

Francis Hunger

**EPISTEMIC HARVEST: THE
ELECTRONIC DATABASE AS
DISCOURSE AND MEANS OF
DATA PRODUCTION**

**APRJA Volume 7, Issue 1, 2018
ISSN 2245-7755**

CC license: 'Attribution-NonCommercial-ShareAlike'.

Introduction

The following discussion of computational capital takes the electronic database, an infrastructure for storing information, as vantage point. Following a brief look into how database systems serve information desires, the notion of ‘database as discourse’ by Mark Poster is explored and further developed. Database as discourse establishes a machinic agency, directed towards the individual in a specific mode of hailing. This mode of hailing in turn leads to a scattered form of subjectivity, that is identified with Manuela Ott and Gerald Raunig as *dividual*. How does *dividualization* emerge from database infrastructure? What is the specific quality of data, that is produced by and being harvested from *in/dividuals* into databases, and what are the consequences of such a shifted view?

In media theory (the reflection on the historicity and modernity of media), there have been several approaches to dealing with what an electronic database is. In *The Language of New Media*, from 2001, Lev Manovich used the term database to explore the meaning of the database from a film-based and narration-oriented perspective. He uses a very broad notion of databases along their (graphical) user interfaces, basically at one point describing the complete World Wide Web as a database (Manovich 212–237). In *The World as Database*, from 2012, David Gugerli described the operation of querying as the basic user approach to databases, providing a perspective in which databases are understood as re-combination machines of the early 21st century (Gugerli 304).

Markus Burkhardt provided the first really convincing in-depth media theoretical exploration of database models, namely the CODASYL-Network Model, the relational

model and the Big Data approach in *Digitale Datenbanken – Eine Medientheorie im Zeitalter von Big Data*, from 2015. His argument explored a surface–depth metaphor, the techno-logic (apparatuses, architectures and operations) and phenomeno-logic (media practices) of databases across different logical-mathematical models, leading to a notion of database as a “cultural technique of symbolic formation *in* and *with* computers” (Burkhardt 329).

Historically electronic database technology can be traced back to techniques of collecting, sorting, saving, and exhibiting information in libraries, museums, company and government files, and similar collections. In media theory the term ‘database’ is in broad use, addressing both electronic databases and any other themed collection. A narrower definition is “an electronic database is an infrastructure for the structured storage of information.” An electronic database is a set of software applications, that – as most infrastructures – does not exist by itself, but consists of different infrastructural stratifications. It most commonly consists of a *query language*, usually oriented toward a natural language such as English and able to manipulate sets of information in the realm of mathematical-logical symbols. This query language can be embedded into a higher programming language and in host-systems, which often provide the *user interface*. A *translation* and *optimization* component transcribes queries into machine code and names of fields and tables into memory addresses and vice versa. It optimizes complex queries to reduce the amount of read-write accesses, since each memory access prolongs the period between query and reply. Another part of the system is responsible for logging, transactional security, concurrency control, and user access rights. A basic electronic database system addresses *memory*, often on hard disk, but increasingly also on

flash memory. For the argument to be made here, the term database will refer to all kinds of electronic databases, independently of the logical model (e.g. network model, relational model, graph model etc.) they employ.

However, I will shortly introduce today’s prevalent model to organize data in-formation in databases. It was developed through the 1970’s at IBM (Codd). The relational database model is based in mathematics and logic, and more specifically in set theory. Since set theory turned out to be a difficult subject for the non-mathematician, the inventors began to use the metaphor of the table. The most important feature of these tables is that the contained information objects can get inter-connected through mathematical formulas. The order of the rows is irrelevant, since each row has an address – the ID-Number, or key. If you had a simple, non-relational table containing authors and another table containing books, the problem was always, how to relate a specific author to the various books, which she or he has written, while at the same time any single book could have not just one, but many authors. Using individual IDs/keys you can realize these relations and make them available to querying.

ID	Book Title	ID	Author Name
42	Demons	214	Miller
46	Bright Light	215	Jones
48	Desire	216	Rich

Contrary to a spreadsheet the sequential order of the entries is immaterial. In this example we see two tables and one cross-table, that allows to join rows. A query for the author Miller (ID 214) would bring up the two books which s/he has written (ID 42, 46) and a Query for the book Desire (ID 48) shows that at least to authors wrote it, Jones (ID 215) and Rich (ID 216). Since each row can have many attributes, that belong to an entity (e.g. the author), data is brought in formation (hence: in-formation)

in a flat, non-hierarchical structure that is potentially endless. The quality of this model is the ability to join ad hoc, with each query, in-formation. However, it is a closed world, for which only what is included in the model, in this case the book title, the author, and the relation between both, is existing. Anything else is excluded from the database’s reality.

Historically the term ‘data base’ emerged around managerial processes in the U.S. military in the mid-20th century, namely in a symposium entitled “Development and Management of a Computer Based Data Base”, organized by the Advanced Research Project Agency on June 10–11, 1963. This might lead to a reading in which the emergence of database technology – as a specific German branch of media theory prefers to ascribe to all computational technology – is attributed to the field of war, whereby developing new technologies was part of a war effort. However, even if war provides a comparably more exciting narrative framework, it is more productive to identify the managerial, logistical dimension of information storage and processing in and with database technology, because, in large part, it reflects the broader need to deal with an ever growing amount of information in industrial capitalism from 1900 onward.[1]

Database as discourse

American media studies professor Mark Poster has authored one of the overlooked texts in cultural theory that can be productive for media theory. He develops a longer chapter “Databases as Discourse, or Electronic Interpellations” in his 1995 book *The Second Media Age*. Poster establishes a perspective that departs from 1980’s Marxian notions, which, as he puts it, were only able to address databases from the perspective of a means

of production. He seeks to complicate the Marxian dichotomies of the 'bad suppressor' on the one and the 'good suppressed' on the other hand, employing a post-Marxian understanding of subjectivity. This understanding rejects the image of a centered, coherent self to which both liberal and Marxian theories of Poster's time adhere.[2] His arguments shed light on the bio-political dimensions of databases. After examining these aspects, this paper will extend its discussion to notions of transactional meta-data production and computational capital.

Poster draws on Foucault and Althusser to delineate an understanding of the constitution of power from both action and knowledge. Poster stresses, that the analytic task should no longer be situated in action, but in language, since databases occupy the symbolic field and are *representations* of something. Extrapolating on Foucault and Althusser, he describes how language inscribes human beings with subjectivity. Language's bearing on subjectivity is also dependent on the influence of institutions. In a complex interaction, the process of *interpellations* (Althusser) inscribes ideology on subjectivity. It does so through the process of major social and political institutions "hailing" individual subjects in social interactions in a specific way. The addressee – supposedly voluntarily – subjects themselves to the internalized constraint of accepting their position of subjecthood as ascribed by the institution. However, the subjects position is never final; it stays open and up for re-negotiation and, as such, also opens the horizon for resistance and re-orientation. The hailing of the subject is embedded within a larger discourse/practice that emerges as a technology of power operating mainly through language. The establishment of this discourse/practice remains hidden from the subject and is thus prerequisite to its ability to manifest power.

From this vantage point, Poster describes the database as a "discourse of pure writing that directly amplifies the power of its owner/user" (Poster 5). Here he maintains that, in contrast to spoken language, a database is authorless in the sense that it has too many authors for their identification. Their power is mediated through the database's belonging to a specific entity, such as an institution, company, military body, library, or university. As such, this institutional affiliation produces acts of hailing the subject. Poster implies that the dissolution of authorship leads to a situation in which nobody can be held accountable for what is collected in databases and how. But authorship (or its absence) in databases does not only manifest through the collection and inscription of information. Authorship also emerges in the super-structure of a database, which pre-configures how information is stored. Authors are actively involved in all aspects of this pre-configuration, be they *administrators* who define the users' access rights; *data scientists* who sculpt the data model and allow or deny certain information to become part of the database reality and thus eventually introduce bias; *managers* who give orientation in a generalized way for the intended use; *programmers* who translate requirements into code or programming language and in this act of translation introduce their own interpretations; *engineers* who invent, install, and maintain the technological infrastructure; *user interface designers* who, on the surface, provide an interface that pre-mediate the user's intentions within the machinic depths of the computing machine where the data is stored; and *politicians* who negotiate the framework of information collection and usage in a juridical, social, and political sense.

Though I agree with Poster that the large amount of different data contributors leads to an dissolution of authorship in electronic databases, it is important to note,

that there is a limited number of identifiable creators within a respective institution. They are potential addresses of political demands that can shape institutional and discursive change.

Another agent in the field of database discourse has entered our consciousness recently – the *user*. Today with each query on a search engine, with each spatial movement (recorded by smart phones), with each act of consumption, users voluntarily and involuntarily produce data, which is *transactional meta-data*. At first glance this appears to be a phenomenon identified in the early 2000s, namely of Big Data, the promise to record *everything*. Poster reminds us that, even when credit card payment became a working infrastructure in the 1970s, tracking of (consumer) actions took place on a regular basis. The computerization of American Airlines in the 1960s gives important insight into how the use of transactional meta-data around users was deployed for an economic, competitive advantage.[3] It established a new epistemic regime that helped the company to move from yearly to monthly, to weekly, and eventually daily ticket price changes and to balance capacity utilization versus the competitors' ticket prices. (Copeland and McKenney)

In example this special mode of fragmented hailing can be observed at a driver for Uber, who has learned to trick the Uber algorithm to get more wanted drives but who is also subjected to drivers reviews, so s/he has to invest in subjectivity, such as offering a smile. Part of his/her abilities is saved in a database and can lead to different modes of hailing the particular driver (Scholz). This would be one fragmented individual part of a particular individual stored in a database. Whereas the same person on OK Cupid expresses another part of subjectivity, mostly related to sexual and relationship desires and gets addressed through this specific

partiality (Rudder). Again, the same person can be a Facebook user, getting addressed through Facebook advertisement and to Facebook is a persona with those attributes stored that make sense for Facebook in order to efficiently serve advertisements (Fisher).

Data production and extraction

The term 'surveillance' is misleading. Where liberal consciousness identifies data collection as an act of control directed toward the individual, this paper introduces the argument that data collection on the whole is not about surveillance, but about the *production and extraction* of data.[4]

What we face is a new regime of data production: a documented interpellation and recorded extraction from every participant in the social field. Each action, even the seemingly non-productive action – for instance in querying navigation systems, in acts of consumption and payment, in infrastructure usage like recording water usage in households, or even reading a book on an electronic device – has turned into an act of data production. A massive production of epistemic value has evolved from the extraction of transactional data from human actions.

In his 1489 treatise *Summa de arithmetica, geometria, proportioni et proportionalita*, Fra Luca Pacioli published a chapter called "Particularis de Computis et Scripturis" that prominently discussed techniques of double entry book keeping based on three distinct recording procedures. The first procedure was a *memorandum*, a diary with daily notes of all kinds; the second a *journal* recording single transactions; and the third a central and indexed *register/general ledger* (Lauwers and

Willekens 296–299). A succession of actions was applied to these records, effectively producing an algorithm, which ensured that each transaction was recorded *twice*. Any amount which was recorded in a specific account as debit had to be simultaneously laid out in another account as credit.

The double recording of one and the same transaction created a new semantic relation, a relation between a *periodical logic of entry and exit* and a *topical logic of goods and capital*. This meta-data production allowed for the auditing of business activities immediately and whenever needed, compared to single-entry book keeping, which may have happened monthly, quarterly, or annually. The receipt-based recording system created a paper-trace basis of trust and enabled for an increase in capital borrowing (Lauwers and Willekens; Fischer). The transaction was made explicit as subject to commercial conduct. It evolved into a datum, which produced new epistemic value.

Historically the socialization of capital was an important passage to a capitalist economy. Before this, property was bound to a single person and reflected their *individual* situation. In capitalism, capital lost its 'identity'; it became depersonalized, it became *Kapital-an-sich* (proper capital). No longer was it bound to the family bonds of the *fraterna*, rather it was collected in the new form of the *compagnia*, which functioned independently of personalized relationships. Behind the backs of the participants, a new social, abstract principle emerged: the purpose of commodity exchange was no longer the immediate consumption but the reproduction of *Kapital-an-sich*. Theoretician Michael Heinrich elaborates on this process as a specific form of movement:

The purpose of this process is a quantitative reproduction of the original sum of money. The money is not exhaustively spent. Rather it is spent in advance; it is only spent in order to subsequently acquire more of it. The value sum that executes this movement is capital. A pure value sum in itself, be it in form of money or in form of goods, is not yet capital. Also, a single exchange process does not create capital from a value sum. Only the chain of events in the exchange processes with the purpose of enlarging the original value sum creates a typical capital movement: capital is not simply value, it is self-reproducing value [sich verwertender Wert]. (Heinrich, Kritik der politischen Ökonomie 83).

The transactional recordings using double-entry bookkeeping thus enabled a complex and de-personalized commercial practice. Business knowledge, which until then was implicit and bound to a specific owner-individual then became explicit and independent of that person. This new transparency blended in with a larger trend of depersonalization of capital in the early Renaissance age. As economics professor Rob A. Bryer notes:

Every transaction can also be judged according to its effect on the rate of return on capital (profit divided by opening capital). [...] double-entry bookkeeping emerged, as capital became socialized, in response to a collective demand from investors for the frequent calculation of the rate of return on capital as the basis for sharing profits. (Bryer 114f.)

What we witness from that period on is the steady production of meta-data (in relation to profit and property) as a means of generating knowledge, which provides the individual merchant with an advantage against his competitors and which allows donors without any family relations to invest in trade.

Given the impact of transactional meta-data on economic processes from the 13th century on, the current expansion of transactional recording appears in a different light. What is currently perceived as excessive expansion of data collection (or as surveillance), is in fact the expansion of the production of a specific kind of data – of transactional meta-data. OLAP (On-Line Analytical Processing) and Big Data approaches have become these processes' machinic agents. We currently witness the early traces of another incarnation of the capitalist economy. Which new semantic relations have become established? How does surplus data change the political, social, and economic spheres? How does it change culture? How does this new epistemic quality change social mediation and media?

Scattered, decentered subjectivities

Where is 'the subject' then situated? If we follow Foucault with Poster, the subject is continuously reconstituted through acts of interpellation (hailing).

When a teacher calls upon an elementary school student to answer a question, the position of the student as an autonomous rational agent is presupposed, a position that student must 'stand into' first in order to be able to answer, in order to be a student. The

operation of linguistic interpellation requires that the addressee accept its configuration as a subject without direct reflection in order to carry on the conversation or practice at hand. (Poster 80)

Since gender, race, ethnicity, class, or other categorical distinctions may adapt interpellations, database technology is absolutely suited for inscribing difference. Consequently, since a database belongs to an institution, organization, or corporation, its discourse is able to amplify the power of its owner. But here the situation of hailing is different from that of a direct teacher-student face-to-face. "With databases, most often, the individual is constituted in absentia, only indirect evidence such as junk mail testifying to the event" (Poster 90). To this example we might add the individualized, targeted advertisement and content suggestions based on former acts of consumption, or the display of algorithmically similar content based on individuals' former choices of products.

Interpellation in this technological setting means that the subject needs to be *addressable*. There are three major forms of assigning an address in database systems. While addressing the citizen of a nation through identity cards or social security numbers has long existed as a biopolitical instrument (now simply updated to its electronic potential), two other forms of address assignment have just recently emerged. The second form of address is through the user's self-announcement "I am here", by providing a login name and password. This is usually tied to some sort of previous one-time identity check, such as verifying an e-mail address, mobile phone, credit card, or home address. The third possibility of addressing is through passive means, e.g. detecting trace information that an individuals' device provides, such as browser name and version

combined with fonts installed and websites last visited. Once the subject is addressable, the database system can hail or interpellate them into its discourse. One cannot stress enough the function of the query. While the information objects (as partial or biased as the database may lay them out) represent the potentiality of being related to each other, it is the query that updates the information request.

The subject is interpellated in a discursive way that significantly departs from the modernist notion of rational autonomy. Instead, databases construct “additional social identities as each individual is constituted for the computer”, depending on algorithms and data scientists grouping identities into sexual orientation, sexes, and ethnicity along commercial and governmental perspectives. The resulting discursive construction of subjectivity is formalized through an informational-mathematical model, along which the database is organized. It necessarily splits off non-formalized aspects of subjectivity. Subject constitution in database systems operates in a way that “refutes the hegemonic principle of the subject as centered, rational, and autonomous” (Poster 87) – the major resonating point in Poster’s text.

If subjectivity is decentered and multiplied along its fluctuating modes of access and interpellation, how has the modernist construction of the autonomous individual then shifted and changed? At first glance we can observe a duplication of the individual in the database by the way of *reconstruction*.^[5] To be more precise, we can observe the duplication of specific aspects of the individual scattered across several databases. This means, in turn, that every single database applies a different mode of hailing to the individual it references, thus constructing a scattered multiplicity of parts of the individual. Neither the database owner nor the individual knows which part of subjectivity the particular

database has saved. Therefore, it is a decentralized, fragmented, potentially always combinable tool of biopower concerning the subject, driven by *computational capital* – the control over resources of computation and transactional data. Poster argues, that the individual of modernity was conscious of their own self-constitution. Now, he asserts, “subject constitution takes an opposing course of ‘objectification’, of producing individuals with dispersed identities, identities of which the individuals might not even be aware” (Poster 93). The fact that our bodies are always connected over networks to databases calls for another politics of the body; a body that no longer can hide from the public eye in some private mansion and that no longer is able to leave the regime of production by attending a place called ‘leisure time’.^[6]

Dividual praxes

Manuela Ott and Gerald Raunig have recently proposed to use the term *dividual* for better grasping the scattered, fragmented individualities.

From Ott’s perspective automated systems of suggestion and preselection^[7] cause a passivation of the subject. In turn this passivated subject tries to compensate through membership in multiple virtual communities, through participation in different platforms, and through the re-distribution of their manifold expertise. The dividual participates and actively and intentionally decenters the user-subject actively and intentionally. The constant hailing for participation by database-driven platforms, according to Ott, causes the expense of time and occupation of proficiencies, which in their intensity can only be described as ‘addiction’.^[8] It leads to growingly dividual identities that become “consciously and subconsciously connected

with such multiplicities and co-created by such multiplicities, that the shape of one's own individuality is less and less perceivable" (Ott, *Es lebe die Dividuation!* 4).[9] The dividual thus appears as the reflection or re-investment in time and attention of the other platform participants. Since many platforms offer their services free of charge, but still need to be profitable, it is a necessity for them to address those dividual aspects that appeal to consumption and monetization. They thus foster a world view where the economic exchange between bodies is the preferred mode.

Gerald Raunig describes platforms such as Facebook as shaped by an expressive practice of confession. This self-expression turned self-propagation is fed by the desire for visibility (originally a sign of the desire for sociability), which again brings the private into the defensive. Sharing as a mode of existence bans the danger of invisibility. According to Raunig in many cases the developing relations can't be called *social*, rather the social appears in the negative: shit storms, revenge porn, fake news.

The concentration of Raunig and Ott's arguments on entertainment platforms may however lead to one-sided conclusions. If we equate Poster's 'scattered subjectivity' with the 'dividual', and recall the discussion about the production of transactional data, there are many more fields where to observe the dividual more closely. The dividual does not only appear during the use of entertainment platforms, such as Facebook, Twitter, Instagram, and such. The database discourse/practice also interpellates the dividual when working in such different sectors like health care, logistics and delivery, industrial production, design, software programming, management, insurance, or when studying.

Interestingly Raunig turns to the database when explaining the new mode of hailing the dividual. Big Data, the collection

of massive data sets relating to everything, shows little interest in the individual (aka surveillance) or in a totalization of data, "rather more in utmost floating and detailed records, which it [Big Data] can traverse dividually – as open immanence field with potentially infinite extension. The vast amounts of data aim to create an epistemic horizon, which depicts the complete past and present and thus tries to catch the future as well" (Raunig 160). Governmental actors try to reduce risk in the future by detecting deviations from expected mass behavior and base decisions of how best to police it. Commercial actors aim to minimize market risks in general and to optimize the consumption potential by hailing the dividual. While the governmental aim seems to be situated closer to 'surveillance', the economic aim fosters and makes use of the production of transactional data.

Computational Capital as transformation belt

Computational Capital means the disposition over data and computing infrastructure. Computational capital aims at generating epistemic value in a specific form that is translatable into economical capital.[10] Akin to the medieval merchant's double entry bookkeeping practice, computational capital makes use of an epistemological practice – the ability to record transactional data and act upon the information generated from that data.

Computational capital has grown from a historical movement that for centuries has been closely tied to human computing. For only in the mid-20th century, when demand for information processing became pressing during industrial capitalism, were machines invented to do the calculation. It is often

overlooked that modelling these calculation machines followed the paradigm of human calculation. This was specifically oriented along the segmentation of labor, both physical and mental, as Adam Smith describes in *The Wealth of Nations* (Smith 12f.). “[P]hysicists and electrical engineers had little to do with the invention of the digital computer – the real inventor was the economist Adam Smith, whose idea was translated into hardware through successive stages of development by two mathematicians, de Prony and Babbage” (Simon and Newell 2).

The electronic computer became a machine capable of processing information. At the same time this machine has no understanding of the meaning of what it processes. “In the depths of digital media technologies lies, however, no natural truth, but an invisible machine calculating signals” (Burkhardt 81). Computational capital insofar only extends to the *computable* – that without meaning. This suggests that the human ability to grasp meaning cannot be separated from computational capital. Computational capital is able to work only when humans produce expressions that can be made symbolic and processed, and only when humans set up rules (algorithms) as to how the machine shall record, process, and store these symbols.

At this point it is necessary to differentiate between several kind of data in regards to, how it enters into computational capital. 1.) *Data Production* implies an active deed of creating new artefacts by combining machine or human labor with the transformation of matter. It can be observed directly in the practice of self-quantification, e.g. the measuring of a person’s daily itineraries with a ‘health’ app, that calculates the calorie consumption from it. Data production occurs when users upload original content, i.e. their images to Flickr *and* tag it, so the abstracted data can be further used.[11] 2.) *Data*

Extraction or Harvesting invokes the notion of rent, similar to the profit that is extracted from land and real estate ownership.[12] Experimentally this could be called *epistemic rent*. In platform capitalism it differs from the classical example of land or real estate rent. User-created posts on a communication entertainment platform, users’ comments to a newspaper website, user-generated imagery added to a navigation application create value that is *indirectly* extracted after recording transactional meta-data associated with the subsequent content consumption by others. This form of extraction draws on novelty to maintain a stream of inter-actions that translate into transactional meta-data flowing. Involving databases, it appears as involuntary by-product to cultural creation, or more generally spoken, to human communication and interaction. Extraction also takes place, when actions, mostly in the commercial area, get digitally recorded as meta-data and produce an epistemic rent reminding of double-entry book keeping. Examples for the harvesting of transactional meta-data are platforms that broker services between different users, such as eBay or Airbnb. New with them is, that acts of exchange, which formerly have not been subjected to database record-keeping on a massive scale, such as selling used clothing or subletting an apartment, now become formalized and recognized as micro-economic transactions. Computational capital, however, mixes here with more traditional economic strategies, for instance in the charging of a service fee.

This emerging mode of production and extraction of data is dependent on database infrastructure and would not work without it.

In order to further progress this argument, with the above differentiation of data being tentative at least, I temporarily resort to the use of metaphors. If data is a raw material and information is a product, then it needs human labor and machine labor to

transform one into another. Like the steam engine to the spinning jenny, computational capital (calculation machines and the knowledge how to apply it) adds *productivity*. In contrast, human labor embodied in data in microscopic doses adds *value*. Data labor is barely visible from an in/dividual perspective and goes largely unacknowledged, because of the microscopic and fragmented nature of each information object for which it has been spent. This may be one of the reasons why public discourse is aligned along notions such as the most current technology (Blockchain, AI and Big Data), or questions of privacy, and not along the labor aspect. However, once it has been accumulated in large databases and been harvested using algorithms, it creates epistemic value for the owner. Human subjectivity then is the soil on which the corn of the 21st century grows. Computational capital is the machinery to harvest the corn, and just when the collected grain has been grinded in the database mill, it has the potential to be sold (that is to realize its value) or to be invested.[13]

I have explored, how the discursive power of database systems lies in their ability to interconnect pieces of information, put them in relation to each other and constantly re-arrange this epistemic arrangement according to a query and it shows, that *querying becomes a dividual practice itself*. Querying here takes on a double identity: The query is a discourse/practice in itself and translates into power. Yet since it often produces transactional data, the query is at the same time subject to discourse/practice and thus power.

From this perspective, not being subject to the recording of transactional data may be a strategy of empowerment. There have been quite a few attempts to empower the in/dividual that is subject to transactional data extraction. Hacking, proxying, digital detox, pattern-smudge, these are all strategies directed towards an in/dividual solution of a

problem that is perceived as 'surveillance'. Shifting the perspective from surveillance towards *transactional data extraction/production in electronic databases* shifts the thinking around countering strategies, because it shifts the perception of the problem from an in/dividual to an institutional one. This shift may turn out as relevant to policy and activist actors who deal with issues such as privacy, data in general, and data production in particular. It also re-positions the thinking about electronic databases towards a tool of machinic agency, infrastructurally embedded into institutional and organizational contexts, which are far from unalterable. They can be politically addressed and challenged. However, in the course of my research it became obvious, that the state of data from the perspective of epistemic value is precarious. Further research into these modes (production, extraction and possible others) and the notion of rent is necessary.

I have shown that demystifying databases means interpreting them as institutional or organizational tools of hailing, addressing, agency and data production. Databases and algorithms are not first and foremost technology, they rather represent human ideas about potential (inter)actions. Databases amplify institutional power, since they are able to address the dividual on an individual level. They do so based on the transactional recordings of former acts of the addressee. Databases make up an infrastructure for the recording, extraction and production of data and meta-data, transforming human interactions from a perspective that seeks to generate epistemic value. A critique of Database systems – understood as a set of agency praxes – does not begin with the demand for privacy or the deletion of data. It begins with addressing the query and its institutional context, which represents the shaping of an informatory request as a dividual practice.

Notes

[1] For instance, one of the papers at the 1963 data base conference stresses its objectives: “1.) Meet manpower requirements with personnel; 2.) Maximum utilization of skills; 3.) Improve career management; 4.) Interrelate personnel activities” (Swanson 2).

[2] For the sake of his argument, Poster simplifies the matter. At the time of his writing, more complex subject constitutions were already available, both in post-Marxist theory, such as in Deleuze/Guattari or the Post-Operaists, and in liberal theory such as the Actor-Network theory of Callon/Latour.

[3] An in-depth discussion of the notion “transaction” and how it is embodied in databases, is developed as a chapter in the author’s Ph.D. thesis. Castelle discusses transaction in relation to database technology (Castelle).

[4] The framing of data recording as surveillance is a *strong narrative*, adhering to libertarian ideology, both in theory, pop culture, and politics. It provides a vulgarized, digestible explanation on an individualistic, narcissist level for the black-boxing of database systems, or more broadly speaking, of calculatory infrastructure. Of course, surveillance takes place as policing different societal levels and in that form has become a biopolitical practice (Foucault).

[5] On entertainment sites such as Facebook, Twitter, or Instagram, fake profiles are worth significantly less than identity-verified, aged profiles with three or more years of online activity.

[6] Leisure time today has turned into an intensified period of transactional data production when using any kind of electronic networked devices for entertainment media consumption (Facebook, Twitter, Instagram, YouTube, Pornhub, Tinder, Netflix) and other recreation that involves acts which can be electronically recorded.

[7] For instance, this includes the search field that makes proposals for how to complete a query, the auto-completion function for typing in smart phones or product suggestions derived from former purchases.

[8] Notorious for this addictive behavior is the attention seeking ‘dark design pattern’ of a red circular surface with a continuously updated number, which signals the number of unattended messages. The darkness of this design pattern lies in the fact that each attended message leaves a data trace of either being ignored or deleted or of new activity generated when one replies.

[9] Actually Ott’s impressive development of the notion of the »dividual« is much more complex. Departing from Spinoza and Deleuze she discusses dividuations as a theory of participation from several perspectives such as bio-technological, socio-technological and aesthetic-artistic dividuations. (Ott, *Dividuationen: Theorien der Teilhabe*)

[10] The notion here is developed in a similar way to Pierre Bourdieu’s social and cultural capital, which builds on Marx’s notion of capital (Bourdieu). To denote capital related to the economic sphere and differentiate it from cultural, social, and aesthetic capital, the notion of ‘economic capital’ is used in relation to Bourdieu, but as reconstructed by Michael Heinrich and Moishe Postone. (Heinrich 1999; Postone)

[11] By tagging an image, users ascribe meaning. That's something the machine can't do by itself. Machine learning training sets like *Faces in the Wild* are composed from these open sourced tagged images.

[12] An indication of the rocky theoretical territory discussing epistemic 'rent' that lies ahead can be found in (Haarmann) and (Fisher and Fuchs).

[13] This analogy throws further questions regarding the legal status of data. While the ownership of land, which was established centuries ago by dispossessing the commons through the Aristocracy, under current conditions is secured juridically through proprietary law, individual data is only juridically addressable through privacy law.

Works cited

Bourdieu, Pierre. "Ökonomisches Kapital, kulturelles Kapital, soziales Kapital." *Soziale Ungleichheiten*. Hg. von Reinhardt Kreckel. Sonderband 2. Göttingen: N.p., 1983. S. 183–198. Print. Soziale Welt.

Bryer, R.A. "Double-Entry Bookkeeping and the Birth of Capitalism – Accounting for the Commercial Revolution in Medieval Northern Italy." *Critical Perspectives on Accounting* 4.2 (1993): 113–140. Print.

Burkhardt, Marcus. *Digitale Datenbanken – Eine Medientheorie im Zeitalter von Big Data*. Bielefeld: Transcript, 2015. Print.

Castelle, Michael. "Relational and non-relational Models in the Entextualization of Bureaucracy". *Computational Culture* (2013): n. pag. Web.

Codd, Edgar F. *Derivability, Redundancy, and Consistency of Relations stored in Large Data Banks*. San Jose/California: IBM, 1969. Print.

Copeland, Duncan G., und James L. McKenney. "Airline Reservations Systems – Lessons From History." *MIS Quarterly* (1988): 353–370. Print.

Fischer, Michael J. "Luca Pacioli on business profits." *Journal of Business Ethics* 25.4 (2000): 299–312. Print.

Fisher, Eran. "Audience Labour on Social Media: Learning from Sponsored Stories." *Reconsidering value and labour in the digital age*. Hg. von Eran Fisher und Christian Fuchs. Palgrave Macmillan, 2015. Print.

Fisher, Eran, und Christian Fuchs. *Reconsidering value and labour in the digital age*. Palgrave Macmillan, 2015. Print.

Foucault, Michel. *Überwachen und Strafen – die Geburt des Gefängnisses*. Übers. von Walter Seitter. Frankfurt am Main: Suhrkamp, 1976. Print. Suhrkamp-Taschenbuch.

Gugerli, David. "The World as Database – On the Relation of software Development, Query Methods and Interpretative Independence". *Information & Culture* 47.3 (2012): 287–311. Print.

Haarmann, Petra. "Copyright und Copyleft – Vermittlung im Falschen oder falsche Unmittelbarkeit." *Exit*. Berlin: Horlemann, 2014. Print.

Heinrich, Michael. *Die Wissenschaft vom Wert – Die Marxsche Kritik der politischen Ökonomie zwischen wissenschaftlicher Revolution und klassischer Tradition*. Münster: Westfälisches Dampfboot, 1999. Print.

—. *Kritik der politischen Ökonomie – eine Einführung*. Stuttgart: Schmetterling-Verl., 2005. Print.

Lauwers, Luc, und Marleen Willekens. "Five hundred years of bookkeeping – a portrait of Luca Pacioli". *Tijdschrift voor Economie en Management* 39.3 (1994): 289–304. Print.

Manovich, Lev. *The Language of New Media*. Cambridge, Mass.: MIT Press, 2002. Print.

Ott, Michaela. *Dividuationen: Theorien der Teilhabe*. Berlin: b_books, 2015. Print.

—. *Es lebe die Dividuation! Zur Notwendigkeit anderer Denkkonzepte angesichts zeitgenössischer Teilhabepraktiken*. 2016. Berlin.

Poster, Mark. *The Second Media Age*. Cambridge, UK; Cambridge, Mass.: Polity Press; B. Blackwell, 1995. Print.

Postone, Michael. *Zeit, Arbeit und gesellschaftliche Herrschaft*. Ca Ira, 2003. Print.

Raunig, Gerald. *Dividuum. Maschinischer Kapitalismus und molekulare Revolution*. Bd. 1. Wien: Transversal Texts, 2015. Print.

Rudder, Christian. *Dataclysm: who we are when we think no one's looking*. New York: Crown Publishers, 2014. Print.

Scholz, Trebor. *Overworked and underpaid: how workers are disrupting the digital economy*. Cambridge, UK ; Malden, MA: Polity Press, 2017. Print.

Simon, Herbert A., und Allen Newell. "Heuristic Problem Solving: The Next Advance in Operations Research". *Operations Research* 6.1 (1958): 1–10. Print.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London: W. Strahan & T. Cadell, 1776. Print.

Swanson. *A Computer-Centered Data Base Serving USAF Personnel Managers*. Santa Monica, CA: System Development Corporation, 1963. Web. 13 Okt. 2017.