DATA VISUALIZATION: MATERIALITY AND MEDIATION [1]

Sara Diamond

In 1997 I had taken up a performance persona of a zebra an animal that seemed an ideal cathexis for debates about art and science with its clonal mosaic patterned pelt, evolutionary stamina and creature stubbornness. Turing's calculations of patterns and intelligence through reactiondiffusion patterns were inspired by the zebra. The performances embraced and critiqued anthropomorphism, or more appropriately zoomorphism, engaging in ritual clowning in which the zebra could make statements without reprisal about the emerging culture of ArtSci and the desire for bonding between artists and scientists. When Peter Ride commissioned me to build a web site dedicated to ArtSci dialogs, I realized my site needed social media, at that time 'chat.' The inert and linear qualities of IRC and the impossibility of building dynamic dialogs with this tool prompted the decision to create a visual and nonlinear chat tool.

Erving Goffman emphasized the ways that conversation a form of everyday performance establishes a "participation framework" for sociality, an insight borne out by the growth of social media in this decade. [2] I was struck by the emotional (affective) qualities of individual posting and interactions and the ways that emotions appeared to impact participation and relationships. I sought a means to visualize the emotional nuances of, as Goffman puts it, "movements, looks, vocal sounds and speaking." [3] At the same time, social media could act as a collective memory, providing a viable tool for both realtime and archival use. These interests led me to create the CodeZebraOS, an online tool for collaboration, initially between the arts and sciences. CodeZebraOS brought a lateral collaborative structure to the linear hierarchy of chat. The system included casual language games that users could defer to when conversation became heated or overly somber.

The aesthetic for the software selfconsciously imposed a zoological and decorative beauty to online discourse, seeking links between data and their sources in biological sciences, zoology, human physical experience and overall materiality. Lushness was meant to connect the software to the physical world, eventually quite literally, because the larger project created interactive sensor laden, soft wear garments, that responded to the OS dance performances and habituation cages (a joke about Pierre Bourdieu's notions of the habitus) [4] in which artists and scientists were locked up together for twenty four hours under rigorous surveillance and expected to produce new discoveries or theories.

To represent multiple affective states I built eight other distinctive creatures (loosely grouped using a system similar to the Plutchik model of emotions) [5] all with voronoi patterned skins in a relational ecology. The structure of voronoi patterns (conveniently similar to early JAVA prototypes) meant that every time a posting was added the visual chat world readjusted the world view. Extensive user testing showed that grasping a reading of each animal pattern was not intuitive for users, but once they learned how to play with CodeZebraOS, to manipulate and tone conversations, they embraced it.

Invoking reactiondiffusion inevitably led me to considerations of scientific aesthetics that privileged simplicity and symmetry. I knew CodeZebraOS was cognitively challenging but I liked it that way. In examining a wide number of visualizations created by artists, designers or computer scientists, I found on one hand a shared data structure, but also tropes and metaphors. However, there was also a resistance within some scientific discourse to the affective, 'beautiful' qualities of these metaphors. There was a resounding imposition of natural metaphors, even when the source code like mine had no relationship to nature.

Infinitely precise scientific instruments, the digitization of analog material, the explosion of digital storage capacity, processing power and the Internet have, for the first time in human history, combined to produce massive quantities of data. The amount of data produced; which includes consumer buying patterns, text corpora, genomic and cosmological data; poses us with unique challenges. In this era of big data, access to data sources and appropriate tools to analyze data is as important as water and oil. The growth of cloud computing, visual search engines, penetration of fast broadband and wireless networks have all created an ideal environment for an explosion of capacity in data visualization. Data visualization is growing exponentially in scientific, social science and even humanities research, as well as in commercial applications such as social media.

Our expectations of the intelligibility and accessibility of data have shifted with the growth of databases and search technologies. This situation creates an expanding demand for tools and expressions that facilitate finding information and analysis. In the last decade strict prior separations between scientific visualization and information visualization have eroded. Firstly, entire new practices that cross the boundaries of information and science, such as genomics and bioinformatics have come onstream. These fields rely on data visualization to excavate structures in largescale data sets they have no photo realist technologies to fall back on. Secondly, as Lev Manovich has remarked, data visualization allows representations to be mapped onto each other, to compare and overlay vastly different data sets, permitting the representation of infinite permutations and complexity. [6]

For these reasons my university, the Ontario College of Art & Design University (OCAD U), York University and the University of Toronto (UofT), created CIV/DDD, a research hub for the development of nextgeneration data visualization techniques and their underlying information processing and communication technologies; in partnership with communications industry, financial sector, green tech and ITC partners, biomedical researchers and physicists. The range of partners suggests the vast territory of data visualization applications that have emerged. This interdisciplinarity also requires attention to questions of design, aesthetic histories and methods which we have built in the framework of the project.

Aesthetics structure experiences in formal perceptual ways and provide interpretive tools, at times constructing meaning. Given that sensory expression most often visual, but sometimes sonic or tactile is the only means to perceive many contemporary data sets aesthetics are fundamental, not additive to the emerging field of Data Visualization. Data are mathematical: comprised of a set of organized measurements created by instruments that calibrate quantifiable qualities of an original source (natural, artificial or recombinant). Data sets are shaped by prior decisions; such as the instruments chosen to collect the data, the structure of the database, source and sampling methods and software choices. [7] These elements implicate data and put a mediating frame around notions of objectivity. The data source remains present hovering as a ghostly presence and yet is translated by the digital and is always in part lost in translation. Aesthetic realism implies exactitude when there may be ambiguity in the data. Data are a mediation of actual phenomena an immaterial material a contradictory mix of abstract points or numbers and producing phenomena. Data become information only when they are placed into an interpretive context. [8] This requires building algorithms that allow for selection, extraction, organization, analysis and presentation. It is the representational act of transforming both data and the data structure into a visualization interface that allows the user to interact with the data. The resulting images create a bridge between the empirical world and the viewer, revealing patterns of the source data that evoke interpretation.

Visualizations allow the comparison of a set of values; the illustration of relationships between data points; the indication of the parts of a system and the relationship and interaction of these parts; the creation and interpretations of maps; the tracking of change over time; and the analysis of text.

Data Visualization offers the possibility of fundamental new insights, a moment of understanding that reveals hidden processes or complex relationships, breaks through existing barriers and sharpens the focus on knowledge while providing visual pleasure.

RETURN TO REALISM

In an article titled The Petabyte Age (2008), Wired Magazine recently declared The End of Science. Editor Chris Anderson argued that scientists must end hypothesis and experimentation; instead, science must move entirely to data analysis derived from 'big data' sets that lie beyond the natural limits of human comprehension and require "dimensionally agnostic statistics." [9] The more scientists learn about physics and biology, the harder it becomes to create testable models. Instead, Anderson argues, researchers should search for patterns and relate these to the data's source to build a fresh analysis, working with pattern recognition and theorization from the abstract back to the material world. While I agree with his observations on the potential for new discovery, Anderson's view proposes big data as providing the means to rescue science from subjectivity and speculation, and offers a reversion to scientific realism and objectivity, in which data stands in for the real.

The tenants of scientific realism propose that there is a universal shared world of perception that makes up common sense, and discovery manifests this world through shared understanding. [10] The aim of science is to accurately describe reality. The empirical world, including its invisible dimensions and its description through analysis, thus becomes of paramount importance. The rationalist roots of scientific realism suggest that perception leads directly to action, and presupposes the alignment of reality and image. Scientists that I have worked with argue that it is necessary to keep the metaphor close to the look (whether observed or photographic) of the data's source. [11] Yet the aesthetics of scientific realism may create limits to imagination, tying visualization too tightly to "analytic reasoning," which could fail to deploy the transformative power of visual experience. [12] A further challenge to realism occurs when the source cannot be seen, only measured and then imagined.

Debates regarding scientific realism have included some recognition that the observer, whether an instrument or a human, has an impact on the means of expressing the data for an experiment. Karen Barad, a physicist and a philosopher of science, observes in Meeting the Universe Halfway that there is not a onetoone relationship between the ontology of the world and its discovery, as is claimed by "the traditional realist." [13] The 'commonsense' view of Nature is continually entangled in the theoretical and experimental practices that mark its description, as is human society. Yet science still makes meaning of the sometimesinvisible material world and we must pay equal attention to empirical research, as it produces ontological knowledge. These observations are equally true when considering largescale information systems that are hybrid forms of physics, engineering, human and machine interaction.

The field of Data Visualization is compelling because it carries the traces of the empirical world and its instruments of measurement and representation. We need to think of data visualizations as technologies. Data visualizations carry with them the aesthetics and assumptions of their contributing technologies. Data visualization technologies absorb aesthetics of 2D and 3D graphics and animation systems, with their formal styles and malleability. In the past decade a new set of graphics tools some viable for online visualization, others only available through high performance computer networks or in the laboratory are available, as either open source (such as Processing) [14] or proprietary software. The more finished the tool, the more that styles and capacities are embedded. Artists, designers and computer scientists continue to build and adapt tools to their specific needs. Mashup techniques transcode data from more than one source within single integrated tools and search engines allowing the ability to mix what were once discrete structural approaches to data types. Each new source of data adds its structure, aesthetic properties and limits. Whitelaw dubs these practices "databending," [15] as they layer contexts and can allow for the emergence of new imagery or meanings.

Not surprisingly, the most illustrative scientific visualizations draw from or map onto photographic imagery. An example might be an image of a 3D model of a virus structure in which six different proteins are interacting in complex ways. The data was captured using electron microscopy. The visualization is built in Chimera, a C++ and Python software package built to assist in molecular graphics. [16] Scientists had already discovered the symmetrical structure of the virus and had faint images of its form — the

visualization is built on top of the image that the microscope captures through 3D modeling. The visualization extends scientific knowledge by allowing the user to manipulate the virus to understand how its multiple layers might interact and penetrate cell walls. Layering, color and interaction experience were key aspects of the aesthetics that designers brought into play.

The next images represent 3D vector field texture based volumetric flow visualizations of tornados. [17] While the application is specific to storms, the algorithm, data structure and metaphor are of value to multiple disciplines that study flow such as mechanics, physics, meteorology, medicine and geology. [12] Advances in texture mapping graphics capabilities makes these images possible and combines with depth sorting, illumination, haloing and color attenuation to enhance perception and depth. [13] The images are aesthetically compelling, drawn by the computer from the data points and able to capture the dynamics of a storm.

The next image is a 3D visualization of a solar storm that occurred on Halloween 2003. NASA's visualization laboratory created it by combining a model of the earth with "daily averaged particle flux data from the SAMPEX satellite by propagating the particle flux values along field lines of a simple magnetic dipole." [18] By making flux and field lines visible, it is meant to illustrate the ways that energy particles from the solar storm transformed the structure of the Earth's radiation belts. Design decisions are apparent in this quote, "The colorscale on the cross section is violet for low flux and white for high flux. The translucent gray arcs represent the field lines of the Earth's dipole field." [19]

These examples bear close resemblance with what science has previously discovered or represented through photographic media, yet each seeks to extend that knowledge in speculative ways through design decisions that move the image into a field of visual analysis. These images operate within a tradition of scientific description as tools for deductive reasoning.

The use of literal metaphors in data visualization may suggest a level of accuracy that is impossible to achieve. After all, a visualization of Internet packets is many degrees of separation away from the conditions of production of that packet and of its producers. Science itself contains variant views of reality and its analysis is contradictory and chaotic, with different worlds episteme and ontology side by side. New trends in science acknowledge phenomenology, complexity theory and emergence. There is recognition that complex systems are difficult to predict and represent. A further challenge to realism occurs when the source cannot be seen, only measured and then imagined. The strength, not the weakness, of data visualization is its ability to use algorithms to present emergent properties and different points of view.

Case studies of visualizations working with the same data set underscore how data sets are shaped by prior decisions, such as the instruments chosen to collect the data, the structure of the database, source and sampling methods and software choices. Metaphors need not be literally tied to the data structure

to be meaningful, as the following interpretations of two sets of data that analyze emotional expression represent, firstly in components of the Globe and Mail, Canada's national newspaper; and secondly in online service dialogues of Canada's banks. This work is an extension of my earlier research, a cross disciplinary team coled by computer scientists Anjun An and Nick Cercone; myself and Fanny Chevalier as visualization researchers, are creating extraction algorithms, an analysis system, and set of visualizations with the work of student researchers Symon Oliver, Guia Gali and Jarrod Wilson.

The digital version of the Globe and Mail, like its print edition, includes a wide variety of sections, ranging from News, to Business, to Life and Sports. In going digital the publication has added commentary in the form of opinion blogs by its core of writers, as well as ample opportunity for readers to vote and comment. Editorial and business leadership at the Globe and Mail see value in understanding the emotional content of their various sections, writers and readers, and how readers might use tools to choose articles. Editorial leadership is eager to better manage the means for reader commentary. Discovering sentiments, patterns and relationships embedded in articles as well as comments is important for tracking the newspaper's role in shaping public opinion on contemporary issues and the ways that readers interact with these opinions. It can help media analysts better understand the impact of sentiments on news events. Our research analyzes emotional expression in all of the critical features of the publication. We compare sections, blogs and articles and look for underlying patterns in emotional expression using data extraction and visualization. We are using a unique contextbased approach in identifying the underlying emotional tone of text by combining machine learning, semantic analysis, and computational linguistics techniques. The system is being implemented and evaluated on annotated data sets, such as blogs and user comments. The bank data provides an analysis of users' ratings and comments of comparative service offerings for Canada's major banks.

The point is that formal strategies and metaphors differ radically in all examples. Each offer different models of interactivity, whether passive viewing, building one's own visualizations, adding one's own data, or flowing data through the metaphor. Gordon Kindlemann proposes that the very power of data visualization is that objective and subjective views cohere, inspiring new insights. [20] These works prove that very point.

UTILITY AND BEAUTY IN DATA VISUALIZATION

Given that data visualization can assist fundamental discovery, or influence social policy and economics, it is no surprise that in the past the application of data visualization has been motivated primarily by the teleological (willful thinking and predictable outcomes), with regard to both the goals of human activity and the ways that machines or tools can serve these. Visualizations are understood as utilities, translating data into meaningful communication that can represent reality and increase productivity. Edward Tufte proposes that data visualizations are "complex ideas communicated with clarity, precision

and efficiency." [21] Perceptions about realism, common sense and the ways discovery and insight occur have a direct impact on notions of beauty in data visualization.

Despite examples of aesthetically demanding yet instrumental work, such as that of Jer Thorpe, a belief in realism and objectivity leads some scientists to suggest that attractiveness is equal to subjectivity or illegibility. Ben Mathews assigns aesthetics to functionality and ease of interface use. Rapid comprehension is then the goal of this design aesthetic. [22] Simplicity is closely aligned to Occam's razor or lex parsimoniae; the mathematical and scientific view that the simplest solution is the best and most common sense one. Aesthetics are biased towards the symmetrical and highly legible with a spare Modernist look.

The acceptance of beauty in scientific and informatics imagery differs between generations and types of science and designers. As visualizations of the truly imperceptible nanotechnology world proceeds, imagemaking becomes generous. Andrew Vande Moere maintains the blog infosthetics.com, one of the key sites for debates about the practice of data visualization. He argues for lush images, "The best works are those where the aesthetics help people understand the data, where they're almost telling a story." [23] Beautiful visualizations compel not only experts, but also the public convincing them, for example, to adopt energyefficient practices. Legibility, instrumentality and beauty need not be discordant.

Yet currents of thought in art and design argue for the data visualization practices of both fields to be separated. Caroline Ziemkiewicz and Robert Kosara propose differentiation between pragmatic data visualization that allows efficient reading of data, and artistic data visualization that uses data in abstract or metaphorical ways. [24] Kosara feels that creative interpretations can "hurt perception" when fast analysis is needed yet can result in "sublime" or "contemplative" experiences at other times. [25] Mitchell Whitelaw (who has a terrific piece in the exhibition) argues that artists should not allow their data visualizations to become designs, that is, an "aestheticized (and perhaps functionally impaired) form of scientific data visualization." [26] These positions legislate a separation between a teleological usevalue and an intrinsic aesthetic value. Extracting meaning and insight from these representations of data requires powerful aesthetics that balance emotion (such as awe), contemplation and deep analysis.

Recent HCI studies such as those of Noam Tractinsky, A. S. Katz and D. Ikar demonstrate that users pay greater attention to beautiful images and that usability and beauty are viable companions. Jer Thorpe and Christian Flaccus's collaboration tree.growth.redux, from 2011, suggests beauty and visual clarity can align.

This discussion segues into aesthetic discourses about the sublime and the uncanny. Art and literature define the sublime whether nature or immense artificial systems as the threatening unknown that cannot be fully grasped by human understanding. Sublime imagery seeks transcendence, elevating the everyday to godliness. 'Raw' data stands in for nature (red in tooth and claw) and nature is extended to

the vast information web that constitutes the Internet and digital information. Data can be perceived as primary material — not produced concrete and objective, rather than contingent and relational. As I indicated, in the last decade of data visualization a predominate use of natural metaphors (treelike, floral, etc.) appears in art, design and programmercreated works, regardless of whether the source of the data was natural or artificial. The prominence of natural metaphors may indicate the merging of scientific and information visualization; it may represent mystification, the correlation of sublime nature and sublime data or an ironic stance towards mystification; it may suggest a growing sense of concern about the biological world, its extraction into data and the need for an ethos of responsibility towards the empirical world. Issues of aesthetics and ethics are present, if not visible, in the tools we build and use. Here are two different examples ironic versus earnest use of natural metaphors.

Some artists, such as Lisa Jevbratt, Barrett Lyon, or Christophe Viau shown below [27] describe emergent properties and systems as an evolutionary living force. Viau proposes that, "Morphogenesis art is investigating synthetic life by focusing on geometrical models of growth and pattern formation 'in silico.'" In the tradition of Artificial Life he seeks to build models that are so lifelike that they become life itself.

Jevbratt argues that genetic code melded with computer code signals a new sublime or unknowable, uncanny beauty. Jevbratt intertwines the materiality of data with programming (coding) as a material and conjuring practice. She says,

To write code is to create reality. It could be likened with the production of artificial DNA, of oligonucleotides a process where life is written. Or it could be seen as a more obviously physical act of generating and moving around material, an act that has dimensionality, which is nonlinear. [28]

Coding does act as a means of bringing a virtual world into being through the manipulation of mathematics (and its aesthetics) as manifested through data points and computation. [29] The study of data as a material with distinct properties (mathematical and indexical) must not throw away the constructivist wisdom that has allowed an analysis of the intertwined relationship between knowledge and its mode of production. While human intervention is required to produce meaning from the originating data (e.g. weather patterns, plant growth, or mobile phone use), the transformation process should not return to romantic notions of alchemy, affected only by a cognoscenti of programmers, artists and designers. The notion of an unconscious and shared 'natural' aesthetic is a problematic construction, as any survey of contemporary international art practice quickly suggests, art is bound by differentiation. In this view, perception is relational and contextual, constructed through the complex intertwining of object, maker and viewer. [30] These arguments require a located maker and viewer to militate against totalizing notions of beauty. Historical references to 'nature,' its relationship to culture and various past expressions, whether domestic chic or formalism, can serve as a double entendre, reminding us of the tension between the ontological and epistemic.

For some artists, an attraction to data visualization stems from the challenge of excavating hidden patterns and structures from the obliqueness of a data set, at times reconnecting these with the social or political conditions of their production. In his 2002 Data Visualization as New Abstraction and AntiSublime, Lev Manovich argues that visualizations of data by artists may create synthetic meaning rather than support mystification. [31] Manovich indicates links between early Modernist abstraction and contemporary artists' data visualization. However, the complexity and form of the structures that artists disclose have changed since the Modernist era, as have the conditions of belief: skepticism characterizes art, not early twentieth century optimism and essentialism. The formal properties of the database are lateral and associative. It privileges the paradigm (perception of the structure or theorization) over a narrative hierarchy.

Understandings of how to treat data as a material play out in the making of visualizations. Two approaches to design arise representing a bottom up/top down process. Edward Tufte argues that data visualization requires choosing data sets that are of value to the researcher, mining the data, creating a structure for the data, analyzing that data set to find meaningful ways to represent it, analyzing patterns, translating the analysis through aesthetic representation, refining the representation to better communicate, and creating means of manipulating the data. [32] In Tufte's view, data enunciate their own structures. There is no base case with data: it is inductive reasoning that pulls out knowledge. Through this process they find form, and sometimes also find metaphor or narrative. This may be viewed as data naturalism, structuralism, bearing a truth to materials approach, or, in working with largescale data sets representing phenomena that cannot be viewed, datadriven design.

Ben Fry proposes a procedure that begins with a narrative or story form. He argues that the designer must start not with the data set but with the empirical question asked by the researcher. Fry then works his way back to data. He considers the nature of the data to be obtained, finds data to fit the question and parses them to provide a structural fit for their meaning, then orders them into categories and filters out all but the data of interest. This approach maintains the role of the scientist in producing theory (a base case), illustrating, testing and deducing. It also offers an opportunity for metaphor, design variation and the recognition of multiple interpretations of the same data set by different disciplines. Both approaches need comparative testing to see how each impacts discovery in the fields where they are applied. In both instances a challenge for artists and designers is to sustain a constructivist understanding of imagery while openly exploring the indexical properties of data.

In developing the research methods for our visualization network we have chosen to contrast these methods.

Data is assembled and extracted > Effective algorithms are produced to represent data qualities over time; to represent data qualities over dimensions; to represent change, as appropriate > Data analysis is

used to create a patternrecognition system and a structure > Analyze the qualities in data that can be represented and choose a metaphor > Protocols for participatory design in relation to datadriven inquiry and visualization are developed and tested for groups > Participatory design exercises with user groups (as above) > Aesthetics are developed and tested that fit the structures and patterns > Dataled visualizations of extracted features/patterns occur > Userled visualizations of features, and aesthetics occurs, comparing data sets > Effective means are sought to create interaction with the data as represented > Modeling methods for complex phenomena across different data sets are applied (social media, biological, physics) > Visualization tools are created that allow the analysis of differentiated data sets and these are compared for similarities and differences > Results are contrasted and compared results > Use and usability testing is undertaken for tools > Selection, adaptation; development of analysis tools > Visualization tools tested with commercialization partners.

CONTEXT AND COGNITIVE SCIENCE

Data visualization requires both the awareness of cognitive aspects of human visual apprehension, such as color theory and the need to make the visualization meaningful to a user's context. In a muchquoted statement, Edward Tufte describes graphical excellence as "that which gives the viewer the greatest numbers of ideas in the shortest time with the least ink in the smallest space." [33] Ware proposes that Data Visualization is the scientific study of "distributed cognition" between pattern mechanisms in the human brain and the algorithms that map data to the computer connecting human cognition, computer memory and its related algorithms, and the physical actions of the user. [34] Indeed, successful design requires attention to the physiology of the brain, hand and eye. However, these formulae describe a mechanism at work in the perception of visualizations but are bereft of understanding the ways that human experience differs from machine, encompassing the nonlinear as well as inductive processes at work. Unfortunately, because of the focus on treating data visualizations primarily as utilities, much cognitive science research in the field studies techniques of performance enhancement, that is legibility and speed, rather than breakthrough discovery or the play of poetics or insight.

Ideas about nature, reality, culture and common sense play out in the field of Cognitive Science. Like scientific realism, cognitive science has gravitated towards a Kantian notion of 'common sense,' which encompasses logic, morality and aesthetics. [35] Immanuel Kant promotes logic, equating it with purposefulness and demotes aesthetic judgment as mere taste. Of more value may be Kant's proposal that aesthetics are a transaction between the artist, the object and the audience, suggesting that the viewer completes the image. This is a process wherein an embodied subject is in constant formation, in a state of 'momentariness' akin to Gilles Deleuze's notion of becoming and allowing insight and awe. [36]

Valuable lessons from cognitive science can help designers and artists to understand the differences between reading and viewing, and the ways that visuals can allow pattern recognition and text can act to lock down meaning and context in visualizations. These lie in parallel to current theories about the

image that reside within visual culture studies, locating aspects of cognition outside of conscious grasp. A cautionary note is required here as well, for poetics makes use of language to create patterns and graphic fonts, such as Forte or Bauhaus 93, or indicate that stylistic signifiers overcome the content they may contain. These boundaries further dissolve in the popular field of text visualization, where semantic and social networking relationships are discovered through visual and textual patterns.

Equally problematic is the tendency of many twentieth century cognitive scientists to universalize perception and cognition. Contrary research from other strains of cognitive science suggests that context and culture effects perception, and that viewers have different experiences in relation to what makes the same data visualization effective. Rather than a normative notion of cognition, Francisco J. Varela, Evan Thompson and Eleanor Rosch draw on evolutionary biology to reject notions of fitness and optimal adaption. They adopt a "proscriptive model" in which diversity is "woven into the basic constraint of maintaining a continuous lineage" and "the evolutionary process both shapes and is shaped by the coupling with the environment." [37]

Hence learning and difference play key roles. Varela, Thompson and Rosch show that because understandings are culturally learned, categories, such as color perception, are not assumed to be objective; hence, "lexical classifications of colour can affect subjective judgments of similarity." [38] This formulation links perception and aesthetic categories together. Such an approach to cognitive science requires a mix of intrinsic and extrinsic factors in understanding the mind and allows a better understanding of cultural diversity. Sensory cognition remains of critical importance in forming judgments, and hence aligns with the need for aesthetics in the field of Data Visualization that take these processes into account. Providing different users with varied metaphors, even shifting color templates in the interface, can allow perception and analysis of the visualization.

Even when taking diversity into account, cognitive science primarily focuses on individual perception, rather than the emergence of hybridgroup experiences and collective identities as result of the new sociality produced by Internet communication. Warren Sack states that "aesthetics for the Internet needs to concentrate on producing the means for visualizing and understanding how social and semantic relationships intertwine and communities and common sense emerge." [39] He observes that new identities overcome cultural difference, although difference is the starting point. Perhaps it is more accurate to state that rather than a new universality, new particular and contingent identities form.

Visualization systems that represent collaborative efforts or discourses require an aesthetic that allows the emergence of common and collectively constructed experiences and identities. It is logical that designs with high degree of interactivity would facilitate the creation of new identities or "intersubjectivities," [40] a term coined by Vilem Flusser, for conjunctures where identities conjoin productively.

INTERACTIVITY AND IMMERSION

Earlier examples have demonstrated degrees of interactivity in data visualizations. Interactivity appears to be an important part of cognitive process, of learning by doing, engaging the body through navigation. The third space that Bruno Latour describes between subject, object and technology is the site of "interactivity, intelligence and creativity." [41] Ron Burnett offers the explanation that part of the power of the "third space" of technologymediated experience for the participant is the opportunity to gain agency by learning the system and aggregating knowledge through play. [42] The same may be true of gaining visual understanding while navigating data sets. This leads to an aesthetics that allows users to exert agency through learning a system and even to adapt and change outcomes.

There are different levels of interactivity within digital media, so are there in relation to data visualizations. Some data visualizations simply provide navigation capacity such as the ability to click on or mouse over material that the user chooses. In 2006, Fernanda B. Viégas and Martin Wattenberg created Many Eyes [43] in order to popularize the use of data visualization and provide a tool kit for building visualizations. They hoped for at least three uses of data visualization: to interpret textual data, to analyze complex objects and to use visualizations to initiate "social data exploration." [44] This is a highly interactive site where participants can add their own data and they or other participants create visualizations from that data from a set of given templates. Users then export their visualizations to their social media sites.

Other forms of interactivity privilege the impact of the information flowing through the site in this instance data acts as an agent interactivity is the flow of data issuing from a stock market feed, a geological phenomenon or a conversation. Stock market feeds have been a fecund source for projects, such as Joshua Portway and Lise Autogena's Stock Market Planetarium depicted below. [45] The elegant and ironic installation plays off the scientific trope and information metaphor of cosmology visualizations, suggesting a new astrological universe of corporations and their stocks, as artificial life creatures mutate, propagate and die in the market, feeding off of its movements and making graphical transitions, clumping and influencing the weight of the depicted universe.

Interactivity and related cognitive processes imply a timebased experience. Navigating 2D and 3D visualizations often requires rapt attention. Building on Deleuze's writings on cinema, Hansen argues for an aesthetics that is appropriate to the temporal experiences of digital media. [46] Digital media create opportunities for humans to experience time and space in ways that stretch and extend their existing physical apparatus. Data visualizations of large and multidimensional data files occur on 3D screens and at times in 3D CAVE environments. These are full body experiences, where the user is navigating data in real time, performing discovery simultaneously or with retrospective thought. Aesthetics is mediated between the body and its object in a continual flow or becoming.

Data visualization can also occur as an illustrative sidebar to highly interactive social media activity. Social media companies commission visualizations to allow users to catalogue their resources and to better understand and organize their relationships with others.

Bruno Latour [47] and John Law and John Hassard [48] describe technologies as invisible nonhuman actors, affecting the performance of a social network or process. Visualization is a compelling strategy for some artists, a less disruptive and more aesthetic means to excavate technological structures that hold hidden hierarchies of power. In data visualization formalism and politicized deconstruction merge, by creating visualizations that reveal sociopolitical relationships within the data. Hence of importance are artists' initiatives to challenge the set of formal techniques and conventions that have emerged linking data extraction methods, structures, metaphors or metonyms. Artists offer a critique of the aesthetic norms of scientific and information visualization. In 1995 Simon Pope and Matthew Fuller created the Web Stalker, one of the first tools to crawl the Web and build a visual diagram of hidden relationships between domains and their hierarchical ordering, discrediting any notion of search engine neutrality. Its form is now common to many data visualization tools in social media.

An End of Modernity by sculptural artist Josiah McElheny is a tenbyfifteenfoot accurate artistic version of the Big Bang, which itself is much more than an explosion: it is the origin of space and time itself, initiating an expansion that occurs everywhere and has no center. A Thousand Points of Light, a visualization by Naeem Mohaiemen as part of The Disappeared in America Project by the Visible Collective/DanBergman is an animated map of mass detentions that occurred in the United States after September 11, 2001; providing information about the detainee and their country of origin. Viewers can update the map with their own data. Such attempts to enforce transparency onto technoculture and offer an overt critique of power relationships may be described as data deconstruction.

In The Secret Lives of Numbers (2002, 2008) Golan Levin and his collaborators Jonathan Feinberg, Shelly Wynecoop and Martin Wattenberg seek an understanding of which numbers reoccur more than others, "in order to determine the relative popularity of every integer between 0 and one million" [49] and to surmise about why this takes place; as well as finding links to the functioning of human memory, social rituals and the structure of commerce. In the face of our society's belief in the objectivity and power of mathematics Levin instead argues for the subjectivity of numbers and, by implication, data, stating:

Humanity's fascination with numbers is ancient and complex. Our present relationship with numbers reveals both a highly developed tool and a highly developed user, working together to measure, create, and predict both ourselves and the world around us. But like every symbiotic couple, the tool we would like to believe is separate from us (and thus objective) is actually an intricate reflection of our thoughts, interests, and capabilities. [50]

He is also playing with the conventions of data visualization, drawing on Edward Tufte and Colin Ware's rules of simplicity of display to comment on the aesthetic and practices of scientific visualization, and at the same time develop a malleable, beautiful and interactive visualization from data sets pulled from a wide range of search engines over a fiveyear period.

One of the most dynamic growth areas of data visualization is text visualization, whether the massive quantities of scientific texts, social media output, chat, or descriptive metadata. Artists with an interest in linguistics and conceptualism now turn to Data Visualization as a digital trajectory to linguistic intervention, semiotics and conceptualism.

Temporal structures define how textbased relationships emerge in the Internet, with synchronous and asynchronous experiences providing very different feelings, intimacies, and forms of consciousness. These pile on top of each other in layers, allowing social relationships and expressions to feel like a thick texture of condensed time. We Feel Fine by Jonathan Harris and Sep Kamvar bears an interest in affective expression and uses the measurement of text data to find it. [51] We Feel Fine builds emotional portraits of specific online populations by extracting expressions of feelings from Weblogs. The project provides six movements (like a symphony), driven by statistical analysis and data aggregation, and then reshaped by users' paths through the data. Feelings accumulate in mounds on the screen, quivering when the mousecursor passes over. The site is poignant and amusing.

Other artists are drawn beyond structural analysis to poetics. Data visualization becomes a means to write concrete poetry. In 2003, Brad Paley created Textarc, a tool that allows text to be processed; where key words are quantified and brought to the foreground. It has been applied to literature, bodies of conference data, calendars and other corpus. Stephanie Posavec explores differences, "in writing style between authors of modern classics" through her project Writing without Words. [52] An example is shown below.

Posavec parses text in an expressive and poetic manner to create works such as Sentence Drawings and Sentence Length as part of her series.

To conclude: data visualization aesthetics are always contextual, depending on the data source and equally, are read in context, whether by scientist, social media user or art audience member. Data do not exist in themselves, and data risk mystification. Many of the sources of data are already structured. The assumptions of these structures, the material impacts of underlying technologies and particular software and the pervasive presence of tropes and metaphors need continual unraveling.

Cognitive science has a key role to play in developing visualization aesthetics that privilege pattern recognition. Viewing complex 3D images and navigating through these requires eyehand coordination

and focused perception. Designing to facilitate, or at times disrupt, cognition requires that artists or designers draw from this body of knowledge. At the same time, cognitive science needs to recognize that visual expression carries with it the aesthetics and aesthetic traditions of its source technologies and the subjectivity of visual images at the most fundamental level.

Data visualization approaches deriving from the art world are of value in their own right, producing compelling works of art, and valuable as a means to raise new questions and approaches to data. Art's deconstructive tendencies are helpful in unfolding assumptions that are built into data collection and structure. Experimental, abstract, multidimensional, highly interactive works can be immersive and provocative; perhaps more so than simplified visualizations that are illustrative of prefigured assumptions. Aesthetics that can evoke and provoke other disciplines, yet draw from the formal and critical values of art, bear great promise. This is a field where art and design practices can be engaged in multiple layers of discovery of new forms of expression and of new realizations in the fields that are aligned with the source data be these genomics, physics, economics, or information theory prompting insights and reflection uncommon to contemporary practices of data visualization.

References and Notes:

[1] Some components of this lecture have been previously published as "Lenticular Galaxies: the Polyvalent Aesthetics of Data Visualization," Code Drift, http://www.ctheory.net/articles.aspx?id=651 (accessed September 2011).

[2] Erving Goffman, Forms of Talk (Philadelphia: University of Pennsylvania Press, 1981), 3.

[3] Ibid., 3.

[4] Pierre Bourdieu, Outline of a Theory of Practice (UK: Cambridge University Press, 1977).

[5] Richard Plutchik, The Emotions: Facts, Theories and a New Model, Revised Edition (London: University Press of the Americas, 1991).

[6] Lev Manovich, "Data Visualisation as New Abstraction and AntiSublime" (unpublished Paper, Berlin, 2002), http://www.manovich.net/ (accessed January, 2009).

[7] Evaluations of data visualizations should raise concerns about the quality of 'source' or 'raw' data, and challenge the assumption that once the data have been 'cooked', that is, digitized and standardized, they guarantee accuracy. Sara Diamond and Susan Kennard, Banff New Media Institute Web Archives (Banff: Banff New Media Institute, 19932009).

[8] Mitchell Whitelaw, "Art Against Information: Case Studies in Data Practice," in Proceedings, Fibreculture, ed. Andrew Murphie, (Perth: Digital Arts and Culture Conference, 2006), 2, http://journal. fibreculture.org/issue11/issue11whitelaw.html (accessed March, 2010).

[9] Chris Anderson, "The Petabyte Age: The Power of Big Data," Wired, July, 2008, 1413.

[10] For a thorough overview of the history of scientific philosophy, see Peter GodfreySmith, Theory and Reality: An Introduction to the Philosophy of Science (Chicago: Chicago University Press, 2003).

[11] P. Boulanger, S. Diamond, and T. Erickson, "A Boom with a View" (keynote lecture, Powering Innovation Conference, Toronto, ORION/CANARIE, 2008).

[12] Fernanda B. Viegas and Martin Wattenberg, Artistic Data Visualization: Beyond Visual Analysis (Unpublished paper available from Cambridge, Massachusetts: Visual Communication Lab, IBM Research, 2005), 2.

[13] Karen Barad, Meeting the Universe Half Way: Quantum Physics and the Entanglement of Matter and Meaning (Durham and London: Duke University Press, 2007), 41.

[14] Casey Reas and Ben Fry, Processing: A Programming Handbook for Visual Designers and Artists (Cambridge, MA: The MIT Press, 2007).

[15] Mitchell Whitelaw, "Art Against Information: Case Studies in Data Practice," Proceedings, Fibreculture, ed. Andrew Murphie (Perth: Digital Arts and Culture Conference, 2006), 12, http://journal. fibreculture.org/issue11/issue11whitelaw.html (accessed March 2010).

[16] T. Ferrin, "Interactive Protein Structure Visualization," http://www.apple.com/science/ insidetheimage/ferrin/ (accessed March, 2010).

[17] Sung W. Park, Brian Budge, Lars Linsen, Bernd Hamann, & Kenneth I. Joy, "Dense Geometic Flow Visualization," Eurographics IEEE VGTC Symposium on Visualization, eds. K.W. Brodie, D.J. Duke, K.I. Joy (2005), 18.

[18] Tom Bridgman, James W. Williams, and Greg Shirah (animators), Daniel Baker, and Shrikanth G. Kanekal (scientists), "Earth's Radiation Belts Tremble Under the Impact of an Electrical Storm: Halloween 2003 Solar Storm," 2004, http://svs.gsfc.nasa.gov/ (accessed March, 2010).

[19] Ibid.

[20] Gordon Kindlemann, "Is There Science in Visualization?" IEEE Compendium IEEE Visualization, in T.J. JankunKelly, Robert Kosara, Gordon Kindlemann, Chris North, Colin Ware, and E.Wes Bethel (transcript, London, IEEE, October, 2006), 4, http://www.cse.msstate.edu/~tjk/ (accessed December, 2009).

[21]Edward Tufte, The Visual Display of Quantitative Information, 2nd ed. (Cheshire, CT: Graphics Press, 2001), 51.

[22] Ben Matthews, "When stakeholders represent others in design conversations" (paper presented at Nordscode Seminar, Lyngby, Denmark, April 2830, 2004).

[23] Andrew Vande Moere, Information Aesthetics Weblog (2008), 3, http://infosthetics.com/ (accessed December, 2009).

[24] See Caroline Ziemkiewicz and Robert Kosara, "The Shaping of Information by Visual Metaphors," Transactions on Visualization and Computer Graphics 14, no. 16 (New York: IEEE, November/December, 2008): 12691276.

[25] Robert Kosara, "Visualization Criticism The Missing Link Between Information Visualization and Art," Proceedings of the 11th International Conference on Information Visualization (North Carolina: IV, 2007): 631636, 634.

[26] Mitchell Whitelaw, "Art Against Information: Case Studies in Data Practice," Proceedings, Fibreculture, ed. Andrew Murphie, (Perth: Digital Arts and Culture Conference, 2006): 13, http://journal. fibreculture.org/issue11/issue11whitelaw.html (accessed December, 2009).

[27] See the website of Christophe Viau, http://christopheviau.com/ (accessed June 28, 2011).

[28] Lisa Jevbratt, "The Infome The Ontology and Expressions of Code and Protocols," (presentation at Crash, London, 2005), 1, http://journal.fibreculture.org/issue11/issue11_whitelaw.html (accessed March, 2010).

[29] The alignment of code and graphics with mathematics, rather than text, is a point made by both Ron Burnett, How Images Think (Cambridge, MA: The MIT Press, 2005); and Lev Manovich, "The Database as Symbolic Form," in Database Aesthetics: Art in the Age of Information Overload, ed. Victoria Vesna, 3960 (Minneapolis, MN: University of Minnesota Press, 2007).

[30] Indeed Hegel postulated the need for a qualitative understanding of beauty, rather than a quantitative one, challenging some aspects of scientific realism or essentialism. See Georg Wilhelm Friedrich Hegel, Aesthetics. Lectures on Fine Art, trans. Thomas Malcolm Knox, 2 vols. (Oxford: Clarendon Press, 1975).

[31] Lev Manovich, Data Visualization as New Abstraction and AntiSublime (unpublished Paper, Berlin, 2002), http://www.manovich.net/ (accessed January, 2009).

[32] Edward Tufte, Beautiful Evidence (Cheshire, CT: Graphic Press, 2006); and Edward Tufte, The Visual Display of Quantitative Information, 2nd ed. (Cheshire, CT: Graphic Press, 2001).

[33] Edward Tufte, The Visual Display of Quantitative Information, 2nd ed. (Cheshire, CT: Graphics Press, 2001), 51.

[34] Colin Ware and Robert Bobrow, "Supporting Visual Queries on Mediumsize Nodelink Diagrams," Information Visualizations 4 (London: Palgrave, 2005), 4958.

[35] Immanuel Kant and Allen Wood, eds., Basic Writings of Kant: 17241804 (NY, Toronto: Random House Modern Literary Classics, 2001); and Immanuel Kant, Critique of Pure Reason, Critique of Judgment, Trans. Norman Kemp Smith (London: Palgrave, MacMillan, 2000).

[36] Giles Deleuze, Cinema 2: The timeimage, Translated by Hugh Tomlinson & Robert Galeta (London: Continuum, 1989).

[37] Francisco J. Varela, Evan Thompson, and Eleanor Rosch, The Embodied Mind: Cognitive Science and Human Experience (Cambridge, MA: MIT Press, 1993), 71.

[38] Ibid., 51.

[39] Warren Sack, "Network Aesthetics," Database Aesthetics: Art in the Age of Information Overload, ed. Victoria Vesna, 205 (Minneapolis, MN: University of Minnesota Press, 2007).

[40] Vilem Flusser, "Memories," in Ars Electronica, ed. T. Druckery (Cambridge, MA: The MIT Press, 1999), 203.

[41] Bruno Latour, "Visualization and Cognition" [Originally Published as "Les 'vues' de l'espirit"], Culture Technique, no. 17 (June, 1985), www.bruno.latour.fr (accessed December, 2008); and Bruno Latour, "Science Po Project," (keynote address, The Future of Objectivity Conference, Toronto, 2008) http://www.sciencespo.fr/portail/ and http://research.ischool.utoronto.ca/objectivity/abs.html (accessed February, 2010).

[42] Ron Burnett, How Images Think (Cambridge, MA: The MIT Press, 2005).

[43] Fernanda B.Viegas and Martin Wattenberg (2007), http://manyeyes.alphaworks.ibm.com/manyeyes/page/Visualization_Options.html (accessed March, 2010).

[44] Catalina M. Danis, Fernanda B. Viegas, Martin Wattenberg, and Jesse Kris, "Your Place or Mine? Visualization as a Community Component," CHI 2008 Proceedings, April 510, 2008 (Florence Italy: ACM, 2008), 275284, 804.

[45] Joshua Portway and Lise Autogena, Stock Market Planetarium (2002), http://www.blackshoals.net/description.html (accessed December, 2009).

[46] Mark B.N. Hansen, New Philosophy for New Media (Cambridge, MA: The MIT Press, 2004).

[47] Bruno Latour, "Visualization and Cognition" [Originally Published as "Les 'vues' de l'espirit"], Culture Technique, no. 17 (June, 1985) www.bruno.latour.fr (accessed December, 2008); and Bruno Latour, "Science Po Project," (keynote address, The Future of Objectivity Conference, Toronto, 2008) http://www.sciencespo.fr/portail/ and http://research.ischool.utoronto.ca/objectivity/abs.html (accessed February, 2010).

[48] John Law and John Hassard, eds., Actor Network Theory and After (London: Blackwell Publishing, 1999).

[49] Golan Levin, Jonathan Feinberg, Shelly Wynecoop, and Martin Wattenberg, The Secret Lives of Numbers Statements (2002), 1, http://www.flong.com/ (accessed December 2009).

[50] Ibid.

[51] Jonathan Harris, & Sep Kamvar, We Feel Fine, http://www.wefeelfine.org/movements.html (accessed December, 2008).

[52] Stephanie Posavec, Writing Without Words, http://www.itsbeenreal.co.uk/index.php?/writingwithoutwords/aboutthisproject/, 2008, (accessed August 2008).