

The Ragged Manifold of the Subject: Databaseness and the Generic in Curating YouTube

Olga Goriunova

This is a draft chapter that will appear in *Media After Kittler*, eds. Eleni Ikoniadou and Scott Wilson, London: Rowan & Littlefield, 2015.

Introduction

Recent exhibitions, such as *We Are All Photographers Now!* (Musée de l'Élysée, 2007), *YouTube Play: A Biennial of Creative Video* (Guggenheim, 2010) and *The World Beach Project* (Victoria & Albert Museum, 2007-2012) engage the current condition of “total creativity” where being a unique subject as best achieved through creative expression is both a social imperative and a financial routine. Creative, aesthetic work is traditionally regarded as both individually and collectively liberating and an enabler of a better society and it is also a forced mode of the available precarious work. All these shows invite works from general public, which are then exhibited following a selection procedure, performed either by a human committee or the jury of software (pieces picked by an algorithm enabling random selection), or both and more (software overlays photographs onto maps, brings works together, co-creates works, etc). The resulting aesthetic space of such exhibitions, more often than not thought not always, is not a collective transindividuating work as hoped for in digital art and theory, but generic unremarkable mass of cultural production, rendered such, first of all, by the gargantuan volume of submissions (reportedly, over 20,000 in the case of the Guggenheim show) and by the way in which every individual piece is locked upon itself. The singularity strived for and encoded in the very medium of an art museum exhibition and a display, finds itself entangled in a paradoxical embrace with indiscernibly repetitive creative effort and a extensive multiplicity rather than intensive collectivity, of pieces.

The tensions at the heart of this conundrum are shadowed by an overhanging Platonic dyad of an original and a copy, model and image, good and bad art. Added to it are the meaning-making, subject-making, mediating machines of art institutions, but also networks, platforms such as YouTube, and desktop or cloud-based editing software. More than that, it is more precisely the mediation of the database as the foundational symbolic form and computational organizational mechanism of both YouTube and other platforms that cut to the core of the changing conditions of possibility of both singular and multiple, unique and the generic and the oscillation in between.

In this paper I would like to focus on the aesthetic generic, and the cultural generic as one generated in the operation of the database management system. Some pairings of concepts that I would like to contemplate include, first, generic and original / copy; generic and meaninglessness / nonsense / noise and reality, generic and seriality, generic as measure. I would like to prod generic for its links to sorting, lists, tables and databases. Most of all, I take up the generic as the concept through which to approach a new understanding of subjects and subjectivities. Here, I undertake a cultural theory and aesthetics' approach to the way in which subjects are formed in contemporary culture, accounting for the computational. My subjects arise from art projects and operations of software, as human, posthuman or nonhuman and are mainly characterized by certain consistency and unity that hold cross-cutting biological, symbolic, empirical and transcendental planes together, and in the production and maintenance of which the operations of the cultural and the technical partake. It is these operations towards the subject that undergo changes that I am focusing on.

My field of exploration here is digital media culture which today expands to encompass various forms of cultural expression as they acquire a computational backbone. It is not the structural migration to computer-enabled media channels, forms, devices and genres ranging from social networks to smart phones, databases, digital TV and photography, e-books and

educational apps, YouTube, memes, emails, Skype and online registration forms, but the logic of invention, visibility, accountability, objectivity, power as well as fun, collectivity, and protest that come to rely on computational execution. Culture today works computationally, in its forms, in its genres, in its contents, infrastructures, circulatory techniques, meaning-making and evaluation systems and larger structures it is plugged in.

Such a change in mediatory mechanisms is usually discussed as revolutionary by philosophers of technology, such as Stiegler, and media theorists, such as Kittler; though, for both, it is seen as only the next step in the long history of mediation. Mediation, an outdated term itself, here is both biological (for Stiegler, assuming an erect position is technological), metaphysical (for Kittler, the production of the soul is based upon the dominant media of literature and handwriting) as well as covering other areas core to the construction of the history of Western dominance (alphabetical writing, postal systems and the printing press) and of the subject (with enhanced memory, perception, rationality and then irrationality). As the change to Turing machine-based society is now obvious, its far-ranging consequences are explored in various disciplines and spheres by employing different vocabulary and concerns. Art is exemplarily amongst the foremost of these.

The work I shall refer to throughout my paper is Curating YouTube¹ (2007-ongoing). Curating YouTube is a project by Berlin-based curator Sakrowski. It is a curating platform, a software tool and a set of exhibitions curated by himself and by invited artists. The focus of the project is the immense volume and repetitiveness of YouTube videos; it is an exploration of YouTube as an art platform and an attempt to learn to appreciate and imaginatively comprehend the genre of YouTube videos en masse and as a mass cultural phenomenon. In a way, it is an exercise of staring into and striving for the generic. YouTube videos that become viral are well-known (often selected, for instance, for showcase on the pages of the Guardian newspaper as the most representative of a moment or a topic, or at exhibitions such as the ones mentioned above) whereas the most uninteresting videos are a subject of the project NoTube² (a contest for the worst or most nondescript and uninteresting YouTube video ever) and First-Viewer-Television³ – a channel for videos that received no views (both projects by IOCOSE). Curating YouTube aims at the mediocre, nor the best and successfully viral, neither the worst, but continuously generic.

The Grid, part of Curating YouTube, is a tool that allows the selection and laying out of videos in a grid; furthermore the videos in the grid can be played one by one, or in an orchestrated fashion, each launched with a delay. The Grid is often used to bring individual work into a collective enterprise more pronouncedly (for instance an orchestra made of people playing the same tune on individual instruments in separate videos), as every single work only functions collectively in the space of the YouTube as social media (the individual exercise of playing an instrument and putting a video online is precipitated by the collective practice and shared technical means), but is framed, technologically, as an individual work. Here, the same tension as outlined above, is at play: striving for singularisation, but only existing plurally.

The above is not entirely correct though: video response function and YouTube channels promoting each other are all genre and technical mechanisms that feed upon this plural strain of singularisation and provide gradually differentiated techno-aesthetic forms to spin off new and blurrier arrangements. Where a project such as 50 50⁴ by Oliver Laric (2007) (a video assembled from consecutive fragments of peoples' individual videos in which they perform a song from the rapper 50 Cent) pointedly - and as one-off - achieve what Curating YouTube explores (but also produces as a platform), such work is already changing with a sweeping commercialization of YouTube, where channels are monitored for the number of followers, with successful channels being offered a possibility to regularly film in studios scattered around the hubs of capitals of the

¹ Curating YouTube, <http://www.curatingyoutube.net>, accessed 21.03.2014.

² NoTube, <http://www.curatingyoutube.net>, accessed 21.03.2014.

³ First Viewer Television, <http://firstviewer.tv>, accessed 21.03.2014.

⁴ Oliver Laric, 50 50, <http://vimeo.com/16024741>, accessed 21.03.2014.

world. A new middle-media (as opposed to mass media) TV production machine will expose an altogether different set of behaviours in the near future.

So, generic, meaning-lacking aesthetic and cultural acts, what are they? Often only regarded in the context of virality, memetics, and infectious stupidity, something being done outside of the subject, without a subjectivity, they nevertheless call for consideration. In Kittler's terms, they would be akin to language speaking itself to fill a gramophone record; culture exercising itself to be circulated in computational networks in the absence of a subject, or rather in its presence – one that is now the Turing machine.⁵

Kittler

According to Kittler and his interpreter Winthrop-Young, from the later part of the 18 century and with the emergence of a bourgeois model of the family, as mothers were charged with bringing up children with a “sufficiently developed psychic centre of resonance and reflexivity, commonly referred to as spirit or soul,”⁶ new child-rearing practices emerged that mainly worked by drawing upon language as the medium. As Winthrop-Young summarises: “new language acquisition practices taught children to merge minimal signified into words, the new writing lessons taught them to merge basic strokes into lessons and then words which are ‘naturally’ understood and heard... the humans inscribed by these techniques (Goethe, Hoffman) will come to believe that nothing is meaningless because everything is always on the threshold of meaning.”⁷ With the reform of writing and reading lessons, the changing roles of aesthetics and philosophy, and the new roles of the self-reflexive civil servants of nation states, it was most of all the mother's love that promised and provided meaningfulness.

Here, Kittler joins the Foucauldian enterprise: if Foucault focused on the discursive construction of subjects as the function of modernity, Kittler focused on “inscription techniques for creation of the modern subject”.⁸ For Kittler, the creation of the modern subject with a metaphysical soul was enacted through the particular materiality of communication media. Kittler wrote: “Whoever wrote in block letters would not be an individual... the great individual entities invented in the age of Goethe – the developmental process of *bildung*, auto-biography, world history – could be seen as the flow of the continuous and the organic simply because they were supported by flowing cursive writing... to develop handwriting as out of one mold means to produce individuals.”⁹

In this way, Kittler completes his analysis of Discourse Networks 1800. His work on Discourse Networks 1900 is focused on electric and electronic media. As the material ploughing through and working communication up radically changes, he asks: “What happens to the soul in the age of Edison?” As the work of the Mother as the discursive machine weaving the meaningless noise into significationally-promising utterances is coming to an end,¹⁰ noise takes over. The effect of the Edison machine, according to Kittler, is about obtaining direct access to “reality” (which is noise). Language is no longer constructed out of the semantically promising “minimal signifieds”; now communication units are only “momentarily...arranged into allegedly significant patterns”, they become computable elements, individually meaningless, a data stream, an exchange of noise.¹¹

Kittler writes: “In the discourse network of 1900, discourse is produced by random generators. Psychophysics constructed such sources of noise, the new technological media store their output.”¹² For Kittler, the difference between sense and nonsense is shattered, and meaning

⁵ Friedrich Kittler, “Unconditional Surrender” in *Materialities of Communication*, ed. Hans Ulrich Gumbrecht, K. Ludwig Pfeiffer (Stanford: Stanford University Press, 1994), 319-334.

⁶ Geoffrey Winthrop-Young, *Kittler and The Media* (New York: Polity, 2010), 32.

⁷ Ibid.

⁸ Winthrop-Young, *Kittler*, 39.

⁹ Friedrich Kittler, *Discourse Networks 1800/1900* (Stanford: Stanford University Press, 1990), 83-84.

¹⁰ Kittler, *Discourse Networks*, XXIII, 71.

¹¹ Kittler, *Discourse Networks*, 81, 68, 73.

¹² Kittler, *Discourse Networks*, 106.

becomes a secondary phenomenon. The condition of chaos is here the originary condition, in which an “absolute homogeneity of material” prevails and which is organized algorithmically, through “randomness and combinatorics.”¹³ Cultural practices working with these discrete units rely on functions that are not governed by meaningfulness; they are “automatic and autonomous.”¹⁴ Kittler elaborates: “...the noise that precedes every discourse becomes at once theme and method,” - and further - “discourses are ‘eclectic combinations’ of noise spectra.”¹⁵

Thus, the change in the materiality of communication provided access to a raw versions of reality: “for the first time in history, writing ceased to be synonymous with the serial storage of data. The technological recording of the real entered into competition with the symbolic registration of the Symbolic... this delight brought to light discourses that previously had never passed a recording threshold.”¹⁶

Now, if we turn back to the cultural practices of today, meaningless and idiotic in their automatism and randomness, are these the media inscription techniques of a contemporary – 21st century subject? As language, in Kittler’s account, starts speaking to itself to fill in the new recording machines with originary noise, now, in the age of the Turing Machine, and, for Kittler, with the primacy of this machine, noise is arranged into patterns to acquire momentary significance. Numerical relations and technical calculations are now where the dialogue of souls was before; and the aesthetic of numerical noise is physiological.¹⁷ So, what are these patterns?

People filming themselves enacting similar gestures (lying with their face down in the craze called planking) - is this human originary noise recorded and computed to generate informationally significant patterns to be analysed further by computational techniques? Is the noise or the pattern here generic, singular or multiple? Or is the generic found in the operation of the computational machine that produces and perpetuates work upon these phenomena and the oscillation between them?

Is noise as a nonsensical, originary plane (in Kittler) generic? The generic ecological condition, the generic cultural platform, as read as: ‘Of a genus, family’ - is here something one can belong to, defined by the same characteristics you have. You exercise generic functions and actions on the basis of your specific genetic allowances and ecological interferences. Here, of course, the generic and the noise it entails is highly patterned, with carefully applied randomness. Pattern and randomness, and pattern and noise do not exist outside each other. If pattern is information, noise is not non-information; information encompasses and functions in relation to both pattern and noise.

Here, the generic is unlikely to be either pattern or noise and, using Kittler, we can say that the generic of our era refers to the loss of meaning-making practices orientated towards an end product (poet, poetry, uniqueness) and instead functions as an open-ended computationally-arranged process of drawing upon noise and patterning, and working upon the derived data and information.

For Kittler, noise (and nonsense) precedes discourse networks and acquires meaning through the media / discourse network of the age.¹⁸ But one could say that meaninglessness is not noise, but is itself a product of the discourse network of the Turing Machine. The computational machine produces the generic, but then the question is whether the generic is always constructed. Is the generic always mediated? Generic seems to touch upon both the vernacular, the immediate, the unmediated, – and the abstracted, the processed, the operated upon. We can talk about a generic artifact, - a generic chair, - but also about a generic landscape, a generic lake and a generic tree, too. A generic tree is not noise, neither it is a pattern. But even as a work of abstraction, the generic does not lose a connection to nonsense.

¹³ Kittler, *Discourse Networks*, 209, 210.

¹⁴ Kittler, *Discourse Networks*, 214.

¹⁵ Kittler, *Discourse Networks*, 218, 219.

¹⁶ Kittler, *Discourse Networks*, 229-230, 223.

¹⁷ Kittler, *Discourse Networks*, 225.

¹⁸ Kittler, *Discourse Networks*, 218.

Computational machines produce the generic through their processing of data, where things are operationalised and normalized to become data and work as information. They are enacted and performed through sorting, listing, classification, visualization, automatic configuration, and similar forms of articulation.

Generic, copy and seriality in Deleuze

A very different route to take to consider how the generic is produced today is to travel through Deleuze. In *Difference and Repetition*, Deleuze focuses on overthrowing the Platonic reign of the copy over the original. There is some kind of solidarity between Deleuze and Kittler here: they both work towards the reign of the copy, but with very different outcomes. In Deleuze's words, "Infinity of copies... allows neither original, nor copy to subsist..."¹⁹ "The model collapses into difference, while the copies disperse into the dissimilitude of the series which they interiorize."²⁰ As both the notion of the model and that of the copy are challenged, the copies become series, take on the originary as the original seriality. Seriality in Deleuze denounces the original, the unique idea; the unique idea is itself a series of differentiating copies.

Is seriality a methodology for carrying out the generic? Is the generic a unit of measure and a way of arranging things to be measured? The generic arises as the process of measuring, an operation, or, in the case of Curating YouTube, arranging videos onto a grid; the generic here is the mediation of the operation of measuring the similarity. The tool sorts and measures. Here the tool is both a grid and a curatorial idea, or the "similar videos" function of YouTube; it is a conceptual-computational mechanism to both locate and produce the generic, analyse it and act upon it.

Is the generic itself a way to produce pattern and noise in the process of the separation of pattern out of noise by measurement? This is the operation of the sorting of the similar and the production of the relation of association by putting it into a database, table or a grid. The generic though is complex; it can't be described or arrived at by following a mathematical function. One can calculate the mean (of, for instance, a column in a dataset), but the generic has been produced by the operation of producing and formatting the dataset itself (for instance in "normalization").

So, the generic is somewhere in the operations of the database, the generic is produced as the dataset is assembled and structured, and as it is formatted according to a model, and implemented in a database management system.

To use Kittler, the magic of the maternal mouth promising the meaningful world, where the unique is achieved by working on the chaotic, gives way to computational measurement and operation upon matter, which is not random and which is itself a measure of uncertainty, or is fully automated and thus lacks conceptuality. The generic in the discourse networks of 1800 is that that hasn't been cultured enough, hasn't literarily been worked on enough to produce the soul, the idea. The generic of discourse network 2000 is in the sets of operations. Not only in the operation that comes after (Curating Youtube after YouTube), but also in that which differentiates (in making a video, in YouTube itself). If the generic of 1800 is an operation of comparison against the original, the generic of 2000 is an operation of producing systems and operations that act the generic out.

YouTube's aesthetic measurement makes the generic; Curating YouTube is itself a pure databaseness. The art here is in the exploration of the database as a system of operations, as an aesthetic concept.

Database strand

It is now worth looking into the generic as a kind of operation (of generalisation), one of whose origins is statistical. The generic as an operation is profoundly transformed today as the

¹⁹ Gilles Deleuze, *Difference and Repetition* (New York: Continuum, 2008), 80.

²⁰ Deleuze, *Difference*, 156.

operations in the relational and other types of databases allow for new kinds of generic to emerge, where the operation of such emergence itself is transformed.

Databases (as a symbolic form, in Manovich),²¹ relational databases in particular as a dominant kind of a database used today and newer kinds of databases (NoSQL movement) attract significant scholarly attention. They are material forms that enact certain ways of thinking and producing the world, where such ways are not necessarily known to people outside of computer science and often not subjected to critical humanities or social sciences thinking.

One route for looking for generic in the history of databases is in the problem of generality in statistics. Emma Uprichard (together with Burrows and Byrne) wrote on SPSS (software package for statistical analysis) in the following manner: “This [change] has knock-on effects for *generalization* – that is, the ability to infer observations from a sample at one point in time to the population from which it was taken at another. The concept of generalization ... is arguably defunct. Understanding how cases are ‘generally’ requires an explicit recognition of how they are ‘specifically’. The logic of predictive analytics ... aims instead to know ‘enough’ to make... ‘moderatum generalisations’ about particular types of cases. [which] relates to tentative claims that are only relatively enduring; they are not intended to hold good over long periods of time, or across ranges of cultures, and they ... need to be frequently ‘updated’ to ‘keep up’ with the ways in which cases may become different kinds of cases.”²²

As the datafication of things and their subjection to computational operations takes hold, the individual and specific and the generalized are redistributed. Evelyn Ruppert, analyzing a management information system that works across case management systems holding data on young offenders, (and following Ann-Marie Mol’s concept of the “body multiple”), calls for a subject multiple as it is enacted and performed by the data multiple.²³ The multiple, the composite, is different in nature to the old kinds of the generic, as it was constructed, in statistics, through the operations of sampling, establishing causality and explanatory machines. (Uprichard and others instead focus on complexity, multiple causality, visualization and exploration as the new kinds of quantitative operations.)

Ruppert analyses the relational processes of a database at work that enacts the multiple, which is neither the specific nor the abstracted generic, but has features of both, bringing us back to the introductory passages of this paper. The individual and unique as well as multiple, multiplied and repeated at the same time, - these are the tensions that are supported and carried out at the levels of data structures and algorithmic operations upon data sets within databases. It is indeed such computational ensembles that sustain the relationality core to the construction of the new types and forms of the generic.

Adrian MacKenzie and Ruth MacNally write about the thickening rather than flattening of the worlds produced through database operations.²⁴ For Ruppert, the computational, database system maintains the multiplicity, the uniqueness of each individual while maintaining them as generic people with multiples of attributes: “He/she is instead a different, variable and unique composition and the system maintains this multiplicity.”²⁵

²¹ Lev Manovich, “Database as a Symbolic Form,” (1998) <http://www.nettime.org/Lists-Archives/nettime-l-9812/msg00041.html>, accessed 21.03.2014.

²² Emma Uprichard, Burrows, Roger and Byrne, D. “SPSS as an ‘Inscription Device’: from Causality to Description” in *Sociological Review: Anniversary Issue*, 56(4), 606-622; https://www.academia.edu/151125/Uprichard_E._Burrows_R._and_Byrne_D._2008_SPSS_as_an_Inscription_Device_From_causality_to_description_Sociological_Review_Anniversary_Issue_-_From_causality_to_description_56_4_606_-_622, accessed 21.03.2014.

²³ Evelyn Ruppert, “Not Just Another Database: the Transactions that Enact Young Offenders” in *Computational Culture, A Journal of Software Studies*, issue 3, 2013; <http://computationalculture.net/article/not-just-another-database-the-transactions-that-enact-young-offenders>, accessed 21.03.2014.

²⁴ Adrian Mackenzie and Ruth McNally, “Methods of the multiple: how large-scale scientific data-mining pursues identity and differences” in *Theory, Culture & Society* 30, no. 4 (2013): 72-91.

²⁵ Ruppert, “Not Just Another,” unpaginated.

This is not only found in statistics and bureaucracy, it is poetry, diaries, writing exercises that previously individuated people into singularities – and it is now the database, among other computational algorithmic systems, that enables the creation of individuated subjects while maintaining them as multiples.

Ruppert suggests that the collective offender is enacted by a database as a composite, similarly to a composite portrait that is the multiple with all the singulars present at the same time. Ruppert writes: “So while different practices decentre the ontology of the subject, the MIS re-centres it and at the same time maintains it as a composite.”²⁶ To use the visual form of a composite portrait, the today’s generic, counter intuitively, would not be the print of the composite, but the dynamic and living tension between the singular portraits and the multiple composite, and would be articulated in the construction of the composite, in the operation of relating, measuring and – in the software gesture of layering, overlaying, hiding, and saving.

Relational Databases

A database is a collection of organized data; a table in relation to the database can be a result of query or data itself can be stored as tables. One of the most discussed features of the now dominant relational databases is its maintenance of so-called “semantic independence” of data. As the originator of relational databases claimed: “Future users of large data banks must be protected from having to know how the data is organized in the machine.”²⁷ Here, the logical structure is divorced from the physical structure of data, and in our operations we deal with the logical structure only. Castelle relates to this in the following manner: “The relational model differs primarily in its wholly symbolic and tabular representation to the user, as opposed to the explicitly encoded referential relations of the hierarchical and network models [of databases]. This fundamental semiotic difference produces a highly valued effect recognized more typically as ‘data independence.’”²⁸

Relational database operates on the basis of formalization (creating a conceptual model) that works through normalization and composition. Normalisation ensures that all data objects are irreducible objects of the same type without duplicates; whereas composition “always happens through data in the relations themselves (integration at runtime)” by either modeling or query (so it is either modeled on or derived from data).²⁹

Bernhard Rieder says: “Normalization reduces expressive power on the level of the unit and extends expressive power on the level of composition.”³⁰ The expressive power thus belongs to the query which establishes relations. There is a “model of knowing” within the model of the relational database which is about integrating data from different sources and establishing relations between data.³¹ Castelle explains the same in the following way: “Where the network model enforces referential (i.e., pointing) links between entities at the logical level, the relational model enforces the absence of such reference. In this way, we can say that what the relational model allows for is a sort of freedom in recontextualization of the (entextualized) database artifact... This freedom is realized in the so-called “expressiveness” of relational query

²⁶ Ruppert, “Not Just Another,” unpaginated.

²⁷ Codd, E. F. and C. J. Date. “Interactive Support for Non-Programmers: The Relational and Network Approaches,” in *Proceedings of the 1974 ACM SIGFIDET workshop on Data description, access and control: Data models: Data-structure-set versus relational*, (Ann Arbor, Michigan: ACM, 1974), 11–41.

²⁸ Michael Castelle, “Relational and Non-Relational Models in the Externalization of Bureaucracy,” in *Computational Culture, A Journal of Software Studies*, Issue 3, 2013; <http://computationalculture.net/article/relational-and-non-relational-models-in-the-entextualization-of-bureaucracy>, accessed 21.03.2014.

²⁹ Bernhard Rieder, from the presentation given at the workshop “The Lived Logics of the Database Machinery,” 2012, <http://computationalculture.net/cfps-events/the-lived-logics-of-database-machinery>, accessed 21.03.2014.

³⁰ Ibid.

³¹ Ibid.

languages like SQL, which (given an appropriately normalized design) allow one to relate — via joins and projections — new entities with every interaction.”³²

Coming back to the Curating Youtube project whose main operation is a relating exercise (comparing, pulling videos and fitting them into the grid) and putting those in relation to the database “in general,” we can usefully employ the following observations from Castelle: “Codd’s relational model was very distinctive from then-existing database models in having only one formal user-level conceptual data type: the relation, which can be thought of as a simple table.... The distinction between the two models is thus fundamentally semiotic; where the network model’s pointers mimic the indexical real-world physical relationship between part and supplier, the relational model represents that relationship in an explicitly symbolic, tabular form.”³³ Extracting the relationship into the symbolic realm and operationalising it, the relational database model supplies structures to inform and maintain cultural and social practices, in which establishing a relation is, to an extent, freed from the material to which it pertains – the basis for Manovich’s proposal of database becoming the new dominating symbolic form.

Castelle extends the genealogy of this form to “commercial and administrative bureaucracy”, while Rieder points out the added “software performativity” and “extended calculative agency” of the model. Formalization, normalisation and calculability are all the main principles of operation of the relational database model where the relationality between data objects can be queried, established and described – a process that would result in the production of a new table, and that would be bound by a pre-existing data schema.

BigTable and YouTube

YouTube, however, runs on BigTable – Google’s own data management system, which is not a relational database.³⁴ BigTable supports “sparse semi-structured data” and allows for impressive scalability “up to petabytes” and “thousands of machines”,³⁵ it belongs to a relatively new NoSQL movement of database management systems that is varied in itself and includes DBMS that still have a local relational database and a number of others built on different data models (graph, key value, column, etc).

Systems, such as BigTable and others, seem especially fit for big and semi-structured data and “support thousands of operations per server per second.”³⁶ To simplify, if a relational database is fit for supporting, classically, financial transactions, and doesn’t allow for data duplicates and concurrent operations, BigTable is fit for networks, with dynamically changing webpages and updated links, searches and posts.³⁷

As such, BigTable doesn’t normalize data and doesn’t abstract relations from the data objects themselves: it is a “sparse, distributed, persistent multi-dimensional sorted map,” that is “is indexed by a row key, column key, and a timestamp.”³⁸

Dourish in particular brings attention to the radicality of the change inherent to NoSQL databases. Relational database radically separates content from structure, where the schema must largely be defined before any data can be entered and is rarely altered afterwards; the openness of establishing new relations “lies within content,” which is itself pre-selected by fitting a much “more rigid schema” (Dourish). The expressiveness of the query and the specificity of relationality provided by the relational database are simultaneously open and wide-ranging and

³² Castelle, “Relational and Non-Relational Models,” unpaginated.

³³ Castelle, “Relational and Non-Relational Models,” unpaginated.

³⁴ Fay Chang, Jeffrey Dean, Sanjay Ghemawat, Wilson C. Hsieh, Deborah A. Wallach Mike Burrows, Tushar Chandra, Andrew Fikes, Robert E. Gruber, “Bigtable: A Distributed Storage System for Structured Data” (2006), https://www.usenix.org/legacy/events/osdi06/tech/chang/chang_html, accessed 21.03.2014.

³⁵ Ibid.

³⁶ Ibid.

³⁷ Paul Dourish, “No SQL: The Shifting Materialities of Database Technology,” in in *Computational Culture, A Journal of Software Studies*, Issue 4, 2013, forthcoming.

³⁸ Fay Chang et al. “Bigtable,” unpaginated.

also limited. Dourish: “narrative and archival provisionality are primarily associated with content. The structure of a relational database is much more resistant to change. So the database is provisional, but only with limits; it is open-ended, but only in the terms originally laid down in its structure.”

BigTable relies heavily on Google infrastructure and is compatible with MapReduce (a model for parallel, distributed computing on computer clusters which is characterized by “loose consistency” and storage in “informally-organized collections of data objects linked together by broad associations”³⁹): BigTable uses the Distributed Google File System.

In Dourish’s opinion, the implementation of the NoSQL databases over clusters connected via networks instead of being held in a single mainframe computer radically changes the requirements and performance of the database management systems. A generation of databases relying on key-value structures (which is not the only building block used by BigTable) are the ones where, in comparison with relational databases, data objects and keys (that link value to an object) are not structured. An unstructured collection of data objects (e.g. an email) are linked to an unstructured collection of data values (e.g. “written by me”), through keys belonging directly to objects rather than being separately defined (key – owner of email). If the key is not there, finding the value could be impossible (who is the owner of this email?); and the main operations in such DBMS are, according to Dourish, adding value to a key, updating a key, retrieving a value, as opposed to selecting from tables and ordering into a table of relational databases.⁴⁰ It is easy to see how such relatively new systems have a larger allowance for failure, for missing or loose connections as they are performed in a distributed fashion. Objects can also be linked into “chains” (through values), and become a kind of “networks.”⁴¹

Dourish draws attention to the fact that the analytical power of such models is reduced, and there are currently other very powerful models that retain some of the benefits presented by the relational database systems and associated with firmly described data structures. The expressiveness of the query and the specificity of relationality afforded by relational databases and enacted in cultural material forms (for an insufficiently detailed account of which Dourish reproaches Manovich) co-exist in the socio-political and cultural domain with other kinds of affordances: less expressivity (a focus on system performance at the expense of reduced representational clarity and consistency), a focus on objects rather than relations and “low integration.”⁴²

Dourish suggests that there are four main characteristics of such new data structures and forms of “data processing” that are comprised of assemblages formed by hardware infrastructures and software applications: granularity, multiplicity, associativity and convergence.⁴³ The following is particularly useful for thinking about the generic: “In exploring the problems of granularity and associativity, then, we found ourselves exploring the way that data objects become ‘live’ inside the system... Multiplicity – the ability to act coherently upon multiple manifestations of objects that, from another perspective, are ‘the same’ – reflects this concern with the particular, and the different constraints associated with different approaches to database form... In particular, approaches to multiplicity speak to the opportunities for partition – that is, for a system to operate as multiple independent entities, either briefly or for an extended period.”⁴⁴

Here, the allowance for and the lure of the same, and slightly different, the tensions between the particular object and the multiplicity of relations that arise from its operation – in relation to others, relating, being a duplicate, differing - is encoded, performed at the level of the database management system and implementation designs that work beyond social media

³⁹ Dourish, “No SQL,” unpaginated.

⁴⁰ Dourish, “No SQL,” unpaginated.

⁴¹ Dourish, “No SQL,” unpaginated.

⁴² Dourish, “No SQL,” unpaginated.

⁴³ Dourish, “No SQL,” unpaginated.

⁴⁴ Dourish, “No SQL,” unpaginated.

platforms, such as YouTube. The object multiple or a subject multiple that Ruppert focuses on is a different multiple here: the relations holding it together are much looser; it is less consistent and less integrated. It is a porous, leaky and inconsistent multiple that is itself a loosely related chain of individual objects with differing characteristics.

Manovich's observations on database as a symbolic form seem to pertain to the relational database model, allowing for the expressivity of query, but within a defined structure. Here, the classic net art project by Olia Lialina *Myboyfriendcamebackfromthewar* (1997) exemplifies the cultural materiality afforded and enforced by the relational database model. In this project, and in many other similar endeavours, a viewer can select her path through the dividing frames with images telling a story and thus create an individual narrative, - though bounded by the constraints of what has been defined by and loaded into the database structure behind the interface. The narrative is abstracted from the material that can be expressively re-arranged enabling a variation on the narrative; however, the conditions of the narrative are pre-defined and its combinatorial finitude of options can be pre-calculated. The long-standing criticism of certain strands of interactivity in interactive art and other spheres allowing for freedom of interaction and, thus, co-authorship, within a very limited set of choices laid out by the *real author* pertains here. But the real author here is not even the database manager or designer, but in some way the database model itself. Kittler, positioning the Turing machine in the place previously occupied by the human subject, as outlined at the beginning of the paper, would be happy now.

NoSQL database management systems offer a somewhat different set of allowances and conditions. Using Manovich's language, the narrative is not subtracted but embedded in the material objects themselves – but the objects are much more loosely connected than, for instance, sentences in a story by Chekhov.⁴⁵ Objects stand on their own and connections / relations arise from their operations or operations on them, as if the words, sounds and sentences were dancing around, moving chaotically, forming chains, and assembling into a certain ragged and unbounded multiple (surrealist poetry), rather than a smooth fabric (realist novel) or relatively ordered interactive story (whether works by Alan Robbe-Grillet or Milorad Pavič or digital storytelling projects). The multiple itself, as a dividable, might not be the most useful term here. It is rather a mix-aliquant, data and subject manifold.

Making sense of an unbounded ragged manifold is the purpose of Curating Youtube. The operation of selecting and putting videos into the grid doesn't produce a coherent unique statement (videos in the grid are neither unique nor collective, neither singular, nor multiple), but the curated page is the place where the power of the generic arises and manifests.

Curating YouTube puts selected videos in relation to each other (like attributing values to keys and establishing connections between values), but without a single overall structure. Rather, there is a multitude of structures – the curatorial gesture of browsing and selection, the grid, histories of art, YouTube, the machinery of cultural institutions, BigTable, the Distributed Google File System, networks and protocols, and many more. A mass of semi-structured data, neither patterns nor noise (at times Grid curators look for patterns and at other times they orchestrate an organic ecology comprised of incommensurable objects), the multitude of copies of slightly different video files, are potent of a very specific relationality, the relationality of the generic, which in turn, sustains the oscillation of both and between the singular and the multiple, producing the ragged manifold of the subject.

Curating Youtube

Curating YouTube works on the infrastructure of YouTube as a database management system as it is plugged into many other structures, and its characteristics as an art project lies in its operating in those material structures. In some way, Curating YouTube reveals the databaseness of the new types of databases, abstracting the database model and material affordances from the YouTube database. The project releases the data from the way in which it exists on YouTube,

⁴⁵ There is a certain irony that I exemplify my proposal with literary examples, and Kittler as a fellow philologist would have appreciated the double bind of this act.

but also in a way that YouTube allows views into it. YouTube has a similar “suggestions” function, but CYT cleanses this function and subjects it to aesthetic work.

To speak metaphorically about something technically precise, Curating YouTube attributes new values to data objects and creates new keys, enacting a specific relationality. As Curating YouTube operates on top of and in multiple systems, many of which call for unique individual digital gestures while linking them to a multitude of other, similar but slightly different gestures, in a boundlessly scalable loosely-ordered messy mass of networks and data, it articulates a version of the problem of the individual versus the collective and re-positions the question of the subject.

Sakrowski writes: “In my opinion, an ordering, which creates identity, can only be performed individually and on excerpts and in this way, maybe reconcile the individual with the mass... [it] enables to deal with the multitude of phenomena in terms of aesthetic[s] and associations: selecting and arranging videos in the grid.”⁴⁶ Whether CYT is about reconciliation or not, producing new kinds of relationality between available data via aesthetic means becomes generative of something that is not immediately obvious when we watch videos on YouTube or think about the databases behind it.

The databaseness and relationality of Curating YouTube directly relates to, creates and explores the generic as the site of individuation as it is produced as an oscillation between the singular and the manifold, in clouds, chains and connections and where constructing such an oscillation is an individuating process. Curating YouTube releases the performativity of the database as a thinking model, and as a material method of enactment, and articulates its medium-specificity in an epistemologically-charged aesthetic gesture. Here, a curatorial device and artwork becomes a thickening, a saturation of relations which are intensified, doubled, made similar, made serial, - and it is from such operations that the subject arises.

The Grid and the Curating Youtube Grid

Finally, the question worth asking is what the Grid is in Curating Youtube. In visual design the grid is a means of applying systematicity, overlaying a structure to ease certain operations, and through that, of operationalising. It is used for leveling, alignment, and equalization, - or the arrangement of information in terms of structure. In such a way, the grid as an object already includes operations on the material to be presented in it.

The grid as a formal device is related to that of the *table* and can function as a visual representation that helps demonstrate systematic relations. It can be used a means of arranging data to infer relationality, systematicity and causality; a fundamental quality of a grid or of a table is that it can be re-arranged to repeat the operation of articulation. Structure adds information to the data; there is intelligence in the structure. The grid is not only a visualization of the data or processes, it is also an operation of articulation. The grid is operative.

It has been argued that a grid is not so much “the image of an emerging modernity” imposing uniformity, regularity or homogenization, but a “universal space” of scalability.⁴⁷ Describing the evolution of the grid from hand-woven nets to the punch cards of the Jacquard loom, and to Babbage analytical engine’s use of punch cards, Hannah Higgins says: “The punch card ... is the mechanism of transition between the soft grids of textile technology and the hardware of the information age; it translates the net from its physical expression in textiles to a modeling form that would tabulate, sort, and integrate... The von Neumann architecture [of the computer] is a grid that forms the very core of modern information technology. ...[I]t is an organized set of modules that allows for manipulation and creativity, making new kinds of searching ... and connection building... possible.”⁴⁸

⁴⁶ Sakrowski, “An Acoustic Journey through YouTube,” (2011), <http://www.curatingyoutube.net/acoustic>, accessed 21.03.2014.

⁴⁷ Hannah B. Higgins, *The Grid Book* (Massachusetts: MIT Press, 2009), 12, 15-17.

⁴⁸ Higgins, *The Grid Book*, 249, 254.

A table, and tabular representation in a database has values assigned to rows and columns - something that the grid lacks. If the grid may, in cases such as these, be a first step in establishing and assigning computability, and enacting relationality through a certain kind of measurement, the table, especially in the relational database design, is a mathematical term for relation. Here, the grid can be a procedure of abstraction, whereas a table can be described with mathematical functions in relational algebra (a foundation of the relational database).

The Grid of Curating YouTube establishes a rupture between the liquid relationality of the YouTube database and the process of production of the art project Curating YouTube itself. Curiously, there is some violence done to the material of YouTube by putting a square number of videos into a fixed structure. A measurement of measurement, the establishment of relationality between the relational, - all in all, a collision of formal structures produces the project. Thus, the individuation of Curating YouTube and its subjecthood emerges through overimposition, a conflict of a few computational models.

A distributed collection of semi-structured data, which is algorithmically integrated and scaled up and down; where the connections made are a result of algorithmic, automatic functions is the material form in which culture operates today. Working at different levels, Curating Youtube deals with technocultural expression as it is specifically “databased,” while teasing out the specificity of these operations by colliding conceptual and formal gestures through the use of the grid.

Conclusion

The focus on the generic via the processes of “re-ordering of classification, naming, numbering and ordering, individuation and generalization”⁴⁹ weaves in measurement, statistics and culture, sciences and humanities, commonplaceness, mundanity and singularity, and refers to copy, uniqueness, and value. I’ve tried in this paper to reflect upon the relationships that generic has to seriality, similarity, relationality, and the operation of relating itself. Singularity and multiplicity, and the production of the subject are all touched upon if the generic is followed in-the-making.

For Kittler, in the 18th and 19th centuries, the subject - and the generic – was produced by the discourse network of cursive writing. The reflexivity of the soul at the centre of cursive writing, and of Platonic enterprise, for Deleuze, produces common sense too: “The ‘sameness’ of the Platonic idea which serves as a model and is guaranteed by the Good gives way to the identity of an originary concept grounded in the thinking subject. The thinking subject brings to the concept its subjective concomitants: memory, recognition and self-consciousness. Nevertheless, it is the moral vision of the world which is thereby extended and represented in this subjective identity affirmed as a *common sense* [italics in the original].”⁵⁰ The inextricable connection between originality and common sense today gives way to new kinds of relations, between copies, patterns, data, and computational operations, in which the generic resides.

It is now the discourse network of the Turing machine and its apparatuses – various models of databases - that contributes to the maintenance (rather than the full Kittlerian replacement) of the subject. The subject is produced not through an exercise in writing but through the relationality of data which, in the example of an art project, such as Curating YouTube, allows for a re-articulation of the database’s operability in aesthetic terms and for an exposure of its methodological dominance across culture and not only administrative bureaucracy. The schema of the unique and singular versus that of the generalized and collective is rearticulated and produces new forms of multiplicity and relationality.

The form of production of the subjectivity takes place in this operation of relationality, itself attached to the database; as the database handles both singularities and establishes relations between them as a multiple or as a ragged manifold. It is somewhere in this computational

⁴⁹ This is from a workshop on the generic, organized in December 2013 by Celia Lury, Sophie Day, and Nina Wakeford as part of the Intel project on Number and hosted by Cory Hayden at Berkeley, CA. I am grateful to Celia Lury for inviting me to speak at the workshop which helped me develop my ideas on the generic.

⁵⁰ Deleuze, *Difference*, 334.

maintenance of the operationalised singular and multiple and the movement between them that the new kinds of subjectivity are produced.

In Curating YouTube neither individual videos, all videos together, nor the grid are inherently generic, but it is in the process of articulation and relation that takes place through operating upon the YouTube database, the kind of operations made possible for the public, and the operations superimposed on those, with the grid and the curatorial gesture, that the generic, and with it, singularity and plurality, and subjectivity have their spaces enunciated. The new kinds of subjectivity and the new generic are produced as videos are 'pulled' and forced into the grid, and they arise too somewhere between the multiplicity of those videos, the databaseness of the project's operability, and the database management systems and architectural ensembles running them.