Development Informatics

Working Paper Series

The Development Informatics working paper series discusses the broad issues surrounding digital data, information, knowledge, information systems, and information and communication technologies in the process of socio-economic development

Paper No. 63

Data Justice For Development: *What Would It Mean?*

RICHARD HEEKS & JACO RENKEN

2016

ISBN: 978-1-905469-58-1

Published by: ***Centre for Development Informatics***
Global Development Institute, SEED
University of Manchester, Arthur Lewis Building, Manchester, M13 9PL, UK
Email: cdi@manchester.ac.uk Web: http://www.cdi.manchester.ac.uk

View/Download from:
http://www.gdi.manchester.ac.uk/research/publications/working-papers/di/

Educators' Guide from:
http://www.gdi.manchester.ac.uk/research/publications/working-papers/di/educators-guide/

Table of Contents

Data Justice for Development: *What Would it Mean?*

Richard Heeks & Jaco Renken
Centre for Development Informatics, University of Manchester, UK

2016

Abstract

This paper looks at the intersection of two growing trends in international development – use of justice in development theory, and use of data in development practice – and asks what data-justice-for-development would mean. The rationale for this can be the presence of current data injustices, and different framings for data injustice point to three different mainstream perspectives on data justice: instrumental, procedural, and distributive/rights-based. These three perspectives are explained but they are also subject to small data, sustainability, Senian, and structural critiques.

A full understanding of the mainstream perspectives and conceptualisation of the critiques is largely the task of a future research agenda. However, the paper does particularly argue that a structural approach should be the foundation for understanding data justice in a development context. It offers three potential ways to conceptualise structural data justice – through the ideas of Iris Marion Young, of political economy, and of the capability approach – and ends with some thoughts on the practical agenda when seeking to deliver structural data justice for development.

A. Introduction

From use of justice as the root of the capability approach in development (Sen, 1999) through the rights-based approach in development (Cornwall and Nyamu-Musembi, 2004) and the extension of ideas on justice to a global scale (Brock, 2009), there has been a continuous and growing engagement between justice and international development in recent decades. Separately, and much more recently, data is seen to play a far greater role in international development both now and in future than has previously been the case (Hilbert, 2016); heralding emergence of what we call "data-intensive development". Both "justice" and "data" therefore have a substantially greater prevalence in the Sustainable Development Goals documentation compared to that for the Millennium Development Goals (Heeks, 2014).

This paper locates itself at the intersection of these two growing phenomena. Specifically, it asks what "data justice for development" would mean. If justice is a – arguably the – foundational concept in development, then it is valuable to investigate its particular implications for the expanding field of data-intensive development. This is especially so given data is completely fundamental to international development as it is to all human life. Data is a primary, public good and no decisions and no actions can be taken without data, so development justice cannot be delivered without data justice. The relevance of a justice perspective for analysis and guidance of data systems has already been demonstrated through recent work on "data justice" (Newman, 2015) and from slightly more extensive discussion of the integrally-related notion of "information justice" (e.g. Johnson, 2014; Smith, 2001)[1].

This paper is based on analysis of literature; focusing mainly on the very small set of work on data/information justice; making some reference to the again-tiny corpus on technology justice (Practical Action, 2015; Practical Action, 2016; Trace, 2016); but also drawing to some degree on the extensive literature on global justice and social justice. Much of the work on data / information / technology justice is practice-oriented. That is, it focuses on actual harms, and it focuses on remedial interventions. But this is typically done in the absence of reference to underlying theories or principles of justice. This paper is different, in founding itself in – and focusing mainly on – different theories of justice, particularly developing the ideas of Johnson (2016a). This strong conceptual foundation adds new insights to the few writings on data/information justice, and it represents – as far as we are aware – the first attempt to discuss data justice in a development context; albeit many of the concepts could be universally applied.

The paper has four main sections: the first, looking at the rationale for data justice driven from the presence of data injustice in developing countries; the second, understanding the different conventional approaches that could be taken to data justice; the third, critiquing mainstream views of data justice including the argument for a more structural justice

[1] Data and information are, of course, distinct. We may, for example, define them in relation to one another: data as raw, unprocessed information; information as data that has been processed to make it useful to its recipient (Heeks, 2015). However, in practice, discussions do not make any clear or valid differentiation between data justice and information justice, thus enabling both literatures to be drawn upon.

approach; and finally, some reflections on implications of structural data justice for both theory and practice.

B. Data Injustice and Development

We can follow Sen's lead in order to explain the need for data-justice-for-development: ""In Sen's project, the place of the conception of justice is occupied by a sense of injustice' (Pogge and Alvarez, 2010)" (Martinez-Villa and Machin-Mastromatteo, 2016: 1286). In other words, it is the presence of data injustice within international development which prompts the requirement for data justice: a requirement both to define and to promote data justice for development.

We can find many examples of data injustices in a development context. These could be categorised in terms of particular harms e.g. surveillance, reinforcement of monopoly, algorithmic profiling, loss of privacy (Newman, 2015). However – and trailing ideas which will be developed further in the next section – it may be better to understand what is written about data injustice in terms of framings; as different ways of understanding data, data processes and data systems in international development.

Some framings relate to the way in which the data is used and the negative impact of that use. For example, the government in China has been criticised for using online data surveillance including hacking and tracking to identify potential political opponents and then harass and arrest them (MacKinnon, 2012). Or a report about customs officials in Haiti – bribed by a US rice distribution corporation – focused on the way they manipulated import data to reduce the import taxes the corporation had to pay (Smith and Jorna, 2011).

An alternative framing relates to the way in which data is handled. For example, mobile call detail records of large numbers of callers in West Africa were released to NGOs and others during the Ebola crisis and then used as the basis for data processing; but this was done without the consent of the individuals involved (McDonald, 2016). In India, the way in which public service data are handled has been criticised – particularly by poor citizens – as being inconsistent, biased and corrupt (Bhatnagar and Singh, 2010).

Thirdly, one finds framings that are related to more basic principles. For example, a principle of privacy of personal data was seen as being violated when individuals in Thailand found their personal photographs and details being circulated online for anyone to see (Kitiyadisai, 2005). A principle of access to data was seen as being violated when destitute communities in India were denied access to data about the way in which famine relief funding was being spent (Madon and Sahay, 2002). A principle of inclusion was breached when land records in India were digitised (Johnson, 2016b). The data the new records held was only quantified and formal. This was seen to exclude traditional, qualitative, informal data about land held within local communities; thus creating a bias in the type of data held.

Although there are material differences between these examples, their presentation is significantly a question of framing. For example, the Chinese government case could have been framed in terms of violation of data privacy. Or the Thailand example could have been

framed in terms of the way that personal data was being (mis-)used. This therefore suggests we may have choices between different perspectives on data injustice – hence choices between perspectives on data justice.

C. Mainstream Perspectives on Data Justice for Development

Data justice for development is required in order to address these various data injustices relating to developing countries; but what does data justice mean? We can define it in simple terms, modifying a general quote about justice (Johnson, 2016a: 2): "Data justice is the primary ethical standard by which data-related resources, processes and structures are evaluated". However, this is close to tautologous and does not provide much deeper insight.

Instead, and building on the idea of different framings discussed above, we can see that there are different types of justice which build on different conceptual foundations: what are sometimes called "different theories of justice" (Stumpf et al., 2016: 2). There are various categorisations of these different theories or types of justice, so for example, there are authors writing about types such as "interactional justice" (e.g. Colquitt et al., 2001) or "corrective justice" (Cohen, 2016). Here, we identify three conceptualisations of justice that are of particular relevance to data-intensive development: "instrumental justice", "procedural justice", and "distributive justice" (which we will interpret mainly in terms of "rights-based justice"). This categorisation adapts and develops particularly the work of Johnson (2016a) but this in turn echoes threads from much wider justice literature that utilises these categories (e.g. Sousa and Vala, 2002; Tomas, 2005; Wenzel, 2000; Young, 2011).

Instrumental data justice means fair use of data; it therefore focuses on the outcome of use of data (Johnson, 2016a). From this perspective, there is no justice inherent to the data domain; instead justice is defined outwith that domain. For example, this would argue that there is no inherent justice or injustice about who owns data in developing countries or in development projects; concerns about justice only relate to the impact of the use of that data. From this perspective, the injustice in the Haiti case above was not related to manipulation of data but to the fact that the outcome of that manipulation was unjust in denying the Haitian government its legally-owed taxes. In the case of the Chinese government, this perspective sees no injustice in the way data on opponents was accessed nor in the fact that it was accessed, but in unjust use of that data for the purposes of harassment and arrest. At its strongest, this view would argue against any notion of data rights – such as data privacy – for their own sake, with constraints on access to data only relevant to the extent that data might then certainly be used for unjust purposes (*ibid.*). In a way, then, there is no such thing as data justice *per se* in this view: we do not need any fresh or specific consideration of justice from a data perspective. Instead, justice is defined from a broader perspective; for example by reference to broader definitions of what is legal and illegal such as discrimination on grounds of race and gender; or, given the likelihood that the

Chinese government is not doing anything illegal according to its own laws, by reference to wider codes of social and political justice.

Procedural data justice means fair handling of data. Building from the literature on procedural justice and on data, this can be conceived as consisting of two dimensions. One dimension relates to the scope of data-handling processes: a narrow interpretation would encompass only processes that handle data – its capture, input to a data system, processing, storage and output. These are the 'upstream' components of the information value chain (see Figure 1, adapted from Heeks (2006)). A broader interpretation would include the use of that data (as information) for decisions and actions: so also including the 'downstream' components of the information value chain and incorporating both human and technology-based elements (thus covering algorithmic decision-making, and smart devices that take autonomous decisions and actions). The latter interpretation would then clearly overlap with instrumental data justice views, given they both involve use of data.

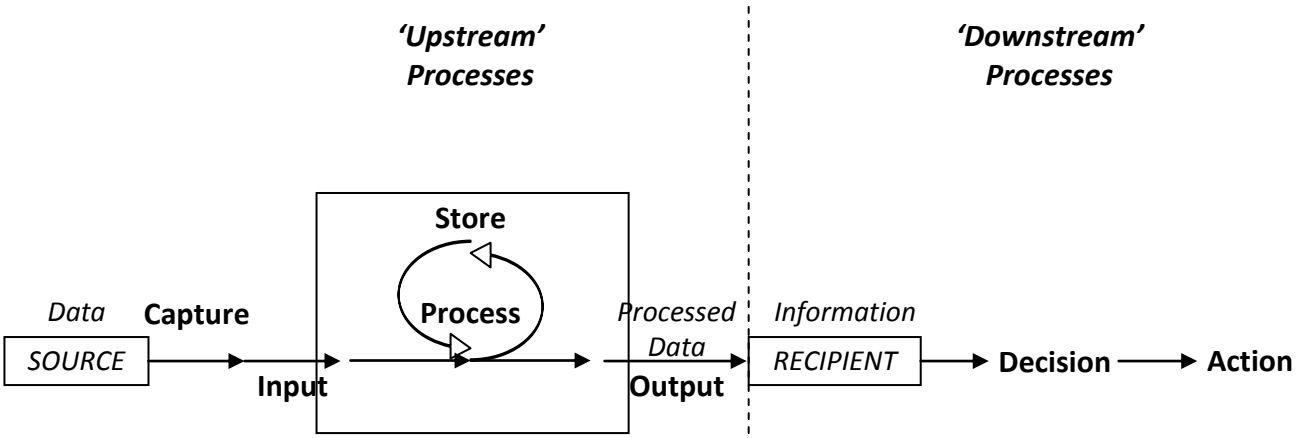

Figure 1. The information value chain and scope of data-handling processes.

The second dimension relates to fairness of data-handling processes: a narrow interpretation would encompass only control – individuals feel a process is fair if they have some control over that process (e.g. Thibaut and Walker, 1975). This is sometimes seen as meaning some direct involvement in a process but that is not practicable for modern-day data handling. One approach is to understand this in terms of consent: a data process is fair if an individual has given consent to that process (Johnson, 2016a). In practice, this would restrict procedural data justice to processes that make use of data about an individual, as opposed to the wider spectrum of all data processes that might affect an individual. The idea of consent could mean an explicit consent to each specific act of data communication, or an implicit consent based on giving over data distribution to other parties in exchange for services provided; as when we enter our details into an e-commerce web site (*ibid.*). From this perspective, the Ebola-related example above would be unjust if mobile phone callers provided neither explicit nor implicit consent for their data to be used for health-related purposes. Conversely, some forms of state surveillance might be just if citizens see them as a fair exchange for the security that the state provides.

A broader interpretation of this second dimension would judge fairness of data processes not only in terms of control but in terms of other perceptions that individuals are found to use when asked to judge processes (Colquitt et al., 2001; Tyler, 1988). These could include: consistency (that the process is always performed in the same way); correctness (process accuracy and freedom from bias); and correctability (that any errors in the process can be put right). For example, the public service case from India cited above provides little or no control for poor citizens over the way in which their data is handled by government but, further, they perceive these processes as unfair due to their inconsistency and incorrectness.

The dimensions and their different interpretations are summarised in Table 1.

		Scope of Data-Handling Processes	
		'Upstream' only: CIPSO	Whole value chain: CIPSODA
Scope of Process Fairness Definition	Control/Consent only	*Narrow*	*Mid-Range*
	Control/Consent; Consistency; Correctness; Correctability	*Mid-Range*	*Broad*

Table 1. Breadth of interpretation of procedural data justice.

Distributive data justice means fair distribution of data. This could be understood very directly from a resource-based perspective in terms of who has what data, given the negative socio-economic impacts of data distribution asymmetries (Newman, 2015). However, it can also be understood from a rights-based perspective (Johnson, 2016a). The justification for this – given distributive justice and rights-based justice are generally treated as two separate approaches – is that data rights determine distribution of data. The most frequently-cited right is data privacy: typically seen as the right for particular items or forms of data about oneself to be withheld from particular others in particular circumstances or for particular purposes (Barker et al., 2009). Hence, privacy is a determinant of distribution: as a right, it partly determines who does not – or at least should not – access certain data. As already noted, the Thailand example and also the Chinese government example above can be understood as a violation of the right to data privacy. But there are other distribution-related data rights to consider. Some of these, as with privacy (which derives from Article 12), can be derived from the UN Universal Declaration of Human Rights (UN, 1948) (Smith, 2001):

- Right of data access (from Article 19): directly underpinning the distribution of who can obtain different types of data. It would be difficult to make this universal i.e. that we have a right to all data; especially as that would clash with the right of data privacy. But it can be understood as a right to data of particular importance; for example, the right to information about the workings of government that underpins the current open government data movement (Gonzalez-Zapata and Heeks, 2016), and which was exemplified in the injustice of the Indian famine example.

- Right of data ownership (from Articles 17 and 27): the original meaning would be a right over data that we produce as "authors" of that data. This could be extended to cover data produced as a by-product of our actions – our "data exhaust" – and perhaps also to data that is produced about us given we are ultimately the origin of that data. What constitutes protection of our "moral and material interests" is also a matter for debate but would presumably include the way in which that data is used, and also the issue of how the benefits from our personal data are distributed (given the potential asymmetry of benefits between data platforms and their users: Newman (2015)).

We could also add:

- Right of data representation/inclusion: one might argue this as an extension of the notion of freedom of expression (Article 19 again). It relates to a right to be represented in datasets: for one's opinions but also other data about oneself to be included. From this perspective, the Indian land records case is an example of rights-based injustice: a violation of the right of the local community for their views to be incorporated into digital land records.

These suggest it is not just the patterns of distribution of data that matter – as partly determined by rights of data privacy and access – but also the patterns of the benefits from that data, and the patterns of distribution within data.

Before moving on, we note two emergent issues. First, the question of what constitutes "fair", given it was contained within all our definitions of the different understandings of data justice. Fairness can be defined interpretively and bottom-up: fairness is what those involved perceive fairness to mean. Although discussed in relation to procedural data justice, this could equally be applied in an inductive manner to find out, e.g. from a group of citizens, what they would regard as fair distribution of data or fair use of data. Alternatively, fairness can be defined as adherence to the law, on the assumption that fairness has been a guiding principle in the creation of laws. Again, though discussed in relation to instrumental data justice, one could also see data processes and distributions as fair if they legal and unfair if they are illegal.

Laws vary between countries, so a third way to define fairness – seeking to move right away from interpretive, bottom-up notions – is in terms of universal principles (Johnson, 2016a). One such principle is equality; most-associated with a rights-based approach, i.e. that everyone should have an equal right to data privacy, access, ownership, representation. This could be extended to procedural data justice, for instance considering equal data treatment for all. And it could be extended to instrumental data justice by invoking the notion of utility – economic or other measures of individual benefit – and seeking equal utility for all from the uses of data. Alternative principles could be utilitarianism: that fairness in relation to data means creating the greatest utility for the greatest number; or more complex formulations such as Rawls' maximin principle of seeking to maximise the minimum level of data-related utility that any one person must endure (*ibid.*; Rawls, 1974).

Second, we have largely set aside the "for development" component; assuming that there are universal perspectives on data justice that must first be identified and conceptualised; and that the "for development" component means the application of those universal perspectives within a development context – typically taken to mean in the context of developing countries. It remains a question for future research – albeit we will touch on it

slightly in the next section – whether there are types, even theories, of data justice that could in some way be specific to development. At present, it is hard to identify the basis for such particularism.

D. Critiquing Data Justice for Development

Data justice for development could be based on one of the three approaches to data justice already discussed: instrumental, procedural and distributive rights-based. As noted, these are "mainstream" views that appear in many discussions of justice (albeit more rarely for instrumental justice). However, we can also identify four critiques of these approaches, which could guide us to alternative perspectives.

First, that the mainstream views take a big data perspective. By and large they assume we as individuals are sources and producers of data that may then be aggregated and used by others, particularly by the state or by large corporations. An alternative approach is the small data justice perspective that can be derived from Practical Action's "technology justice" initiative (see also UNU-CS (n.d.)). This reorients to the livelihood needs of individual citizens as data users more than data producers; driven especially by the needs of those in developing countries. For example, we could rewrite one of their statements (adapted from Practical Action (2015: 1)):

> "Data Justice for Development works to achieve a state of *data justice*: where people have the right to decide, choose and use data that assists them in leading the kind of life they value, without compromising the ability of others and future generations to do the same.
>
> The goal of *data justice* demands universal access to the data and data services needed for a reasonable quality of life. It also demands that consumers and developers of data systems consider their impact on other people, on the planet, and on our future, to guide more just innovation and use of data."

Though unexplored to date, we can start to see an outline of how small data justice would adapt mainstream views. It incorporates a distributive rights-based approach given the centrality of access to data and related services for everyone, but this is a right of access with a particular focus: not to all types of data but homing in on the data that people need in order to live the life they value. It incorporates an instrumental approach given the emphasis on use of data; but use by individuals themselves – perhaps as groups or communities – rather than by others: this is the essence of small data. Finally, it has a procedural element that covers not only the upstream and downstream elements discussed above, but what we might call an "extreme upstream" position by considering not just processes of data-handling and use but the prior processes of data systems design. This connects with the idea that system design processes affect system processes and outcomes (e.g. Heeks, 1998): in other words, that you can only produce a just data system through a just data system design process (Eubanks, 2014).

The adapted quote above incorporates one other criticism of mainstream views on justice: their lack of consideration for the issue of sustainability (Dobson, 2003). While there has

certainly been work discussing the integration of justice and sustainability (e.g. Ikeme, 2003), there are theoretical and practical challenges to this (Klinsky and Golub, 2016). This would be especially true for data. Given the particular features of data – its immateriality, its roughly zero-cost handling processes including reproduction – it is not clear what data sustainability would actually mean in the context of data justice. In relation to rights, one might invoke either a right for one's data to be preserved or the opposite: the "right to be forgotten" (Ambrose, 2013). But the greater weight leans towards an instrumental interpretation and the longer-term impact of data and its uses, perhaps including the impact of data systems such as their carbon footprint. All this in encapsulated in use of some variant of the well-known phrasing of the Brundtland Report, that use of data and data systems "meets the needs of the present without compromising the ability of future generations to meet their own needs" (UNWCED, 1987: 41). Overall, though, the particular connection between data justice and sustainability requires further specification.

In speaking of "the kind of life they value", the adapted quote above give a nod towards the work of Amartya Sen, and a third criticism of traditional views on justice that arises from a Senian perspective (Britz et al., 2013; Martinez-Villa and Machin-Mastromatteo, 2016). This argues that justice must be understood not as a set of theoretical principles but as justice-in-practice, including consideration of human agency and thus allowing for individual differences including capabilities and motivations. In relation to motivation, this would allow that some people choose what might otherwise be seen as less-just options; for example, choosing to give up their data privacy in return for services and benefits, or choosing not to access particular types of data even though they have a right to do so.

In relation to capabilities, this view would explain that – while people may have access to justice in theory – they cannot access that justice in practice because they are unable to do so; because they lack the capabilities (Johnson, 2014). Users who lack a high level of literacy, for example, cannot give informed consent given the complexity of most consent agreements. Or, picking a specific example, when government data in Chile was opened up with many datasets being made available online, only a few specialist NGOs were able to make use of the data because individual citizens lacked the capabilities to find, to understand or to use that data (Gonzalez-Zapata and Heeks, 2015). In all these examples, what we move towards is a more differentiated and agency-oriented view of data justice based on what individuals value being and doing, what they choose to be and do, and what they are able to be and do.

But discussion of the capabilities approach brings us to the fourth criticism of mainstream views: their failure to consider social structure. Capabilities and choice are enabled and constrained by broader social structures, though Sen himself has been criticised for underplaying this (Deneulin and McGregor, 2010). With data justice, the structural criticism is harsher: that social structure is absent – that the instrumental, procedural, and distributive rights-based approaches do not encompass the social structures which at least partly determine data uses, processes, distributions, and rights (Johnson, 2016a).

Why is access to data maldistributed in the global South? Why is participation in data processes unequally distributed? Why do the benefits of data systems in developing countries include some and exclude others? In significant part because of the structural

conditions that shape the distribution of data resources and practices (Burns, 2015; Johnson, 2014). If citizens of India are prevented from accessing information about government, a key determinant is the structural relations between government and citizens in India (e.g. Thomas, 2010). If China's government unjustly uses online data to harass opponents, a key determinant is the structural relations between government and citizens in China (e.g. Zhao, 2008). This is arguably reflected in article 28 of the UN Universal Declaration of Human Rights (UN, 1948) which entitles everyone to "a social and international order" in which all other rights and freedoms can be fully realised.

Hence the argument that the foundation of data justice must be structural data justice, which we can define as "the degree to which society contains and supports the data-related institutions, relations and knowledge systems necessary for realisation of the values comprised in a good life" (adapting Young (1990: 37) from Johnson (2014: 265)). As with many definitions and discussions of justice it requires a further definition; in this case of what a good life comprises. But, if we accept this is the foundation for data justice, then it already provides us some implications because we have to look beyond instrumental, procedural, and distributive rights-based understandings.

E. Ways Forward for Data Justice for Development

The previous section's discussion provides just a first sensitisation in relation to the "4Ss" of critique – small data, sustainability, Sen, and structure – and some initial directions for future research. In this final section, we consider further the conceptual and practical implications of these critiques for data-justice-for-development; looking particularly at the last, which argues that the foundational perspective on data justice should be structural. We begin with three different ways to conceptualise structural data justice.

One could follow the argument of cosmopolitanism. This might begin with a Rawlsian approach given the "dominance" of John Rawls' ideas on justice (Dower, 2006: 494) and since his is a quasi-structural perspective that equates justice with social structure being "the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation" (Rawls, 1971: 7). However, Rawls is only quasi-structural. Beyond the basic idea of institutions reflected in the quote above, "Rawls says little more about what the concept of structure refers to" (Young, 2006: 111). His is therefore an abstract, "transcendental" view of social structure that emphasises "finding the ideal institution for justice, ignoring existing relations of power" (Martinez-Villa and Machin-Mastromatteo, 2016: 1286). Rawls is therefore unlikely to provide the necessary scope of insight needed for a theory of structural data justice, nor address the "justice-in-practice" critique.

However, extensions of his work in the cosmopolitan tradition building outwards to global justice may have something to offer. In particular, the work of Iris Marion Young who takes a network view of social structure – what she calls her "social connection model" – seeing individuals occupying particular social positions and that structure "consists in the connections among these positions and their relationships, and the way the attributes of positions internally constitute one another through those relationships" (Young, 2006: 112).

As per the earlier critique, she also folds agency into her approach, seeing structure "as a process", "not as a state" (*ibid.*), and arguing that all those who participate in unjust structures – such as those related to data injustice – including the "victims", share responsibility even if not equal responsibility.

An alternative direction of conceptualisation might align with the emergent notion of "critical data studies" (Dalton et al., 2016). "Critical theorists hold that technological artefacts are the result of political struggles, so they are carriers of specific interests" (Silva, 2007: 171). From this view, data and data systems are the result of political struggles between differing interests. A first task of a critical perspective on structural data justice will then be to understand the conflicting interests – particularly the conflicting interests relating to data rights, processes, distribution and uses – that arise, and the mechanisms of power through which those interests come to either strengthen or weaken.

Institutional logics could be one basis for this but that would require identification of fairly distinct worldviews around data, and would also require further conceptual development of the notion of power within institutional logics; something that is currently a work in progress (Thornton et al., 2012). Simpler at present might be to follow a more traditional political economy approach that defines interests from the explicit and implicit elements within statements and actions, and which defines power in terms of a set of social structures. It fits with the argument that political economy analysis is the central task of development studies (Payne and Phillips, 2010).

A typical set of social structures – reflected also in the work of Young (2006) and summarised in Figure 2 – would encompass knowledge systems (discourse), institutions (including technology), relations, resources and data practices: the latter being the information value chain shown above. One could develop this in a more data-specific direction by using Kitchin's notion of "data assemblages": "all of the technological, political, social and economic apparatuses that frame [*the*] nature, operation and work" of data systems (Kitchin and Lauriault, 2014: 6). This provides detailed examples of data systems-specific social structure incorporating knowledge systems, institutions, resources and practices though saying little about structural relations and instead, given its geography discipline origins, focusing instead on place. Whichever the particular view of political economy and social structures of power taken, this would analyse the way in which structure determines data justice: for example, the way in which structural inequalities of capitalism or of governance determine inequalities in distributions of data or in the implementation of data rights.

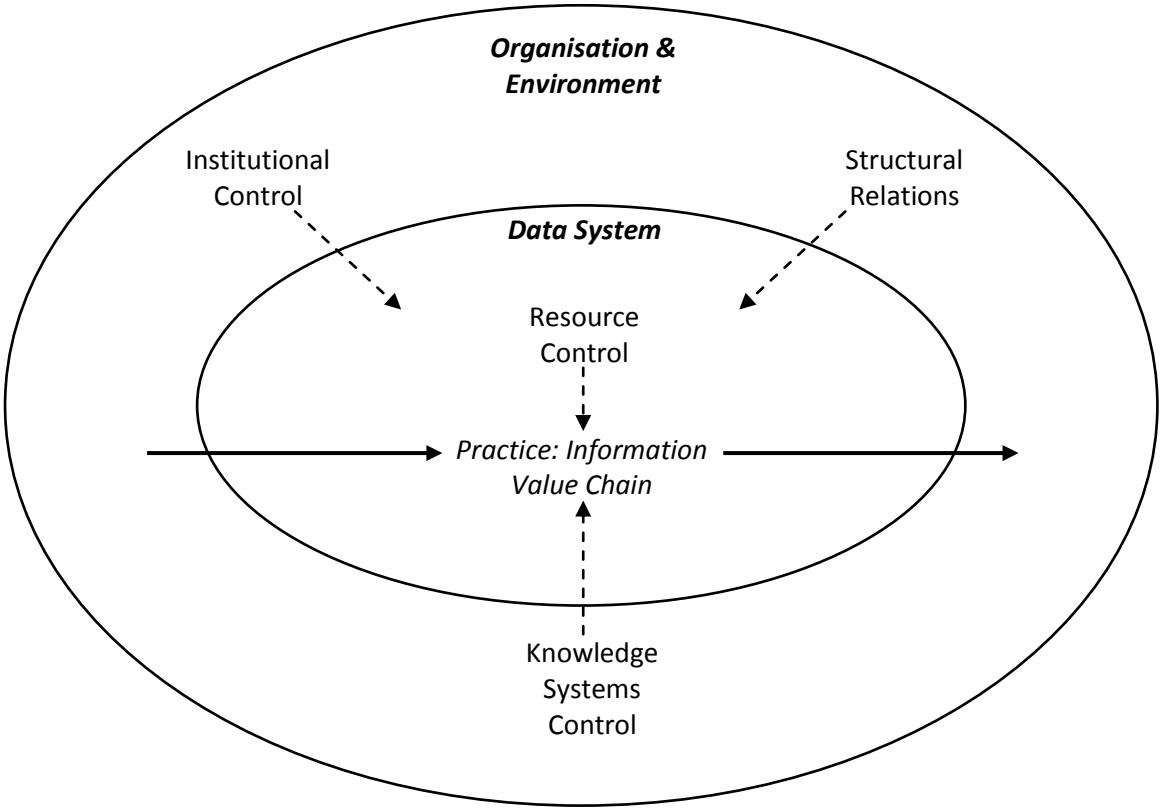

Figure 2. The political economy of data systems.

A potential shortcoming is that such analysis may lapse into structural determinism; thus missing the role of human agency that was brought up explicitly in the Senian critique above and implicitly in the small data critique. A response would be a revised critical approach; for example that of critical modernism which is a small strand of ideas within development studies that acknowledges the role of social structure in significantly shaping development outcomes but also the potential for human agency and choice and outcomes other than simple reproduction of existing structures; and which also holds some contingent optimism about the contribution that science and technology – such as data innovations – can make to development (Hickey and Mohan, 2004; Langford, 2015). Operationalising a critical modernist view of structural data justice would mean incorporating the power of agency – "power to" – alongside structural power – "power over": politics in practice as well as politics in the abstract. Mirroring Young's social connection model, this would potentially be a network view of power given this can incorporate both structural and agentic views of power, and given it reflects the multiple networks – both local and global – into which data systems connect citizens in developing countries (Benkler, 2011; Castells, 2000)[2]. To ensure this forms an analysis of structural data justice, as a reminder, the focus would be on how

[2] Other directions this network power-based investigation of structural data justice could take might include work using actor-network theory to understand how networks assemble and disassemble around particular views or applications of data systems (e.g. Heeks and Stanforth, 2015); and work using the digital systems-specific conception of the power of connective action (Bennett and Segerberg, 2012).

network power emerges around conflicting interests to shape the fairness of data rights, processes, distribution and uses.

A third alternative direction of conceptualisation would be an amended capabilities approach to data justice, deriving from the work of Amartya Sen and others. This has an immediate advantage of being built around – rather than having to be oriented to focus upon – if not a precise definition of justice, then at least an understanding that justice means increasing the freedoms of individuals to be and to do what they value in life (Sen, 1999). It also has the potential to be an all-encompassing view of justice – as Figure 3 (adapted from Robeyns (2005) and Wilson-Strydom (2011)) suggests – rather than just addressing the structural data justice critique.

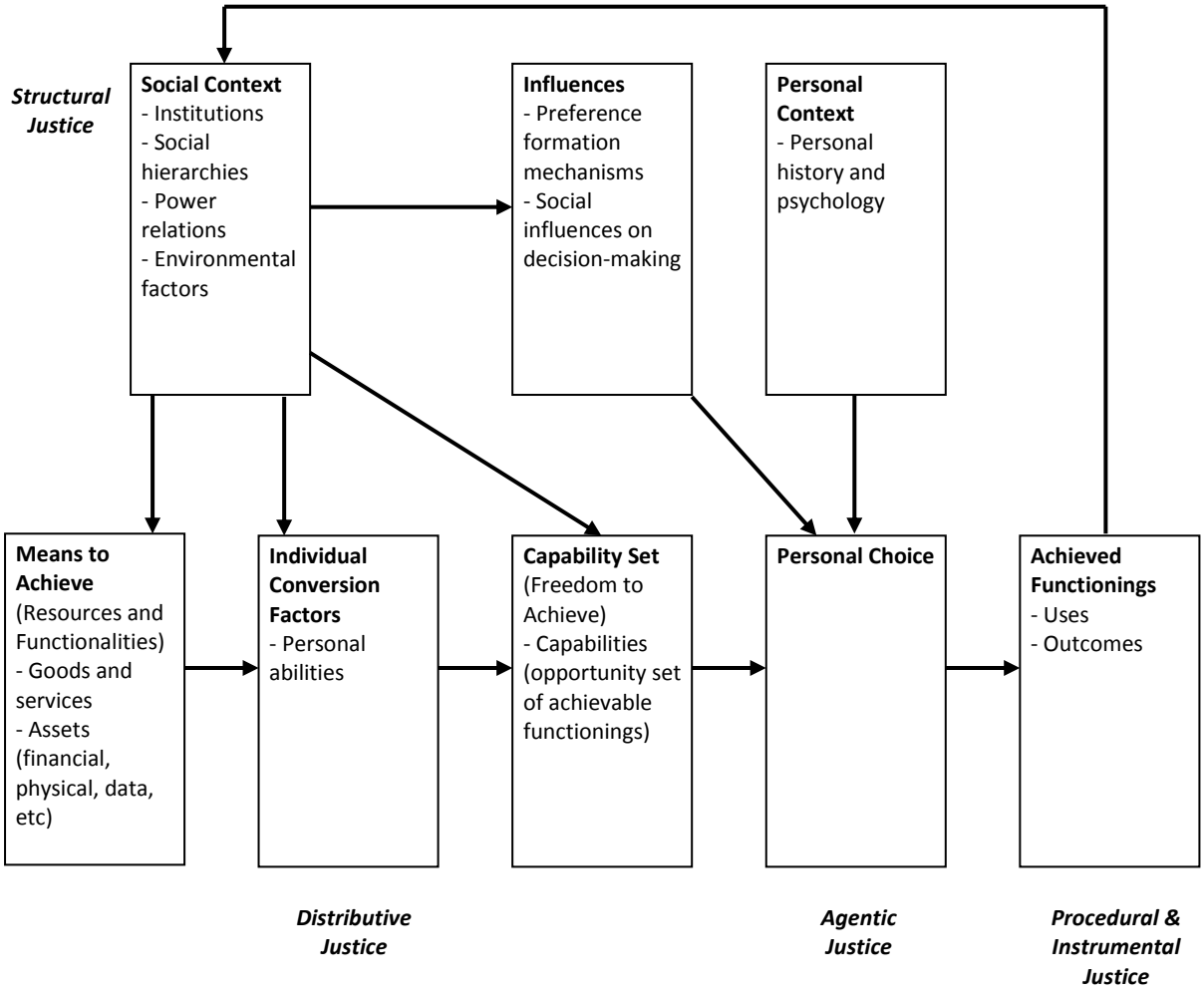

Figure 3. Components of justice within the capabilities approach.

As noted above, Sen is typically judged as under-emphasising social structure, with a liberal individualism that sidelines collectivities and would tend to reproduce rather than challenge unjust structures (Carpenter, 2009). However, the capabilities approach can be redefined from a more critical perspective to give greater weight to the role of social structure; in line with the ideas of structural data justice. As with all three routes to conceptualisation,

operationalisation would require identification of the particular social structures that impinge on data justice in a development context. For example, the social norms that make education less accessible to girls than to boys; thus constraining the ability of women to participate in data processes and to exercise their data rights (Britz et al., 2013).

The capabilities approach has a strength in acknowledging inequalities and their impact. It can therefore readily incorporate the simpler view of distributive data justice: not just mapping the unequal distributions of data itself and data-related resources such as technical infrastructure and data systems, but also the capabilities associated with the information value chain (*ibid.*). This would of course be an especial strength of a capabilities perspective on data justice: helping understand why data justice for development means more than just making data freely available. Making data freely available is just one element required to produce just data functionings: there must also be fair distribution of other means to achieve and conversion factors, plus a conducive social/personal context and influences.

Emphasis on the personal – individual conversion factors, influences and context – is another prominent feature of the capabilities approach, and a reminder that it highlights agency and justice-in-practice. This is reflected most clearly in the idea of personal choice and the conversion of potential capabilities into actual functionings. To acknowledge this feature, we add the "agentic justice" label to make manifest the role of individual agents in the practice of data justice. This also points to the potential place of data rights in the model. They would conventionally be understood as institutions – particular norms perhaps with legal or related backing – that might, given presence/distribution of relevant data-related means to achieve and conversion factors, then be converted into rights-as-capabilities: opportunities for data access, data privacy, etc. But those rights would only be achieved through agency: the personal choice to exercise rights-as-capabilities and convert them into rights-as-achieved-functionings. Lastly, those and other achieved functionings represent the processes and uses of data, and thus enable the capabilities approach to include a sense of procedural and instrumental data justice.

Despite the connection with development studies – being arguably the disciplinary origin of critical modernism and of the capabilities approach – we find once again that these conceptualisations of data justice are generic. The "for-development" angle would mostly be interpreted as applying these in a developing country context. For example, understanding the place of developing countries within global and historical structures of political economy, or recognising the resource and capability constraints typical of developing countries; both of which may make data justice harder to achieve. In this, as in many of the ideas discussed above, there is a major research agenda: developing the conceptual foundations of data justice identified, and operationalising them in development contexts.

And finally, we asked above what the practical implications would be for data-justice-for-development of taking a structural data justice view. The specifics might vary depending on the conceptualisation of structural data justice selected. However, the common objective would be to alter existing, or create new, social structures of relevance to data justice in developing countries. For example, creating a Data-Justice-for-Development social

movement to partly act as a counter to the dangers of state or corporate data hegemony; perhaps building out from the nascent interest in "data4good" (e.g. ICML, 2016).

This would have an activities agenda – perhaps this might be developed into a Data-Justice-for-Development Manifesto – drawn from the mainstream views (Johnson, 2014): demanding just and legal uses of development data; enabling data consent that was truly informed; building upstream and downstream data-related capabilities among those who lack them; promoting rights of data access, data privacy, data ownership and data representation. It would draw from critiques: supporting small data uses by individuals and communities in developing countries; advocating sustainable use of data and data systems. And it would be structural. It could be structurally-just in itself: such a social movement would provide the "data subalterns" of the global South with connective power.

But it would also be structurally-just in its activities: stimulating an alternative discourse around data-intensive development that placed issues of justice at its heart; developing new organisational forms such as data-intensive development cooperatives; and arguing for change in data-intensive development institutions. Conventionally, this would mean new data justice-based laws and policies (including action on data monopolies), but it would also mean recognising data systems themselves as institutions of social structure. Code, algorithms, standards, etc of the data systems used in international development are accumulating parts of the social fabric that may lead to unjust outcomes. A social movement would need to open up, challenge and provide alternatives to such technical structures if data-justice-for-development is to be delivered in practice.

References

Ambrose ML (2013) It's about time: Privacy, information life cycles, and the right to be forgotten. *Stanford Technology Law Review* 16(2): 369-421.

Barker K, Askari M, Banerjee M, Ghazinour K, Mackas B, Majedi M, Pun S and Williams A (2009) A data privacy taxonomy. In: Sexton AP (ed) *Dataspace: The Final Frontier*. Berlin: Springer, pp. 42-54.

Benkler Y (2011) Networks of power, degrees of freedom. *International Journal of Communication* 5: 721-755.

Bennett WL and Segerberg A (2012) The logic of connective action. *Information, Communication & Society* 15(5): 739-768.

Bhatnagar SC and Singh N (2010) Assessing the impact of e-government. *Information Technologies and International Development* 6(2): 109-127.

Britz J, Hoffmann A, Ponelis S, Zimmer M and Lor P (2013) On considering the application of Amartya Sen's capability approach to an information-based rights framework. *Information Development* 29(2): 106-113.

Brock G (2009) *Global Justice: A Cosmopolitan Account*. Oxford, UK: Oxford University Press.

Burns R (2015) Rethinking big data in digital humanitarianism: Practices, epistemologies, and social relations. *GeoJournal* 80(4): 477-490.

Carpenter M (2009) The capabilities approach and critical social policy: Lessons from the majority world?. *Critical Social Policy* 29(3): 351-373.

Castells M (2000) Materials for an exploratory theory of the network society. *British Journal of Sociology* 51(1): 5-24.

Cohen A (2016) Corrective vs. distributive justice. *Ethical Theory and Moral Practice* 19(3): 663-677.

Colquitt JA, Conlon DE, Wesson MJ, Porter CO and Ng KY (2001) Justice at the Millennium: a meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology* 86(3): 425-445.

Cornwall A and Nyamu-Musembi C (2004) Putting the 'rights-based approach' to development into perspective. *Third World Quarterly* 25(8): 1415-1437.

Dalton C, Taylor L and Thatcher J (2016) Critical data studies: A dialog on data and space. *Big Data & Society* Jan-Jun: 1-9.

Deneulin S and McGregor JA (2010) The capability approach and the politics of a social conception of wellbeing. *European Journal of Social Theory* 13(4): 501-519.

Dobson A (2003) Social justice and environmental sustainability: Ne'er the twain shall meet. In: Agyeman J, Bullard RD and Evans B (eds) *Just Sustainabilities*. London: Earthscan, pp. 83-95.

Dower N (2006) Rawls, John. In: Clark DA (ed) *The Elgar Companion to Development Studies*. Cheltenham, UK: Edward Elgar, pp. 494-499.

Eubanks V (2014) Big data and human rights. In: Gangadharan SP (ed) *Data and Discrimination*. Washington, DC: Open Technology Institute, pp. 48-52.

Gonzalez-Zapata F and Heeks R (2015) The multiple meanings of open government data: Understanding different stakeholders and their perspectives. *Government Information Quarterly* 32(4): 441-452.

Gonzalez-Zapata F and Heeks R (2016) The influence of the transparency agenda on open government data in Chile, paper presented at *CeDEM16: International Conference for E-Democracy and Open Government 2016*, Krems, Austria, 18-20 May.

Heeks R (1998) *Information Systems and Public Sector Accountability*. Information Systems for Public Sector Management working paper 1. Manchester, UK: IDPM, University of Manchester.

Heeks R (2006) *Implementing and Managing eGovernment*. London: Sage.

Heeks R (2014) From the MDGs to the Post-2015 Agenda: Analysing Changing Development Priorities. Development Informatics Working Paper no. 56. Manchester, UK: IDPM, University of Manchester.

Heeks R (2015) *Basic Models of Information and Information Systems*. Manchester, UK: IDPM, University of Manchester.

Heeks R and Stanforth C (2015) Technological change in developing countries: Opening the black box of process using actor-network theory. *Development Studies Research* 2(1): 33-50.

Hickey S and Mohan G (2004) Relocating participation within a radical politics of development: critical modernism and citizenship. In: Hickey S and Mohan G (eds) *Participation: From Tyranny to Transformation?* London: Zed Books, pp. 59-74.

Hilbert M (2016) Big data for development. *Development Policy Review* 34(1): 135-174.

ICML (2016) *#Data4Good: Machine Learning in Social Good Applications*. Conference Proceedings. New York, NY: ICML. http://arxiv.org/html/1607.02450v2 (accessed 30 Aug 2016).

Ikeme J (2003) Equity, environmental justice and sustainability. *Global Environmental Change* 13(3): 195-206.

Johnson JA (2014) From open data to information justice. *Ethics And Information Technology* 16(4): 263-274.

Johnson JA (2016a) The value—and limits—of distributive justice in a justice-centered approach to information privacy, paper presented at *Western Political Science Association 2016 Annual Meeting*, San Diego, CA, 23-26 Mar.

Johnson JA (2016b) The question of information justice. *Communications of the ACM* 59(3): 27-29.

Kitchin R and Lauriault TP (2014) *Towards Critical Data Studies: Charting and Unpacking Data Assemblages and Their Work*. Programmable City Working Paper 2. Maynooth, Ireland: National University of Ireland Maynooth.

Kitiyadisai K (2005) Privacy rights and protection: foreign values in modern Thai context. *Ethics and Information Technology* 7(1): 17-26.

Klinsky S and Golub A (2016) Justice and sustainability. In: Heinrichs H, Martens P, Michelsen G and Wiek A (eds) *Sustainability Science*. Dordrecht, Netherlands: Springer, pp. 161-173.

Langford M (2015) Rights, development and critical modernity. *Development and Change* 46(4): 777-802.

MacKinnon R (2012) *Consent of the Networked: The Worldwide Struggle for Internet Freedom*. New York, NY: Basic Books.

Madon S and Sahay S (2002) An information-based model of NGO mediation for the empowerment of slum dwellers in Bangalore. *The Information Society* 18(1): 13-19.

Martinez-Villa BA and Machin-Mastromatteo JD (2016) Four theories to improve justice in Latin America. *Information Development* 32(4): 1284-1288.

McDonald SM (2016) *Ebola: A Big Data Disaster*. Paper 2016.01. Bengaluru, India: The Centre for Internet and Society.

Newman N (2015) *Data Justice: Taking On Big Data As An Economic Justice Issue*. New York, NY: Data Justice.

Payne A and Phillips N (2010) *Development*. Cambridge, UK: Polity.

Pogge T and Alvarez D (2010) Justicia global: Dos enfoques. *Isegoría* (43): 573-588.

Practical Action (2015) *Technology Justice*. Policy Position Paper. Rugby, UK: Practical Action.

Practical Action (2016) *Technology Justice*. Rugby, UK: Practical Action. http://practicalaction.org/technology-justice (accessed 30 Aug 2016).

Rawls J (1971) *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Rawls J (1974) Some reasons for the maximin criterion. *The American Economic Review* 64(2): 141-146.

Robeyns I (2005) The capability approach: A theoretical survey. *Journal of Human Development* 6(1): 93-117.

Sen A (1999) *Development as Freedom*. Oxford, UK: Oxford University Press.

Silva L (2007) Epistemological and theoretical challenges for studying power and politics in information systems. *Information Systems Journal* 17(2): 165-183.

Smith M (2001) Global information justice: rights, responsibilities, and caring connections. *Library Trends* 49(3): 519-537.

Smith RG and Jorna P (2011) Corrupt misuse of information and communications technologies. In: Graycar A and Smith RG (eds) *Handbook of Global Research and Practice in Corruption*. Cheltenham, UK: Edward Elgar, pp. 255-284.

Sousa FH and Vala J (2002) Relational justice in organizations: The group-value model and support for change. *Social Justice Research* 15(2): 99-121.

Stumpf KH, Becker CU and Baumgärtner S (2016) A conceptual structure of justice: Providing a tool to analyse conceptions of justice. *Ethical Theory and Moral Practice*. Epub ahead of print 20 April 2016. DOI:10.1007/s10677-016-9728-3.

Thibaut J and Walker L (1975) *Procedural Justice: A Psychological Analysis.* Hillsdale, NJ: Erlbaum.

Thomas PN (2010) *Political Economy of Communications in India*. New Delhi: Sage Publications.

Thornton PH, Ocasio W and Lounsbury M (2012) *The Institutional Logics Perspective: A New Approach to Culture, Structure, and Process*. Oxford, UK: Oxford University Press.

Tomas A (2005) Reforms that benefit poor people: Practical solutions and dilemmas of rights-based approaches to legal and justice reform. In: Gready P and Ensor J (eds) *Reinventing Development*. London: Zed Books, pp. 171-184.

Trace C (2016) *Rethink, Retool, Reboot: Technology as if People and Planet Mattered*. Rugby, UK: Practical Action Publishing.

Tyler TR (1988) What is procedural justice?: Criteria used by citizens to assess the fairness of legal procedures. *Law and Society Review* 22(1): 103-135.

Wenzel M (2000) Justice and identity. *Personality and Social Psychology Bulletin* 26(2): 157-176.

UN (1948) *United Nations Universal Declaration of Human Rights 1948*. New York, NY: United Nations.

UNU-CS (n.d.) *Small Data*. Macau: United Nations University Computing and Society.

UNWCED (1987) *Our Common Future*. Geneva, Switzerland: UN World Commission on Environment and Development.

Wilson-Strydom M (2011) University access for social justice: A capabilities perspective. *South African Journal of Education* 31(3): 407-418.

Young IM (1990) *Justice and the Politics of Difference*. Princeton, NJ: Princeton University Press.

Young IM (2006) Responsibility and global justice: A social connection model. *Social Philosophy and Policy* 23(1): 102-130.

Young IM (2011) *Justice and the Politics of Difference*, 2nd edn. Princeton, NJ: Princeton University Press.

Zhao Y (2008) *Communication in China: Political Economy, Power, and Conflict*. Lanham, MD: Rowman & Littlefield.