**29th January 2016**

**Socialising Big Data Group Meeting**

**Attending**: Ruth, Adrian, Penny, Celia, Camilla

**Abbreviations:**

SF – social framework

SI – social intelligence

EOC – ethics of care

**Discussing A&R’s pre-circulated piece of writing:**

Adrian – we don’t say what SI means anywhere. We need to clarify this.

Celia – In the SF, SI and EOC were both used. Both terms have their own histories/genealogies which don’t necessarily map on to what we want to say. How much do we want to use them?

Penny – thought we were trying to think about how we describe the ‘social’. We haven’t put our finger on what type of ‘social’ we think we are describing. There is an intrinsic sociality to BD – this is what we are thinking about rather than our practitioners.

Celia – is ‘intelligence’ the right word? Does ‘social’ takes us far enough?

**Defining SI and EOC:**

Adrian – likes the tension between SI and EOC. There is no simple mapping but the friction between them both is interesting. We can draw on the non-moral, non-human centred EOC. This is a non-normative model. But even so do these two map onto each other? Perhaps we just need one? ‘Intelligence’ is linked to knowing – a relational form of knowing but it maybe unamiable to care – it puts the social fabric under tension.

Celia – we do point to the tension in the SF. Because we have said that SI requires an EOC but we haven’t attended to the tensions which are inherent. We just say that ‘a social’ doesn’t easily have normative expectations. It just refers to being relational. We collapse ‘social’ with an EOC. In a paper we might want to pull out the tensions more.

Adrian – yes, or we may just have to do one. If we see BD as a form of SI – as this more accurate, relating occurring in the data - it is not necessarily doing what an EOC would do which is to sustain the social fabric. . In other ways it has other interests – commercial, governmental etc…

Penny – we might want to think about how the SI works. A lot of it is in this instrumental framing. An EoC is not nec relevant. In waste the SI is not necessarily there. The intelligence can be random – that’s what brings out the tension. but we don’t want to reduce the social capacity to the instrumental. We are interested in the other types of attachments/or being social in the data. For Viridor it wouldn’t matter what the data connected to. John is much more worried. he only thinks about targets …

Celia – Theres a Q about the disruptive effects of BD can be exaggerated but one of the possibilities of BD is about being instrumental in different ways – you can use it in a v instrumental way - e.g we may be interested in getting a better bottom line but don’t necessarily know how to do this. It’s potential is in the unforeseen/unexpected ways to connect through looking for patterns. There is a moment of I don’t nec know what the best way to do the bottom line is. unforeseen ways of connecting. correlations not pattern.s

Adrian – I think in terms of structuring the article, that is where the ethical would come in. Given you can explore and find these associations. What do you do then? Like journalists when crowd sourcing a news story. and it goes viral. what do they do then? do they go back and check the voracity of the story … We can look through the ethical implications but what do we do in practice? For example, when does something becoming a story in journalism? Do you just reproduce the data or do journalists have to check the data? What about ethics? If I propagate the link to this isis video, what do I do then? The article could explore this – the question would then be, how do we map this in relation to the collaboratories? newness of potentiation and the ethics of care – what do you do about that? Then the Q would be – can we map that in the stats case?

Celia – national statisticians said that it was government who does politics but on the other hand, they said that only their data could be the basis for government. that’s what they used to think. But now they have to think differently. Their dilemmas are similar to journalists. the epistemological status – the ethic of care is to knowledge. The EOC relates to knowledge.

Penny – really parallel to the journ alists. We could become the new verifiers of the K.

Adrian – The crowd sourced intelligence is one version of SI. social media is kind of SI that is constantly producing K claims but it is problematic bc they are so susceptible to imitation and they can propagate unexpectedly. statistician’s duty of care is veracity. Is veracity their goal?

Penny– when you clean data you make it viable. you notice things that statistically look odd. so you clean it. You make it viable rather than true. You look for things which are odd and then look for what is viable in terms of your knowledge base. They want to make this crowd sourced stuff viable.

Celia –Emma Uprichard’s article on Dirty Data or data cleansing argues the opposite. In relation to the potential of BD, she asks; do we give it a life? How does it move? What does it mean? Viability is interesting – the potential of BD – do we give it a life? and how does it move? and what does it mean?

Penny – does Evelyn have ethnographic data on this?

Celia – viability is also good because it relates to the EOC in care for life, or

Penny – but also care *for* the data because if you leave it dirty it can’t do what it needs to do

Celia – Emma’s argument is that dirty data is often better because it has traces of its composition mixed into it that can then be used to make it do other things

Adrian – tension between dirty and clean data – keeping the mess - is in the genomics discussions. Huge efforts put into preserving the artifactuality rather you might wanna go back and … that’s what the SRA is – maingain the artefactuality rather than the epistemic value – which is also imp for science but you wanna keep the artefactuality so you can go back and create new knowledge claims which take into account whatever made the mess.

**Diagram in the report**

Adrian – somebody cleaned up the diagram. Was originally quite rough, with shaky lines. It now looks much clearer – does it still make sense? Evelyn thought it was too unprofessional.

Celia – we had discussion about circle and circulation … Nicky Gregson – critiqued the notion of the circular economy at the waste collaboratory being too neat - too many socials – that’s why ppl think we don’t say anything

Ruth – finds the diagram difficult to work with – because it separates social effects/composition which made her think of work by Wanda Orlikowski on structuration, who took Giddens and applied to technology in organisations to look at the duality of technology. At one and the same time technology can be a structure but also be structured by social practices. But in fact, it is always possible to unfold it. Helpful for thinking about this. Sometimes there is a design to neatly package data into something but the problem is that data is always unstable. Depends on social practices.

How is BD different from a technology? Is there something fundamentally different rather than technologies more generally?

SI and EOC have already been applied to technology – but does BD play with these concepts differently? I think it does because of the effects of BD connecting. Also, in relation to Orlikowski – she is insistent on the specificities which relates to the EOC – it is not possible to have universal principles. Data will have different affects depending on the social relations.

Adrian –- our claim to the specificity of BD comes from the SI. If we do SI, we focus on relationalities and potential relationalities, and that’s what distinguishes BD from e.g. large scale health records systems which are just meant to admin management – specificity arises in BD bc of the focus on mining, or predicting. .. We need to do some more work on the specificity, in other words …

Celia – the 3 things we talk about in the FW on EOC we talk about are interdependence, responsibility for potential social effects and not universal principles. recognition of specificity.

Ruth – are the specific social effects about targeting, classifying .. ?

Celia – wd be there for stats

Penny – if we are thinking of data as a technology, classifying wd always be an issue, even if we used a card index. with BD it is the capacity to do it for itself. the capacity of the BD to target, classify, discriminate w/o being attached to any intentions. it has a kind of autonomy to make connections and producing correlations. so if you make a big card index in the library, the index cards can’t do anything except sit there. Someone can come in and do things. But if you have a big mass of data it can start producing correlations. it’s this thing of producing correlations that may or may not have something to do with its .. it is the capacity of BD always to produce statistically relevant correlations bc of the size of the datasets that could have nothing to do w any experiential reality – it has the capacity to generate plausible correlations

Adrian – not sure about ‘autonomy’ – but the other bits all fine

Ruth – this is part of the duality which can appear to be an agent, to act on its own but behind that it is actually enacting other forms of social values, preferences, algorithms. Data can’t actually do anything by itself

Penny – It can’t do it by itself – but But it’s something about the way the data itself without any other intervention can create connections

Adrian – we could think of a new term to describe this? inceptionism. This is not to say that it is totally autonomous but then can be capitalised

Penny – I think it relates to the crowdsourcing example. it can intensify itself

Adrian – these technologies are highly organised but are generative. uyou can’t organise crowds in that way. these techs are highly organised yet still have this generative thing.

Celia – what about the example of hacking which has looping effect of different forms of population. data. biosociality. For example, people don’t fit into categories. so they do stuff. so the category changes. but you think I am this or that, even if you don’t quite agree. Evelyn’s work on the census shows this. It didn’t initially include ‘Canadian’ but eventually was adapted to include the term. They are making up entities which no one knows what they are – or maybe they don’t want to know who they are. They don’t want to say this is this kind of ppl It’s categories of ‘people like you’ – that is the category - it doesn’t matter who the other people are. This is a very different type of social actor. Like the idea of a ‘crowd’ – we don’t yet have a term to describe this. David Stark talks about ‘Emergent classifications’. Crowd is our one way of speaking to this entity. Because a crowd is diff to may other kinds of social groupings. But do we have a term yet for these David Stark talks about emergent classifciations. and we don’t know how to be responsible in relation to them.

28 minutes.

Adrian – That’s the EOC. If this is a form of intelligence that is somewhat alien to us because it doesn’t inhabit the same kind of classificatory spaces and the same kind of cognitive forms that we normally work with, then we find it on the one hand incredibly powerful Tbecause they can be governed up or capitalised up in certain ways, they can be leveraged – they are leverable for finance and other things, but on the other hand they are difficult to make sense of. They pose all these sense making problems that, not just us, but all kinds of groups have to find ways of naming and articulating, identifying, embodying, practicing – embodying in a literal sense - becoming part of the habitus. We have to make ways of labelling these groups. Would be great if this could be the focus of the paper

Penny – ‘people like you like this’ – is key. really captures it. data generates that possibility without giving it any content – except that relational content – can be random but

Celia – we think of these classifications as having properties – we have these attributes

Adrian – or can turn it around – ‘things like this like people like you’

Celia – they do! this approach is already taken in marketing – items like this like ppl like that so that’s a good item to produce

Adrian – I was thinking that the thing liking you was more like the algorithms are doing bc they’re doing the liking. Like has two meanings. It is a comparison and a measure of similarity. It is also an expression of desire. The algorthims are doing both. Things like this – similar things. Things liking us is OK with STS – the configured user. An apple computer likes a certain kind of user. one who doesn’t’ want to get inside.

RUTH – the things are liking us – saying we are like – and we don’t necessarily understand why – but it is generative of social effects.

Penny – not sure about this?

Adrian – ‘like’ refers to an affinity so in turning the phrase around we could talk about the agency of the object

Celia – it is precisely that. A core component lean startup methodology. build, measure, learn, feedback loop. figure out the problem to be solved. make a minimum viable product to begin the learning process asap. then learn who your product likes. then inc the number of ppl that like you product and your product will like. …. and that will involve measurement and learning. .. 33.18 and must involved actionable metrics that can predict cause and effect.

data collected in particular ways for partic metrics that can be acted on partic ways.

… if you get those things that emerge and we don’t really know what they are, if they are to be capitalised up or whatever, even if we don’t give them a name or a fixed id, they still have to be evaluated in some way so the data has to be cleansed to be able to go into certain kinds of metrics that can then be compared or not across diff situations.

Adrian – we cd say that that that property …. we see that property is in genomics too – by doing this we can account for more variation around a partic pathology… if only we do more sequences we can account for 50% of the variation instead of 24% – that wd be an example of that – incremental leveraging. is that SI tho? that process – that feedback logic is running right thru the algorithms. all the optimisation algos. all trying to measure – to narrow the gap between what was previously known and what can be known using the model – doing that kind of looping. But is that what we mean by SI?

Celia – that’s one of the possibilities of SI. But what we are saying is that an EOC wd say yes there are these potentials in BD, and optimisation is a really dominant way of doing SI, but really we need to broaden or think differently about the new social entities – the new configs of the social that could come from BD. Bc – linking thru to things like metrics – our ppl across the collabs are interested in the metrics as well as the data. the data in itself – big computational, - but then there are also metrics. not raw data.

Adrian – we have got more than enough for an article to say these issues and develop a way of talking about them.

**Metrics** **(and search)**

Adrian – metrics would be a way in. lots of issues on metrics to explore in the article. We need to think of new configurations of the social which could come from BD through to metrics.

In relation to genomics this raises the following issues: tensions

1. cost and project complexity- all these metrics about costing. we can do more for less. But complexity. how complex it is to do stuff. in a number of talks at the collab – is fine to have things getting cheaper and having more data but what is really difficult is project complexity. Because the bigger projects get more complicated socially to arrange .. and how it all fits together . Project complexity is the thing we are trying to work out how to measure. and we can’t bc we know that the more ambitious we get to know things – and we wanna know all the variations in the population – but to do that we have to run projects that are extraordinarily complex in their structure. So the data themselves might be predictable in that we know that it will cost this much, take this much time and take so many terabytes, but the complexity of the project is something we can’t get a handle on.

Celia – and is that partly bc to begin w we can’t even start by knowing which variables are going to be imp?

Adrian – 2 problems. One is..

Celia – and if you can’t assume that then the possibilities of doing stuff w data are so huge, that even if the algo is learning, its gonna take it forever

Adrian – or it might be learning on stuff that’s not worthwhile. bc its gone off on the wrong track. you find a whole lotta stuff that’s not actually to do with the problem we’re trying to deal with. I think there are a couple of versions of that problem running thru the genomics collab. Some around costs. Some around rareness. Some around the nonhuman entity itself is not stable – as in the wheat genome. What shd be a nice malleable entity turns out to be just full of evolutionary mess that completely bamboozle attempts to dissect up and analyse it. So the object itself is elusive and wily.

So we could, just thinking pragmatically, when we talk about metrics in genomics case talk about there’s this problem that the projects have to operate on a scale that means they have a social constitution – as in many different expertises and institutions globally distributed, variations in the population have to all be brought tog – it’s like a global problem – a global assemblage type problem. there’s a second problem which is that the entity itself is effectively the result of life and it’s evolved and the entity itself isn’t simple. And then there’s the dynamics of commoditisation and capitalism, like how much the scientific work is being picked up and turned into a commodity, and how much remains something that has to be worked thru in a diff way. those tensions between commodification – what we saw with the Amazon ppl saying – ‘here we can give you everything you need’, and the desires of the scientists to have control over what they are doing, build their own things, to do the composition of the data in ways that they think is fully accountable to all their colleagues and preserves trust and epistemic community or whatever .. so they wd be the kinds of things I think we cd do w that, the metrics issue in the genomic bit of the article. trying to be pragmatic.

1. but the more ambitious we are, the more complex the project structures are which is difficult to get a handle on in advance. It often includes many different people working across institutions. We don’t know what the variables are in advance. It can be problematic because we might not necessarily produce the metrics that we want.
2. Non-human entities are not stable. In genomics, the data is full of evolutionary mess. The object is illusive.
3. Commodification – desires of the scientists to have control over the composition of the data so they can preserve trust.

**Target**s - In relation to waste:

* The authority have targets which are set by the EU. They know that their targets are unachievable with the current metrics.
* Waste is also not stable entity – difficult to get accurate measurements. E.g when it becomes wet

Celia – have targets set by the EU. Maybe tonnage is too – debate about how the targets weren’t relevant to their local situation. not a useful metric of them.

Penny – was about them being unrealistic. to meet the targets they wd have to invent a new metric. but they couldn’t wriggle around the targets. once a figure arrives at eu as a target then they have to find a way of meeting it.

Celia – what’s the difference between targeting and searching? if our target is reduce costs. searching for something yu don’t know what it is until you find it. so what you do is target something and hope it is going to be useful. I think there is a tension between search and target.

Penny – the target doesn’t help you to get to your overall goal – like more sustainable environment - but the BD might help you to get to a more sustainable world so you have to follow the targets. When we discussed experimental scenarios such as urban laboratories to think about different collectives the authority could not reconcile this with their idea of targets. John is excited that BD will get to target, bc getting to target will get you to more sustainable world. But what if BD can get you to more sustainable world but not get to the target. But that is too crazy for him.

Adrian – in other areas – Manhattan – waste data is treated n a more experimental way.

49.00

Celia – I am thinking about Herbert Simon talks about search. Search for something when you don’t know what it is. He talks about the process of search in itself is important. Searching and targeting have been used in military campaigns. Bringing targets into the world of the economy/university – what does this mean? Not wanting targets is not that we don’t want to do things. Targets are meant to be destroyed. Strange to valorise them. we never get to destroy our targets. also how do you find things that you don’t’ know.

Interface effect. the history of the interface. predator prey and targeting. whirlpool defence system. shows the diff histories – searching and targeting the site of a lot of military campaigns. finding things and destroying them.

Penny – targets could be specific for our collab. is central to waste. could be a focus of the waste section.

Celia – statisticians – do they really insist our data is not oriented towards any specific target but it has to be ableo to be acted on by gov.

Penny – waste ppl don’t mind how they get to the metrics as long as they get there. Stats are more concerned about the quality.

Waste – they are searching for something and they don’t know what it is.

**Discussion over lunch**

BD is solutions searching for problems. Penny, but so was Internet etc.

Adrian - BD has been constituted as something intrinsically social – BD is K about the social – group formation, publics, - as well as about the structure of the universe – a social message built into it - we are going to give you a new view of the social – a social physics – that gives it a social relevance – inherently social - but once ppl all the treat the social in the same way – the classifiers do the social work in the same way, that becomes socially not very interesting

Celia – is there somebody who has talked about social fiction? in the sense that, this is fiction about what the social could be? this is what’s going on in something like BD. not so much science fiction as social fiction. let’s imagine what the social might be! it is social, we just don’t know what kind of social it is.

Adrian – most of the claims about the inventiveness of the data – we might find a new emergent form of the social - . very soon revert to back to a std set of illustrations (suppositions?). the shopping basket allows us to infer that the teenage girl is pregnant. The constant recycling of these forms of K of pattern and so the claim to recover relationality or be generative of new classification systems is not borne out v often.

Celia – but that’s why our collabs are imp. Bc the ppl we are talking to are trying to move beyond that – the rhetorical examples . they’re really trying to bring these new socials – in a sense – like the Rabinow biosocial - new groupings are gonna emerge from these kinds of practices and they are struggling with them and they are trying to produce them but at the same time they are really trying to reduce waste and cure diseases, They are not trying to make [social fiction?}… they are trying to d their various things but. LOOKS ON INTERNET. there’s a social fiction series. maybe that’s not what I mean. Speculative fiction.

Can we get the history of SBD?

Adrian – maybe we cd just go thru waste, genomics, stats and say what those sections wd need to do. Some cross cutting and some unique themes.

Ruth - EoC in genomics is about preserving the connectedness of the data to allow going back. Sustaining its social composition.

Penny – does each section try to articulate what social intelligence wd mean for that domain. what is diff about BD versus other kinds of data?

Adrian – wd be useful – tendency to homogenise the SI of BD but one of our findings is that the SI of BD has to take on diff forms in order to be viable in diff sessions.

Camilla – waste changes its composition when it gets wet. All of the issues – instabilities - we brought up for genomics wd work for waste. wd work for stats too. in stats one of the problems is the BD is already commodified.

Penny – the notion of a normative version of something. how to translate into anthropology? is it hegemonic?

Adrian – does pose this problem. if is a moral norm, where did it come from? The problem STS has faced which is why John law goes on about it – ANT tries to produce non normative descriptions of things – they cdn’t really say what is a good nw and what is a bad nw – a voraciously capitalist nw can be a fantastic nw in terms of its networking but be v predatory. or annihiliating of difference. The ANTppl wd say we can’t really say that – it just looks like a great nw to us, even tho we don’t feel great about it. So john says we gotta get normative as well. We can’t just map the nws. if you make a normative claim you are saying something is good. you are making a judgement on what is good and what isn’t

Celia – you positively value a norm.

Penny – so you cound’t have a norm that you all agreed was awful?

Celia – normative does recognise that my norm is diff to yours.

Celia – a normative ethic of care – that is the right way for the SI of BD to be organised

Ruth – back to dellacasa – whether we like it or not care always has to be invested in things and their relations in order for them to endure. is almost a pragmatic account. resisting going towards a normative. We refer to her – how do we relate this to an ethic of care.

Adrian – I thought the principle theme is she is trying to move away from a moral version of care which is it is good to take care of ppl – not to say that it’s not but she doesn’t want to limit it to this – she wants to expand it to include things. And bc things sustain social fabrics – there is no social fabrics without things and no things without a social fabric – but what more are you asking?

Ruth – linking to Penny’s Q about norm and how we are situating our EoC in relation to her work,

Adrian – we wd have to say that we think that BD is intimately involved with what counts as the social fabric today. Our claim is that. There may be a version of the social fabric in this stuff. Not that it’s monolithic or … so if EoC is always about social fabrics, and that’s what we are concerned with, then we have to work out what the EoC in relation to Bd wd be. Bd is to do with reforming or transforming the social fabric. it changes what counts as social.

Celia – the dictionary, non sociological definition of normative is establishing, relating to, deriving from a std or norm, especially of behaviour. I think you cd hold onto that. We define our norm in relation to an EoC in the short version of the paper. We say it’s about recognising interdependence, exercising responsibility, and recognising and working with specificity. We give 3 things that we say are an EoC. Whether we meant to or not is another matter. We can change them. if we want diff norms or stds in our ethic of care we can .. We initially thought these wd be the normative dimensions of an ethic of care.

* Interdependence – which is a bit of a worry because interdependence is a characteristic of all of them. Everything. it must be a positive valuation of interdependence. .
* responsibility – whatever we mean by that
* specificity – not about universal norms but about the Specific.

I think we prob got them in part from the literature on EoC. But one of the things we might do at the end of the article is to say we can see from these 3 cases both the possibilities of SI but we can also see certain things by which we cd expand what we think an EoC in relation to BD would be. Bc in genomics this Q gets … in waste … And part of it would be that we might have some general principles for an EoC but some of them are going to be specific …. that’s a v boring way of doing an article. And then we are contributing to the lit on what an EoC might be. or how you develop one.

Ruth – the tensions between SI and EoC. Can we articulate more about what those tensions are at this point.

Celia – we shd do that and presumably that is what would come out in each of the pieces.

Penny – so if we took those 3 things you cd contrast that in each case – how SI of a partic kind of data does or doesn’t seem to fit. The positive valuation of interdependence seems to be unproblematic. But responsibility seems v problematic. And specificity is interestingly problematic. So interD isn’t really aproblem. it falls on exactly.

Celia – but is it? Because at variojs points don’t our ppl working w BD want to make certain kinds of cuts?

Penny – that’s stopping it being BD then, isnt’ it?]

Celia – no, not always. one of the tensions in waste is that on one hand we want to know that this is an old lady who needs special support, but on the other hand …… When we’re …. What do we make the data connect to? And those are the issues of privacy and anonymity. They’re quite big questions. And they cut all sorts of ways and often there isn’t a simple …

Penny – so when you ask ‘what do we make the data connect to?’, isn’t part of the problem is it’s not what I make the data connect to, its’ what the data connects to by itself? Because data is gonna do it.

Celia – but the data will have been formatted in certain kinds of ways which mean it can connect to this and not to that …

Penny – but it may connect to something else that you hadn’t thought of …

Celia – yes but all the time – and sometimes you don’t know the consequences of formatting it in this way rather than that way, but I am still not sure that makes it autonomous. I mean, it’s not all deliberate

Penny – you can say I am goiong to format it so it can connect to A and not to B, but then I am interested in how it connects to C. Bc by formatting it so it can cnnoect to A, I totally unknowingly enabled it to connect to C. THIS POINT CONTRASTS WITH ORLIKOWSKI’S MODEL. IS NOT UNIQUE TO BD THO. WD BE TRUE FOR OTHER TECHNOLOGIES. BUT OF SOME TECHNOLOGIES MORE LIKELY THAN OTHERS. IS THE SI OF BD THAT MAKES THIS A PARTICULAR RISK

Adrian – in a lot of those deanonymisation scandals that has been what’s happened. that’s like, we set it out so nobody cd really work out what it’s about, but then, when its’ put tog with this other data it …

Ruth – the crowd makes the potential for unforeseen couplings v big bc the interferabiity of Bd is much bigger than other kinds of technology

Penny – maybe that’s all we need to say. What abou Theres that whole massive haraway discussion about being aware of the suffering you might cause

Adrian - The cui bono thing. which wd be one of the things in Q here – who is?

1.44

Celia - what if the unintended consequences of this are disastrous, is there any entity that can take responsibility? if there ia a multiplication of interdependence and unintended consequences, it is difficult sometimes to work out who or what is responsible. For responsibility do you have to have a who? I don’t know, which may be another way this is not part of BD

Adrian – I guess that’s one of the core things that’s going on in ethic of care debates. as it’s expanded from this moral duty to take care of others, which nobody really disagrees with – the Q is who? Who shd you be taking this relation of care to? So the Q of care is in a way tautologous because there is not ethic of care that doesn’t bring responsibility but the q of responsible to who is at the core of it. Are you responsbiel to the water or the sky [HANS JONAS THE LAND ETHIC] as well as this or that person? all those debates around non humans.

Celia – we said in the paper to take resp for potential social effects whetehr the negative ones of …. or the positive ones of empowerment. So it has to be who. And also what kind of responsibility? is it a legal respon? if we say you are responsible, do you go to prison? do you pay somebody some money? Do you say sorry/? do you stop doing it?

Penny – sin’t it about how you take responsibility for something that can never be placed at your door? the whole point is that bc the connections are arbitrary, …. I thought it was more about how do you ? Say the waste authorities did a massive bd thing and it all went horribly wrong, but they wd say before they did iwe know this can go wroing so we will stand up and say we are respon? or will we say that can never be tracked back to us bc of contingency and so they don’t hav eto be resonsible?

1.49

Celia – no the privacy policy is all about trying to exercise responsibility. Or ticking a box to say you agree to these terms and conditions. It’s like saying, we are responsible for this, but If you tick this box you are responsible. They’re attempts to distribute responsibilities and sometimes you tick a box and you say, anything that happens, is OK. …. It’s like saying I accept that you are not responsible for this, for any of this. You can tick that box. But other attempts are …. but you will be responsible for this, this and this bc this is a known side effect. We know this can happen and we’re telling you and … those stupid ads where they say, do it responsibly! ….. Responsibility has to be about acknowledging that you have had some kind of role in the consequences of what is happening – even tho nothing has a single cause, nothing …

Penny – what interests me is how it amplifies the possibility of being responsible. One you xxx the first principle that everything is interdependent …then it’s never your fault - because you move one step away from causality which in a way is what responsibility relies on

Celia – But they do make cuts. Not everything is interdependent – ‘so you should have anonymised better – you removed names but you didn’t remove unique identifiers’ – don’t you think?

Penny – No, because it’s all about the unknown connections [??]

Adrian – in a way we are saying that BD brings with it a kind of impossible responsibility. If it can generation hitherto unseen forms …

Penny – and that’s the whole point of it … but nobody can be responsible for it

Adrian – but you versions of hospitality, or whatever,

Celia – but of course somebody can be responsible for it – you can be responsible for things that you didn’t mean to do – I didn’t mean to kill that person? Also, I don’t think BD is complete interdependence of everything. It’s different kinds of dependence and interdependence.

Penny – so your argument seems to imply that from the very beginning we shd be making the cuts and not letting it – so it’s about constraining the capacity …?

Celia – it always is. BD isn’t the whole world, is it? it is a cut of the world. it’s not everything

Penny – but then we shd say that upfront – that the ethical approach to BD is paradigmatically about constraining it? Cutting it?

Celia – data cdn’t be data unless it was constrained

Adrian – yes - but that seems to me is the fundamental tension between SI which implies this horizon that moves back xxxx things move over into relationality, and on the other hand, what you might think about that, how you might value that. You might say, actually that relation is destructive. or that relation is damaging.

Penny – so it’s about acknowledging that more strongly?

Celia – but the gov takes responsibility for flooding – it’s not responsible for the rain –we are globally responsible for the rain – but we ask the gov to take resp for all sorts of things that aren’t in its control

Penny – but that’s different, I think …

Celia – that’s the social contract – take responsibility for natural disasters

1.53

Penny – but I don’t think that’s the same thing! The gov doesn’t take responsibility precisely bc it doesn’t have to.

Celia - … it has ineffective flood prevention …

Peny – or it says we cdn’t control that

Celia – it does sometimes, but then we have a democratic debate about whether it cd have controlled it – bc we say gov shd control it – you shd make every effort

Adrian – we are talking about diff kinds of ethics here. There are consequentialist ethics that judge good and bad entirely on the consequences, but the EoC is not a consequentialist ethic. It’s a kind of virtue based ethic that says the way in which we ourselves skill ourselves is the measure of value, it’s not so much that we weigh up the good and bad consequences – that’s not the measure of it. ethic of care ppl don’t do that

Celia – No but maybe – just as we are using EoC to be careful about SI, well there are these new forms of intelligence that pose problems for ppl like us without an ethic of care because it might not be adequate to this kind of intelligence – and that does come down to those political – the notion of a social contract. it also the social, it is about whether the forms of politics as well as the notions of care that we have are adequate to this world that we live in, with its potential to make connections that aren’t easily recognised at the moment.

Celia – the journalism example is interesting then. Because it’s like, there’s veracity and we can check all those things – but doesn’t journalism also say, we will not show these things bc they are too horrible? It’s not bc these images are not true, or important, they are just too horrible

Adrian – or divisive. Like the Charlie Hebdo thing. Shd they publish the cartoon that shows the naked Mohammed? it’s freedom of speech, but on the other hand it’s incredibly divisive of the social fabric, so there the tensions are being played out. Yes you do have to take into account the nature of the social fabric, but the reason you do that under an ethic of care is because you accept the relational constitution of things, not bc you think you can measure the consequences. That’s the difference, I think. Traditional EoC were based on the subject, and it’s good to take care of others almost as a moral value. The EoC like Bellacasa is a relational one, which is about what you need to do to sustain the forms of relationality that we don’t survive without. that is, things live and differences are sustained. It’s not bc you know the consequences that u do that, it’s not that u can calculate the conseqs and say that this action will lead to this, the ethic of care is based on the relation itself, not K of conseqs. So I don’t think we shd run an argument that says it’s the responsibility of ppl doing BD to calculate the consequences of what they are doing, it’s too hard, in fact it makes it impossible. what we shd do is to say the responsibility is to act in a way that maintains the relationality

Celia – yes but which relationality?

Adrian – the social fabrics in question. I am not saying that there’s a single social fabric, bc that’s really hard to argue as well. But they don’t have to know or calculate the consequences.

Celia – I agree they don’t have to know the consequences, but I am not sure I think that relationality and social fabric is v helpful.

Adrian – this is a human non-human social fabric. it’s not just my interpersonal relations. it’s those relations which materially, corporeally …

Celia – but doesn’t it just collapse into SI bc if it’s *any* relations that currently exist …?

Penny – what about …? if you make it concrete in the waste thing, they do have an ethic of care that’s not that difficult to get at – it’s about environmental sustainability – so that that’s the target of that ethos that you don’t necessarily meet – so that u cd be pursuing targets that allow or almost encourage u to poison the earth by doing things that might end up generating more waste or in such a way that it doesn’t connect w this other value of environmental sustainability. so there is a value that covers everything that they do. The value is environmental sustainability; the targets may or may not have anything to do w that, and if they don’t ..

Celia – I think that’s right but it’s got the same thing as any specificity, like – what about sustaining the social fabric of neo Nazis in Germany? we need to make sure that they are sustained.

Adrian – if you were in a relation to them u might want to do that but are you?

Celia – yes bc we are all in the world.

Adrian – yes, but it’s not saying …. I don’t know the answer if I think about it - if we go for a specificity we don’t have to say that we are saying there’s one form of specificity – the democratic. We don’t have to say that at all. There are some quite non democratic uses of BD and what can u do about that? They will go on, the way BD is used in China will not be democratic and the way it’s used on Facebook may not be democratic, but u still might have an ethic of care going on. Even if it’s not democratic social fabric. The point is that you don’t have to affirm or insist on any particular version of the social. In relation to the fascist use of BD … neo-Nazi use of BD …

Celia – Collabs with Hitler ...

Penny – it just seems to be one of those things that’s so easy to say, but what does it actually mean?

Adrian – Ethic of care, or?

Penny – in terms of this exercising responsibility particularly.

Celia – we can drop that one

Adrian – I thought the point we had was – I put it in the doc - that if we just see the responsibility is a *question* bc we don’t know to whom – we don’t know that. That raises the ethical problem. We are working on the assumption that ethical problems are ones for which there is no solution. When you are forced to make a decision when u can’t decide one way or another very easily. So here the Q is who? We don’t know. And we have to live w the fact that we don’t know the who.

Celia – the who, and maybe, and also the *how* – regulation, responsibility – can we have responsible standards? protocols or? I don’t know …

Adrian – yes, so they wd be like unfolding out of these kinds of questions, about the who, if you try to concretise it, you have to think about what algorithms, what protocols, what aggregations of data? It that what u mean?

Celia – yes – if somebody’s looking at tagging the algorithms that are used in finance, which is partly about the recog that these algorithms have had unfortunate consequences, and that most ppl using them don’t really understand them, so they have algos that can tag algos, the idea is that u can see which algo is responsible for doing what. I think that’s quite an interesting attempt to think that thru – like what is regulation in this situation?

Penny – so that’s a way to try to track the process if it does all go wrong? You’re not trying to make a *person* responsible, you are trying to id a *process* that u cd then improve?

Celia – yes so that u cd say, u can’t use algos that do this, or to understand that …

Penny – so that’s a way of anticipating a problem that .. ?

Celia – or *monitoring* it in some way. But it was partly a response to the economic crisis …

Penny – but then it starts to anticipate the next crisis?

Celia – the idea is that u can begin to anticipate the consequences or the effects or you can

Adrian – yes, algorithmic audits is the sort of thing that’s developed methods of algorithmic audits, so you actually audit in that sense of accountability on algos. it’s another way of talking about it, I guess.

Penny – that makes more sense bc then, in a way, u are making the algo responsible. Which is something u cd track.

Celia – and it also distributes the EoC

Adrian – yes. What about the recognition of specificity? what r we saying about that, other than what we always say, that u shd be specific? is that just a response to the way that everything is thrown into BD? We wanted to say, ‘don’t lump everything together – resist the temptation to see everything as just another instance of BD’

Celia – yes, but it is also this thing of, you cdn’t say ‘u cd always exercise responsibility by doing this’. Like, you can’t always blame the algorithm.

Penny – it’s about recognition of that constantly …. a recognition of the composition. so that it’s that BD isn’t just data swirling around connecting, it’s all kinds of decisions – it is the algorithm, it is the xxx, it is …. so it’s thinking about the compositional aspects

Celia – yes. And it’s also attempting to enhance awareness of the possibilities in a specific – like who can act on things? And in relation of you looking at the specific composition of the specific situation or circumstances, then u … this is just responsibility really …

Penny – they’re not quite the same. I certainly feel clearer because some of these situations and circumstances actually end up being the specific regulatory regimes and institutional practices, as well as the specifics of cuts and algos and the cleaning practices and all that kind of thing – bc that is quite a lot of what we were talking about, which is why are these waste things work in Sweden but not in Manchester – is bc of the completely diff institutional framing, where trust lies and how politics works – and that is also quite imp so that the technology - an obvious STS point – that all those bits and pieces are all part and parcel.

Celia – there’s that thing about contexting, - so what is specificity? where is specificity? We also need to think about specificity – is it BD? is locality? so if u think about the relationality, and then the importance of contexting, but contexting isn’t just something that … maybe I’m just thinking too much about algos, but how do you do specificity? do u do it in terms of situatedness? who can do contexting? how is contexting done? Thinking of that piece by Seaver where he’s – Dana Boyd – BD doesn’t have a context. But BD has loads of context and it not only has them it does them. it is v keen to do contexting. And that we need to talk about cultures of contexting. BD doesn’t do contexting like anthropology does, but it does it in its own ways, and that’s part of the composition as well, it’s about how contexting happens.

Penny – I think those 3 things are different.

Celia – we don’t have to have definitive answers – they have to exercise responsibility like this – we r saying these are qs u need to be thinking about, and ways of approaching these qs or issues that we’ve thought about so far.

Adrian – looking at what we’ve got. That’s a good sketch of what we think in involved in the EoC. We have some stuff under what we mean by SI as associated w BD. I’ve started to lump a few things under there.

* The timing(?) thing, this emergent thing, that isn’t being named or being measured,
* things about crowd sociality with its ideas of similarity and affinity being a particular feature of this sociality, of this SI.

So those 2 sections – SI and EoC – are we gonna go w the idea that the article is about tensions between them? on the one hand we depict this SI taking place with these characteristics, on the other hand, we say in this domain an EoC wd involve the following considerations, which we flesh out concretely. And then, the point of the article is – it will have already covered quite a lot if it’s said those things – to then consider how they fit together. Fit or not fit.

Penny – the crowd bit comes in w this ppl like u like … ?

Adrian – yes – it runs across all of them – emerging classification – u can use a bit of Christian Bolt stuff on em

Adrian – there’s good links between SI and the crowd idea anyway The BD thing in a way is a product of collective behaviours. Altho BD in Genomics and Astro physics …. But the form BD has taken is v much to do w social media platforms and the results of crowd or collective behaviours. So it’s credible to say that SI is connected w crowd.

Adrian – and then there’s the references at the bottom. So the work to be done wd be to go thru each collab

Penny – and think thru how we’ve now spelled out SI and EoC

Adrian – and even almost just following those headings.

Penny – and if we get stuck we can just say – someone might have a better idea.

Ruth – I am unclear about the headings for EoC – interdependence, specificity, responsibility - are they in EoC as it already exists? or are things that BD and SI bring to it? do they things that trip it up?

Adrian – I see them as things that trip it up – we are saying that BD[?] doesn’t do any of this stuff

Ruth – are these things specific for EoC for BD? it is w BD that they become more problematic?

Adrian – problematic for the EoC? or problematic for BD?

Ruth – for articulating an EoC around BD

Adrian – they are probably difficult around all of these things – u know climate change EoC. Which is not to say u shdn’t do EoC. Asylum seekers. I don’t see our article as about trying to revise the theories of ethics of care. I see it more as what do we have to say about socialising BD. If there are implications for the forms of ethical theory then we cd just point to them. The principle contribution is what do we mean by socialising BD? not about problematizing EoC per se.

Ruth – so interD, respon, and recog of specificity are ways of shedding light on socialising BD.

Adrian – yes. BD already has this relation to the social built into it, so the work we are doing on socialising is not just continuing that. it’s about saying how you have to rethink the social along w BD data making claims about the social – to know the social –

Penny – so BD makes us rethink the social. We can rethink it thru these 3 – is that correct?

Adrian – one Q for me tho is, how much to we make of the collabs? Are we doing a more ethnographic description of what happened in those events or are we bracketing that out and putting it as part of methods somehow? saying this is how we arrived at this stuff, it’s all based on events – we are not saying the collab is the site of EoC or anything? it’s not EoC in practice? it cd be but we r not making that claim, as such? we just put it in the background as part of how we got the data?

Celia – I wd tend to go w that bc otherwise wd be too big for an article

Adrian – otherwise we wd have to engage w all the collab lit

Ruth – we can refer to the WP

Penny – in a way the collabs were a XXX [mimetic?] version of - u didn’t know where we were gonna come out … and that isn’t the collaborative process at all. project in XXXX. was never like this. so no point in flogging collab right thru this bc isn’t the point.

Celia – I did collab project w academic designers – supposed to be a collab – but who writes about this? and how much money did we get from Research Council for about 10 hours?

Celia – the 3 sections have to have a bit more in them and then we can loop back thru and see whether our proposed structure and themes work? so first iteration is the 3 sections.

Penny – when it comes to it, it is prob about a day’s work.

Adrian – Evelyn keens to do an article but not lead.

Adrian – if we did those 3 sections by end of March? Then Celia writes a top and tail. Then we pull threads thru the sections.

Penny – this has been a way to bridge the diffs between the 3 cases.

Adrian – go for a 7,500 word article.

BELOW IS FROM CAMILLA’S NOTES

**Template for the paper**

Penny – the challenge is how do we do the paper?

Adrian – temporality of BD in our domains – starting with genomics, then to statisticians and then to waste (who are searching for the promise of data)

Penny – maybe we cd bring in info from other countries on BD and waste.

* Each domain defines SI and EOC

‘EOC’

1. Interdependence - what do we allow the data connect to? It has been formatted so it could connect to something else but what will this be?
2. Responsibility – Haraway you should be aware of the potential suffering that you cause. But who is the ‘who’? Who takes responsibility for unintended consequences? And, what kind of responsibility? What is the consequence of this? Put the responsibility back onto the individual? BD brings within it an impossible responsibility. BD is a cut of the world – it’s always about making decisions and constraining it.

Adrian - EOC is a virtue based ethics – not a consequential one so it is not that we weigh up good vs. bad. Perhaps the notion of the social contract might be adequate? You have to take into consideration the nature of the social fabric. EOC was originally based on individual care but this is different, a relational one. How do you sustain forms of relationality? We shouldn’t say that responsibility lies with people using BD. They don’t have to know the consequences of the social fabric (human/non-human)

Celia – doesn’t agree with ‘social fabric’

Penny – in the waste example, the target ethos and the environmental ethos do not necessarily meet. Their EOC is more straightforward.

Celia – but what about if neo-Nazi Germany was your social fabric?

Adrian –are we talking about a democratic use of data?

Celia – in finance algorithms are tagged onto other algorithms to track the process. Not trying to identify a person but a process which can be monitored. Nathan Coombs describes how this is used to help anticipate the next crisis. Could be useful to think about.

Adrian – Yes in algorithmic audits

Penny – that makes more sense – something that you could track

Celia – yes and distributes the EOC

1. Recognition of specificity

Adrian – resist the temptation to see everything as ‘BD’

Ruth – against universal principles

Celia – we can’t always blame the algorithm

Penny – recognition of the composition and the decisions which have been made in the data (algorithm, cleaning process and the cuts)

Celia – also revealing who can act

Penny – specifics of the institutional regimes/framings, how politics works which are all part and parcel

Celia – contexting – what is specificity? Is it about locality? How do you do specificity? Seaver – arguing against BD having and doing contexts. We need to talk about cultures of contexting. It does it in its own ways –this is what makes contexting.

But in the **conclusion** – we might want to expand on these things – general principles which have emerged from our three case studies

Tensions between SI and EOC – can articulate what those tensions are. This will come out in the three of cases

**Jobs to do:**

1. Go through each collaboratory – and check through the definitions of ‘SI’ and ‘EOC’

Ruth – still unsure about EOC – are those things that we have brought in – from the literature? What about when we bring them together with BD does this trip them up? Are they particularly difficult?

Adrian – don’t see the article as informing the theory on EOC. The principle purpose is to explore what do we mean by SBD?

Q –what format should this take?

* Do we explore our collaboratories ethnographically?
* Or do we talk about the collaboratories as our ‘method’?

Penny- Not enough space talk about collaboraties as a method – so we just explore them in relation to our data. We already have this in our WP.

Plan

* Mid feb – write 2000 words on the three themes
* Then Celia will read through each three together