouette is projected on a screen facing a darkened space with audience seating. The view from the participant head mounted-display is projected on to another screen in the space. The intimate experience of the immersant is thus transformed into a performative gesture. (See Figure 11)



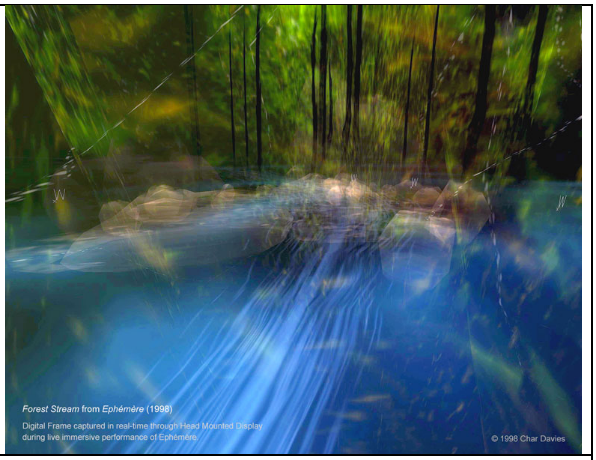


Figure 12: Left: Char Davies. Tree Pond, Osmose (1995). Digital still captured in real-time through HMD (head-mounted display) during live performance of the immersive virtual environment, Osmose. Right: Char Davies. Forest Stream, Ephémère (1998). Digital still captured in real-time through HMD (head-mounted display) during live performance of the immersive virtual environment, Ephémère. Images used with permission.

Source: http://www.immersence.com/osmose/os_treepond.html
Source: http://www.immersence.com/ephemere/eph_forest_stream.html

The creation of *Osmose* pioneered a poetics of virtual reality, as an "immersive spatiality" and the "role of the physical body within its domain" (Davies, 1998, p1). Laurie McRobert proposes that Davies has "manage[d] to 'transubstantiate' 3D space from impersonal to biological space" and that "what is important in Davies' art is that it enables people to overcome the pre-learned bias for geometric space – that abstract mathematical/geometrical realm… and allows them to delve, instead, into a biological experience of space" (McRobert, 2007, p105). Davies writes of generating "semiotic and sensory fluctuations" through the embodied experience. These are "based on a painterly strategy of maintaining a "razors edge" between representation and abstraction, whereby multiple associations or interpretations are deliberately evoked (rather than a single meaning being literally illustrated), our intent was to heighten ambiguity in order to refocus the participant's attention on their own act of perceiving, or rather of *being*" (Davies, 2004, p79).

The strategy of residing purposefully on the edge between representation and abstraction in order to evoke multiplicity of meaning is echoed in Kim and Park's (2013) proposal for the artistic development of new sensoria in the context of potential multiform meanings arising within big data. It is also echoed in the artistic research within dataremix, and each of the data mappings it generates, the creation of the immersive virtual environments, the participant interaction, and ultimately resulting in the positioning of each work along an edge between, and at the interface of the *subjective | objective.*

Complementary to the edge between representation and abstraction, and the edge between subjective and objective, Recombinant Poetics' generates emergent meaning through inter-authorship through the dynamic recombination of media elements and processes. Seaman's works take the form of immersive generative virtual environments, as well as mixed-media/projection enabled installation and performances. He characterizes recombinant poetic works as consisting of: viewer interaction, a system of meaning, carrying "compressed potential meaning constructed of language, image and sound elements" and a computer mediated, technological environment. The environment makes available elements and processes "to the viewer as emergent content of an experience engendered through interaction" (Seaman, 1997, p2). Construction of the environment(s) encodes, or "re-embodies" elements and processes into the work in ways that reflect the artist's poetic sensibility, while simultaneously yielding a generative system (Seaman, 1997) in which a "plethora of new "poetic" technological relations" can be explored (W C Seaman, 1999, p36). Going beyond mathematical combinatorics yet simultaneously operating on OULIPO principles, poetic constraints serve as processes and generative methods (Seaman, 2001b) within recombinant poetic works. They enable user interaction to generate limitless juxtapositions, recontextualizations and recombination's of image, sound and text elements within generative virtual environments. Conceiving of each media object as bringing a unique "meaning force" to the new configuration/re-configuration, and constellating this with participants historical experiences with media relations, each interaction results in a new set of meaning(s) / meta-meaning(s) within a novel "Field of

Meaning." Seaman views this infinitely iterative and generative process as inter-authorship (Seaman, 1998; W C Seaman, 1999; B Seaman, 2014).

The World Generator/The Engine of Desire (1995 – ongoing) is part of a series of Seaman's Recombinant Poetics works that explore immersive virtual environments as the site for inter-authorship of emergent meanings through dynamic recombination of text, image and sound media elements from databases. Functioning as artistic systems the generators allow participants to construct virtual worlds dynamically in real-time, and through that process, generate emergent meaning. The interface metaphor within World Generator/Engine of Desire is comprised of a set of rotating media containers that contain the media elements and their content. Each media element contains a potential field of meaning, which when poetically recombined with other objects generates emergent meaning.

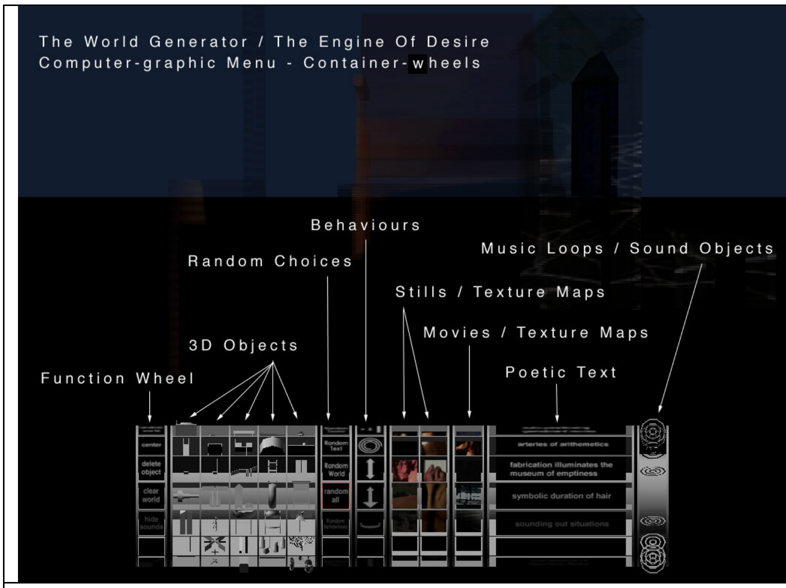


Figure 13: Menu system for The World Generator / The Engine of Desire, (1995-ongoing) Bill Seaman, Gideon May

Source: https://billseaman.com/images/portfolio/menu%20system.jpeg

Figure 14: Generative poetic recombinant construction within the World Generator interactive generative immersive virtual environment.

Source: https://billseaman.com/images/portfolio/menu%20system.jpeg

In addition to generative virtual environments, Seaman has articulated Recombinant Poetics through works in performance and embodied interface strategies. *Exchange Fields*, creates an embodied recombinant interface into an installation comprised of multi-channel video, text, sculpture, and music. The work "sought to develop a novel interface strategy by eliciting culturally determined environmental 'behaviour in relation to objects' as a grammar of gestures that could be used as input to the reacting system" (Seaman, 2001a, p1). The design of a series of furniture sculptures developed for the work incorporates an "implied 'suggestion' of how the body might be positioned in relation to that object" and embodied in the form of the object.

Figure 15: Exchange Fields (2001-2010) installation

Source: https://billseaman.com/works/2000_Exchange_Fields.html

Multiple participants move about the installation space. When positioning their body in the gesture embodied in a given sculpture, the position of their body in that gesture functions as an input system to the computer that is performing the recombination of multiple pre-recorded performance/dance videos, sound and texts. Multiple dance performances were choreographed in relation to each gesture embodied in the sculptures. The participant physically positioning their body into the sculptural gesture triggers a recombination of video, text and audio elements. The work elicits an 'action in relation" to "evoke a universality of experience" (Seaman, 2001a, p68). The participants physical position in relation to the visual, and auditory output that results from their embodied position generates the content and emergent meaning in the work.

Seaman writes, "The sensate body as a model implies a multi-sensory approach to responsive environments"(Seaman, 2001a, p69). Elements in design of the work include selection of gestures, how they enabled the interaction, communicating the use of gestures to users, and design of the furniture as physical interfaces that suggest their own use and the installation space. In this work, the "formal characteristics of the interface suggest interactions with the environment" (Seaman, 2001a, p70) with the aim of developing a "means of intelligent feedback that would enfold different layers of artistic content, with embodied experience" and as a means of "extending embodied poetic interactivity" (Seaman, 2001a, p71) with "each furniture/sculpture present[-ing] a field of potential readings and behavior evocations" (Seaman, 2001a, p78).

Exchange Fields provides useful context for the embodied interaction in *ATLAS in silico*, *INSTRUMENT | One Antarctic Night* and *Ecce Homology*. Interaction is one of multiple nested dataremix layers in each work. The participant(s) physical motion and gestures are tracked, utilizing various types of sensors, in the physical installation space. Tracking data translates their movements and gestures from physical world space into the immersive virtual space. As they interact within the artworks, the participants' physical movement and gestures are evoked through responsive characteristics of the virtual environments. The dual aspect of *Exchange Fields*, in which the "sensate body" is an approach to responsive environments and simultaneously enacts recombination is extended in the author's research, and this extension creates part of the new knowledge within the works in *Volume 2*.

Analogous to the ways in which the furniture/sculptures in *Exchange Fields* position the body in gestures and this positioning functions as an input that performs recombination of media elements, participants' full-body, physical movements during interaction in the author's artworks operate as aesthetic (subjective) gestures that in turn serve to recapitulate the scientific (objective) and analytical processes that generate (and analyze) the big (scientific) dataset utilized to create each work. The features of this *objective | subjective* embodied interaction in each of the author's works, and the elements of the virtual environment that it operates upon, are created through dataremix, and in some instances also operationalize dataremix dynamically within the environment. This enables the intertwined layers comprising each artwork to resonate in the multi-form creation of a "sensate abstraction" to enable aesthetic experiences of big data and data abstraction. The interaction and additional dataremix layers within each of the artworks are detailed in section 3.0.

A material reflection of nature's limitlessness, grandeur, and unpredictability as reflected in the aesthetics of the classical Romantic sublime, is reflected in the shifts in scale of data collections experienced in the period of scientific expeditions that explored and classified the natural world. Examples relevant to the artistic research are expeditions that inspired the trajectory of the Global Ocean Sampling Expedition (GOS) that generated the world's largest metagenomics dataset. This metagenomic dataset is utilized in the creation of *ATLAS in silico*. The oceanic expeditions include the HMS Beagle (1831-1836) and HMS Challenger (1872-1876). As an indication of the scale and scope of the data collection, The Challenger expedition alone returned with specimen collections the analysis of which generated a report spanning 50 volumes and which took 20 years to complete (Corfield, 2003). The transition from hypothesis-based to big (scientific) data discovery-oriented science that gained momentum in the early 2000s, is also in one sense, a matter scale that activates the sublime as a concern in artistic visualization approaches to big (scientific) data.

The strategies of reduction, noted earlier in this section by Manovich and Jevbratt, employed in data mappings to tame the "informational chaos" resulting from this transition in scale, are a foundational element in the analytical processes utilized to derive insight from big data and make meaning and knowledge in science. In effect, selection of a set of data dimensions and how to map them to formal elements (e.g. visual, auditory, haptic or behavioral) form part of the lower-level, foundational choices (e.g. sampling resolution, digital format etc.) that when combined result in the higher-level, overarching framing narratives of big (scientific) data creation, analysis and representation, as well as data abstraction. The mappings also serve as a foundational construct for Corby's new "objective and subjective image" that produces "critical, aesthetic and affective experience" of big data, as an entry point to Kosara's sublimity scale or Gaviria's classification of artistic visualization, as a non-visual frame in Sack's aesthetic of governance, as pivotal to Manovich's anti-sublime and collective "data-subjectivity", as intuitive and esthetic reasoning in Jevbratt's inverted sublime, or as steps towards Kim and Park's envisioned novel big data sensoria.

2.4 Conclusion

This section contextualizes the contributions to knowledge embodied in the exhibited artworks and associated peer-reviewed publications in *Volume 2* through a discussion of aesthetic experiences of big data, several iconic artworks that employ large scientific data, virtual environments, and poetic recombination.

The artistic research explores potentials for destabilizing the framing narratives of big (scientific) data through development of novel aesthetic and high-dimensional approaches to data mapping that resonate the scientific (objective) content of the data and metadata utilized in each work with the (subjective) context of the data. As described in section 2.1.2, resonance conceptually facilitates the constellation of diverse elements in multi-form engagements within the authors transdisciplinary x-resonance artistic practice. The conceptual framework, dataremix, described in Section 3 (below) is a novel approach that emerges within x-resonance*,* and through the development of the works in *Volume 2*. It creates aesthetic experiences of big data and data abstraction (section 1.2, research question 1), that are utilized in investigating how one might go about development of an artistically-informed (scientific) discovery (section 1.2, research question 2).

3.0 Artworks and Publications

This section discusses the exhibited artworks and associated peer-reviewed publications in *Volume 2* in the context of dataremix, the research questions (section 1.2), and the themes in the contextual review (section 2.0). The discussion of the works in this section is complemented by discussion of *Ecce Homology*, which while outside the scope of this thesis commentary, is a relevant context for the development of dataremix, *ATLAS in silico* and *INSTRUMENT | One Antarctic Night*.

3.1 From Poetic Mappings of Large Scale and Local Data Features to ATLAS in silico and INSTRUMENT | One Antarctic Night

An unanticipated encounter with the BLAST algorithm began the journey of seeking to foster an awareness of black box tools and algorithms that are used globally to make meaning and knowledge in science. BLAST (Basic Local Alignment Search Tool, (Altschul et al., 1990; Altschul, 1997)) is a statistical pattern matching algorithm. It performs multiple passes over genomic data. A user submits a query sequence to a BLAST server. It is matched against one or more target genome datasets using an iterative process. With each pass, the algorithm applies heuristically derived scoring matrices to the putative matches to iterate alignments between the query and one or more sequences in the target dataset. It reports out alignments along with scores that indicate how likely it is that the alignment or match is random. The lower the score, the more likely it is a non-random match, and an instance of sequence homology. While teaching a seminar on genetics and culture[33], the author's students were taught to isolate a DNA fragment, from the un-sequenced genome of a bacterium that had been cloned into a plasmid, and sequence it. Once they had obtained the sequence for their unique DNA fragment, they were then to run a BLAST analysis on it using the NCBI[34] BLAST servers. That evening, with DNA sequence in hand, students ran BLAST from the comfort of their dorm rooms and results started coming in to the online student assignment portal around 2 AM. Students applied the default scoring matrices supplied by the NCBI BLAST server and the system returned results almost effortlessly at high speed, without the students needing to attend to the algorithm or have an in-depth understanding of how it functioned. That raised the question: does this also happen with practicing scientists? Seeking an answer to this question took the form of speaking to

[33] Genetics and Culture: From Molecular Music to Transgenic Art http://geneticsandculture.com/

[34] The US National Center for Biotechnology Information https://www.ncbi.nlm.nih.gov/ provides online resources for genomics research. It is one of several global public genome repositories and online BLAST servers.

scientists from every lab at the UCLA DOE Molecular Biology Institute, consulting with the Manager of the Keck Bioinformatics Institute, emailing with the National Center for Biotechnology Information (NCBI) and ultimately reading and re-reading Altschul's papers (1990; 1997) on BLAST to seek an understanding of how the algorithm became a black box. Ease of use makes BLAST, and its many iterations, ubiquitous in the creation of meaning and knowledge in the life sciences. Seeking to understand this ubiquity led to encounters with the many visualization and representational tools used to present structure of genomes, gene sequences and the results of BLAST used in the sciences.

Pattern and pattern matching are a primary modality for knowledge discovery in bioinformatics, and other big data analytical processes. In the early 2000s, the timeframe in which the author taught the course, genomics was pioneering methods for big (scientific) data discovery-based research, and the genome sequence viewers, sequence alignment visualizations, and BLAST results visualizations were primarily alpha-numeric using a combination of A, C, T, G to represent nucleotides, and a one (or 3 letter), set of codes for the 20 amino acids (see examples in Figure 16).

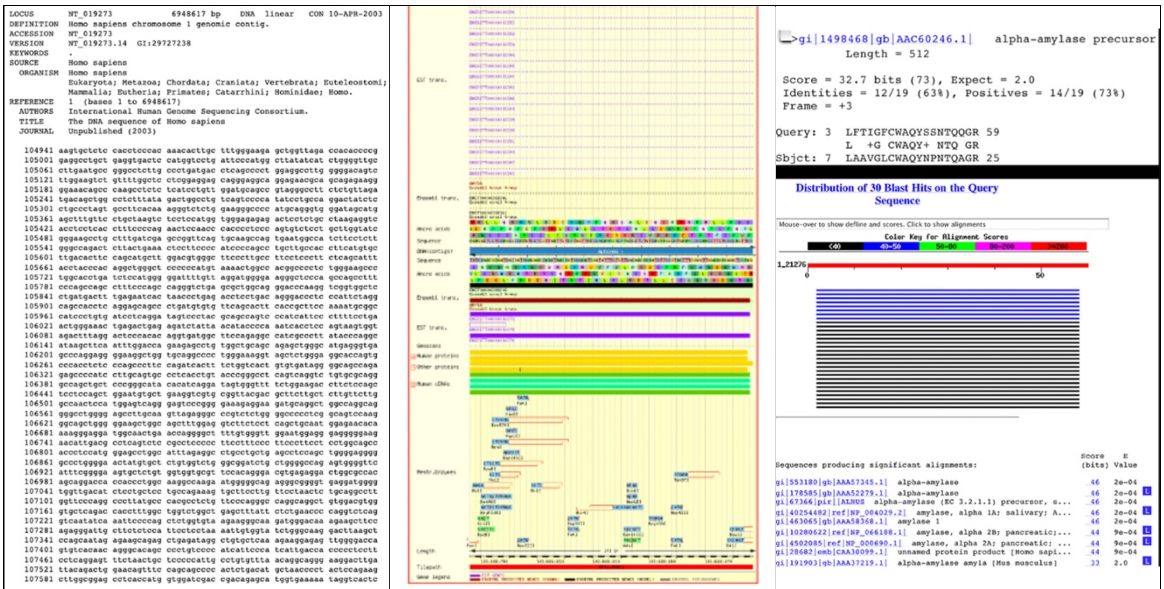


Figure 16: Varied visualization and representational formats for genomes, sequences, and BLAST results

The three panels in this figure are examples of genome sequence and BLAST results viewers at the NCBI and ENSEMBL genomics centers at the time the author was developing the concepts for Ecce Homology and the series of questions that ultimately developed into this thesis research. Left: Portion of the sequence for human chromosome 1, Middle: genome sequence and alignment viewer, Right: BLAST results viewer.

Source: screenshots are from online searches and BLAST runs conducted by the author during the creation of Ecce Homology at either the NCBI (https://www.ncbi.nlm.nih.gov/) and Ensembl (http://www.ensembl.org/)and are part of the part of the artistic working notes/sketches for the project.

Traditional representations of DNA and protein sequences in genomic databases use strings of letters spanning up to hundreds of thousands of characters. These representations make it difficult, if not impossible for human perception alone to discern internal organization, structural elements, features of biological interest and patterns of similarity or difference amongst sequences, or larger patterns and structures within and amongst genomes. Pattern matching algorithms such as BLAST overcome our perceptual limitations, but their existence and operation as a black box creates yet another layer of opacity to the generation of meaning and knowledge. This series of encounters and observations is presented here to convey that the path to the artistic research and its outcomes was (and is) an emergent process, resulting in a wide-ranging, multi-form, creative engagement with big (scientific) data, data abstraction, and the research questions.

Genomic calligraphy, developed during the creation of *Ecce Homology* represents the author's initial exploration for creating poetic experiences of big data and data abstraction. It prototyped several key features of dataremix's aesthetically-impelled, data-driven, yet non-hypothesis, nor problem-driven mappings between data and formal elements of representation.

The genomic calligraphy, poetically and aesthetically relates information about large scale genomic structural and sequence relationships to the representation of individual genes and their resulting proteins. The holistic mappings are an analog to the ways in which protein structures, as specified by DNA, reflect their function, and the way in which the form, and visual structure, of pictographic languages is related to their meaning.

Figure 17: Ecce Homology interactive installation

Left: Interaction through an aesthetic of slowness, Right: Genomic calligraphy and BLAST algorithm running in real-time (human perceivable time) in the exhibition at the UCLA Fowler museum.

Intersecting large scale features along multiple data dimensions, combined with local, small scale, and data detailed features, is a dataremix strategy for design of high-dimensional graphical and sonic data mappings. This is applied to multi-scale, multi-dimensional qualitative and quantitative data to create three new representational forms: algorithmic (meta-shape grammar) objects as natural specimens and scalable auditory data signatures for *ATLAS in silico* and signal objects for *INSTRUMENT | One Antarctic Night*.

3.2 Dataremix: resonance at the interface, on the edge, of *objective | subjective*

Dataremix is a "living " concept. It continuously evolves as part of the artistic practice. This evolution and operation are intertwined with the development of the artistic research, the works in *Volume 2*, and the research questions: "*How can a visceral and poetic experience of big data and data abstraction be created?" and "How would one go about generating an artistically-informed (scientific) discovery?*" Since dataremix and the new knowledge contributed via this artistic research threads throughout the artworks and publications in *Volume 2*, this section first presents the current working definition of dataremix, followed by a summary (Table 1 section 3.3 below) of how dataremix operates, and generates contributions to knowledge, in each of the works.

3.2.1 An evolving, working definition of dataremix

Dataremix is a unique appropriation (of scientific and other data) and poetic recombination practice for data in any and all states along the continuum of its transformation from raw to processed. It constellates and resonates along an *objective | subjective* interface varied objective, empirical and/or discovery-oriented scientific objects, and practices, with subjective, poetic, aesthetic, artistic, objects, and practices. It is both a mode of poetic/artistic construction and an approach to transdisciplinary collaboration.

Informed by x-resonance, the initial working definition of dataremix[35] also positions it as aligned within and potentially extending "Eduardo Navas's concept of regenerative remixes: "juxtaposing two or more elements that are constantly updated, meaning that they are designed to change according to data flow" (Navas, 2010, p8). Navas's (2010) classification includes extended, selective and reflexive remix in addition to the regenerative. These are respectively: remixing to extend the duration in time, to add or subtract original content while retaining the essence of the content, or to create an allegory through adding and/or subtracting material such that the remixed work stands on its own while at the same time remaining dependent on the original. The remix (recombinant) operations performed through dataremix within the works fall within all categories to some extent, yet given the generative and algorithmic processes that create them, they align predominantly with the regenerative category.

[35] PDFs of West, R. et al. 2013, 2015 are in *Volume 2*

The dual aspects of dataremix (construction and collaboration) are united and made possible by a view of data as existing along a continuum from raw to processed with transition states that can be accessed, appropriated, and utilized for various purposes, including poetic recombination via multiple objective or subjective transformations or analyses of the same data. Along this continuum, data is envisioned as a flux, flow, fluid or element with state changes and transitions triggered by artistic, poetic, algorithmic, analytic, statistical or other manipulations, transformations, and recombination of data and decisions by the artist, analyst and/or designers of the system of representation. Externalization of state changes and transitions along the continuum make partially processed data at intermediary states available for poetic recombination to create different outcomes than those envisioned within the original domain or problem specific data-to-knowledge pipeline. This externalization is manifested through use of multiple sensory modalities in creating the poetic recombination to make the abstract experiential. Data is viewed as a substrate that is acted upon by both systems and users at any point and in any direction along the continuum. This omni-directional view destabilizes the prevailing concept of the nature or behavior of data as a material resource that is processed along a unidirectional *data-to-information-to-knowledge-to-wisdom* hierarchy (Ackoff, 1989; Rowley, 2007). It serves as a way of disrupting conventions of seeing and knowing with data.

As an approach for transdisciplinary collaboration dataremix is characterized by the existence of multiple entry points, inputs and objectives, each relevant to one or more disciplines or collaborators, that result in a variety of outputs with different finalities. In these collaborations joint group goals can be created collaboratively alongside work to progress individual goals, both simultaneously and at different temporal rates, as part of a strategy to support engaging in extraspecialist activities as elicited by the evolution of the work.

Developed as part of the author's artistic practice, this collaborative approach inherits attributes of x-resonance. For example, dataremix also functions through the continuous reinterpretation and elaboration of elements from distinct disciplines to achieve integrative outcomes purposefully situated at the *objective | subjective* interface. Transdisciplinary working methods include developing a shared collaborative language, novel methods, and integrating multiple world views and value systems in response to the evolution of artefacts and the collaboration itself. The collaborative process functions as a path to multiple *objective | subjective* finalities, each of them valid within some contexts and not within others.

3.3 Operative dataremix layers and contributions within the works

The first three rows of Table 1 (below) represent layers within each work. These correspond to representational elements, spatiality (immersive virtual environment) and embodied interaction. The layers relate to how contributions to knowledge arise from the first research question: "*How can a visceral and poetic experience of big data and data abstraction be created?"* within and across the works. The fourth row relates to how the artworks function, in part, as "experimental systems" for exploration of the second research question: "*How would one go about generating an artistically-informed (scientific) discovery?*".

Dataremix Layer	Ecce Homology	ATLAS in silico	INSTRUMENT | One Antarctic Night
Representational Elements	Genomic calligraphy: calligraphic / pictographic representation of DNA and protein sequences (Figure: 7, and 16)	Meta-shape grammar objects (SGO) (Figure 19) Scalable auditory data signatures (SADS) (Figure 20)	Astronomical signal-object visual and auditory glyph (Figure 21) Signal object light trails to highlight similarity query results (Figure 22)
Spatiality	Projection embedded in architectural space (Figure 9, 10)	Interactive immersive generative virtual environment 4D scalable and reconfigurable metadata environment	Interactive immersive generative virtual environment (Figure 24)

		(Figure 23)	Physical-virtual-online continuum of interaction (Figure 25) Astronomical signal-object field (aka star-field) as visual and auditory collaborative remix instrument (Figure 26)
Embodied Interaction	Participants' body and hand motion tracked and analyzed by custom computer vision system for pattern matching. Movement inscribes human gesture traces at intersection between the human genome (culture) and rice genome (nature) (Figure 8, 27) Physical movement and gestures create light traces into the projection and pattern matching of motion and gestures selects a human gene for BLAST analysis against a gene from the rice genome (Figure 27) Aesthetic of slowness: the slower the movement the more consistent the effect of the light traces (Figure 27)	Participants' body and hand motion tracked by varied sensors: (Figure 32 and 33) Aesthetic of slowness: Gestures evoked by responsiveness of the virtual environment recapitulate the scientific and analytical processes: sampling, comparative bioinformatics analysis, and detail in context viewing and analysis (Figure 32 and 33)	Participants' body and hand motion tracked by VR HMD system room-scale tracking (VIVE lighthouses) (Figure 29) Collaborative, multi-player interaction with astronomical signal-object field as visual and auditory dataremix instrument (Figure 29(Aesthetic of slowness: Gestures evoked by the responsiveness of the virtual environment recapitulate the scientific and analytical processes of searching for and characterizing signal in context of "noise," and exploration of signal similarity/ difference (Figure 21, 22, 27, 29, 30) Interaction releases data from signal objects and continuously overwrites the virtual environment as participants continue to interact within the virtual space through the data embodied in the object-field (Figure 30) Externally visible, and readable/interpretable physical interactive gestures in-world become performative gestures in physical space (Figure 29)

Discovery	NCBI implementation of BLAST algorithm runs in the installation / exhibition and is visualized at run-time (Figure 31) Genomic calligraphic representation of DNA helix bending and curvature differences from ideal helix bending and curvature along DNA sequences (Figure 18 and 35) Genomic calligraphic representation demonstrates structural and biophysiochemical features along a DNA sequence and distinguishes gene introns from exons, and protein stroke representation encompasses primary and secondary structural features along the protein sequence. (Figure 18 and 35)	BLAST algorithm comparative genomics runs from CAMERA servers returned into *ATLAS in silico* metadata environment elucidates nuanced differences in BLAST results based on GOS metadata. Demonstrates differences between three sequences that are labeled as the same functional protein domain in the context of GOS metadata (Figure 36) BLAST results returned into a 4D virtual environment constructed from metadata and that contains complete target data repository (Figure 36)	Cross-correlation of experimental AST3 robotic telescope data with GAIA DR2 data and known astronomical objects from SIMBAD catalogue to annotate AST3 data and identify putative variable candidate objects (Figure 37) Compute Lomb-Scargle periodogram for all 817,373 AST3 astronomical signal objects, followed by RMS and SNR using published threshold values to identify putative variable candidate objects (figure 30, 37)

Table 1: Summary of dataremix layers and contributions to knowledge within of the works

3.3.1 Dataremix representational elements across artworks

The dataremix representational elements in each of the works (layer one) are new forms of artistic visualization created from informational components. This new knowledge includes genomic calligraphy (West et al., 2005b, 2006; West, Lewis, et al., 2013), algorithmic objects as natural specimens (West, Lewis, et al., 2009), scalable auditory data signatures (Gossmann et al., 2008) and algorithmic signal objects(West et al., 2018; West, Mendelowitz, Thomas, Poovey, Hillard, Wang, et al., 2020). These operate at the artistic end of Kosara's sublimity scale. They function as poetic representational elements for varied dimensions within large and multiscale data and its use in creation of the works.

The genomic calligraphy for *Ecce Homology* is presented in Figure 7 (above), and Figure 16, (below). Figure 7 presents a comparison of the genomic calligraphy for human and rice plant amylase proteins to a three-dimensional model of human amylase generated from the protein data bank. Protein genomic calligraphy strokes are generated from a combination of data features at multiple scales that generate the calligraphic brush profile and features that generate the movement of the brush. The calligraphic markings/strokes incorporate multiple data dimensions that are biophysiochemical features calculated from the sequence along a sliding window. The left-hand profile for the brush is determined by the proportion of amino acid mass to volume calculated along a sliding window throughout the sequence. The right hand profile is determined by the degree of hydrophobicity, or amino acid's tendency to be buried inside the protein. The curvature of the stroke is determined by the percentage of ionizable amino acids along a sliding window throughout the sequence. These are most commonly associated with interactions with other amino acids within a protein. Sequence "chunks" are segmented by a turn prediction algorithm (Kaur and Raghava, 2003). Each segment along the sequence is connected to its neighbor by a connector whose shape is based on a secondary structure property of the segment as computed by the turn prediction algorithm. The length of the protein stroke is determined by the length of the amino acid sequence. Figure 16 presents DNA calligraphic strokes. The brush profile for nucleic acid strokes is determined by nucleotide composition (A, C, T, and G percentages), length of the sequence, directionality of the sequence (designated as five-prime (upstream) or 3-prime (downstream) end of the sequence), GC percentage composition, helix bending and helix curvature. Nucleotide and amino acid strokes are rendered by a custom calligraphic brush model. The brush model is implemented by combining several methods from Sousa and Buchanan (2000) and Chu and Tai (2002). DNA helix bending and curvature along the strand is computed utilizing the BEND algorithm by Goodsell and Dickerson (1994), GC content (percent along a sliding window) is calculated according to (Vinogradov, 2003), the sequence is obtained from Genbank (NIH, no date).

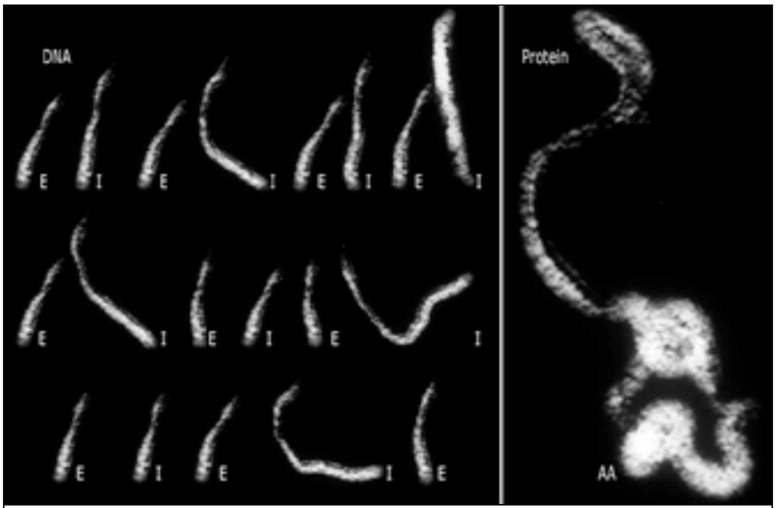


Figure 18: Calligraphic / pictographic representation of DNA and protein sequences from Ecce Homology

Left: Calligraphic / pictographic rendering of the human alpha amylase gene DNA sequence as a series of introns (I) and exons (E)

Right: The amino acid translation of the DNA sequence for human alpha amylase.

Combined, these attributes demonstrate the ability of genomic calligraphy, a novel representational form, to create poetic experiences of big data and data abstraction, and simultaneously present primary and secondary structural information and biophysiochemical features along DNA and amino acid sequences. The DNA sequence strokes also disambiguate nuanced differences in DNA helix bending and curvature along regions of genes that are specified as either introns or exons. The more curved an intron or exon stroke is, the more it varies from the predicted ideal helix bending and curvature. The left panel of Figure 17 presents intron (I) and exon (E) nucleic acid strokes. This demonstrates new knowledge generated by genomic calligraphy in relation to research question 2, how one might go about artistically informed (scientific) discovery. The meta-shape grammar objects (SGO), and scalable auditory data signatures (SADS) developed in the creation of *ATLAS in silico* are new forms of representational elements for artistic visualization. They create new dataremix strategies departing from those developed in the creation of genomic calligraphy to a far wider range of scales and diversity of data types. All three forms incorporate data that spans informational/molecular scales, biophysiochemical scales, temporal, spatial, physical, ecological, socio-economic and environmental scales. (See *Volume 2*, Data & Mappings section for ATLAS in silico for detailed data dimensions).

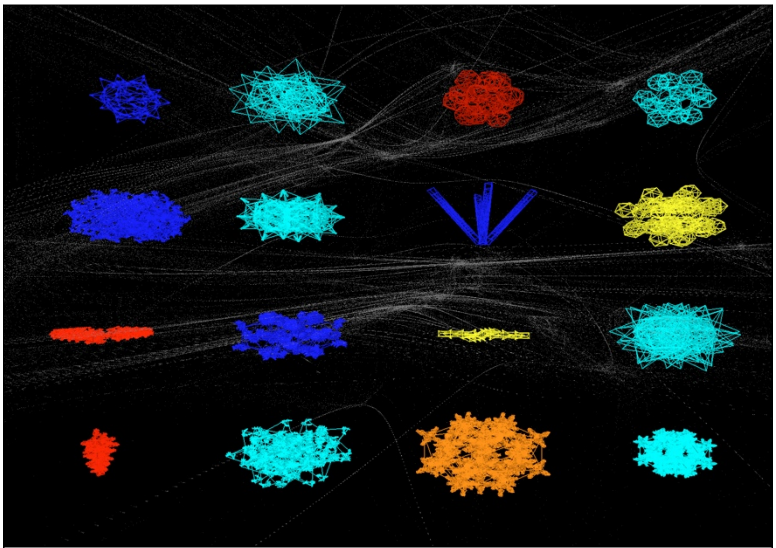


Figure 19: Meta-shape grammar objects (SGO) from ATLAS in silico

Shape grammar objects selected for visual and auditory comparative analysis within ATLAS in silico. Each object contains a unique scalable auditory data signature. Using physical gestures participants "scrub" across the set of objects to magnify the object and reveal its sonic signature. (See video in *Volume 2* ATLAS in silico section and Text section, West, Lewis et al. 2009)

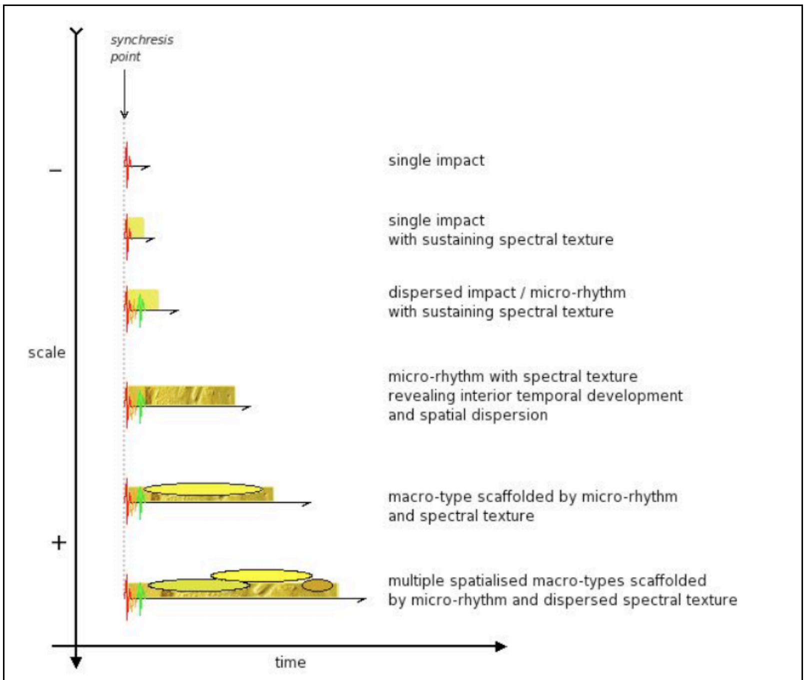


Figure 20: Scalable Auditory Data Signature (SDADS) from ATLAS in silico

Scalable auditory data signatures layer dimensions of sound that map to multiple dimensions of data from the Global Ocean Sampling expedition. See *Volume 2,* texts section, Gossman et al, 2008.

The shape grammar objects are n-dimensional glyphs that represent millions of individual records from the GOS. Each record contains a nucleic acid and amino acid sequence, and contextual metadata. Scalable auditory data signatures are sonic analogs to the shape grammar objects. Both are poetic and aesthetic representations that reflect differences in the underlying data and metadata such that an untrained, non-specialist observer can distinguish differences between visual and auditory objects. This requires an aesthetic approach that generates a large variety of distinctive patterns. The method for shape grammar generation is presented in (West, Lewis, et al., 2009). And the approach for generating scalable auditory data signatures is presented in (Gossmann et al., 2008). These approaches represent new knowledge generated through the artistic practice and are contributions towards the generation of poetic and experiences of big data and data abstraction. The shape grammar objects and scalable auditory data signatures are holistic representations of the relationship between phenotype and genotype, conveying multi-dimensional biophysiochemical information about amino acids and molecular properties of nucleic acids simultaneously in the context of temporal, spatial, physical, ecological, socio-economic and environmental metadata. These algorithmic objects function as natural specimens within *ATLAS in silico's* virtual environment.

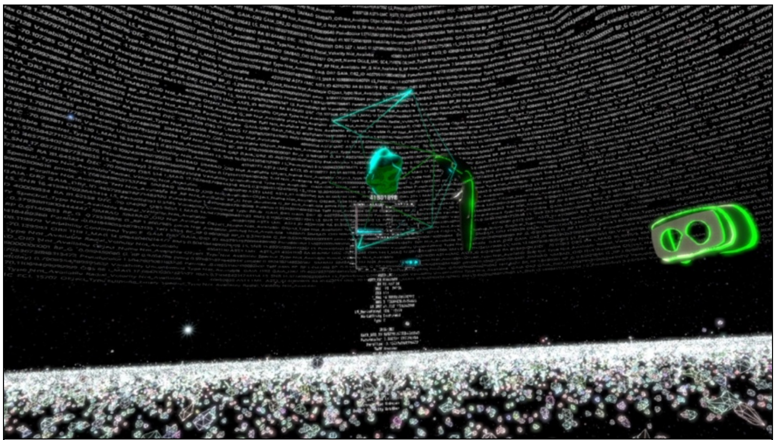


Figure 21: A participant activates an astronomical signal object from the object field within INSTRUMENT | One Antarctic Night

Activation of the object from the field releases its data and sonic signature into the virtual environment. The data released is visible to the participants and is inscribed into the environment to overwrite the virtual world. The sonic elements are activated and modified by participants through their interaction to generate a dynamic soundscape. (See video in Volume 2, INSTRUMENT | One Antarctic Night section)

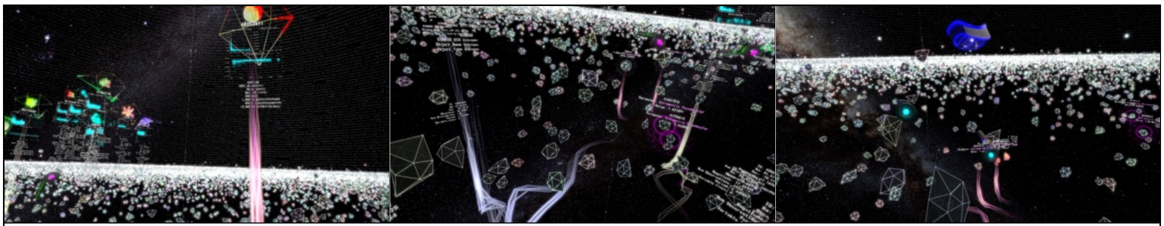


Figure 22: Light trails emitted by signal objects as they release their data into the virtual environment identify objects with similar characteristics nearby other participants

Activation of a signal object from within the field by a participant releases its data visually and sonically into the virtual environment. It also initiates a query across the dataset for an object with similar data attributes as the one activated, but nearby other participants. Those objects are highlighted by light trails that emit from the originating object.

3.3.2 Dataremix Spatiality Across Artworks

Development of these new forms extends the *objective* | *subjective* multi-scale, high-dimensional mapping strategies of dataremix to the creation of new forms of poetic virtual spatiality in which the operative metaphor is “metadata as world”. *ATLAS in silico’s* immersive and generative virtual environment is an abstract visual and auditory pattern. Context is an organizing principle for the environment and the experience through construction of the virtual world from metadata describing the GOS dataset, in addition to additional socio-economic and environmental metadata collected from online data repositories by the author. The virtual environment is both dynamic, and coherently structured. The pattern reveals its characteristics through its disruption as participants interact within the virtual world.

Figure 23: A participant interacting within ATLAS in silico’s virtual environment on the Varrier™ autosterographic display.

The development of scalable metadata environments is presented in (West et al., 2014). The virtual environment contains multiple coordinate systems that sub-divide the unbounded virtual world into regions. Metadata captured at each GOS sample location includes: habitat type, geographic location, sample location, country, latitude, longitude, sample depth, water depth, chlorophyll density, salinity, temperature, pH, start/stop time of collection. To this socio-economic metadata of infant mortality rate per thousand, and internet users per capita, and environmental data of CO2 emissions per capita in metric tons is added for the countries nearest the sampling sites. Each region contains all possible values of the metadata attribute it represents. The virtual environment is placed within a computed fluid force. Each record in the GOS dataset is placed within the environment as a color-encoded particle that is initialized at a random location. It's movement within the virtual world, within the fluid force and in relation to all other particles (database records), reveals the values for all metadata annotations contained within the specific GOS database record. This dataremix strategy resonates quantitative and qualitative data and high-dimensional aesthetic mappings to create visceral and poetic experiences of big data and data abstraction.

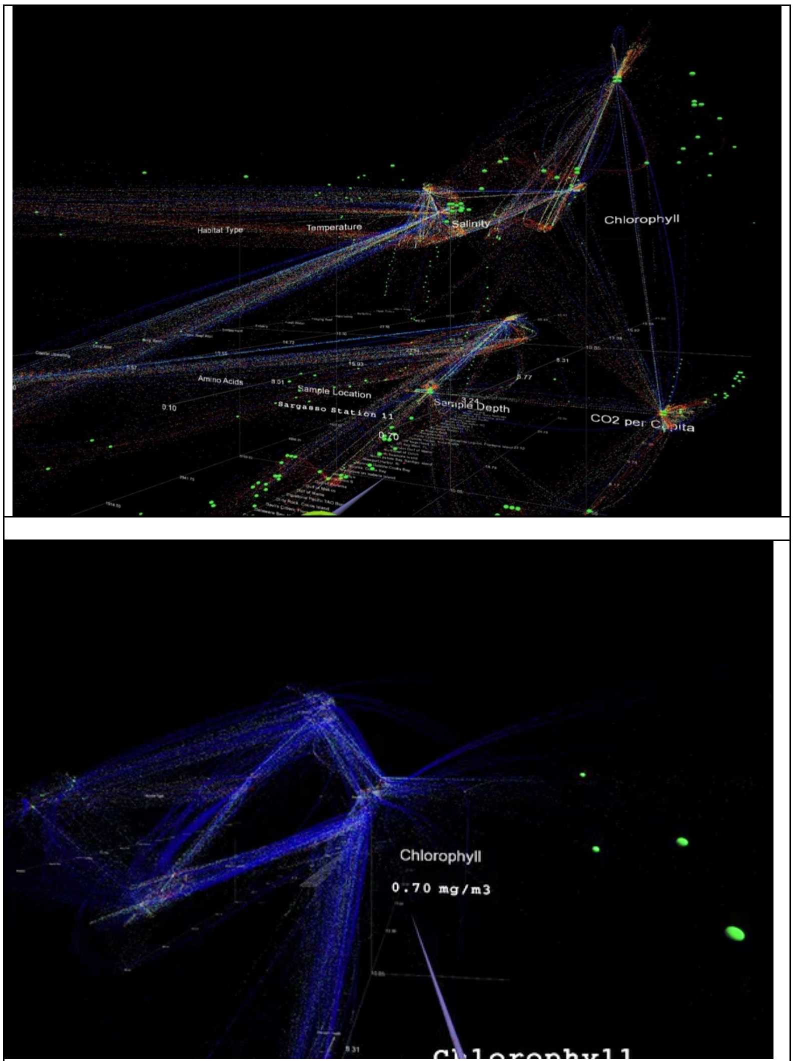


Figure 24: ATLAS in silico's reconfigurable scalable metadata environment

Top: View of the default configuration of the scalable metadata environment for *ATLAS in silico's* virtual environment.

Bottom: View *of ATLAS in silico's* environment reconfigured by a participant during interaction.

Reconfigurable and scalable metadata environments (Figure 24) are contributions to knowledge developed in the creation of *ATLAS in silico*. The non-mimetic, 4D, multi-scale, multi-dimensional and unbounded nature of reconfigurable metadata environments, coupled with multiscale high-dimensional meta-shape grammar objects, and scalable auditory data signatures, create a novel and poetic virtual spatiality that also provide an alternative to cinematic zooms for representing and spanning vastly differing spatial and temporal scales within data.

The virtual spatiality in *INSTRUMENT | One Antarctic Night* is the result of experimentation with aesthetic approaches to the AST3 robotic telescope dataset's size, density and rectangular dimensions. Astronomical data is often presented as a dome that represents the sky as we naturally view it from Earth. The dome configuration is reminiscent of the full hemisphere view of the sky, which embodies the relationship of scale between the human body and the universe – an expanse that is at once infinitely vast and intimate in our individual experience of it. Embodying this sensation of vast disparity of scale within the IOAN's virtual environment is central to the artwork conveying our physical, embodied relationship of scale to the immensity of nature, and the disparity in scale between the body and the abstraction of nature into vast and abstract digital data.

The AST3 dataset from the Large Magellanic cloud (LMC) utilized in the artwork is dense, and contains time series observations for 817,373 astronomical objects. Yet this is still a very small "rectangular slice" of the sky as seen from Earth, and an even smaller "slice" of the Universe.

Figure 25: a0329.160.fit, mjd 56048.0479629 – 817, 373 objects, LMC – image from the AST3 telescope dataset from the center of the LMC

Tests for possible dome-based structures (Figure 26) constructed from the data within virtual reality were unsuccessful in conveying an evocative and poetic experience of the vastness and scale of the universe as seen from Antarctica in relation to the human body. This resulted in seeking an alternate metaphor for presenting and interacting with the data.

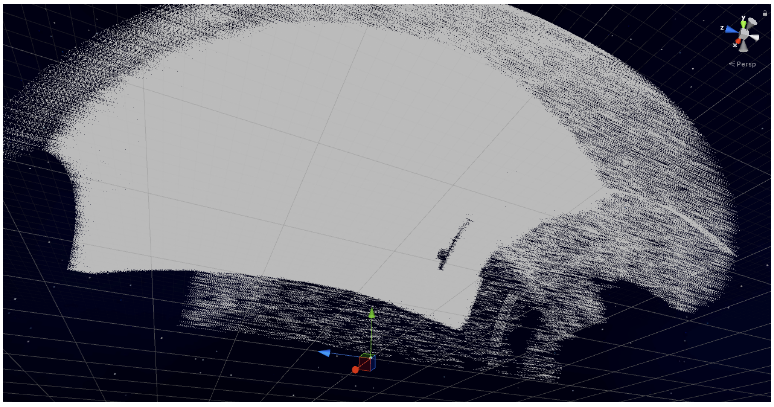


Figure 26: Screenshot from tests to explore a dome-based metaphor and structure. for the data

Top, middle and bottom images: dome-like morphs of the AST3 LMC dataset

There are a total of 4183 time stamps (image frames) in the dataset. Through image registration and partial overlaps 817,373 objects were identified from the 4183 time stamps within the data at LMC center. With an average of 927 measurements per object there are 758,457,209 magnitude readings in the dataset (Table 2 row AST3-LMC). Magnitude is the observed brightness (luminance) of an astronomical object as it is seen from Earth. From the magnitude values, we calculated secondary measurements, such as mean intensity, binned magnitude, and periodicity(Zechmeister and Kürster, 2009) (Table 2 row AST3-LMC-RA,Dec).

Distinguishing whether or not an object's signal is variable, and what type of variable object it is, constitutes one of the research aims for AST3. It is a significant aspect of the work done by astronomers in characterizing astronomical objects. Exploring the AST3 data, in search of evocative potentials a set of 758,457,209 magnitude values, each with a position on a CCD sensor, or galactic coordinate and a time stamp, established the notion of "signal" and "signal object" as a primary metaphor for the project. Objects emit characteristic signals at multiple wavelengths along the electromagnetic spectrum.

What is “signal” and what is “noise” depends on what you are looking for, and the wavelength (region of the electromagnetic spectrum) that you are viewing the Universe through e.g. radio, infrared, optical, x-ray, or gamma-ray, for example. A given object emits a characteristic signal at each of the wavelengths of the spectrum. The notion of “object” and “signal”, and the relationship of the data to the inherent nature of the objects, and how we perceive objects in context of the broader universe as we seek to understand the ultimate nature of reality, became a central theme around which the poetics, generative processes and interaction for IOAN are organized.

Ultimately, a “field” metaphor was selected for the AST3-LMC data. The “field” denotes a spatial relation to the body that is beyond any boundary that we can physically reach, it extends to the horizon, and beyond. It evokes a relationship to nature that is on the boundary between wild nature and cultivated (artificial) nature, such as agriculture.

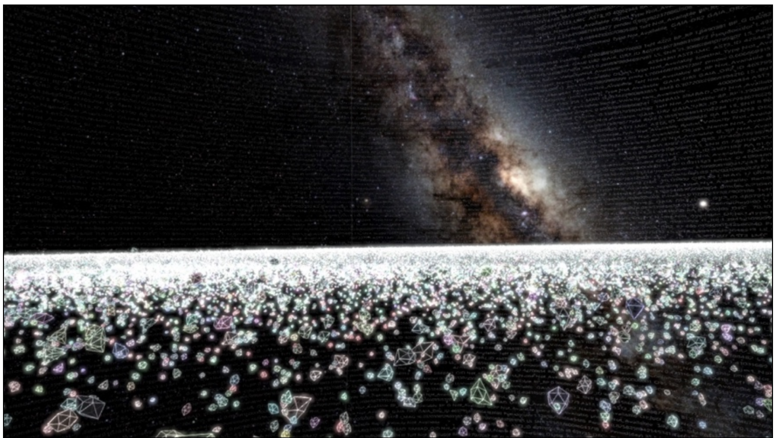


Figure 27: The signal object field within the interactive generative virtual environment for INSTRUMENT | One Antarctic Night

In *INSTRUMENT | One Antarctic Night*, the astronomical data is instantiated as a “field” that one floats within, to explore, activate and perform the data collaboratively as a soundscape with two other immersed participants. The field is comprised of 817,373 geometric objects, one for each astronomical object in the AST3-LMC dataset, positioned in relation to each other according to their galactic (RA, Dec) coordinates. Each object is individually manipulable. Multiple data attributes from an astronomical object are utilized to procedurally render the corresponding geometric object. The astronomical objects are not all stars, some are larger phenomena, such as galaxies, or star clusters etc., yet they are all characterized as a source of "signal" in relation to "noise" within the astronomical analysis of the data. Within the field, and in the virtual environment, each geometric (astronomical) object is a source of signal. The high-dimensional multi-scale mapping creates a novel representational form, the signal object.

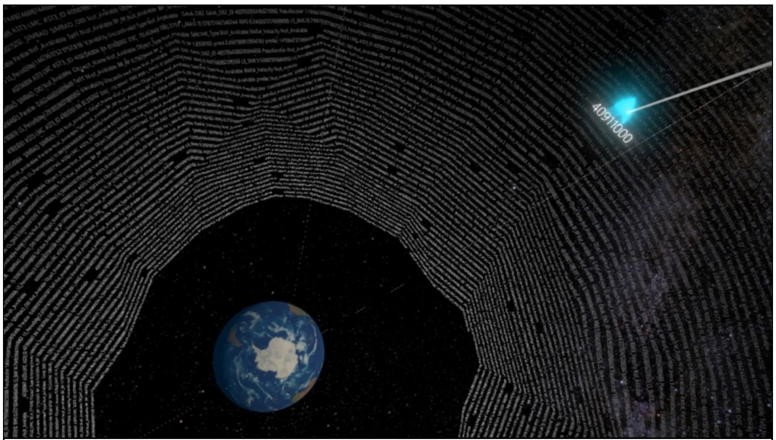


Figure 28: Data released from signal objects overwrites the virtual environment

Performing the data as a generative soundscape releases data from within the activated signal objects. This data streams into the virtual space and overwrites the virtual environment.

AST3 data is mapped visually and sonically into an explorable signal object field instrument designed for collaborative multi-participant interaction. AST3 data, its GAIA/SIMBAD annotations and calculated statistics drive the generative sonic characteristics of the objects and their response to participant interaction.

INSTRUMENT | One Antarctic Night creates a new hybrid spatiality that extends beyond the in-world virtual space, to the physical limit of the VR installation's physical setup, where one can first see and hear the VR hardware. Development of this continuum of experience, that includes transitioning from the physical to the virtual space, synchronous, local-multiplayer interaction, and returning to the physical environment and asynchronous online interaction is presented in (West, Mendelowitz, Thomas, Poovey, Hillard, Hays, et al., 2020).

Figure 29: INSTRUMENT | One Antarctic Night physical-virtual spatiality and continuum of interaction: interactor – spectator – passerby - online

Left: The continuum of experience with fluid transitions between participating as an interactor, spectator, passerby or via remote participation
Middle: Participants collaboratively interacting within the virtual world, as viewed from the spectator zone.
Right: Online and mobile web visual score and signal object field that replays and makes available for remixing and sharing via social media.

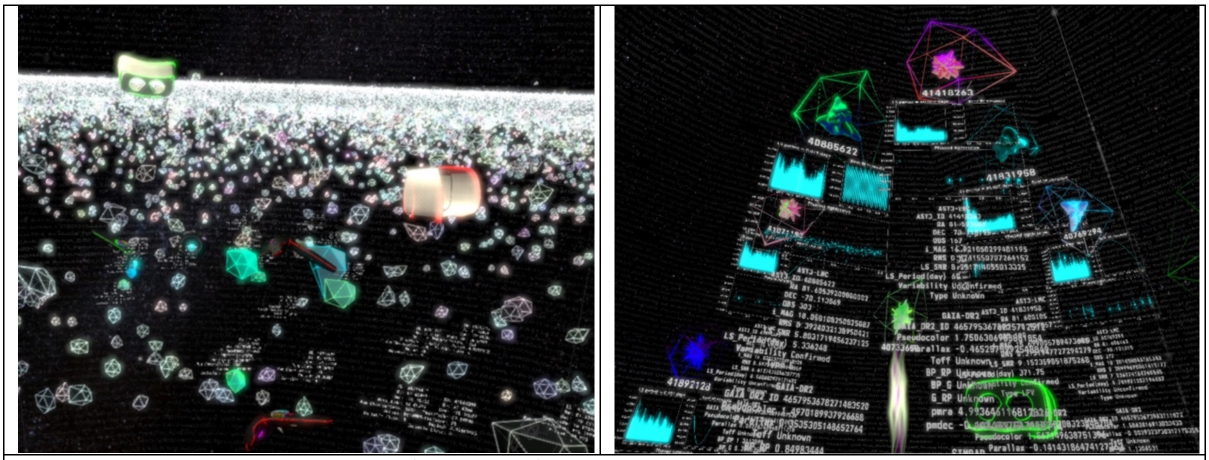


Figure 30: Astronomical signal-object field (aka star-field) as collaborative visual and auditory remix instrument

3.3.3 Dataremix embodied interaction across artworks

The aesthetic of slowness operating within *Ecce Homology,* in response to the "looking fast" aesthetic forces of the BLAST algorithm, creates the opportunity of "looking slow" through the embodied interaction that recapitulates the pattern matching and reveals the operation of the BLAST algorithm within the artwork. This aesthetic of slowness operates in the embodied recapitulation of the scientific processes of the Global Ocean Sampling Expedition within *ATLAS In silico*, and similarly in the recapitulation of scientific processes for the characterization of signals from AST3 data for astronomical objects within *INSTRUMENT | One Antarctic Night.*

Figure 31: Aesthetic of slowness within Ecce Homology triggers a BLAST comparative genomics analysis in the gallery

Figure 32: Interaction within ATLAS in silico on the Varrier™ 100M pixel autostereographic display.

Interaction within the environment generates responsive and data-driven graphics and spatialized audio to create a generative virtual world and soundscape.

Figure 32 presents the recapitulation of the scientific sampling and analytical process for the Global Ocean Sampling Expedition as embodied interaction within the Varrier™ display. In the left panel, the participant reaches into the virtual fluid that is filled with millions of data particles. The data fluid is positioned partially in front of the display surface and partially extends behind the surface of the display system. Participants reach into the fluid to "collect" or "capture" multiple particles, in gestures similar to collecting or sifting sand from water at the ocean's edge. This aesthetic, embodied gestural sampling, parallels the sampling of communities of organisms, and their abstraction into millions of sequences that are used to enact comparative operations seeking structural and functional homology in a new view of life as based on the sequence instead of the species.

The collection of a subset of data particles triggers a state transition in the interaction. Particles from the data fluid enlarge and become more detailed, revealing their structure as meta-shape grammar objects. The objects move to the surface of the display, as shown in the middle panel. They rotate on their individual axis behind and in front of the display surface. Participant movement and gesture reveals the unique scalable auditory data signatures for each object. The experience continues for however long a participant explores and compares the visual and sonic objects. This comparative operation recapitulates pattern matching comparative analytical and bioinformatics operations performed upon the GOS data. Pausing ones movement and remaining stationary at the position of a given object, enlarges that object, and brings it forward, in front of the display, into the tracked volume to envelop the participant's body. All other objects recede. Movement and gesture enable exploration of the object as the data values that generate the shape grammar object and scalable auditory data signature are recited by a text to speech engine. This recapitulates the detail-in-context viewing that is part of analysis of individual sequences. At any point the participant can gesture to return the object into the data fluid and resume their exploration or exit the tracked volume of the display.

The comparative operations enable a view of an individual sequence in the context of the overall diversity of sequences – those sequences supporting life under the given environmental circumstances at the time of sampling. In the selection of an individual sequence, as the analytical operation focuses in on structure and function, reducing and narrowing, the subjective aesthetic operation restores its subjective context. The GOS data is combined with metadata from its physical sampling location, geospatial, habitat, sample depth, temperature, salinity, location, and sociodemographic, CO2 per capita, infant mortality rate per capita, internet use per capita. All made invisible by the reduction of the objective science, as a way of reducing the sublimity of big data through structured framing narratives of analysis, such as clustering algorithms, statistics and BLAST.

Figure 33: Aesthetic of slowness in ATLAS in silico: embodied interaction

Full-body, naturalistic, gesture based interaction with ATLAS in silico on a rear projected stereoscopic preserving projection setup for the installation. Left: the participant is moving to position their body inside and underneath one large shape grammar object. Right: the participant is moving to position their body so as to collect a subset of data particles from the data fluid for subsequent comparative analysis.

Interaction with ATLAS in silico is focused on a single user that is directly interacting with the environment, within the tracking volume in front of the screen. Physical coordinates are translated in to virtual world coordinates and the virtual environment, both graphics and sound, extend to the physical space in the tracked volume to encompass the participant. Multiple participants can stand nearby and view the world and the interactions in stereoscopic 3D.

In contrast, the interaction within INSTRUMENT | One Antarctic Night is designed for three players positioned locally within a room scale tracked area (20-feet by 20-feet) to interact. The collaborative and social embodied interaction recapitulates astronomers recapitulations analytical operations performed to search for and characterize signals from astronomical objects.

Figure 34: Participants' collaborative and physical gestures evoked by the design of the virtual environment become performative to spectators and passersby.

Externally visible, and readable/interpretable physical interactive gestures in-world become performative gestures in physical space

Participants tap or scrub across objects to preview their data and sonic signatures. They pick up and position signal objects to release their data into the world. This generates a SQL query that searches for objects with similar signal characteristics near other participants. This happens simultaneously with the data streaming from the object into the virtual world, and with a light trail emerging from the object. The light trail traverses the signal object field to locate and highlight the object with greatest similarity for the other participant (Figure 22, above). Additional analytical gestures mapped on to aesthetic physical gestures include filtering operations that are activated by participants double-striking the surface of the object field. Participants can exchange signal objects with each other. Signal objects release both data and sound into the virtual world when activated by participants. This creates an evolving sonic composition with layered sonic textures serving as metaphors for the different signal emitted from objects at different wavelengths on the electromagnetic spectrum. The sonic characteristics of the signal objects are data-driven, yet non-hypothesis and non-problem driven mappings of the astronomical data to sound attributes. Objects that are confirmed as variable have structured sonic motifs as compared to the sonic attributes of objects that are not confirmed variables. The interaction is continuous and non-stop when participants change. There is no start or stop to the interaction, reflecting the continuous nature of the signal emitting from astronomical objects. The signal object field, virtual environment and embodied interaction combine to create a poetic experience of big data and data abstraction.

3.3.4 Investigating artistically informed (scientific) discovery across artworks

To investigate how one might go about generating artistically-informed (scientific) discovery, the artworks and contributions to knowledge created through exploration of the first research question, are utilized as experimental systems for exploration of the second research question. This occurs via running experiments using analytical tools from the scientific domain within the process of creation of the artwork and "returning" the results into the artwork. This approach makes use of Borgdorff's principles of artistic research in which artworks function as "experimental systems" (Borgdorff, 2012, p207–208).

Genomic calligraphy for Ecce Homology is created by a custom computational brush model. Ink is deposited in the locations where the computational brush intersects a virtual textured paper surface (Sousa and Buchanan, 2000). Development of the brush model is presented in (West et al., 2005b).

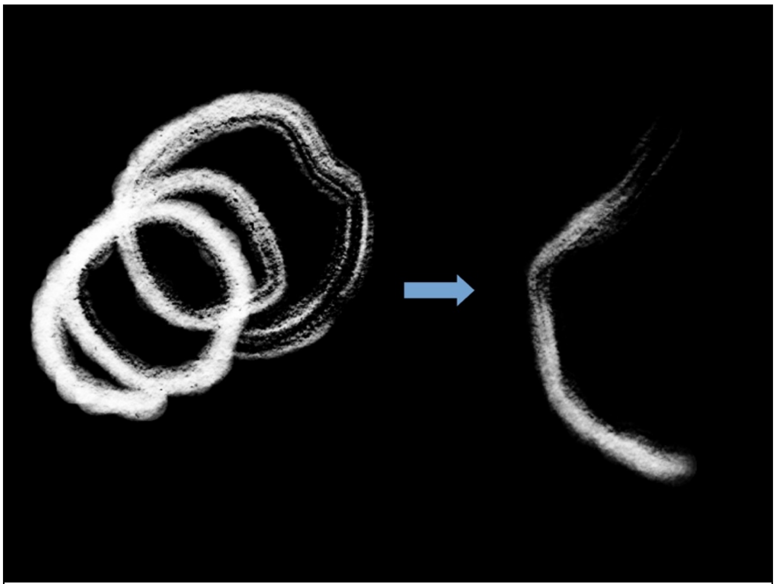


Figure 35: Mapping the output from DNA helix bending and curvature algorithm

Left: Direct mapping of the output of the DNA BEND algorithm onto a DNA calligraphic stroke

Right: The difference between ideal helix bending and curvature and sequence-specific helix bending and curvature along the same sequence as in the left panel, mapped onto a DNA calligraphic stroke. The greater the difference from an ideal helix bending and curvature, the more curved the path of the stroke. Lower differences result in less curvilinear the stroke paths

Through the artistic practice, experimentation with the helical strokes resulting from the BEND algorithm, which predicts ideal DNA helix bending and curvature along a DNA sequence, resulted in contribution of new knowledge towards the second research question. The aesthetic approach adopted resulted creating DNA sequence strokes from the value of the difference between the DNA sequence helix bending and curvature and an ideal helix bending and curvature. This serves to elucidate nuances in DNA helix bending and curvature along DNA sequences specified as either introns or exons. The result of this high-dimensional mapping of the difference between the ideal helix bending and curvature along a DNA sequence to a genomic calligraphy stroke and its ability to elucidate nuanced differences between the intron and exon strokes is presented in Figure 16 (above).

Running BLAST comparative bioinformatics analyses on GOS sequences and returning the results into ATLAS in silico creates new views of BLAST.

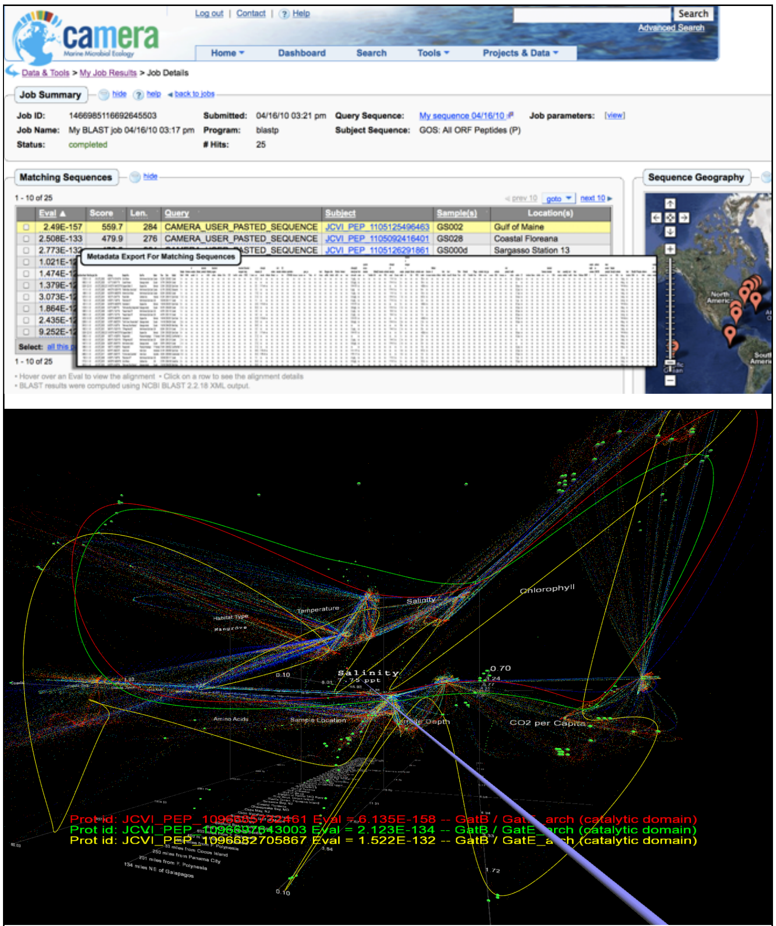


Figure 36: BLAST results for a sequence from the Global Ocean Sampling Expedition are returned into ATLAS in silico.

Top: the online interface for BLAST analyses via the CAMERA metagenomics portal. The analysis returns the BLAST results separately from a spreadsheet containing the corresponding metadata.

Bottom: a BLAST analysis returned into ATLAS in silico's virtual environment. The metadata constructs the environment. The color coded sequence IDs (red, green, yellow) are the top three results from the BLAST survey. The red, yellow and green curves in the virtual world comprise the values of the respective sequences and their metadata attributes within the virtual.

In the top panel of Figure 36, a web-based interface to BLAST returns results to a query sequence from the GOS dataset. The results are ranked and contextual data is returned in a separate spreadsheet. In the bottom panel, the BLAST results are returned into ATLAS in silico's virtual environment. The color coded sequences represent the top three results, with red being the highest ranked, then green then yellow. The results are returned into the world with all of the data particles in the fluid simulation. BLAST results are presented in the context of the entire GOS target dataset. The red, green and yellow curves are positioned in three-dimensions throughout the virtual environment. They are a holistic representation of all values of metadata for each of the color coded sequences presented in the ranked results. This view shows that the top ranked result (red) has metadata attributes more similar to the next ranked result (green) than it does to the lowest ranked result (yellow) and that the lowest ranked result has metadata attributes that differ significantly from those of the red and green (higher ranked) BLAST results. These visible differences in metadata attributes exist despite the fact that all three sequences are labeled as having the same functional designation. This view of metadata in relation to the entire target dataset, and nuanced disambiguation of differences in metadata amongst sequences labeled as identical functional products is not available in the scientific interface (top panel) .

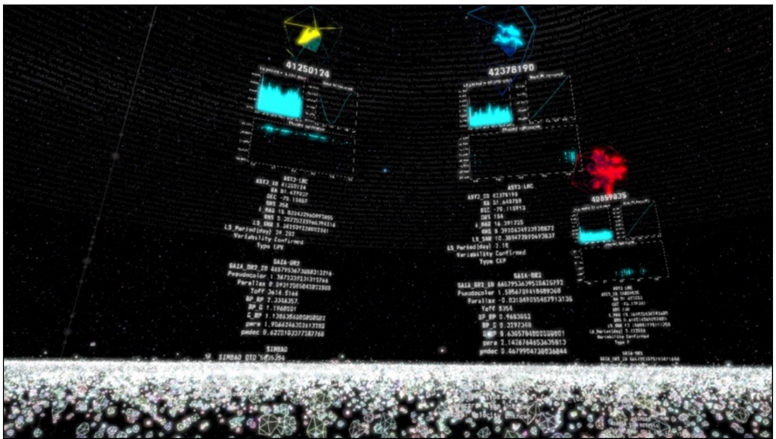


Figure 37: Cross-correlate AST3 experimental data with catalog data from GAIA DR2 and SIMBAD displayed on a per-object basis in-world

To connect the AST3 unpublished scientific data with open astronomical data, it is annotated with additional data by cross referencing to data from open repositories. This enriches the AST3-LMC data with additional dimensions for each object. RA, Dec coordinate queries into the GAIA DR2 archive and SIMBAD database were used to retrieve data for objects within less than 1.5 arc seconds from the queried coordinates[36]. This yields a set of attributes for the AST3 LMC objects, beyond those which were within the time series photometry readings from the telescope (Table 2 rows GAIA DR2 and SIMBAD).

Database	Records	< 1.5 arc second x AST3-LMC	Fields
AST3-LMC	817,373	--	7: ID, ts, x, y, magnitude
AST3-LMC-RA, Dec	817,373	--	24: ID, ts, x, y, ra, dec, magnitude + 17 computed values
GAIA DR2	> 1 billion	745,514	30: ID, ra, dec, + 27 attributes (spectral type, other bands etc.)
SIMBAD	9,235,962	67,311	24: ID, ra, de c + 21 attributes

Table 2: Cross-referenced annotation of AST3 data for the Large Magellanic Cloud

Row 1: data from telescope sensor. Row 2: Celestial coordinates added to the data from the telescope and statistics computed. Row 3: AST3-LMC object celestial coordinates (RA, Dec) were utilized to query the GAIA DR2 catalog and SIMBAD database. Coordinate queries for the 817,373 AST3 objects identified 745,514 matches within less than 1.5 arc seconds within the GAIA DR2 catalog of over 1 billion objects. Row 4: Coordinate queries for the 817,373 AST3 objects identified 67,311 matches within less than 1.5 arc seconds within the SIMBAD database of 9,235,962 records.

In addition to annotating the data with cross-references into public catalogs, a modest set of statistics is calculated for each object. Lomb-Scargle (LS) periodograms (Zechmeister and Kürster, 2009; Lomb-Scargle Periodograms — Astropy v2.0.4, no date) LS periodograms are routinely used in astronomy to reveal periodicity in irregularly sampled data. The data for AST3 is irregularly sampled. Periodicity is used to characterize the variation in brightness of an observed object to determine if the variation is an intrinsic characteristic of the object, or if it is due to other causes, such as effects of the Earth's atmosphere, satellite trails or interstellar dust etc. Some objects have repeating patterns of variability such as binary stars, or pulsating stars, others, such as supernovae have characteristic variability that is not repeating. The LS periodogram is calculated for all 817,373 astronomical objects and returned into the virtual environment as part of the graphical and sonic characteristics of each signal object.

[36] This distance is selected based on guidance from the astronomers collaborating on the project.

3.4 Conclusion

The author's subjectivity in developing the high dimensional mappings that bring together *objective | subjective* elements within each dataremix level (representational elements, spatiality and embodied interaction) creates a multi-faceted sublimity in each work. Each successive layer of the experience accesses an underlying layer, this both reveals, and subsequently obscures the nature of the underlying layer by transforming the nature of the virtual environment. This transformation then at the same time operates as the gateway to the next successive layer of the experience. The mapping of analytical gestures onto physical aesthetic gesture conveys an embodied aesthetic experience that expresses a form of data subjectivity both in the creation of massive data that abstracts nature, and in responding to what Manovich identifies as a "fundamental new dimension of being "immersed in data" (Manovich, 2002, p11). It also engages Kim and Park's (2013) vision for development of new sensoria for big data.

4.0 Conclusion and Future Directions

This PhD by publication discusses the contributions to knowledge in two recent transdisciplinary artistic research projects: *ATLAS in silico*, and *INSTRUMENT | One Antarctic Night* and their exhibited and published outputs. As an exploration of data abstraction, the artworks utilize novel big (scientific) data as the point of departure for the artistic research. A transdisciplinary and collaborative artistic practice, x-resonance, provides a broad framework for approaching the research questions. In the course of the research, a novel methodology emerges, dataremix, which is employed and iteratively evolved through artistic practice to address the research questions: 1) *How can a visceral and poetic experience of data abstraction be created? and 2) How would one go about generating an artistically-informed* *(scientific) discovery?* The artworks and dataremix co-evolve through the artistic practice, exploration, and discovery to embody a creative engagement with the research questions. The contributions to knowledge resulting from this artistic research, and discussed in the thesis, are summarized in this section.

4.1 Conclusion

How can a visceral and poetic experience of data abstraction be created?

The research makes several interconnected contributions to knowledge in the creation of visceral and poetic experiences of data abstraction. In the creation of the artworks, contributions occur through a dataremix approach to *objective | subjective* data and metadata mappings for varied representational elements, the design and construction of virtual environments, and design of embodied interaction within the virtual environments (West, Malina, et al., 2013; West et al., 2015). At the data mapping level within the works, the contributions extend the conceptualization of representational elements for artistic visualization that engages big (scientific) data to include four new forms: genomic calligraphy(West et al., 2005b, 2006; West, Lewis, et al., 2013), algorithmic objects as natural specimens(West, Lewis, et al., 2009); scalable auditory data signatures (Gossmann et al., 2008); and signal objects (West et al., 2018). These scalable representational elements operate at the artistic end of Kosara's (2007) sublimity scale of visualization. The research also proposes an extension to Jevbratt's (2004) forces of looking down and looking in operating in the inverted sublime. To the forces of looking down and in through big data, the research proposes inclusion of the forces of looking fast and looking slow. Looking fast refers to the velocity of big data, to the CPU/GPU compute time scales that generate and algorithmically analyze big (scientific) data, and looking slow refers to making these dynamic and algorithmic processes experiential and observable in human time scales through artistic visualization. This aesthetic of slowness operates in the artworks through affordances for embodied interaction designed to recapitulate scientific data creation and analysis processes pertaining to a given dataset or scientific domain. This aesthetic first arose in the embodied interaction that enacts the algorithmic processes of BLAST within *Ecce Homology*. It is extended in the recapitulation of the sampling and analytical processes of the Global Ocean Sampling Expedition metagenomics data/metadata as embodied interaction within *ATLAS in silico's* immersive generative virtual environment. And it is extended further in the recapitulation of the data capture and characterization of signals from astronomical objects and phenomena by the AST3 robotic telescope from Antarctica as embodied collaborative multi-participant interaction within the virtual environment of *INSTRUMENT | One Antarctic Night.*

The artistic research also extends Corby's (2008b) concept of a new type of *objective | subjective* image consisting of "informational and aesthetic components" to virtual environments and experiences. This new extended form is a poetic *objective | subjective* generative virtual environment and embodied interactive experience that are created from informational and aesthetic components. Through this new form, the research also contributes an extension to Davies' poetic virtual space, and Seaman's embodied Recombinant Poetics within his work *Exchange Fields*. The extensions to poetic virtual space and embodied recombinant poetic interaction operate synergistically with the *objective | subjective* within the artworks. The synergy results in the creation of a novel poetic, non-mimetic, 4D, reconfigurable, abstract and unbounded virtual space created from multi-scale/multi-dimensional big data and metadata: the Scalable Metadata Environment (West et al., 2014). It is developed in the creation of *ATLAS in silico*. This novel poetic virtual environment, originally constructed for/from the GOS dataset and metatdata, has the potential to be generalized and applied to other datasets and in the creation of other works, either by the author or by others. The synergistic operation of the contributions to knowledge also results in a novel physical-virtual poetic spatiality and continuum of interaction that bridges synchronous, local multi-player-spectator-passerby on-site, and asynchronous online interaction in *INSTRUMENT | One Antarctic Night*. This contribution in (West, Mendelowitz, Thomas, Poovey, Hillard, Hays, et al., 2020) is also generalizable and can be applied to other datasets and artworks created by the author or others.

Extension of representational elements for artistic visualization to include new forms, an aesthetic of slowness through embodied aesthetic recapitulation of scientific (objective) processes, extension of Corby's *objective | subjective* image to virtual environments, extension of Davies' poetic virtual spatiality and extension of Seaman's recombinant poetics through embodied interaction holistically combine in the artworks to create visceral poetic experiences of big data abstraction. The integrative synergy amongst the various contributions throughout the development of the research in all the works creates a sensate abstraction (West, Gossmann, et al., 2009). This poetic and artistic approach that strives towards a visceral, aesthetic, and abstract sensuality which can serve as an index to the abstraction of reality as vast and digital data contributes one possible avenue for creating the new sensoria for big data proposed by Kim and Park (2013). This synergy also facilitates the manifestation of the *objective | subjective* through art arising through the transdisciplinary artistic research.

How would one go about generating an artistically-informed (scientific) discovery?

Contributions towards the first research question develop artworks that are visceral and poetic experiences of data abstraction, and which manifest the *objective | subjective* through art. The contributions of new knowledge towards the second research question occur through application of Borgdorff's principles of artistic research in which artworks function as "experimental systems" (Borgdorff, 2012, p207–208). Three contributions to knowledge arise from this investigation, one within each of the artworks as discussed in section 2. These contributions are: elucidating differences in DNA helix bending and curvature along regions of gene sequences specified as either introns or exons, revealing nuanced differences in BLAST results in relation to metadata, and cross-correlation of astronomical data to identify putative variable signals from astronomical objects for further scientific evaluation. Within the artwork *Ecce Homology* as an experimental system, an aesthetic and artistic approach was developed that reveals the difference in DNA helix bending and curvature, from an ideal helix curvature, to sequence-specific curvature, for intron versus exon regions along a gene sequence (Figure 7 section 2 and Figure 16 section 3). This approach was utilized to create the genetic calligraphy used to interactively run BLAST bioinformatics comparative analyses between genes from the human and rice genome in real-time during exhibition of the artwork. The high-dimensional, holistic mapping methodology developed *to create Ecce Homology's* genomic calligraphy was extended in developing the holistic data/metadata mapping and detail-in-context representational schemas for *ATLAS in silico*. Within *ATLAS in silico* as an experimental system, BLAST bioinformatics comparative analyses are performed and the results returned into the metadata environment in the context of the entire GOS dataset. This returning of results within the target dataset and an environment structured from its metadata reveals nuanced differences in BLAST results for homology amongst open reading frames (amino acid sequences) that are identified as being the identical functional domain, yet have differences in metadata attributes. To the author's knowledge, this is a unique instance of returning BLAST results into a virtual environment constructed from the complete target dataset and its metadata that the BLAST query sequence is compared against (West et al., 2014). The "traceback" representational elements developed for this form a part of the embodied interaction that recapitulates scientific processes within the *ATLAS in silico* artwork. Within *INSTRUMENT | One Antarctic Night* as an experimental system, unpublished, unique experimental data from the AST3 robotic telescope on Dome A, Antarctica, is cross-referenced and correlated with data from the GAIA

DR2 catalog release, and published astronomical catalog data from SIMBAD repositories. Lomb-Scargle periodograms are computed along with a modest set of descriptive statistics in order to identify putative variable signals within the 817,373 objects and 758,457,209 magnitude readings that comprise the AST3 Large Magellanic Cloud dataset (West, Mendelowitz, Thomas, Poovey, Hillard, Wang, et al., 2020). This is utilized in a holistic, high-dimensional mapping to create the signal objects, the object field (star field) instrument, the virtual environment, and the collaborative, multi-participant embodied interaction within the artwork.

4.2 Reflection and Future Directions

Undertaking this PhD by Publication has provided an opportunity to reflect on transdisciplinary research and artistic practice that evolved over 10+ years. As is common in reporting results of research, the narrative of performing the research, as told in the context of the structure of scholarly writing, often presents what appears to be an ideal path from motivation, to research questions, hypotheses, methods discoveries and solutions. As a transdisciplinary practitioner the path is often uncertain, and in the process of being constructed as it is being traversed. Analyzing the works and contextualizing their contributions to knowledge offered a view of the terrain that has been traversed. Consolidation of the contributions to knowledge across this body of work brought forward a series of unanswered questions that relate to the untold narrative of the process of the construction of the path. They question the sustainability of transdisciplinary practice in an environment that cannot "count" or "account" for things that resist disciplinary definition. Walking the path over a decade ago began with the question: "Can there be, and how would one go about, artistically informed (scientific) discovery?" This question was, and continues to be, as much a search for an individual resolution as it is an inquiry about the construction of knowledge and meaning.

In regards to future directions for research, the concept proposed by Kim and Park's (2013)of a need to create a new sensoria for big data, so we can work with big data at its full, native resolution, offers avenues for continued future work. This is a natural and timely extension of the work initiated through intersecting the representational, embodied interaction, and virtual space layers of dataremix within the projects discussed in this thesis.

5.0 Appendices

Appendix 1

Materials in Volume 2 provided for context yet not within the scope of the thesis commentary

Artwork

Ecce Homology

Publications related to *Ecce Homology*

West, R. et al. (2013). High-Dimensional Calligraphic Visualization of DNA and Protein Sequences. Proceedings of ACM Eurographics/Expressive Conference Visual Showcase. 2013. Anaheim, California.

West, R. et al. (2005). Both and neither: in silico v1. 0, Ecce Homology. Leonardo, 38 (4), 286–293.

West, R. et al. (2004). Ecce Homology. ACM SIGGRAPH 2004 Sketches. SIGGRAPH '04. 8 August 2004. Los Angeles, California.

6.0 Bibliography

Ackoff, R.L. (1989). From data to wisdom. *Journal of applied systems analysis*, 16, 3–9.

Altschul, S. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25 (17), 3389–3402. Available from https://doi.org/10.1093/nar/25.17.3389.

Altschul, S.F. et al. (1990). Basic Local Alignment Search Tool. *Journal of Molecular Biology*, 215 (3), 403–410.

Arioli, G. and Gazzola, F. (2015). A new mathematical explanation of what triggered the catastrophic torsional mode of the Tacoma Narrows Bridge. *Applied Mathematical Modelling*, 39 (2), 901–912. Available from https://doi.org/10.1016/j.apm.2014.06.022.

Arlander, A. (2020). Diffraction, Mixture and Cut-Ups in Performing with Plants. *Performance Research*, 25 (5), 31–38. Available from https://doi.org/10.1080/13528165.2020.1868835.

Barad, K. (2007). *Meeting the Universe Halfway: Quantum Physics and the Entanglement of Matter and Meaning*, Second Printing edition. Durham: Duke University Press Books.

Barad, K. (2014). Diffracting Diffraction: Cutting Together-Apart. *Parallax*, 20 (3), 168–187. Available from https://doi.org/10.1080/13534645.2014.927623.

Beer, D. (2019). *The Data Gaze: Capitalism, Power and Perception*. 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd. Available from https://doi.org/10.4135/9781526463210 [Accessed 1 November 2020].

Blais, J. and Ippolito, J. (2006). *At the edge of art*. London: Thames & Hudson.

Borgdorff, H. (2010). The Production of Knowledge in Artistic Research. In: Biggs, M. Karlsson, H. and Riksbankens jubileumsfond (eds.). *The Routledge companion to research in the arts*. New York: Routledge, 44–63.

Borgdorff, H. (2012). *The Conflict of the Faculties. Perspectives on Artistic Research and Academia*. Leiden University Press. Available from https://doi.org/10.26530/OAPEN_595042 [Accessed 5 July 2021].

boyd, danah and Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15 (5), 662–679. Available from https://doi.org/10.1080/1369118X.2012.678878.

Brent, R. (2000). Genomic Biology. *Cell*, 100 (1), 169–183. Available from https://doi.org/10.1016/S0092-8674(00)81693-1.

Candy, L. (2011). Research and Creative Practice. In: Candy, L. and Edmonds, E. (eds.). *Interacting: Art, Research and the Creative Practitioner*. Libri Publishing, 33–59.

Carlile, P.R. (2002). A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development. *Organization Science*, 13 (4), 442–455.

Chu, N.S.-H. and Tai, C.-L. (2002). An efficient brush model for physically-based 3D painting. *10th Pacific Conference on Computer Graphics and Applications, 2002. Proceedings.* October 2002. 413–421. Available from https://doi.org/10.1109/PCCGA.2002.1167885.

Corby, T. (2008a). Landscapes of Feeling, Arenas of Action: Information Visualization as Art Practice. *Leonardo*, 41 (5), 460–467. Available from https://doi.org/10.1162/leon.2008.41.5.460.

Corby, T. (2008b). Landscapes of Feeling, Arenas of Action: Information Visualization as Art Practice. *Leonardo*, 41 (5), 460–467. Available from https://doi.org/10.1162/leon.2008.41.5.460.

Corby, T. and Bailey, G. (2005). Cyclone.soc [image]. Available from http://tomcorby.com/cyclone/ [Accessed 10 August 2021].

Corby, T. and Baily, G. (2007). Cyclone.soc: an interactive artwork visualizing internet newsgroup postings as cyclonic weather conditions. *Journal of Visualization*, 10 (4), 339–339. Available from https://doi.org/10.1007/BF03181886.

Corby, T., Baily, G. and Mackenzie, J. (2012). Southern Ocean Studies [image]. Available from http://tomcorby.com/southern-ocean-studies/ [Accessed 11 August 2021].

Corby, T. et al. (2021). The Data Imaginaries of Climate Art: The Manifest Data Project. *Leonardo*, 1–10. Available from https://doi.org/10.1162/leon_a_02136.

Corfield, R.M. (2003). *The Silent Landscape : The Scientific Voyage of HMS Challenger*. Washington, D.C.: Joseph Henry Press. Available from https://libproxy.library.unt.edu/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=101963&scope=site [Accessed 7 August 2021].

Cox, D. (2002). Inspiring the renaissance person. 21 July 2002. New York, NY, USA: Association for Computing Machinery. Available from http://doi.org/10.1145/1242073.1242077 [Accessed 2 February 2021].

Cox, D.J. (1991). Collaborations in art/science: Renaissance teams. *The Journal of Biocommunication*, 18 (2), 15–24.

Cox, D.J. (2008). Astral Projection: Theories of Metaphor, Philosophies of Science, and the Art of Scientific Visualization. University of Plymouth. Available from https://pearl.plymouth.ac.uk/handle/10026.1/1634.

Davies, C. (1998). OSMOSE: Notes on being in Immersive virtual space. *Digital Creativity*, 9 (2), 65–74. Available from https://doi.org/10.1080/14626269808567111.

Davies, C. (2001). Espaces Entrelacés: VR as Poiesis. *Art in Technological Times*. 2. Available from http://www.immersence.com/publications/char/2001-CD-Espaces_entrelaces_SFMOMA.html.

Davies, C. (2004). Virtual Space. *Space: in science, art and society*. Cambridge ; New York, 69–104.

Davies, C. et al. (1998). Ephémère [image]. Available from http://www.immersence.com/ephemere/ [Accessed 14 August 2021].

Delatte, N. (2010). Failure literacy in structural engineering. *Engineering Structures*, 32 (7), 1952–1954. Available from https://doi.org/10.1016/j.engstruct.2009.12.015.

Frayling, C. (1993). Research in Art and Design. *Royal College of Art Research Papers*, 1 (1).

Frazier, M.E. et al. (2003). Realizing the Potential of the Genome Revolution: The Genomes to Life Program. *Science*, 300 (5617), 290–293.

Freeman, J. (2016). Rodent Activity Transmission (RAT) systems [image]. Available from https://culture.theodi.org/rodent-activity-transmission-rat-systems/ [Accessed 12 August 2021].

Frick, L. (2012). Sleep Drawings [image]. Available from https://www.lauriefrick.com/sleepdrawings [Accessed 11 August 2021].

Gantz, J.F. et al. (2008). *The Diverse and Exploding Digital Universe*. International Data Corporation. Available from http://www.idc.com.

Gaviria, A.R. (2008). When Is Information Visualization Art?: Determining the Critical Criteria. *Leonardo*, 41 (5), 479–482.

Gieryn, T.F. (1983). Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, 48 (6), 781–795. Available from https://doi.org/10.2307/2095325.

Goodsell, D.S. and Dickerson, R.E. (1994). Bending and curvature calculations in B-DNA. *Nucleic Acids Research*, 22 (24), 5497–5503.

Gossmann, J. et al. (2008). Scalable Auditory Data Signatures for Discovery Oriented Browsing in an Expressive Context. *Proceedings of the 14th International Conference on Auditory Display (ICAD 2008)*. 2008. Paris: International Community for Auditory Display, ICAD08 1-6. Available from https://aragorn.library.gatech.edu/handle/1853/49951 [Accessed 21 June 2017].

Hansen, M. and Rubin, B. (2002). The Listening Post [image]. Available from https://www.youtube.com/watch?v=Rzfnndd9fCk [Accessed 12 August 2021].

Harris, J. and Kamvar, S. (2005). We Feel Fine [image]. Available from http://www.wefeelfine.org/ [Accessed 11 August 2021].

Helmer, K.G. et al. (2011). Enabling collaborative research using the Biomedical Informatics Research Network (BIRN). *Journal of the American Medical Informatics Association : JAMIA*, 18 (4), 416–422. Available from https://doi.org/10.1136/amiajnl-2010-000032.

Hey, T., Tansley, S. and Tolle, K. (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. Available from https://www.microsoft.com/en-us/research/publication/fourth-paradigm-data-intensive-scientific-discovery/ [Accessed 24 November 2020].

Hughes, P. and Brecht, G. (1979). *Vicious Circles and Infinity: An Anthology of Paradoxes*, 0 edition. Harmondsworth: Penguin Books.

International Human Genome Sequencing Consortium. (2004). Finishing the euchromatic sequence of the human genome. *Nature,* 431 (7011), 931–945. Available from https://doi.org/10.1038/nature03001.

Jankun-Kelly, T.J. et al. (2006). Is There Science in Visualization? *IEEE Visualization Conference Compendium*, 4.

Jevbratt, L. (1999). 1:1 (2). *1:1(2)*. Available from http://128.111.69.4/~jevbratt/1_to_1/index_ng.html [Accessed 10 August 2021].

Jevbratt, L. (2004). The Prospect of the Sublime in Data Visualizations - YLEM Journal. *Ylem*, 24 (8), 2–10.

Kaur, H. and Raghava, G.P.S. (2003). Prediction of β-turns in proteins from multiple alignment using neural network. *Protein Science : A Publication of the Protein Society*, 12 (3), 627–634.

Kim, H. and Park, J.W. (2013). Topics on aesthetic data visualization: viewpoints, interpretation, and alternative senses. *SIGGRAPH Asia 2013 Art Gallery*. SA '13. 19 November 2013. New York, NY, USA: Association for Computing Machinery, 1–7. Available from https://doi.org/10.1145/2542256.2542259 [Accessed 27 November 2020].

Kitchin, R. (2014a). Chapter 8: The Reframing of Science, Social Science and Humanities Research. *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*. 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd. Available from https://doi.org/10.4135/9781473909472 [Accessed 30 June 2021].

Kitchin, R. (2014b). *The data revolution: big data, open data, data infrastructures and their consequences*. Los Angeles: SAGE.

Koblin, A. (2005). Flight Patterns [image]. Available from http://www.aaronkoblin.com/project/flight-patterns/ [Accessed 11 August 2021].

Kosara, R. (2007). Visualization Criticism - The Missing Link Between Information Visualization and Art. *2007 11th International Conference Information Visualization (IV '07)*. July 2007. Zurich, Switzerland: IEEE, 631–636. Available from https://doi.org/10.1109/IV.2007.130 [Accessed 5 June 2019].

Lander, E.S. et al. (2001). Initial sequencing and analysis of the human genome. *Nature*, 409 (6822), 860–921. Available from https://doi.org/10.1038/35057062.

Lapointe, F.-J. (2012). How I became an art[scient]ist: a tale of paradisciplinarity. *SEAD: White Papers*. Available from https://seadnetwork.wordpress.com/white-paper-abstracts/final-white-papers/how-i-became-an-artscientist-a-tale-of-paradisciplinarity/ [Accessed 6 September 2020].

Legrady, G. (2011). *We Are Stardust Interactive Installation*. Available from https://vimeo.com/18802780 [Accessed 10 August 2021].

Legrady, G., Villegas, J. and Burbano, A. (2009). The 'we are stardust' installation. *Proceedings of the 17th ACM international conference on Multimedia*. MM '09. 19 October 2009. New York, NY, USA: Association for Computing Machinery, 1087–1090. Available from https://doi.org/10.1145/1631272.1631516 [Accessed 10 August 2021].

Li, Q. (2018). Data visualization as creative art practice. *Visual Communication*, 17 (3), 299–312. Available from https://doi.org/10.1177/147035721876820.

Lomb-Scargle Periodograms — Astropy v2.0.4. (no date). Available from http://docs.astropy.org/en/stable/stats/lombscargle.html#basic-usage [Accessed 12 February 2018].

Malina, R. (2010). What are the Different Types of Art Science Collaboration. *Roger Malina*. Available from http://malina.diatrope.com/2010/08/29/what-are-the-different-types-of-art-science-collaboration/ [Accessed 6 September 2020].

Manovich, L. (2002). Data Visualization as New Abstraction and Anti-Sublime. 12.

Martone, M.E. et al. (2003). The cell-centered database: A database for multiscale structural and protein localization data from light and electron microscopy. *Neuroinformatics,* 1 (4), 379–395. Available from http://dx.doi.org.libproxy.library.unt.edu/10.1385/NI:1:4:379.

Martone, M.E. et al. (2008). The Cell Centered Database project: An update on building community resources for managing and sharing 3D imaging data. *Journal of Structural Biology*, 161 (3), 220–231. Available from https://doi.org/10.1016/j.jsb.2007.10.003.

Mayer-Schönberger, V. and Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston, UNITED STATES: Houghton Mifflin Harcourt Publishing Company. Available from http://ebookcentral.proquest.com/lib/unt/detail.action?docID=3304666 [Accessed 5 July 2021].

McRobert, L. (2007). *Char Davies's Immersive Virtual Art and the Essence of Spatiality*, 1 edition. Toronto: University of Toronto Press, Scholarly Publishing Division.

Miebach, N. (2019). WEATHER SCORES [image]. Available from https://nathaliemiebach.com/weatherscores.html [Accessed 12 August 2021].

Mobjörk, M. (2010). Consulting versus participatory transdisciplinarity: A refined classification of transdisciplinary research. *Futures*, 42 (8), 866–873. Available from https://doi.org/10.1016/j.futures.2010.03.003.

Navas, E. (2010). Regressive and Reflexive Mashups in Sampling Culture. In: Sonvilla-Weiss, S. (ed.). *Mashup cultures*. Wien [Austria] ; New York: SpringerWienNewYork.

Nicolescu, B. (2014). Methodology of Transdisciplinarity. *World Futures*, 70 (3–4), 186–199. Available from https://doi.org/10.1080/02604027.2014.934631.

NIH. (no date). GenBank. Available from https://www.ncbi.nlm.nih.gov/genbank/ [Accessed 13 August 2021].

Nowotny, H., Scott, P. and Gibbons, M. (2003). Introduction: `Mode 2' Revisited: The New Production of Knowledge. *Minerva*, 41 (3), 179–194. Available from https://doi.org/10.1023/A:1025505528250.

Polli, A. (2016). Particle Falls [image].

Rheinberger, H.-J. (1997). *Toward a History of Epistemic Things: Synthesizing Proteins in the Test Tube*, 1st edition. Stanford, Calif: Stanford University Press.

Rothko, M. and Rothko, C. (2004). *The artist's reality: philosophies of art*. New Haven [Conn.] ; London: Yale University Press.

Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33 (2), 163–180. Available from https://doi.org/10.1177/0165551506070706.

Sack, W. (2011). Aesthetics of Information Visualization. In: Lovejoy, M. Paul, C. and Vesna, V. (eds.). *Context Providers: Conditions of Meaning in Media Arts*. Bristol, U.K.: Intellect, 23.

Scrivener, S.A.R. (2011a). Reflections on Interactive Art and Practitioner Research 1 - Establishing a Frame. In: Candy, L. and Edmonds, E. (eds.). *Interacting: Art, Research and the Creative Practitioner*. Oxfordshire, UK: Libri Publishing, 60–72.

Scrivener, S.A.R. (2011b). Reflections on Interactive Art and Practitioner Research 2 - Interpretation. In: Candy, L. and Edmonds, E. (eds.). Oxfordshire, UK: Libri Publishing, 305–323.

Seagate and International Data Corporation. (2020). *Rethink Data Report 2020*. Seagate. Available from https://www.seagate.com/files/www-content/our-story/rethink-data/files/Rethink_Data_Report_2020.pdf [Accessed 9 June 2021].

Seaman, B. (1997). Models of Poetic Construction and Their Potential Use in Recombinant Poetic Networks. Available from https://www.researchgate.net/publication/237677598_MODELS_OF_POETIC_CONSTR

UCTION_AND_THEIR_POTENTIAL_USE_IN_RECOMBINANT_POETIC_NETWORKS [Accessed 27 January 2021].

Seaman, B. (1998). Emergent constructions: Re-embodied intelligence within recombinant poetic networks. *Digital Creativity*, 9 (3), 153–160. Available from https://doi.org/10.1080/14626269808567120.

Seaman, B. (1999). Recombinant Poetics / Emergent Meaning as Examined and Explored in a Specific Generative Virtual Environment.

Seaman, B. (2001a). Exchange Fields: Embodied Positioning as Interface Strategy. *Convergence: The International Journal of Research into New Media Technologies*, 7 (2), 68–79. Available from https://doi.org/10.1177/135485650100700206.

Seaman, B. (2001b). Oulipo | vs | Recombinant Poetics. *Leonardo*, 34 (5), 423–430. Available from https://doi.org/10.1162/002409401753521548.

Seaman, B. (2003). The Hybrid Invention Generator : assorted relations. *Technoetic Arts: A Journal of Speculative Research*, 1 (2), 103–115. Available from https://doi.org/10.1386/tear.1.2.103.18696.

Seaman, B. (2007). Recombinant Poetics and Related Database Aesthetics. *Database Aesthetics: Art in the Age of Information Overflow*. Electronic mediations. University of Minnesota Press, 121–141.

Seaman, B. (2012a). Four transcultural case studies: Transmedial walks, drives and observations. *Technoetic Arts: A Journal of Speculative Research*, 10 (2/3), 157–165. Available from https://doi.org/10.1386/tear.10.2-3.157_1.

Seaman, B. (2012b). The Engine of Engines – Toward a Computational Ecology. In: Simeonov, P.L. Smith, L.S. and Ehresmann, A.C. (eds.). *Integral Biomathics*. Berlin, Heidelberg: Springer Berlin Heidelberg, 297–309. Available from https://doi.org/10.1007/978-3-642-28111-2_29 [Accessed 11 January 2021].

Seaman, B. (2013). An examination of a specific network of poetics from the realm of language-image-sound relations /.

Seaman, B. (2014). Generative Works: From Recombinant Poetics to Recombinant Informatics. *2014 International Conference on Cyberworlds*. October 2014. 5–11. Available from https://doi.org/10.1109/CW.2014.10.

Seaman, B. (2015). Neosentient Architecture Generator. *2015 International Conference on Cyberworlds (CW)*. October 2015. Visby, Sweden: IEEE, 8–13. Available from https://doi.org/10.1109/CW.2015.55 [Accessed 30 December 2020].

Seaman, B. and Rossler, O. (2008). Neosentience a new branch of scientific and poetic inquiry related to artificial intelligence. *Technoetic Arts: A Journal of Speculative Research*, 6 (1), 31–40. Available from https://doi.org/10.1386/tear.6.1.31_1.

Seaman, W. (2014). A multi-perspective approach to knowledge production. *Kybernetes,* 43 (9/10), 1412–1424. Available from https://doi.org/10.1108/K-07-2014-0145.

Seaman, W.C. (1999). Recombinant Poetics: Emergent Meaning as Examined and Explored Within a Specific Generative Virtual Environment.

Smith, J. (2006). *Charles Darwin and Victorian visual culture*. Cambridge: Cambridge University Press.

Sousa, M.C. and Buchanan, J.W. (2000). Observational Models of Graphite Pencil Materials. *Computer Graphics Forum*, 19 (1), 27–49. Available from https://doi.org/10.1111/1467-8659.00386.

Stalbaum, B. (2006). Paradigmatic Performance: Data Flow and Practice in the Wild. *Leonardo Electronic Almanac*, 14 (07), 8.

Star, S.L. and Griesemer, J.R. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19 (3), 387–420.

Steelman, T.A. et al. (2019). Identifying transformational space for transdisciplinarity: using art to access the hidden third. *Sustainability Science*, 14 (3), 771–790. Available from https://doi.org/10.1007/s11625-018-0644-4.

Stevens, H. (2018). Globalizing Genomics: The Origins of the International Nucleotide Sequence Database Collaboration. *Journal of the history of biology*. Available from https://doi.org/10.1007/s10739-017-9490-y.

The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet. (2007). Washington, D.C.: National Academies Press. Available from https://doi.org/10.17226/11902 [Accessed 25 September 2017].

Thompson, J. and Craighead, A. (2009). Horizon [image]. Available from http://www.thomson-craighead.net/horizon.html [Accessed 12 August 2021].

Vinogradov, A.E. (2003). DNA helix: the importance of being GC-rich. *Nucleic Acids Research*, 31 (7), 1838–1844.

Wakefield, G. and Ji, H.H. (2021). Creative Artificial Intelligence within the Artificial Life Installation "Infranet". July 2021. Available from https://doi.org/10.14236/ewic/EVA2021.34 [Accessed 4 August 2021].

West, R. et al. (2005a). Both and neither: in silico v1. 0, Ecce Homology. *Leonardo*, 38 (4), 286–293.

West, R. et al. (2005b). Both and neither: in silico v1. 0, Ecce Homology. *Leonardo*, 38 (4), 286–293.

West, R. et al. (2006). ACTG to Calligraphy: Genetic Visualization for Ecce Homology.

West, R., Lewis, J.P., et al. (2009). Algorithmic Object As Natural Specimen Meta Shape Grammar Objects From Atlas In Silico. *Leonardo Electronic Almanac*, 16 (6–7), 1–5.

West, R., Gossmann, J., et al. (2009). Sensate abstraction: hybrid strategies for multi-dimensional data in expressive virtual reality contexts. *IS&T/SPIE Electronic Imaging*. 2009. International Society for Optics and Photonics, 72380I-72380I. Available from http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1334749 [Accessed 21 June 2017].

West, R., Malina, R., et al. (2013). Dataremix: Designing the datamade through artscience collaboration. *Proceedings of IEEE VISAP*. Available from http://visap.uic.edu/2013/papers/West_DataRemix.pdf.

West, R., Lewis, J.P., et al. (2013). High-Dimensional Calligraphic Visualization of DNA and Protein Sequences. 2013. Anaheim, California: ACM.

West, R. et al. (2015). DataRemix: Designing the Datamade. *Leonardo*, 48 (5), 466–467. Available from https://doi.org/10.1162/LEON_a_01060.

West, R. et al. (2018). Experiencing a Slice of the Sky: Immersive Rendering and Sonificaiton of Antarctic Astronomy Data. *Electronic Imaging*. 2018. Burlingame, CA.

West, R., Mendelowitz, E., Thomas, Z., Poovey, C., Hillard, L., Hays, K., et al. (2020). Designing a VR arena: Integrating virtual environments and physical spaces for social sensorial data-driven virtual experiences. *Electronic Imaging*. Available from https://doi.org/10.2352/ISSN.2470-1173.2020.13.ERVR-360 [Accessed 5 June 2020].

West, R., Mendelowitz, E., Thomas, Z., Poovey, C., Hillard, L., Wang, L., et al. (2020). Immersive Rendering and Sonification of Large Scale Antarctic Astronomy Data in Virtual Reality: The Making of INSTRUMENT | One Antarctic Night. *Scientific Committee on Antarctic Research (SCAR): Astronomy and Geo-Space Observations from Antractica Session*. 3 August 2020. Hobart, Tasmania, Australia: International Science Council: Scientific Committee on Antarctic Research, 30. Available from www.scar2020.org.

West, R.G. et al. (2014). Scalable metadata environments (MDE): artistically impelled immersive environments for large-scale data exploration. 28 February 2014. International Society for Optics and Photonics, 901205. Available from https://doi.org/10.1117/12.2038673 [Accessed 22 December 2017].

Wolf, G. (2011). Quantified Self: Self Knowledge Through Numbers. *Quantified Self*. Available from https://quantifiedself.com/ [Accessed 5 July 2021].

Woodward, K. et al. (2015). One Sinister Hurricane: Simondon and Collaborative Visualization. *Annals of the Association of American Geographers*, 105 (3), 496–511. Available from https://doi.org/10.1080/00045608.2015.1018788.

Zechmeister, M. and Kürster, M. (2009). The generalised Lomb-Scargle periodogram: A new formalism for the floating-mean and Keplerian periodograms. *Astronomy & Astrophysics*, 496 (2), 577–584. Available from https://doi.org/10.1051/0004-6361:200811296.

Zuboff, S. (2015). Big other: Surveillance Capitalism and the Prospects of an Information Civilization. *Journal of Information Technology*, 30 (1), 75–89. Available from https://doi.org/10.1057/jit.2015.5.