**McNally, Notes for Socialising Big Data article, January 2016**

# Overview

How can we ‘do’ Social Intelligence and an Ethic of Care? Or to put it another way – Where did the idea of Social Intelligence and Ethic of Care come from? Answer to both questions: Collaboratories

# Collaboratories

**Collaboratories as a method for the Big Data Problem Space – Isomorphism between Big Data and Collaboratories**

‘*Big Data involves a redistribution of data collection and research methods expertise and the restructuring of infrastructures, which necessitates engagement with a wider range of collaborators’* (SBD WP pp. 28-29 Summary of Key Conclusions).

*‘When problematized, Big Data requires particular forms of collaboration between different stakeholders and practitioners’* (SBD WP pp. 28-29 Summary of Key Conclusions).

*‘In this regard, collaboration provides a useful way to understand the problems of working with Big Data. Through this approach, for example, we can identify who the important players are and ask questions about this burgeoning topic. The Final Collaboratory has revealed some of the key players in the field, but certain stakeholders were absent, such as, the users and producers of Big Data*.’ (SBD WP pp. 28-29 Summary of Key Conclusions).

**Collaboratories and the role of Provocations**

*‘Many genomic researchers are well-versed in the main ethical and social issues associated with genomics. ….. This familiarity with ethical and legal discussions can make it difficult to initiate other topics of dialogue. …. We found it necessary to put discussions on a different footing by working in visual terms (graphics, tables), making use of genomics researchers own databases and software tools, and generally trying to re-purpose genomic researchers own data literacy in the conversation by showing them data gathered from databases about their own data. This approach leads to mixed results. On the one hand, it certainly overcomes some problems of distance and unfamiliarity. …. On the other, this data is now presented with a view to challenging them to think about their own metrics and their own ways of talking about the value of sequence data. Some robust discussion usually arises’* (SBD WP Genomics and Collaboratories pp. 25-26)

‘*That said, more provocation and controversy could have been introduced’* (SBD WP Final Collab p. 29)

**Collaboratory as an agonistic space for the articulation of differences and productive misunderstandings**

Evelyn makes the point that the (NSI) collaboratory was not organised to share skills, develop methods or analyse Big Data (i.e. not for instrumental purposes), but to pose critical questions. She asks:

* *how can collaboratories be left open to ‘productive misunderstandings’? That is, rather than seeking consensus, how might we state the values or benefits of bringing together and allowing for the tensions and different perspectives and interests to be expressed and engaged?* (SBD WP pp. 24-25 National Statistics and Collaboratories).

This reminds me of Barry et al’s logics of interdisciplinarity. The authors identify 3 logics that motivate interdisciplinary collaboration: the logic of innovation (linked to the economy); the logic of accountability (linked to democratisation); and the logic of ontology. It is the latter that Evelyn seems to be appealing to in asking for how collaboratories can be left open to ‘productive misunderstandings’. The mode of interdisciplinarity that is associated with the logic of ontology is that of agonism/antagonism which also seems to align with the notion of ‘productive misunderstandings’. Thus the goal of a collaboratory could be to create an agonistic space where different kinds of stakeholders can enter into debate with each other rather than speaking to themselves.

**But is there a will to participate in such collaboratories – and who would fund them?**

* *in the face of economic and practical constraints such as time, the relevance and value of collaboratories that are more exploratory and conceptual rather than directly instrumental (e.g. developing specific applications/methods) need to be outlined rather than assumed* (SBD WP pp. 24-25 National Statistics and Collaboratories).

# Social Intelligence

Social Intelligence is awareness of the ways in which BD is constituted through and deeply implicated in social relations in both its composition and in its social effects.

**Against the fallacy of Data-Driven Anything**

Narratives about Big Data as the driver of socio-economic change – whether for the better or the worse – tend to attribute Big Data with intrinsic properties – the ability to act in a deterministic way – to have impacts on individuals and collectivities with positive or negative outcomes.

Social Intelligence problematizes the concept of data-driven anything – data-driven economies, data-driven science, data-driven societies.

**Collaboratories as windows onto the Social Intelligence of Big Data**

By bringing together a diversity of stakeholders, collaboratories are spaces in which it is possible to foreground the articulation of counter-narratives instead of over-optimistic and over-pessimistic narratives of Big Data that tend to be ballistic, deterministic and unrealistic.

In terms of social intelligence, our genomics collaboratory tended to focus on the social composition dynamics of Big Data rather than on their societal effects. [Was this also true of our other collaboratories, which also lacked or under-represented participants representing individuals or groups who are affected by, rather than composers of, Big Data?]

The genomics collaboratory brought into focus the various collective efforts that are necessary to bring Big Data into being and into a state in which it could potentially have impacts on the social.

As Hannah stated, collaboratories provide windows on the Big Data problem space – the barriers and hurdles to data composition. These insights are antidotes to the ballistic narratives of Big Data autonomously spouting into being and self-propelling itself throughout society.

Even though those ‘impacted’ upon were not direct participants in our collaboratory, the collaboratory on waste in particular problematized the ability of Big Data to have a predetermined effect or impact on society. Whilst chips in bins had a positive effect on householders in Sweden, there was no guarantee that this technology and the associated Big Data would have the same impact in the UK. In other words, even social impact is not an inherent property of Big Data that can be baked into Big Data by careful composition; it too is relational.

# Ethic of Care

**Genomics already operates an ethic of care**

‘*A minor academic and professional industry has developed around the ethical, social and legal implications of genomics as data-intensive science. In Europe, UK, North America and several other countries, government-funded research has extensively researched ethical issues associated with genomics, mainly in the interests of protecting patients, citizens and the public in general from either losing control of their own data, or in the interests of helping various social groups manage potential disadvantages or discrimination associated with genetic data. … We would suggest that many framings of ethical issues associated with genomic data have been narrowly individualistic, and they have paid little attention to ethics already implicit to data practices’* (SBD WP pp. 22-23 Genomics and Ethics).

The above translates into a framework for interrogating our existing data on genomics from SBD and also into questions we could ask of future studies / collaboratories on Big Data and genomics such as:

* *how deeply are ethical concerns carried into data practice?*
* *in what ways does an ethic of care already operate in data curation?*
* *do not assume that ethics only relates to human research subjects or patients but also plays out in many different forms of relationships* (SBD WP pp. 22-23 Genomics and Ethics).

# PROTEE, Collaboratories, Social Intelligence and an Ethic of Care

The PROTEE indicators provide a framework for relating what can be learnt from collaboratories to enrich our social intelligence and develop understanding of what constitutes an ethic of care.

Using this approach, Social Intelligence focuses on taking into account the myriad entities that are implicated in the social life of Big Data, whilst an Ethic of Care focuses on how relations between entities are, and are not, being taken care of.

Realism, Strategy and Falsifiability are the names of the classes of PROTEE indicators. Whilst related, each class provides a different perspective and each draws out a different dimension of Social Intelligence and an Ethic of Care.

**Realism –** as an antidote to the ballistic fallacy of data-driven anything. Realism asks for attention to be paid throughout the entirety of the social life of BD. It looks at the social life of BD throughout its life course. In place of fragmented narratives of Big Data emerging like water from a fire hose and then proceeding to have utopian or dystopian socio-economic impacts, the objective here is to enrich collective Social Intelligence and understanding regarding the multiplicity of entities and their actions and interactions that are involved in the composition of Big Data and its ability to have social effects. It seeks a thick description – or rather – recognising the impossibility of consensus - a multitude of descriptions.

Here, an ethic of care would be to highlight the myriad actions that are, and can be, undertaken to care for the social relations of Big Data – to make it fit and available for others to use. It highlights the actions of curation that attend to the qualities of Big Data. Who undertakes such actions of care? And who benefits from them? What is made possible through them? What is impossible without them? And what about the social effects of Big Data. They do not just happen by themselves. How are they being taken care of, when and by whom?

**Strategy –** collaboratories also add to our Social Intelligence by highlighting what entities are *outside of* the dominant narratives and collectives – the stakeholders who are absent from the narratives (or the collaboratories) because they choose to be (e.g. commercial companies), or because they have been excluded. Yet although absent, they may still be implicated. Some of them deeply. Collaboratories enrich our Social Intelligence by attending to what kinds of entities and relations Big Data is excluding – highlighting the antagonistic relations and entities that cannot coexist with Big Data, or are being driven into extinction by it – or which exist in a different data economy.

Regarding an ethic of care, collaboratories are occasions to imagine the collaboratory from the outside – to imagine what strategies might involve the entities that are not participating? What could be done to relate to them – to include them – to entice them to participate?

Such imaginings would highlight which kinds of Big Data actions tend to be inclusive and which exclusive, and to which kinds of stakeholders.

It would cultivate recognition of what kinds of social relations Big Data is replacing and consideration of their worth.

It would also invite consideration of how their participation could alter the socio-economic dynamics of Big Data – for better or for worse. What would be displaced by their inclusion? Should they – can they – be involved at any cost? What are the opportunity costs of bringing them in and leaving them out?

‘*As applications of sequencing have broadened, uses and techniques for of analysing sequence data have expanded, but often in tension with existing scientific expertise (for instance, plant breeding vs. genetic modification; ecological field study vs sequence-based studies) (SBD WP Genomics Economies p. 18).*

**Falsifiability –** Falsifiability is about not closing down opportunities. It is about resisting homogenisation and taking responsibility for the future. It is about daring to keep innovating and resisting blindly falling into hegenomic ways of knowing and being. It is about constantly seeking alternative ways to truth through Big Data rather than relying on a few algorithms to dominate how we constitute our future worlds.

*‘Needless to say, algorithms and techniques developed initially for bioinformatics and genomic research have filtered out into other data-intensive sciences’ (SBD WP Genomics Economies p. 18).*

*‘Need to ensure that a single form of data practice does not homogenise or dominate domains to the exclusion of other ways of acting and knowing’ (SBD WP Genomics Economies p. 18).*

From a Social Intelligence perspective, the collaboratories would be occasions to interrogate how social meaning is being attributed to Big Data? Are new ways being sought? Are old ways being challenged? Are the social meanings of Big Data being challenged as time progresses, or is a settlement being quietly reached without contestation?

From an Ethics of Care perspective, this class of indicator would highlight the risks of not being vigilent, and the need for responsibility – for the need for critical voices that are a constant ‘thorn in the shoe’.