



Upon Not Opening The Black Box

Simy Kaur Gahoonia

IT University of Copenhagen
Copenhagen, 2300, Denmark
sgah@itu.dk

**Pedro Ferreira,
Marisa Cohn,
Line Henriksen,
Katrine Meldgaard Kjær,
Michael Hockenhull,
Baki Cakici,
Marie Blønd,
Rachel Douglas-Jones,
Cæcilie Sloth Laursen,
Sonja Zell**

IT University of Copenhagen
Copenhagen, 2300, Denmark
{pedf,mcoh,lihe,
kakj,michh,bakc,mblo,
rdoj,cael,soze}@itu.dk

Abstract

On the eve of the General Data Protection Regulation (GDPR) coming into effect we, a university laboratory, marked the occasion with an interactive installation called Compliance. Data traces from Compliance were subsequently processed by the lab, here enacted in the form of a play. While much discussion has centered around modern 'black-boxed' processing of data, less attention has been paid to the value of the data itself, and whether it merits use. We draw on dramaturgical methods for both analysis and presentation [15], allowing for readers to imagine staging their own, different, versions of the event. Drawing on the ambiguous ontological status of (yet unexamined) data, we offer a discussion on the value of data, its use and non-use, as well as how to live with this ambivalence, continuously negotiating social contracts about our further conduct with the data.

Author Keywords

gdpr; data non-use; dramaturgical methods

CCS Concepts

•Human-centered computing → HCI theory, concepts and models;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

© 2020 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6819-3/20/04 ...\$15.00.

<http://dx.doi.org/10.1145/3334480.3381813>



Figure 1: The Compliance installation consisted of an inflated plastic igloo bubble arranged in the public atrium of the university. Compliance emphasized the experiential qualities of becoming a "data subject." "Anyone becomes at some point a data subject... [when] they disclose some personal data" [1]. Compliance thus staged an extended moment of disclosure and discretion of the data subject. An act of deletion subtracts a piece of personal data, diminishing what is subject to compliance, but an act of writing down a trace of that deletion creates new kinds of data that are ambiguously in relation to compliance. The bubble was a private space (only one person could enter at a time) but transparent and on display. Many people approached and asked about the installation but only few entered. Photo: © ETHOS Lab.

Character gallery¹

LAB MANAGER; DATA CONSULTANT, DATA ANALYST, LAB DIRECTOR, LAB MEMBERS 1-4,²

Act I: Unboxing the Black Box of Data

A university laboratory in the age of Big Data³. A dimly lit room, a disembodied voice of a female IoT device can be heard quietly reading the text of the General Data Protection Regulation (GDPR), fading to silent as the show begins. The members of the laboratory are invited to gather around a meeting table upon which, spotlight, sits a sealed black box. A lab manager speaks to the room who have convened to unbox the data collected within.

LAB MANAGER: Welcome. You see here the black box containing the data traces from our Compliance event. Compliance was an experiment the lab did over a year ago around the launch of GDPR. We had long planned to make an intervention into the ongoing GDPR debate, so people who were members of the lab at that time, some of which

¹While HCI researchers have used methods from drama and performance to explore use-scenarios for design of interactive systems, here we applied such an approach to explore our own uses of data. We adopted a drama workshop format to analyze (or not) the data collected from the Compliance installation. As Mehto et al [15] note, this format allows participants to playfully embody positions within a problematic scenario through metaphor and fiction, enabling controversy to emerge without necessarily aligning to the stakes of the outcome.

²During the workshop we (authors 1, 2 and 3, playing the roles of lab manager, data analyst and data consultant) created a narrative structure of three acts which shifted the group towards dramatic conflict between the imperative to open the box and the stakes of (not) analyzing the material within. Each act began with a dramatic reading followed by improvisation by the workshop participants. The characters in the play below are fictional pastiche characters distilled from the workshop that bear only passing resemblance to actual lab members.

³in the age of Big Data': the term 'data' bears a very specific meaning. This is a kind of data whose value is often presented as self-evident, an unavoidable step towards progress where extractionary metaphors abound (cf. the new 'oil', the new gold, etc... [17])

are in this room today, conceptualized an interactive installation around the theme of compliance, that differed in tone from the institutional communications focused on compliance and legal considerations surrounding GDPR rollout.

Among the group, some nod. They seem to recall the event.

LAB MANAGER: The installation we thought up was an igloo-shaped plastic bubble, blown and held up by an air pump. In it, we fit a couch, table and a speaker reciting the GDPR text with a soft voice, creating a meditative atmosphere in the bubble, where one was encouraged to contemplate one's relationship to data.

LAB MANAGER: It was put up here in our own public university. And in order to enter this coveted space of compliance, you had to delete a personal data item and document this deletion by writing it down on a piece of paper provided and slot it into the box.

LAB MANAGER: The installation was taken down after the event. But it produced traces (the box and contents). I have held on to this box for so long with the idea that we would do something with it and the data it contains. So now, we are gathered together to decide what to do with this data. Are there any questions?

Act II: Upon Not Opening the Black Box

The group is found seated around a table already in the midst of heated discussion about the nature of this unboxing. A member of the audience reveals themselves to be a visitor from the Danish Data Inspectorate.

DATA INSPECTOR: It has come to our attention that your lab has come into possession of this black box. Let us begin by considering what we already know. While there was personal data involved at the event, this was deleted, and

what you have in the box. Well, I highly doubt whether what you have is in fact of meaning whatsoever. To be honest, this object stretches the imagination of the inspectorate to the farthest consideration of our role. Nonetheless, a certain kind of . . . presence . . . might take shape. I have heard that this lab uses unusual methods for writing [9, 10, 16], collage [5, 7], vignette [12], ways of acting at a distance upon material known as data. (She hands out a form) To help you ascertain the nature of what is in the box and how to handle it, I have provided you this form. (See figure 3). *(The group begins to peruse the form as the data inspector reads it aloud)* Discovery: How do we first determine if we are indeed processing data? Is this data? In what way is it data? How is it data? How do we know? On what basis do we possess this data? Datafication: Have you been in proximity to the data? Have you considered how it was collected, stored, retrieved, transmitted, compiled, deleted, or annihilated? Processing... *(She drones on. The lab members sit quietly filling in the form.)*

Simply asking these questions has a way of making what is in the box into data. Filling out this form, new traces are made, "traces being folded back into data, an endless cycle of 'informating'" [18] (citing [8]). Datafication is not merely a step in the evaluation of compliance but a becoming.

DATA INSPECTOR: *(Continuing her explanation of the form)* When do we "process" personal data [4]? How do we make ourselves aware that we are processing data? The idea of personal data is so broadly defined that you are perhaps processing personal information about others as you sit here looking at this box. Just by being in proximity to it. Certainly if you have touched it. *(She pauses, troubled)*.

The room is silent, the group hesitates, mumbling questions half to themselves and to each other. Do we have consent? Is there any value to looking in the box? Under what condi-



Figure 2: The workshop in progress. The lab and its members, consisting of researchers and management staff, met up on November 27, 2019 for an 'unboxing' workshop, in which they took action on the box containing trace data from the Compliance event. The discussion that took place during the workshop is presented in this paper in the format of a play. In this image we can see the setup of the discussion described in ACT II. Photo: © ETHOS Lab.

tions might we look? Maybe we are here to ask whether to open the box at all.

LAB MEMBER 3 *(Becoming frustrated with the group debate, interrupts):* We should get rid of it. The fact that we have held onto this object at all was an error, an oversight. We asked people to delete something, and we have accidentally sustained a trace of those deletions. It is not data that has been collected, it is merely data that has not been fully deleted [8].

How do you feel about the material and its potential use as data?
Excited, inspired, irrelevant

What might be inside the box: (check all that apply)

traces <input checked="" type="checkbox"/>	meta-data <input checked="" type="checkbox"/>	rubbish <input checked="" type="checkbox"/>
inscriptions <input checked="" type="checkbox"/>	intentions <input checked="" type="checkbox"/>	garbage-in-garbage-out <input checked="" type="checkbox"/>
discards <input checked="" type="checkbox"/>	raw data <input checked="" type="checkbox"/>	data trash <input checked="" type="checkbox"/>
data sweat <input checked="" type="checkbox"/>	media <input checked="" type="checkbox"/>	toxic data <input checked="" type="checkbox"/>
by-products <input checked="" type="checkbox"/>	nothing <input checked="" type="checkbox"/>	detritus <input checked="" type="checkbox"/>
personal data <input checked="" type="checkbox"/>	echos <input checked="" type="checkbox"/>	sheddings <input checked="" type="checkbox"/>
fragments <input checked="" type="checkbox"/>	monsters <input checked="" type="checkbox"/>	harvest <input checked="" type="checkbox"/>

other: *paranoid*
 other: *release*
 other: *secrets*

Figure 3: Excerpt of the form used by the DATA INSPECTOR.
 Photo: © ETHOS Lab.

LAB MEMBER 1: Why are we even discussing whether we can look at the material? Do we not need a purpose?! (LAB DIRECTOR: For the "Pursuit of Knowledge"?) But what is that? So, what, just because we call ourselves researchers it is legitimate for us to look inside even if we do not know what is there or towards what end we are looking [11]? How do we determine if this is worthwhile?

LAB DIRECTOR: It seems we need to first clarify the context of the installation to determine the expectation participants had about what was collected. Did they have a reasonable expectation of privacy [6, 14], or that this material would be reused? It was collected in a public space.

LAB MEMBER 1 (*frustrated*): There is a difference between having a reasonable expectation that the material collected would be used in some way and, and seeing it show up in an academic paper!

LAB MEMBER 2: So little is known about what is on the pieces of paper in the box. We do not know that what is on the cards constitutes personal data in the first place, the kinds of data to which such GDPR compliance rules apply.

We cannot know the degree to which it is personal data unless we look. (*He leans the box towards himself, peering inside.*) We must open it to know how we might use it.

LAB MEMBER 1: Why should we put things to use? Is our curiosity enough? Just because we are researchers in such an... an institution?! (*She waves her arms, indicating the university building.*) Let us be real. What we are talking about — masking as curiosity and pursuit of knowledge—is the making of academic publications. Something that is a product that is valuable to us, but not to others.

LAB DIRECTOR: But this depends on what the paper is about. What if the paper is not about "them", the participants, but is rather about this process, or about what we find out as we open and make sense of what is there.

DATA INSPECTOR: (*Trying to break the silence*) Does the uncertainty [3] about what is in the box affect our ideas of what to do with it? Does it prompt our curiosity further, does it inflate its sense of value, does it release us from responsibility, does it require us to do nothing?

LAB MEMBER 3: (*Emphatically*) But we DO have an idea of what is on the papers! I attended the event. But even if I had not we can reasonably conjecture what is on the pieces of paper: "I deleted an email from my boss", "I deleted a text from someone who ghosted me". They are irrelevant. We should destroy the box. If we know the material is not useful, we have no good purpose for holding on to it. We have in fact held onto it far too long already.

LAB DIRECTOR: (*Gently.*) So then, you are suggesting that our knowledge of the local situation and context trumps the general status of uncertainty and rational position on knowledge production? It seems we have a problem of memory, memories of data, memories of data collected that conflict,

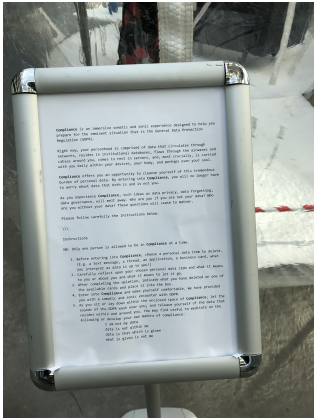


Figure 4: The instructions provided to those entering into Compliance. Photo: ©© ETHOS Lab.

⁴ different stories of the collection and its conditions. We have people present here, who participated in the event, we can ask about consent. (*Turning to a member of the group, who participated in the Compliance event*) You, do you give your consent retroactively for us to use your data? (*Everyone turns their attention to the lab member being questioned. He becomes uncomfortable. The lab director redirects*) There are qualifiers to consent. We can, as researchers, working in the public interest, work with the material without consent [6].

LAB MEMBER 1: Who are we to engage with this data? Those who participated were unknowing that we would look at it. That was not clear! (LAB MANAGER: But it was voluntarily given.)

LAB DIRECTOR: Or was it? It was part of a transaction, an exchange, a payment demanded. In order to enter you must produce some text, you must create some data.

LAB MEMBER 2: This is the social media loop hole — in order to enter you must comply, you must give data.

LAB DIRECTOR: Is this a form of re-use? It has been so long since the original event. If this constitutes re-use then consent was not given.

LAB MEMBER 1: We have been in error to keep this material for so long. We are mandated to delete as soon as

⁴Loukissas calls for awareness to data's "enduring local ties" - to the local context, rituals, and cultural practices that shape data in its sites of collection. While he speaks of big data, we can see that even when data has a single origin in a particular local event, that there are nonetheless "a range of possible local ties" and with them differing expectations and imaginations of what may happen with this material. Such a staging of prolonged unboxing helps to surface "conflicting practices of data production" and explore "alternatives to universalizing discourses of... data" [13]

possible unless there is a specific purpose. [6] This material has had no specific purpose since its inception.

LAB MEMBER 4: Except perhaps to be lying in wait.

LAB MEMBER 2: But we do not know if it is data. For all we know, they wrote nothing.

LAB MEMBER 1: We do not know, but what might we know? Is it sensitive, is it identifiable, did they sign it? They could have put their full name, address.

LAB MANAGER: Can I just do a count?

LAB MEMBER 1: But why?

LAB MANAGER: I do not know. I want to know how many people were there, how many attended.

LAB MEMBER 4: This is Schrodinger's data!

The lab manager turns aggressively to the data inspector.

LAB MANAGER: And you! You are collecting more data, meta-data, new data about the data about the data!

LAB MEMBER 1: This reminds me of the archive I am working with. I spend so many hours with the legal department here at the university discussing how to approach an archive if I do not know what is in it until I look. Internet archives are like that. We do not really know what they contain. We have to make a set of agreements about what we will do with the material before we look into it. Everything could be in there. We must act as if everything could be in there. This is an ever-present tension: how we can know without looking, and how can we look without knowing.

LAB DIRECTOR: If this material speaks to the soul of the participants, maybe we do not address each piece one by

one, but only as an aggregate. How might we look in the aggregate without looking in the particular? Could we count without reading? Could we read without analyzing? Could we sample without knowing the whole? (LAB MANAGER *(interjects)*: But to what end?)

LAB DIRECTOR: What we have is proof that we have undertaken an act of deletion. We needed to collect this material in order to be able to give auditable proof of deletion.

LAB MANAGER: It seems like we come to an impasse that draws out the sense that we cannot agree. Some of us will be made more uncomfortable by one option or another. We are not a uniform group. At the same time, we can make an agreement about what we will and will not do with it, to handle our different sentiments of discomfort.

LAB MEMBER 4: Is the box itself not a piece of data... (one that we have already been handling, processing, interpreting, analyzing) should we not focus on that? It is abject, liminal, and has a life before this project. Does it get to have a life after the project? Do we not owe something to it?

LAB DIRECTOR: Could a research paper be its afterlife [2]?

LAB MEMBER 1: But what kind of afterlife is that? One where we profit from its destruction? It has been established by now that if we empty it of all the data, it loses its life, its status and aura.

LAB MEMBER 2 knocks the box to the ground. Everyone gasps in horror at the transgression.

Act III: An Afterlife for the Black Box

Enter a data specialist, a consultant the lab has hired to help open the box and analyze the data.

DATA CONSULTANT: (*Arrogantly*) Your discussions do not matter much to my practice. I extract value from data and I can assure you that you are deeply misguided. Using data is not an option, it is a moral imperative. I see you are failing to make progress so I am here to assist you. How would you like to proceed?

LAB MEMBER 4: The box itself has gained mythic status. Its contents have not changed but our perception of the box has changed.

DATA CONSULTANT: The box? With data? Do you mean the database?

LAB MEMBER 4: Perhaps we should not try and resolve the ambivalence that this material raises for us? Maybe we should engage it? Write a paper! What if there is a moral value, a benefit to society, of exploring data that is given under ambivalent conditions?

LAB MEMBER 1: But we should not profit from this data. Not even academically!

LAB MEMBER 2: But publications are not just currency, it is a way of forming collectives and collective understanding. Would this not be a way of doing justice to the material inside? To talk about the GDPR as a continuum in a way that is of no expense to the participants.

LAB MEMBER 1: Yes, I agree, but there are aspects of the ambivalence that we cannot consider if we do not open it. That is, have we exhausted already the value of this experiment by discussing whether to open it? Are there justifiably any expectations that we might find more ambivalence by reading the material within? Are we doing enough justice to the material by leaving it unread?

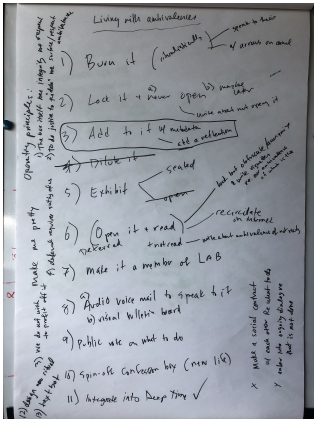


Figure 5: This image shows the board where the different possibilities were plotted, discussed and eventually ruled out. Photo: © ETHOS Lab.

The sense of a need for agreement is present in the room. We are at a standstill. Because some of these possible actions are mutually exclusive. We do not want to simply defer the decision. Today is the time for a decision. Yet we cannot decide.

LAB MANAGER: We are stuck in a contradiction here. What is of value here? Is it the box? Or is it what is inside the box? The box has been a disturbing presence in the lab, we cannot perpetually defer this decision. We need to consider our sanity!

LAB MEMBER 4: Let us consider that what is inside is connected to people’s soul. Those who have confessed—it has been severed from them and we could work against anonymization to reunite the souls with the data that has been severed from them. Fix what has been broken. A re-linking of souls.

LAB MEMBER 1: But if we did that the box would still cease to be. Would that not be avoiding the real issue: the difficulty of living with ambivalence? The box is in fact neither open nor closed and it is affectively pulling us to open it... by not opening it we leave the ambivalence and that affective attachment in place. How can we do justice to that living with ambivalence?

LAB MANAGER: We should defer the opening for another year. Not as a form of inaction, but as a deliberate act. We add vignettes to the box! Date these and wait for another year, if we decide to open the box then we can use them to reflect back on these same issues. What has changed?

LAB MEMBER 4: But we must not just put it aside. It is like we have been reacquainted with an old friend. We cannot just forget them again. If we are going to defer opening

the box, how do we make sure that we stay connected and linked to it?

LAB MEMBER 3: Can we put the box on display such that people can add to it. Should a sign be placed by it that says "this box has never been opened"?

Concerns emerge on whether there should be a new box. A confessional or a shredder outside the lab. Slowly there is convergence towards keeping the box open for additions. A lab member makes a list of the different possibilities (cf. 5).

Coda: Living with ambivalence

The group examines the list of options and feels the pressure of the meeting time running out. It would be a failure to simply defer decisions and leave the box in a limbo state of waiting for action. Yet among the various options presented are some which are mutually exclusive. The group realizes that what is important here is not the data itself but the social contract they have begun to build together, an ongoing conversation that is not done, yet will proceed through a set of operating principles: 1) The box itself (here representing the unexamined dataset as a set of conditions rather than just what can be extracted/mined from it) has integrity that our process should respect 2) To do justice to this data we choose to surface and live with the ambivalences raised by leaving it unexamined, and 3) Deferral requires something of us — to live with the ambivalence of unexamined data we defer deletion/destruction (or use) of the data but we also embrace deferral as a deliberate act rather than a form of doing nothing.

Our work reflects on the grayzones and ambiguities of deciding to use, or not, data, questioning imperatives of data-centric approaches to research. It shifts these discussions to ongoing processes focused on doing justice to data with deliberate consideration of its potential, or lack thereof. We

discuss how deciding on data use, and therefore consent, is not a linear process that can be decided *a priori* in the light of the GDPR. Instead, we can best think of these decisions as establishing social contracts and living with ambivalence. Working with dramatization as analysis, helps surface these concerns, bringing to light the fact that our decisions to make use of data are always in relation to its non use.

Finally we connect data as an unfolding event with the format of the play as presentation, and the idea of the script and stage instructions, as something that allows for the data to continue to unfold by allowing others to repeat what we did, albeit in different ways each time. In our case, we decided to leave the box on display, offering a way for participants to add their reflections on living with this material in its ambivalent, unopened, state. We offered a way for participation without diluting the material inside thus integrating its prolonged temporality back into the life of the lab. We invite others to do differently.

Acknowledgements

We would like to thank the ETHOS Lab at IT University of Copenhagen and all those who participated in Compliance.

REFERENCES

- [1] 2019. What is a data subject? (2019). <https://eugdprcompliant.com/what-is-data-subject/>
- [2] Christine L. Borgman. 2019. The Lives and After Lives of Data. (May 2019). DOI: <http://dx.doi.org/10.1162/99608f92.9a36bdb6>
- [3] Cheryl Cooky, Jasmine R Linabary, and Danielle J Corple. 2018. Navigating Big Data dilemmas: Feminist holistic reflexivity in social media research. *Big Data & Society* 5, 2 (July 2018), 2053951718807731. DOI: <http://dx.doi.org/10.1177/2053951718807731>
- [4] Alison Cool. 2019. Impossible, unknowable, accountable: Dramas and dilemmas of data law. *Social Studies of Science* (May 2019), 0306312719846557. DOI: <http://dx.doi.org/10.1177/0306312719846557>
- [5] Julie Cosenza. 2014. The Crisis of Collage: Disability, Queerness, and Chrononormativity. *Cultural Studies Critical Methodologies* 14, 2 (April 2014), 155–163. DOI: <http://dx.doi.org/10.1177/1532708613512272>
- [6] Edward S. Dove. 2018. The EU General Data Protection Regulation: Implications for International Scientific Research in the Digital Era. *The Journal of Law, Medicine & Ethics* 46, 4 (Dec. 2018), 1013–1030. DOI: <http://dx.doi.org/10.1177/1073110518822003>
- [7] Peter Elbow. 1997. Collage: Your Cheatin' Art. *Writing on the Edge* 9, 1 (1997), 26–40. <https://www.jstor.org/stable/43158837>
- [8] Mikkel Flyverbom and John Murray. 2018. Datastructuring—Organizing and curating digital traces into action. *Big Data & Society* 5, 2 (2018), 2053951718799114.
- [9] Sarah Gilmore, Nancy Harding, Jenny Helin, and Alison Pullen. 2019. Writing differently. *Management Learning* 50, 1 (Feb. 2019), 3–10. DOI: <http://dx.doi.org/10.1177/1350507618811027>
- [10] Rachel Handforth and Carol A. Taylor. 2016. Doing academic writing differently: a feminist bricolage. *Gender and Education* 28, 5 (July 2016), 627–643. DOI: <http://dx.doi.org/10.1080/09540253.2015.1115470>

- [11] Donna Haraway. 1988. Situated knowledges: The science question in feminism and the privilege of partial perspective. *Feminist studies* 14, 3 (1988), 575–599. <http://www.jstor.org/stable/3178066>
- [12] Michael Humphreys. 2005. Getting Personal: Reflexivity and Autoethnographic Vignettes. *Qualitative Inquiry* 11, 6 (Dec. 2005), 840–860. DOI: <http://dx.doi.org/10.1177/1077800404269425>
- [13] Yanni Alexander Loukissas. 2017. Taking Big Data apart: local readings of composite media collections. *Information, Communication & Society* 20, 5 (May 2017), 651–664. DOI: <http://dx.doi.org/10.1080/1369118X.2016.1211722>
- [14] Yvonne McDermott. 2017. Conceptualising the right to data protection in an era of Big Data. *Big Data & Society* 4, 1 (June 2017), 2053951716686994. DOI: <http://dx.doi.org/10.1177/2053951716686994>
- [15] Katri Mehto, Vesa Kantola, Sauli Tiitta, and Tomi Kankainen. 2006. Interacting with user data—Theory and examples of drama and dramaturgy as methods of exploration and evaluation in user-centered design. *Interacting with computers* 18, 5 (2006), 977–995.
- [16] Alison Pullen and Carl Rhodes. 2008. Dirty writing. *Culture and Organization* 14, 3 (Sept. 2008), 241–259. DOI: <http://dx.doi.org/10.1080/14759550802270684>
- [17] Cornelius Puschmann and Jean Burgess. 2014. Big data, big questions| Metaphors of big data. *International Journal of Communication* 8 (2014), 20. <http://ijoc.org/index.php/ijoc/article/view/2169>
- [18] Nanna Bonde Thylstrup. 2019. Data out of place: Toxic traces and the politics of recycling. *Big Data & Society* 6, 2 (2019), 2053951719875479.