

Hearing Pure Data:

Aesthetics and Ideals of Data-Sound

Mitchell Whitelaw

The basic resources, for sound artists and producers, are now digital. Production tools have for the past decade been moving from hardware to software; this process has recently reached saturation point, such that the computer has completely internalised – virtualised – the studio: the only vestige of hardware is an audio interface, necessary still to convert between data and audible signal. Creative sound culture is restless; casting around for new resources, it appropriates and misappropriates whatever it can. Given that the basic platform for sound culture is the personal computer, it's not surprising that it has begun to draw on *data* as the raw material of that environment.

In some recent sound practices, data is explicit, not implicit, and concrete, not transparent; I refer to such work here as data-sound. This work turns in on that conversion from data to signal, normally taken for granted, which underpins contemporary sound culture. In the process it reveals data as itself an elusive construct, a figure, an idea(l), as much as a commonplace material. Data-sound entails an imagination of data and its milieu – dataspheres or spaces – which is a cultural act in itself. Sound culture, and increasingly culture as a whole, is digital; so the question of how data is, and can be heard, is significant.

Pure Data

Data misappropriation is straightforward: select “file: open any” or “import raw” in your sound editor. Select any non-sound file from your hard drive, and press play: most likely a harsh block of noise, maybe flecked with modulating bleeps, squeals and pulses. Select different file sizes and formats, and observe the results. Import, export and file conversions form a web of potential for modifying files; shifting from one application to another, any data can be displayed and edited as text and/or sound and/or image.

This is “data bending”, both a process and a loose-knit subculture of audio artists working across the interbreeding genres of contemporary electronica.¹ In its most recent form it is a digital relative of “circuit bending”, the exploratory hacking of electronic musical toys and

¹ For a starting point see the databenders group – <http://launch.groups.yahoo.com/group/databenders/>

instruments.² “Bending,” because the instrument must remain functional – like some neurosurgery, circuit bending must be undertaken without an anaesthetic, while the patient is powered-up and playable. Similarly, data bending is a kind of creative, adaptive modification which both subverts and maintains functionality. It comprises a handful of strategies: the first, “open any”, is described above. All files become audio files: the hard drive becomes a sample library. This transcoding can give rise to particular aesthetic and conceptual projects: audio from images, audio from text, audio from text from images, and so on. The transformation is bi-directional, so sound can also be converted back to other formats, and this opens up new modes of manipulation. Open an audio file as text, and begin altering characters: paste in a few tracts of your favourite philosopher, or delete random chunks; reopen as audio. Open audio file as image; apply filters, scribble, save, reopen.

Frequently data bending is used to generate sonic source material, rather than an entire piece: stAllio!’s twelve inch *True Data* consists of edited excerpts from “random data files” sequenced into noisy, abrasive techno.³ Smartelectronix, a collective producing audio plugin software, suggest the same approach in their recent competition. Here the data files of Smartelectronix plugins themselves are to be edited and sequenced into “Pop, hardcore, ambient, noise, electro ...” in order to reveal “the sound of Smartelectronix.”⁴

As a creative strategy, data bending might be explained as simple sound expansion – that tendency, in the musical avant-garde, to constantly seek out new sound materials. A seminal example of sound expansion is Cage’s prepared piano. Yet clearly the act of modifying a piano (however gently) is significant in itself, even apart from the expanded sound palette it generates. So too here; data misappropriations, transcodings and manipulations are more than mechanisms for sound expansion, they are cultural acts and statements, and readable as such.

One of the striking features of this work is the notion of data that it constructs. There is a pervasive notion of “raw” or “pure” data. stAllio!’s *True Data* hints at this (tracks include “open as raw”); Smartelectronix seek sounds that embody the core identity of their work: the data structures of their audio software. On the microsound list, Christopher Sorg writes of

² While the coinage “data bending” is recent, similar practices have a longer history. Ian Andrews reports industrial acts Throbbing Gristle and Severed Heads using audio from computer data-cassettes in the late 70s / early 80s. (Ian Andrews, “re: [microsound] data -> sound examples,” email to microsound list, 8 June 2004.)

³ StAllio!, “stAllio! - True Data 12”.

<http://www.animalswithinanimals.com/stallio/discog/truedata.html>

⁴ Smartelectronix -

<http://www.smartelectronix.com/competitionRules.php?selected=compRules>

“trying to figure out the aesthetics of a pure, sequential data stream.”⁵ One of the preferred programming tools for experimental digital audio/visuals is Miller Puckette’s Pd – “pure data.”⁶ In this last example the software seems to be reminding its users that, despite the structured media artefacts it produces (images, sounds) its internal representations are abstract; purely data. Otherwise this notion of purity reflects a desire to somehow access *data itself*. Christopher Sorg, again:

What I have been particularly interested in is the idea that all data inside the computer is essentially the same, and that it just takes someone to "peel the skin", and peer inside, either with ears or eyes, or whatever senses we care to translate the switching of 1s and 0s into. ... To me, the most interesting thing about "data-bending" ... is letting the data speak for itself, trying to listen to the data stream with as little interference as possible.

In part this notion of rawness, and of “peeling the skin” from digital media, is related to the recent wave of media errors, glitches and clicks that has swept through electronic music. Elsewhere I have described such work in terms of “inframedia” – the technological underside or substrate of media technologies.⁷ From record cracks, to CD-skips, digital glitches and crashing and malfunctioning hardware, this aesthetic points to the physical infrastructure which underpins electronic media. It is a reminder of materiality, a collapsing of representational transparency. Data bending and related practices are an expansion of the “clicks ‘n’ cuts” aesthetic: here the momentary error, skip or crack, opens up, and the substrate itself pours through. As in other inframedia aesthetics, the physicality of sound is significant: we feel and sense the data, rather than reading or understanding a message. Data-bending is a form of anti-content – especially considering its tendencies to self-referentiality (using audio applications and plugins as sound sources). It makes a (doomed) attempt to resist the production of yet more arbitrary media content, and instead reveal what is “true” (about it self). This raw data signifies (by a kind of metonymy) the digital in general, as a sphere or space; what Lev Manovich refers to as “the computer’s own cosmogony.”⁸ It also signifies the cultural status of this cosmogony. If power and agency subsist, now, in the sphere of the digital, then to hear raw data is to hear some trace of that.

⁵ Christopher Sorg, “re: [microsound] data -> sound examples,” email to microsound list, 7 June 2004.

⁶ See for example <http://puredata.info/about/>

⁷ Mitchell Whitelaw, “Inframedia Audio,” *Artlink* 21(3) (Sept 2001) 49-52. Also available: <http://www.ce.canberra.edu.au/staff/mitchellwhitelaw/papers/InframediaAudio.pdf>

⁸ Lev Manovich, *The Language of New Media* (Cambridge, Mass.: MIT Press, 2001) 46.

Mishearing Information

The deliberate misreading of data bending leads us to a basic opposition, and a crucial tool for pursuing an analysis of data-sound practices. Just as data bending is anti-content, it is anti-information. Information is a formatted message; significant difference, as opposed to randomness; it has a sender and receiver. “Information technology” relies on a substrate of formatted symbols, and ultimately binary bits: digital data. But information is not data: information is the content of data, its message. So data bending is an attempt to hear the underneath of information, to deny the (intended) message. Information implies communication and subjectivity; while *the data itself* suggests data as pure (found) object, alien to the subject, unintentional and a-referential.

However this pure object, this separation of data from information, is impossible, and impossible from two sides at once. Semiotically, a message always creeps back in, even if the message is “listen to me, accessing pure data.” The process and performance of data bending always returns to fill in meaning; once again, sound-expansion is never purely itself. From the other side (the underneath), it is impossible because *the data itself* is perceptually inaccessible. The data is always and inevitably ordered, organised, formatted – even if it is mis-formatted, it is re-formatted, and in fact any format is yet another trace of subjectivity and intention: it is a cultural artefact, an agreed convention of form. Not other, not alien, but part of us. In the case of reading in a non-sound file (say an application) as sound: the original file contains certain structures and patterns, with instructions and resources stored in various segments of the file, themselves formatted. Read as audio, these original structures are flattened and traversed, as the binary data is re-formatted into a string of (for example) 16-bit integers. What we hear is not the data in itself, but data in one format, smashed through another; and both formats are cultural artefacts. Format punctures any ontological purity we might imagine, but also inevitably conditions the sonic outcomes. The parameters which the data bender chooses to reformat (and transcode), fundamentally shape the sound.

This is not to dismiss data bending as a creative enterprise. That the romance of hearing pure data is impossible, only makes its pursuit more interesting, and certainly doesn't preclude the possibility of compelling art emerging along the way. It does suggest an alternative route for data-sound practices, though, which is to work with, rather than against, format and information. If we accept that some process of translation, some mediation between data and sound is inevitable, then the question is, what is translated, and how, and how else could it be done? This question is crucial at a time when the social, economic and cultural valency of the datasphere is growing. As Manovich writes,

This is the new politics of mapping of computer culture. Who has the power to decide what kind of mapping to use? Which dimensions are selected? What kind of interface is provided for the user? These new questions about data mapping are now as important as more traditional questions about the politics of media representation.⁹

Sonification and the Data Sublime

Mapping data to sound is the preoccupation of a small but active research community working on “data sonification” and “auditory display.”¹⁰ In one sense sonification is the converse of data bending: where data bending is arbitrary, abstract and aesthetic, sonification is designed, referential and functional. Where data bending seeks out *the data itself*, sonification seeks out meaningful, usable information. Artists and researchers Ben Hansen and Mark Rubin refer to “the use of sound in exploring the information hidden in data.”¹¹ Hansen, Rubin and a group of other sound artists have taken a sonification approach to data/sound aesthetics.

Listening Post is a recent installation work by Hansen and Rubin that shows a highly evolved approach to data/sound mapping, and raises some of the implications of this approach. The work draws its data in real time from thousands of public online discussions, in chat rooms and online bulletin boards. As its name suggests, the work attempts to “listen in” to this discourse, to render this textual chatter audible. More specifically, the work seeks to convey the content of those discussions, their scale (the sheer volume of text data) and a sense of their momentary dynamics or “immediacy.” This is achieved through a sophisticated set of data collection, analysis and sonification processes. In one example that the artists outline, software agents search the text stream, returning posts that match those retrieved by other agents; the result is a topically grouped accumulation of texts, displayed and intoned by voice synthesisers. This is one of four mappings, or display algorithms, that *Listening Post* cycles through in a ten-minute period. So, aesthetic experience is structured here according to mapping; this provides formal variety, but also communicates the point that data and mapping are functionally distinct — that data might be mapped and re-mapped, and that these mappings render the same data in different ways.

⁹ Lev Manovich, “The Anti-Sublime Ideal in New Media,” *Chair et Metal / Metal and Flesh* (Fall 2002), <http://www.chairemetal.com/cm07/manovich-complet.htm>

¹⁰ See ICAD (International Community for Auditory Display), <http://www.icad.org>

¹¹ Mark Hansen and Ben Rubin, “Babble Online: Applying Statistics and Design to Sonify the Internet,” *Proceedings of the 2001 International Conference on Auditory Display*. Also available: <http://www.stat.ucla.edu/~cocteau/papers/pdf/rubin2.pdf>

Listening Post takes on the datasphere in its most prominent manifestation – the internet. It emerged from research on the functional sonification of network traffic, a popular application for sonification that offers system administrators an ambient display of data activity that is (for them) highly significant, but normally intangible.¹² While these systems focus on network transactions, *Listening Post* mines the network's content. Another work of software art, Jason Freeman's *N.A.G (Network Auralization of Gnutella)* elegantly sonifies both transaction and content.¹³ *N.A.G* seems initially to be simply a client application for the Gnutella peer-to-peer file-sharing network. It is a Gnutella client, but a slightly bent one; enter a keyword and *N.A.G* begins not only downloading matching files, but playing back loops and fragments of its find. The aural texture that results shifts as *N.A.G* finds new files, and as the network dynamics (particularly download rates) change. However *N.A.G* works best not as a network sonifier, but a kind of free-associating cultural core sample. *N.A.G* turns the “false positives” of a keyword search to its advantage, creating a sprawling musical collage of unexpected trash and forgotten favourites. It illustrates a kind of mutant sonification, radically open and uncertain in sonic content and correspondingly in the “information” it communicates.

These works, and others like them, share a basic dynamic of revelation which draws attention to the dynamics and structures of networked data. Often, as in *N.A.G* or Andi Freeman's (1999) *<head>banger* browser, they intervene in existing protocols and structures with a few simple connective gestures, re-wiring software in a way that correspondingly re-wires attention and experience. Mapping is a process of linking, joining together; here, as in data bending, the first (and crucial) step is the connection: Manovich's question of what connection, and why, still seems secondary.

Even in *Listening Post*, which makes sophisticated and self-conscious mappings, there is a primary urge for revelation. Hansen and Rubin aim to make a “meaningful rendering of a massive data stream,” and “distill the content and the structure of this collective communication.” The sense of data as object emerges again here, along with a desire to reveal what is inherent to the data. As well the scale of the dataset is significant in itself, and its vastness is a part of what the artists seek to communicate. Manovich identifies the same phenomenon in visual data.art:

¹² See for example *Peep*, <http://peep.sourceforge.net/intro.html>

¹³ *N.A.G (Network Auralization for Gnutella)*, <http://www.turbulence.org/Works/freeman/index.php>

If Romantic artists thought of certain phenomena and effects as un-representable, as ... beyond the limits of human senses and reason, data visualisation artists target the exact opposite: to map such phenomena into a representation whose scale is comparable to the scales of human perception and cognition.

This is the “anti-sublime”: the condensing and collapsing of the unimaginable vastness of contemporary datasphere, into perceivable objects. It applies exactly to *Listening Post*, and perhaps other works as well. *N.A.G* presents a local, specific, momentary aural sample of the vast and shifting pools of a peer-to-peer file sharing network. Some data bending work seems to have the same, evocative sense of scale; the microsound RNDTXT project takes as its dataset a massive 15Mb text file of random text culled from spam email.¹⁴ This mass of text is imperceptible in itself, as text, but a rich subject for visual and sonic data mapping and mashing. UBSB’s release *traceroute* consists of data slabs surreptitiously gathered from a broadband network hub, converted to audio; scale here is bandwidth, rather than network size.¹⁵

But what of the anti-sublime? Manovich is right to suggest that such works take the unimaginable “beyond” of data, and make it available to experience. Yet this seems no different to the way that artists have traditionally evoked the sublime. A painting of a stormy sea never attempts to capture or entirely reproduce that sea or its dynamics, but to evoke an idea and instill a feeling; it makes a finite and specific impression of the vast beyond, and its own limits, its own static smallness, only adds to the pathos. So too, much of this work makes self-consciously limited but evocative impressions of the sublime of data. Perhaps data-sound connects with what Jon McCormack and Alan Dorin call the “computational sublime”: “the instilling of simultaneous feelings of pleasure and fear in the viewer of a process realized in a computing machine.”¹⁶ This feeling is induced by the accelerated symbolic logic of computation, its ability to outstrip human thought and imagination, to seem “beyond us” even though, as the saying goes, the computer’s only talent is to be stupid, very quickly.

Interactivity and Data-Subjectivity

This sublime of data returns us to the notion of data as pure, found object; there is a shared sense of data as other and elsewhere, constituting or inhabiting another realm. This construct

¹⁴ microsound RNDTXT, <http://microsound.org/rndtxt>

¹⁵ UBSB, *traceroute* (Ash 4.7 LP). See

<http://www.kcw70.dial.pipex.com/html/releases/ash4.7.html>

¹⁶ Jon McCormack and Alan Dorin, “Art, Emergence, and the Computational Sublime,” *Proceedings of Second Iteration, Second International Conference on Generative Systems in the Electronic Arts*, (Melbourne: CEMA), 67-81. Also available: <http://www.csse.monash.edu.au/~jonmc/resources/art-2it.pdf>

is comfortable and familiar, for it has deep cultural roots: cyberspace, virtual reality, the Ars Memoria, Heaven and Hell. All the more reason to seek out ways of unthreading it. Interactivity, another favourite new media construct, may help here. When data-aesthetic practices generate closed systems, or aesthetic objects, they play into this mystification of data; data and map become inextricable; we have little sense of how things could be otherwise. Wittingly or not they present intentionality, in the form of a map, as unintentionality, *the data itself*. Interactive or open systems allow us to tweak the map or alter parameters, even navigate or manipulate the dataset itself (as in *N.A.G.*). They begin to reveal the contingency of any particular mapping and the abyssal plasticity of the dataset, the impossibility of the *in itself*. Interactive systems leave the question of intention at least partially unanswered, turning it over to the user: what do you want to know, make, find? What, in this dataset, is information, and what is noise?

A final word from Manovich. He makes the point that disciplines such as scientific visualisation (and sonification) are already effective in extracting meaning, and even beauty, from abstract data. “The more interesting and ... maybe more important challenge is how to represent the personal subjective experience of a person living in a data society.” This is too modest a challenge, I think. We are all already data-subjects, from our GUIs to our ATMs; data-sound and other practices reflect that reality. The challenge now is to transform that subjectivity, to instill in it a pragmatic data-literacy which increases its power. Perhaps the most important lesson from data-sound comes from process, rather than product, for in this domain it is the artists who are prototypical data-subjects. Theirs is not a single subjectivity, but uses diverse strategies and mappings, and these not complete, rational or determined, but arising through mixtures of whim, convenience, insight and chance. They may show us a way, to hear data for ourselves.