An African Relational Approach to Healthcare and Big Data Challenges

Cornelius Ewuoso[1]

Abstract

Big Data has amplified some challenges in the healthcare context. One significant challenge is how to use healthcare big data (HBD) in ways that honor individual rights to informed consent or privacy. Careful analysis from diverse backgrounds will be vital in contributing ethical guidelines that can adequately address healthcare Big Data's growing complexities globally. Especially, the study argues that an under-explored African philosophy of *Ubuntu* can usefully influence big data practices in ways that address this challenge without undermining its benefits. *Ubuntu* emphasizes harmonious relationships. Harmonious relations entail identifying with one another and exhibiting solidarity to each other. One can identify or exhibit solidarity with others through psychological attitudes such as thinking of oneself as part of a "we" and acting in ways that will more likely improve the quality of life of others. The African relational philosophy of *Ubuntu* deserves to be given an audience not only for epistemic justice but also because the continued absence of African perspective in the discourse on ethical use of HBD science represents a missed opportunity to enrich ethical thinking about HBD from diverse backgrounds. Research is, however, required to provide greater specificity on how *Ubuntu* values may be integrated into HBD analytic techniques.

Introduction

Big data (BD) is difficult to define. Precisely the features, scope, purview or threshold of BD continues to sow confusions and generate ethical controversies. Despite the preceding, the emerging consensus (Ekbia et al. 2015: p. 3; Alharthi et al. 2017: p. 286) is that BD is characterized by 5 Vs: Volume (that is, a vast amount

✉ Cornelius Ewuoso
  Cornelius.ewuoso@uct.ac.za

1   Department of Medicine, University of Cape Town, Cape Town, South Africa

of datasets requiring innovative and big tools for capturing, storing and analyzing); variety (un/semi/structured and collected from diverse sources); velocity (rapidly evolving datasets and expanded by actual data streams); veracity (that is, data uncertainty, quality, reliability or predictive force); and value (this is the artificial intelligence that is created either for learning new patterns in vast datasets or offering personalized services).

There is a growing adoption of BD in various fields (engineering, life sciences, business, behavioural studies, online and offline commerce, education and politics). This article focuses on healthcare big data use and access. Healthcare big data (HBD) come from different sources such as sequencing data, Electronic Health Records (EHR), biological specimens, Quantified Self (QS), biomedical data, patient-reported data, biomarker data, medical imaging, large clinical trials; which may be stored in repositories or biobanks (Mittelstadt et al. 2016: p. 306). Other means of expanding HBD streams include healthcare literature databases like PubMed, automated sources like health and fitness devices and volunteered sources such as e-health networks like *patientslikeme*.

HBD analytics integrates, explores, identifies clusters, correlates, analyze and infer (with an unparalleled degree of exactitude) based on datasets from the preceding complex heterogeneous sources to create – HBD – value or (artificial) intelligence for offering a range of personalized health services, support health policies or clinical decisions, or advance science (Michael et al. 2013: p. 22). The HDB value relies mostly on analytic techniques such as algorithms and machine learning that are generated from processed data. HBD may be processed by using graphics processing units or cloud computing. The advancement in the omics studies, patient-contributed online data, imaging processes and the increasing affordability and accessibility of health electronic devices also imply that a large volume of heterogeneous data can now be analyzed to create HBD intelligence at a low cost. Advances have also been made in extracting previously difficult text data (from doctor's notes, millions of books photographs from the past) for data analytics and mining purposes.

Notwithstanding its (potential) benefits, HBD also creates a challenge for all stakeholders such as data utilizers, contributors and beneficiaries of HBD intelligence. New digital technologies can empower, yet they are also intrusive. HBD vividly exemplifies this dual character (Ekbia et al. 2015: p. 27). As an example, digital surveillance can usefully provide support for contact tracing in the event of a virus outbreak such as the COVID-19/Ebola outbreak. It could also monitor an individual's healthcare decisions, behaviors or outcomes with the aim of fostering healthy habits and practices in the individual or the wider population, providing critical insights on health needs or status, assisting with the development of equitable interventions. However, it also introduces a level of oversight that significantly threatens individuals' privacy.

Furthermore, HBD may also be abused or misused. The literature is awash with examples of data misuse. Edward Snowden's revelation of illicit large scale data gathering activities of the United States National Security Agency is one example (Berendt et al. 2015: p. 224). The tool that makes it possible to tailor care to each contributor's specific needs may also be used to profile, discriminate, or harm the

contributor. For example, algorithms and machine learning could be exploited for adversarial attacks that bilk the system or track health care visits of individuals without their knowledge. This has already been reported to be happening by one study (Thesmar et al. 2019).

Therefore, it is imperative to address the challenges this new tool generates and develop means of preventing misuse without sacrificing the indisputable benefits of HBD. Ethics will play an essential role in avoiding HBD abuse without sacrificing its benefits. Regrettably, ethical thinking (including policies and guidelines), as shown by studies (Asadi Someh et al. 2016; Milton 2017; Saqr 2017), still lags behind the mounting HBD issues. Careful analysis from diverse backgrounds will be vital in enriching ethical thinking regarding the growing complexities of HBD and ensure the safe use of this new technology globally. Accurately, in this highly connected world, as people become more and more interconnected through technology, varied and new insights are needed to tackle global communities' challenges. This article contributes an under-explored African *Ubuntu* philosophical perspective to the conversation on the ethical use of Big Data (BD) in healthcare. The article argues that the African relational theory of *Ubuntu* can usefully influence HBD practices and use in ways that address the challenges generated by the same without undermining its benefits. The *Ubuntu* relational philosophy promises to provide insights for balancing different interests and perspectives in one global community. Contrary to Kantian ethics that emphasizes autonomous capacity and independence, relationalism claims that interaction has moral significance and is the primary *datum* that is essential for shaping individuals' identity, society and morality (Crossley 2005).

To forward the argument, I have divided the article into three sections. In the first section, the article describes the specific ways in which the African philosophy of *Ubuntu* is said to be a relational theory by African scholars. In the second section, the article discusses some *Ubuntu*-inspired rules of thumb for contributing towards the ethical use of HBD and addressing its challenges. In the third section, the article provides further explanations on how the relational ethics of *Ubuntu* may be used to justify – in ways that are suitable to the era of BD technology – any potential loss of informed consent, privacy and confidentiality, as well as address potential criticisms to the core arguments.

Relationalism and Ubuntu Ethics

A relational philosophy counts as African if it is informed by values that are more prominent in – although not necessarily unique to – sub-Saharan Africa. The prominent value in this region is communal relationship. In the African *Ubuntu* relationalism, the specific types of relationship which ought to be prized are communal relationships (Ewuoso and Hall 2019). It is useful to point out that African conception of communal relationship also tends to be totemistic such that communal relationship is not limited to relationship with humans alone but extends to the physical environment (the ecosystem, plants, animals, etc.) and the spiritual or invisible

world. One becomes human through communal relationships. As Gessler Muxe-Nkondo observes:

> If you [ask] ubuntu advocates and philosophers: What principles inform and organize your life? What do you live for ... the answers would express commitment to the good of the community in which their identities were formed and a need to experience their lives as bound up in that of their community (Muxe Nkondo 2007: p. 91).

Some leading proponents of the African theory include Thaddeus Metz, Desmond Tutu and Mogobe Ramose, to name a few. *Ubuntu* also has its origin in the cosmology of the southern African peoples but predominantly used in South Africa. In addition to its origin in southern African countries, Broodryk (2002: p. 27) believes that what is implied by *Ubuntu* is also found in languages such as Afrikaans (*mede-menslik-heid* meaning "humanness" or "humanity").

Several formulations of *Ubuntu* exist and African philosophers do not generally conceive communal relationships in the same way. This, in turn, impacts on how this philosophy is applied to address different issues. Accurately, some African scholars argue that community is ontologically prior to the individual, while others believe that community and individual are co-substantively dependent on each other (Ewuoso and Hall 2019). Thus, the specific formulation of *Ubuntu* which this article applies has not been believed by all African philosophers. Precisely, this formulation is the outcome of a descriptive study that aims to systematically outline the core aspects of *Ubuntu*, which cuts across existing and sometimes competing, interpretations of the same (Ewuoso and Hall 2019). This article demonstrates how the philosophical construction of *Ubuntu* (advanced in the review) can usefully influence HBD intelligence creation and use. This relational theory's relevant aspects are described for the western audience, who might be unfamiliar with this theory.

The aphorism *umuntu ngumuntu ngabantu* best expresses the emphasis on communal relationship – "a human is a human being only through their relationships with other humans." The failure to prize communal relationship implies a failure to showcase *Ubuntu,* a failure to be human, or a person. A communal relationship is a relationship in which individuals cognitively, emotionally, conatively and volitionally identify with each other, that is, through psychological attitudes such as thinking of oneself as part of a group, expressing pride and shame in the actions of others and coordinating behaviours to realize shared ends because this is who we are. A communal relationship also requires one to exhibit solidarity through sympathetic altruism, engaging in mutual aid, caring, considerate, or other helpful behaviours that are more likely to improve others' quality of life (Metz 2015). Both identifying with others and exhibiting solidarity with one another are necessary components of communal relationship. One showcases *Ubuntu* when one empathizes with others and considers one's wealth as also the wealth of others, others' salvation as one's salvation (Nussbaum 2003: p. 21).

The combination of identifying and exhibiting solidarity is what is commonly described as friendship, friendliness or love in *Ubuntu*. The combination of identifying and exhibiting solidarity is also what distinguishes *Ubuntu* relationalism from other relational ethics. While other relational ethics such as the

Ethics of Care emphasize exhibiting solidarity, *Ubuntu* relationalism requires a combination of identifying with others and exhibiting solidarity as necessary components of communal relationship. One can exhibit solidarity towards others without identifying with them. Anonymous gifts to homeless shelters is an instance of exhibiting solidarity without identifying with others (Metz 2017). Summarily, communal relationship in the African *Ubuntu* relationalism is a relationship in which one seeks the good of others; sees others as worthy of aid and is not indifferent to their needs; acts for the sake of others; and finally, rejoices knowing others have been aided and is unhappy upon learning that others have been harmed (Metz 2007: p. 336). Given this, a right action is one that promotes social cohesion and relationships of harmony in which individuals can think of themselves as part of a "we" and exhibit solidarity with one another. From an *Ubuntu* perspective, an action is immoral because it represents a failure to relate communally.

Furthermore, communal relationships often occur between subjects and objects. A subject is one who cognitively, emotionally, conatively and volitionally identifies with, as well as exhibits solidarity to, other persons. An object is one with whom others can commune with in relevant ways. Additionally, in a modal account of the relational philosophy of *Ubuntu,* it is not the "actual" relationship of identifying with others and exhibiting solidarity with them itself, nor only the individuals in such relationships who have moral status, but those who can be or have the capacity for such relationships. One has the capacity if one can, in principle – without changes to one's life – share a way of life with others. The deeper the capacity, the greater one's moral status, such that those who can be subject and object of communal relationships have higher or full moral status than those who can only be the subject or only object of such a relationship. Communion is the exercise of this capacity. To this end, in the relational philosophy of *Ubuntu,* actions are also immoral not merely because they fail to maximize the overall benefits (utilitarianism) or degrade one's autonomy (Kantian ethics), but because they undermine one's capacity for friendship.

One reason for the existence of legal systems is to advance the values of the people. If a legal system is to be relevant to the society, it ought to express and promote its values. Therefore, it is not surprising that *Ubuntu*-inspired communion underlies general principles in the South African Constitution (and many other legal systems in sub-Saharan Africa), such as the need for understanding, reconciliation and reparation and not vengeance or retaliation. Courts in Southern Africa also associate it with constitutional rights like privacy and dignity (T Bennett et al. 2018; TW Bennett 2011). For example, it is stated in the preamble of the South African Constitution that "South Africa belongs to all who live in it." It is equally unsurprising that the *Ubuntu*-inspired norm also explains salient thinking about development as mutual empowerment or decolonization as a strategy for inclusion and cohesion (Tavernaro-Haidarian 2019). Finally, the *Ubuntu*-inspired norm explains dominant sub-Saharan practices, such as joint ownership, preference for consensus, combining resources and intensifying communal bonds.

Ubuntu Relationalism and Healthcare Big Data

Ubuntu-inspired norm has many implications for HBD practices (and for HBD collectors, utilizers and data contribution). This section discusses two moral implications. First, the relational ethics of *Ubuntu* implies that there is an obligation to deepen the opportunity individuals have for friendliness. Concretely, this entails that HBD analytics and data mining ought to make individuals – particularly beneficiaries – better-off, for example, by enhancing patient-tailored care and advancing the scientific understanding of health disorders and treatment effectiveness. A failure to make individuals better off roughly implies a failure to seek their good or act for their sake. Given the value of communion, certain types of analytics that aim to cause harm should be made off-limits. For the utilitarian, harm is a failure to maximize welfare benefits. In the relational philosophy of *Ubuntu,* harm occurs when one undermines another's capacity for communion. Thus, certain types of analytics which ought to be made off-limits would include those that may compromise one's capacity to commune, such as those that can damage one's reputation in the society or reveal confidential information.

This first norm also implies that HBD intelligence ought to be used to prevent health conditions that can undermine one's capacity to relate. Concretely, this may imply that health professionals ought to take advantage of the discoveries such as the genetic origins of illnesses that have now been gained through the Human Genome Project, to benefit individuals (data contributors, society and other potential beneficiaries) who are at risk of developing debilitating health conditions. The high- throughput-omics data that has been providing incredible insights into disease mechanisms and effective treatment methods, ought to be leveraged to enhance care. The current effort to map the human brain through the Human Brain Project with the hope of providing more value or intelligence for addressing pressing issues regarding neurodegenerative conditions (Rüping 2015), should be intensified; since this can concretely become another means of helping individuals enjoy a deep communal relationship with others. Community engagements or expanded involvement (particularly with beneficiaries, contributors and/or their representatives and leaders) in HBD access and use can help to foster this goal.

Additionally, this rule also implies that data utilizers ought to use HBD intelligence to facilitate improved public health management, enhance participatory and enhance early warning systems like the public health surveillance that are used to quickly identify potential public health emergencies or diseases that can potentially undermine social harmony. As we have seen with the coronavirus 2019 (COVID-19) pandemic, illness and diseases can undermine social harmony or one's capacity to adequately relate with others (Ewuoso 2019). As an example, digital surveillance channels played a prominent role in the early detection and monitoring of the 2014 Ebola virus outbreak in West Africa (Vayena et al. 2015). Digital surveillance or proximity tracking technologies are potential tools that could also be used to support contact tracing in this current COVID-19 pandemic and reduce secondary transmission of the virus. These are tangible examples of

increasing the opportunity individuals have to enjoy deep communal relationships. In the relational ethics of Ubuntu, the motivation for a communal relationship is not merely to avoid punishment or ostracization from the community. The underlying rationale is that individuals who identify and exhibit solidarity with one another do so consequent to trust. They are willing to make themselves vulnerable in the belief that others will act in their best interest because this is "who we are".

This moral rule also has great significance for HBD collectors; specifically, it requires data collectors to gather data from multiple regions and populations. "Algorithms" as Char et al. (2018: p. 1) correctly observe, "designed to predict outcomes from genetic findings will be biased if there have been few (or no) genetic studies in certain populations."

A second moral rule – given the significant emphasis on "sharing a way of life" that arises from this relational theory and addressing HBD ethical challenge and preventing misuse is that individuals have a moral obligation to contribute to the advancement of HBD science. This moral rule provides a basis grounded in African philosophy, for Knoppers et al. (2017) "human right to science." Here, I conceive this right as right not only to benefit from science but also to share in its advancements through data contribution and sharing. The Global Alliance for Genomics and Health's initiative to build open (cancer) databases is one tangible and precise way of honouring this second moral norm. The initiative acknowledges that BD can both be beneficial and harmful. Individuals may be discouraged from contributing to BD because of the risks, which can be prevented. The aim is to shift the attention on medical BD as something harmful to one that places a significant emphasis on its potential for the public good as an expression of the human right to science. Similar initiatives should be replicated across regions.

In anticipation of potential criticism, it is essential to state here that the relational ethical theory applied here does not imply that one should advance science "at all cost" such as by sacrificing individual rights at will to achieve these ends. Any potential loss of these rights must be thoughtfully and carefully balanced against the HBD value. African philosophy of *Ubuntu* can provide this balance and justification in ways that will foster public trust and confidence in BD technology more broadly. If the public perceives that the information they disclose to enhance public health surveillance or improve individuals' health will be abused, this can erode their trust in this new technology. African relational philosophy of *Ubuntu* deserves to be given an audience not only for epistemic justice but also because the continued absence of African perspective in the discourse on ethical use of HBD science represents a missed opportunity to benefit ethical thinking about HBD from diverse backgrounds.

Furthermore, critics may also point out here that communitarian ethics would permit "routinely" infringing on others' rights to promote communal goods. Infringing on the rights of those who have not done anything to deserve this roughly implies using unfriendliness to promote friendliness. This is immoral based on this philosophy since communal relationships ought not to be promoted however one can or at all cost. This distinguishes *Ubuntu* from consequentialist theory like utilitarianism. Precisely, consequentialism fails to account for the intuition that certain ways

of relating are wrong in themselves apart from the result and individuals have a right not to be harmed, even if this will lead to the maximization of the general welfare. African philosophers tend to think that one has a duty to honor communal relationship and not to maximize it.

Based on *Ubuntu* relationalism, certain ways of relating and promoting desirable ends are unfriendly themselves and thus, immoral. Some of these include xenophobia, racism and coercive behaviours like rape. In these instances, the actor expresses ill-will towards, rather than acting to benefit, the victims. Honouring communal relationships requires one to be friendly and do what one can to promote friendliness without using a very unfriendly means. Furthermore, one also has a deeper obligation – based on the philosophy of *Ubuntu* – to end unfriendly relationships (in which individuals have distanced themselves from others or seek to harm others in some ways), than to promote new ones; to end ill-will or division, than to promote goodwill. Given this deeper obligation, this African relationalism permits one to use unfriendliness only to counter-act another unfriendliness. However, the means of ending unfriendliness cannot be disproportionate to the unfriendly action. This will be using a very unfriendly means to end unfriendliness. As I have argued in one study (Ewuoso 2019), if forcing a knife away from an aggressor can sufficiently eliminate danger, one is not justified, based on this philosophy, to kill the aggressor to achieve this end. Summarily, this theory requires one to treat those who have been friendly in a friendly manner; and those who have been unfriendly in a proportional unfriendly manner.

Areas of Ethical Concerns for Health Big Data Use

HBD raises a taxonomy of technical and ethical challenges. For example, HBD analytics may be complicated by technical issues such as the lack of data management capacity. EHRs may also have missing values or entered incorrectly. Additionally, omics data are highly dimensional, which can make data mining also challenging. Data variety and volume also raise complex technical questions regarding storage, messiness and noisiness of data (Mittelstadt et al. 2016).

However, if, as we have seen, *Ubuntu* relationalism requires one to promote friendliness without using a very unfriendly means, what does this mean for honoring informed consent, privacy and confidentiality, in the era of HBD technology? In the next section and subsequent ones, I will focus primarily on this question, highlighting the acceptable trade-off for any potential loss of individual rights.

Informed Consent

Issues around informed consent often centre on the individual's voluntary decision, the sufficient knowledge to make the decision and freedom from coercion (T. Crow et al. 2000; Metz 2010a, 2010b). Most Kantians would ground the duty to obtain informed consent in one's capacity for autonomous decision. Based on Kantian ethics, individuals with diminished rational capacity, or lack of it, do not enjoy this

privilege. Individuals should be respected because of their inherent dignity. They have this dignity by virtue of their rational capacity. Kant does not think that this same duty applies to others who have no rational capacity or have a diminished capacity. If rationality is the only factor determining whether individuals have dignity or personhood, the loss of it, may entail that these individuals are neither persons nor have any dignity (Kant 1996). This notoriously fails to account for the intuition that individuals with diminished capacity also have certain inherent dignity, deserving of respect not merely because how respecting them would make rational individuals feel. And specific ways of relating with them – such as subordinating or coercing them – will be immoral. *Ubuntu* relationalism overcomes this weakness by grounding the ethical obligation to honour informed consent in relationships. Correctly, from *Ubuntu* perspective, if individuals are to genuinely share a way of life, they ought to be transparent about the terms and conditions of the interaction and voluntarily participate in it (T. Metz 2010a, 2010b: p. 160). The right to informed consent in the philosophy of *Ubuntu* entails the free decision of an individual to commune with others. This right also requires that one should be given sufficient information (about the terms of this communion) to make the decision to commune. Some may fail to exercise their relational capacity at all, or in the relevant ways. Nonetheless, they retain their right to be informed about the nature of the communion and to exercise this right freely. One can hardly "genuinely" interact with others who are not transparent, honest about the terms of the relationship or deceives/coerces one into communing with them. In this way, the *Ubuntu*-inspired view of the right to informed consent advances the Universal Declaration of Human Rights, prohibiting actions that offend human dignity (Robinson 1998). Specifically, by demonstrating how human dignity is violated, that is through actions that harm the free exercise of one's capacity for communion since individuals have dignity by virtue of this capacity.

Respecting one's right to informed consent will be very challenging in the era of HBD. For two reasons: first, this new tool has the potential to lead to surveillance capitalism, whereby individuals – specifically potential HBD beneficiaries – are unable to have or truly weigh all options but become subjects to control of algorithms or automated processes that aim to influence their decisions and in the process, undermine their right to make an informed choice (Asadi Someh et al. 2016: p. 4). As previously stated, HBD value or intelligence relies mostly on algorithms that are often the result of processed data. Data may have been inputted by individuals who may have – unknown to them – systematic cognitive errors. These analytic techniques can reflect these systematic errors. In some cases, individuals will not be able to query the assumptions that fundamentally shape these algorithms and machine learning or hold the same accountable. This can exacerbate the BD divide.[1]

Second, with HBD, it is difficult to know in advance who will access data; or how contributed data will be used in the future, thus making it difficult to determine if

---

[1] This typically a situation where "the data-poor are caught in a position of weakness wherein the ability to understand the data and methods used to make decisions about them as individuals and members of groups is beyond their means" (Mittelstadt et al. 2016: p. 324).

consent (by HBD contributors) can be truly informed. As the American College of Medical Genetics and Genomics (ACMG) rightly observes, BD value in health and medical contexts requires, as well as thrives, on sharing in collaborative networks (Knoppers et al. 2017: p. 55). Third-party data users in these networks may use data to achieve ends, different from the contributor's intention at the time of consent. HBD analytics may be undertaken, not necessarily to harm others, but in ways that contributors would not have consented to; such as to predict private information like health habits of owners, sexual orientation or which health conditions the donor is at risk of developing; thus, raising further questions about data control and ownership. Finally, HBD may have also been scrapped from online communities from unwilling and uninformed participants. Examples include data scraped from wearable medical devices that monitor heart rates and blood pressure or data scraped from health communities like *patientslikeme*. Some end-user agreements policies tend to allow data obtained through these means to be collected, aggregated and analyzed, in ways that may be unacceptable to contributors (Mittelstadt et al. 2016: p. 313).

One way of dealing with the aforementioned is to re-affirm the inviolability of informed consent; as was the case in 2017 in Canada (Knoppers et al. 2017); or in 2016 with the European General Data Protection Regulation (as well as the United Kingdom's Data Protection Act, 1998) that drastically inhibited "information-based research utilizing aggregated datasets to uphold ethical ideals of data protection and informed consent" (Mittelstadt et al. 2016: p. 305). This will require data contributors to reconsent to every data use, which may undermine HBD science's advancement. In an ideal context, data contributors would be adequately informed of every stage of data capturing, storing and usage, as well as approached to re-consent to new uses that were not part of the initial agreement. This would prevent abuse. It might, however, be difficult to recontact donors, who may have relocated or died. A second suggestion is for data collectors, aggregators, banks and repositories to obtain broad consent (for future and unspecified use of donated data). However, it seems intuitive that broad consent is less likely to be truly informed if contributors are asked to provide consent "for an unspecified use".

A plausible approach that will not adversely undermine innovation is for traditional conceptions about informed consent to adapt to big data's new reality. One way of adapting would be to ensure that it is not abused even when consent is not informed. Misuse may be prevented by emphasizing friendliness. Collegiality is required here and entailed by the African relationalism that emphasizes togetherness – to ensure that data is used to achieve the relevant friendly ends. Precisely, Institutional Review Boards can enforce guidelines that protect sensitive data, monitor data use, compensate data donors in the event of a violation. Institutions and companies can also compensate contributors in situations where their explicit wishes are violated. Scientists also share a responsibility to report unethical use of data and shun researchers who violate ethical standards. Data cooperatives (to facilitate access and pooling of data) that are democratically controlled and encourage diverse discursive landscape can be another way of sharing a way of life with others and preventing data misuse. HBD cooperatives will ensure that access to personal data is granted only after a deliberative democratic decision-making process that takes into consideration the diverse views of key stakeholders such as contributors and beneficiaries of HBD intelligence, thereby enhancing transparency

in HBD use, protection from unauthorized or accidental access and ensuring that data is used to achieve ends that the contributors will endorse.

Cooperatives, specifically deliberative cooperations and engagements, reflect the value of sharing a way of life with others in *Ubuntu* relationalism and are known to be the basis of – for example – Southern African legal systems that seek to engage the public (*imbizo*), as well as strive to ensure that the final decision reflects – to the extent possible – the views of all (*indaba*) (T Bennett et al. 2018). In most southern African States, as Bennett and colleagues (T Bennett et al. 2018) observe, "*Ubuntu* is the essence of our democracy – or should be." Summarily, an *Ubuntu*-inspired collegiality and HBD data cooperative entail collective wisdom where stakeholders share joint responsibility in preventing data misuse and where opinions – wherever possible – of contributors, communities and beneficiaries of HBD intelligence are sourced through deliberative engagements.

As previously mentioned, the philosophy of *Ubuntu* requires individuals to honour and respect the value of informed consent and not to maximize it. I have argued in a different study (Ewuoso 2016) that based on the relational philosophy of *Ubuntu,* some form of coercion, such as the violation of the informed consent of research participants in genomic research, may be moral if this is necessary to aid those whose quality of lives are threatened. This is not necessarily degrading to their personhood, which is gained through relationships. Contrarily, it can be a legitimate use of unfriendliness (infringement of right) to end unfriendliness (present or future illness since sickness diminishes one's capacity to relate). In anticipation of a potential criticism in this direction, the reader should observe that Ubuntu relationalism's communal character does not imply majoritarianism or conformity to a group's culture, as is found in cultural relativism or other majoritarian ethics. For example, in cultural relativism, an action (such as slavery) is right if it promotes a group's culture. In the philosophy of Ubuntu, the term community does not refer to a society out there. Rather it is an ideal way of relating or an objective standard that determines what moral rules become dominant in society. A good moral rule is one that can bring people together rather than divide them.

Against this background, data contributors have a right to know – to the extent possible – how their data would be accessed and used. The rights of beneficiaries to have information, at least the material information that a reasonable person would consider relevant, ought to be honored to the extent allowable by collectors. Since the value of friendliness requires subjects and objects to be clear about the terms of their communal interactions, wherever possible, these terms should be made available to third party users of data. Where it is not possible to directly honour these rights, an acceptable tradeoff would be for HBD to foster one's capacities for communal relationship.

Privacy

Contrary to the emphasis placed on respecting the value of informed consent, privacy (and confidentiality) is not as weighty in this relational theory because of the deeper obligation to aid others. Part of aiding involves being involved in others' wellbeing,

playing a role in discussing how individuals ought to treat their illness (Metz et al. 2010: p. 279). It is essential to point out that the right to make an informed decision about one's health is different from whether others know about one's health status (privacy and confidentiality). The duty to respect informed choice ought to be valued even when others have become aware of one's health status(T. Metz 2010a, 2010b); hence, the higher moral significance given to informed consent.

Privacy is described by two former United States Supreme Court Justices (Samuel Warren and Louis Brandeies) a legal right "to be let alone" (Ekbia et al. 2015). HBD will make honouring this right extremely challenging. As Gambs (2018: p. 17) observes, "while privacy risks have existed for a long time due to the sharing of personal data, [BD] magnifies these risks and makes them more difficult to grasp and predict due to the possibility of combining data from different sources while also bringing new risks related to the inference attacks that are possible against models learned from Big Data." The United States-based company Target's unwitting disclosure of a teenager's pregnancy before her father ever found out is an example of the ethical challenges around preserving privacy in the era of BD. By analyzing her customers' shopping habits, Target could predict with a high degree of certainty which of her customers were pregnant and their due dates (Wielki 2015: p. 197). Self-reported health data through systems such as Quantified Self are also vulnerable to similar privacy breaches.

Many guidelines currently require removing 18 HIPAA Safe Harbor identifiers such as name, phone number and residential address before data collected from contributors within the research context are deposited in repositories (Edemekong et al. 2019). However, anonymization or deidentification of data does not necessarily remove the risk of re-identification. HBD are pooled from multiple sources. Some of those data will be de-identified in compliance with HIPAA guidelines. Some other data, such as data registered in online health communities, may contain personally identifiable information since they are not necessarily subject to anonymization rules. Moreover, anonymization does not eliminate privacy risk. Complete anonymization is not always encouraged since it can render the data useless for secondary use. Given this, re-identification of anonymized data can occur through data crosslinking and merging from multiple sources.

Data collected from other sources raise other unique privacy problems. Contrary to information contributed in traditional settings that could be purged, the data registered in online health communities can assume a life of their own. In view of this, personal identifiable information embedded in such data can exist forever. Additionally, HBD analytics and mining can also reveal incidental information. For example, genetic data screening could reveal incidental information about incurable diseases such as Alzheimer's (Krier et al. 2013; O'Sullivan et al. 2018); and thus, cause life-long trauma or render the data contributor ineligible for insurance coverage. Sequencing methods are such that they tend to reveal incidental information, that is, information outside of the research aim or primary reason for the test. Also, health data such as genomic data can reveal (private) information about the contributor and those of relatives and family members, thus increasing the risk of harm to the latter. Most data protection guidelines tend to focus on the individual data donor, thus ignoring the potential harm to relatives, groups, or communities (Ewuoso 2020).

The consequence of the preceding is that the risk of privacy invasion will likely increase with every use of HBD. Moreover, sharing of data in collaborative networks undermines informed consent and the individual's privacy since HBD is data generated by individuals and, thus, directly concerns their privacy. Besides, the technical concern about the safety of data storage systems or databases also raises privacy concerns. Data storage systems are not immune from hacking and hackers, who may target such systems for mischievous ends and publish private information.

Confidentiality

Confidentiality concerns the duty to keep disclosed information secret. Clinicians often promise in The Hippocratic Oath (The hippocratic oath 1998) to respect the fiduciary relationship by protecting confidential information – "Whatsoever in the course of practice I see or hear….. that ought never to be published abroad, I will not divulge…. if I keep this oath….may I enjoy honor…..if I transgress….may the opposite befall me" (Laar et al. 2015: 1). A violation of confidentiality occurs when a piece of information disclosed to the professional in a context requiring trust, or where a reasonable person can expect his/her information will be held in trust, is disclosed without the individual's consent. Medical practice requires trust. Confidentiality allows for some level of trust to exist in medical practice. Though confidentiality is not an absolute duty and there are cases such as where public health is threatened, when the duty of confidentiality may be waived, confidentiality nonetheless is required to maintain a patient's overall well-being. Improper disclosure of highly sensitive information could harm a patient's reputation or public trust in health professionals (Beltran-Aroca et al. 2016).

In a datafied world, it might become challenging to honor confidentiality. As Char et al. (2018: p. 3) correctly observe, "the traditional understanding of confidentiality requires that a physician withhold information from the medical record [for example] in order to truly keep it confidential. Once machine-learning–based decision support is integrated into clinical care, withholding information from electronic records will become increasingly difficult, since patients whose data are not recorded [will not] benefit from machine learning analyses." In other words, the reasonable expectation of confidentiality which is critical to honouring fiduciary relations, will be difficult to honour without mostly undermining beneficial care in the era of HBD. Moreover, it is difficult to guarantee that personal information disclosed in e-health communities will be protected by confidential norms.

In the previous section, I have argued that one way of addressing the challenges generated by HBD is to prevent misuse. In the next section, this article grounds a third moral rule for justifying any potential loss of privacy and confidentiality that may arise because of HBD.

Addressing Privacy and Confidentiality Risks: A Relational Perspective

In addition to the previously discussed *Ubuntu*-inspired moral rules, a final moral rule is that there is a moral obligation to honour the means by which individuals gain their personhood. A critic may argue here that this moral rule is not significantly different from the first moral rule, which requires scientists to use HBD to increase the opportunities for a deep communal relationship. In response, the first moral rule requires individuals to use science to enhance an individual's capacity. This final moral rule precisely describes how this may be accomplished: based on one's informed values. Respect for a person and by extension, their dignity, is hardly complete without also respecting their values. Most scholars will agree that individualism is the salient value in the global north, while communal relationship is core in prescribing moral duties in the global south. Communal relationship is the basis of personhood and dignity in most formulations of African philosophy.

This has implication for how African philosophers think about human rights. The 11th article of the Universal Declaration of Human Rights describes human rights violation as practices which are contrary to human dignity (Robinson 1998). As previously mentioned, a typical human rights violation from the *Ubuntu* relational perspective entails a deliberate and wrongful violation of one's right to commune. This conception differs from the Christian and Islamic views, which generally contend that humans have dignity, thus deserve respect because they have their source in God (T Metz 2010a, 2010b). Albeit *Ubuntu* shares some features with these religious views (since God requires his chosen people to be sympathetic towards strangers, welcoming, hospitable, warm and generous), it is also distinct from them. In *Ubuntu* relationalism, an individual can enjoy communion as a physical creature. This way, the Afro-communitarian view, unlike the Christian and Islamic views, does not require any spiritual basis that suggests that humans have rights and human dignity because they have a spiritual nature that is from God.

In view of moral rule (1) and (3), one can roughly summarize that a plausible way of mitigating privacy and confidentiality risks is for HBD analytics and mining to be guided – to the extent possible – by the values and preferences of HBD contributors. When guided by the values of contributors, HBD usage will more likely foster their (contributors) dignity (since rights violation are typically practices contrary to human dignity). Honoring the values by which individuals have dignity also concretely entails that data is used for the purpose for which they are collected or a related goal. Public spirited discussions and deliberative engagements with community leaders and contributors can ensure that data contributors are empowered to retain control – to the extent possible – of their data, who is accessing their data and whether the data are used in ways that honour their values. Ethics research committees and policymakers can foster such collaboration with contributors by enforcing contributors' rights to review data about them and giving them (contributors) the opportunities to correct inaccurate information or

object to how their data are used. The reader should observe that privacy and confidentiality will be significantly threatened if HBD use is not informed by contributors' values (and beneficiaries). Data utilizers should observe legal and ethical regulations. For example, Kant's categorical imperative demands that individuals are not to be used merely as a means to other people's ends. Using data beyond the scope explicitly consented to by contributors roughly entails subordinating their will and using them as mere means, thereby harming their dignity. Given that unfriendliness is not always impermissible, data may be used beyond the limits explicitly allowed by contributors only if this is necessary to counteract a proportional unfriendliness, such as when it is used to benefit contributor's or public health.

A critic may point out here that it will be nearly impossible for third party users analyzing a large volume of data collected from a repository to know the values of data donors. Dynamic consent can be one useful way of ensuring that data use aligns with the contributor's values. In addition to dynamic consent and expanded involvements/engagements with key stakeholders in the decision-making about data use, another way of honoring participants' values is to ask data contributors to tag their contributed data with their preferences or specify the duration of data use (or opt-out of data use). However, where contributor's preferences cannot be accessed, HBD analytics (or HBD use) ought to be guided by the goal to create valuable intelligence that fosters friendliness, such as public health surveillance. This will be an acceptable tradeoff since the obligation to honour privacy and confidentiality is not the same as maximizing it. Herein, there is less focus on privacy and confidentiality and more emphasis on the responsible use of data. Whilst this suggestion does not directly honour a donor's (human) right, it would at least ensure that data is put into good use.

HBD volume is not valuable if the data cannot be analyzed in a way that generates future insights or good intelligence in the relevant sense. It is equally important – where donors' wishes and preferences cannot be accessed – that the type of analytics permitted should be the minimum necessary to create well-defined intelligence. Data analysts are well trained to act morally. But they may also be tempted by the sheer volume of data and the lowering cost of analytics to undertake undefined analytics. Such practices, if not guided by defined rules, may undermine privacy and confidentiality. Herein, observe that deliberative democracy has a weakness; it may not always be feasible or may fail to involve the views of data contributors, especially if a researcher working in the global north, for example, is requesting to access the data contributed and deposited in a repository that is controlled by a governance structure that is primarily domiciled in the same global north. In these instances where engagement with donors or their representatives is not feasible, policymakers as valuable members in collegial partnership can help (through legislation, clear procedures for sourcing and using data) to formalize moral rules that promote friendliness as specify additional norms. This will encourage more people to contribute to the BD pool. Institutional review boards can also be valuable collegial partners in monitoring data use and protecting vulnerable groups, thereby fostering public trust. As Brad Smith (cited in Richards et al. 2014: 415) rightly observes, "people [will not] use technology they [do not] trust." Organizational training in the

ethical use of HBD can also contribute towards eliminating risks to privacy and confidentiality in the era of BD.

A critic will be right to point out that the precise ways in which the moral rules this article outlines may be integrated into the assumptions that shape algorithms and machine learnings (that are essential for creating HBD value) require greater specificity. HBD raises several multifaceted issues for all stakeholders. I have also not included in this particular article discussions about how the principles of beneficence and justice align with *Ubuntu* relational values to foster sound ethical use of BD. It is impossible to consider all of them in one article. This current article aims to provide broad rules – from an African perspective – which ought to govern the ethical use of HBD. The specific ways these rules can form the basis of the assumptions that substantially shape algorithms and machine learnings deserve an elaborate discussion. More accurately, this will require careful analysis of the implicit and explicit forms of algorithm and machine learning bias, followed by a detailed analysis of how this relational theory may be used to address such bias. I will defer this elaborate discussion to another article. The new European General Data Protection Regulation requires data utilizers to explain decisions taken by algorithms and machine learnings that significantly impact individuals (Gambs 2018). This may be one way of ensuring that sound moral rules are integrated into these analytic techniques. This will not necessarily make data utilizers experts in the field. They are most likely already experts, possessing the required skills. Data scientists can behave ethically, but they also have the capability to design machine learnings and algorithms to cheat or function in very unfriendly ways. Holding these scientists responsible – and sanctioning them – for the wrongful decisions taken by these analytic techniques can discourage other bad scientists or usefully lead to better algorithms and machine learnings that incorporate good values such as friendliness.

Conclusion

HBD science has amplified some major challenges in the healthcare context. These challenges require actionable recommendations and new ways of thinking about how we can live in a datafied world. One significant challenge is how to use HBD to create intelligence without undermining informed consent, privacy and confidentiality. This article has argued that the African relational theory of *Ubuntu* can usefully influence HBD practices in ways that address these challenges without undermining its benefits. Specifically, this article argues three norms. First, HBD should increase the opportunities one has to enjoy a deep communal relationship. This mostly entails making individuals better off. Second, given the potential of this new tool, individuals also have a duty to advance this new technology by contributing towards its development. It is nearly impossible to "genuinely" share a way of life with others if one is only taking from the other without also giving back. Finally, HBD ought to be used in ways that align with contributors' and beneficiaries' values. However, given that HBD will make it ethically challenging to honor rights to informed consent, privacy and confidentiality, an acceptable tradeoff is to use HBD to create valuable

intelligence that fosters public good. Further research is needed to study how these norms can be integrated into HBD analytic techniques.

Declarations

Conflict of interest The author declares that he has no competing interests.

Ethical Approval This study is a concept paper and does not require ethical approval.

References

Alharthi, A., Krotov, V., & Bowman, M. (2017). Addressing barriers to big data. *Business Horizons, 60*(3), 285–292.

Asadi Someh, I., Breidbach, C. F., Davern, M. J., & Shanks, G. (2016). Ethical implications of big data analytics.

Beltran-Aroca, C. M., Girela-Lopez, E., Collazo-Chao, E., Montero-Pérez-Barquero, M., & Muñoz-Villanueva, M. C. (2016). Confidentiality breaches in clinical practice: what happens in hospitals? *BMC Medical Ethics, 17*(1), 52. https://doi.org/10.1186/s12910-016-0136-y

Bennett, T. (2011). Ubuntu: an African equity. *PER: Potchefstroomse Elektroniese Regsblad, 14*, 29–61. http://www.scielo.org.za/scielo.php?script=sci_arttext&pid=S1727-37812011000400003&lng=en&nrm=iso. ISSN 1727-3781

Bennett, T., Munro, A., & P, J. (2018). *Ubuntu: An African jurisprudence*. Cape Town: Juta & Company Ltd.

Berendt, B., Büchler, M., & Rockwell, G. (2015). Is it research or is it spying? Thinking-through ethics in Big Data AI and other knowledge sciences. *KI-Künstliche Intelligenz, 29*(2), 223–232.

Broodryk, J. (2002). *Ubuntu: Life lessons from Africa*. Pretoria: Ubuntu School of Philosophy.

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. *New England Journal of Medicine, 378*(11), 981.

Crossley, N. (2005). *Key concepts in critical theory*. Sage Publications Ltd.

Crow, K., Matheson, L., & Steed, A. (2000). Informed consent and truth-telling: cultural directions for healthcare providers. *Journal of Nursing Administration, 30*(3), 148–152.

Edemekong, P. F., & Haydel, M. J. (2019). Health Insurance Portability and Accountability Act (HIPAA). In *StatPearls*. Treasure Island (FL).

Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., et al. (2015). Big data, bigger dilemmas: A critical review. *Journal of the Association for Information Science and Technology, 66*(8), 1523–1545.

Ewuoso, C. O. (2016). Beneficial coercion in psychiatric care: Insights from African ethico-cultural system. *Developing World Bioethics*. https://doi.org/10.1111/dewb.12137

Ewuoso, C. (2019). Addressing the conflict between partner notification and patient confidentiality in serodiscordant relationships: How can Ubuntu help? *Developing World Bioethics*. https://doi.org/10.1111/dewb.12232

Ewuoso, C. (2020). Ubuntu philosophy and the consensus regarding incidental findings in genomic research: A heuristic approach. *Medicine, Health Care and Philosophy*. https://doi.org/10.1007/s11019-020-09953-4

Ewuoso, C., & Hall, S. (2019). Core aspects of Ubuntu: A systematic review. *South African Journal of Bioethics and Law*, *12*(2), 93–103. https://doi.org/10.7196/SAJBL.2019.v12i2.00679

Gambs, S. (2018). Privacy and Ethical Challenges in Big Data. In *International symposium on foundationsand practice of security,*(pp. 17–26): Springer

Kant, I. (1996). *The metaphysics of morals* (Rev. ed., Cambridge texts in the history of philosophy). Cambridge, New York: Cambridge University Press.

Knoppers, B. M., & Thorogood, A. M. (2017). Ethics and big data in health. *Current Opinion in Systems Biology, 4*, 53–57.

Krier, J. B., & Green, R. C. (2013). Management of incidental findings in clinical genomic sequencing. *Curr Protoc Hum Genet, Chapter, 9*(Unit9), 23. https://doi.org/10.1002/0471142905.hg0923s77

Laar, A. K., DeBruin, D. A., & Craddock, S. (2015). Partner notification in the context of HIV: an interest-analysis. *AIDS Research and Therapy, 12*, 15. https://doi.org/10.1186/s12981-015-0057-8

Metz, T. (2007). Toward an african moral theory. *Journal of Political Philosophy, 15*(3), 321–341.

Metz, T. (2010). An African theory of bioethics: reply to MacPherson and Macklin. *Developing World Bioethics, 10*(3), 158–163. https://doi.org/10.1111/j.1471-8847.2010.00289.x

Metz, T. (2010). Human dignity, capital punishment and an African moral theory: Toward a New Philosophy of Human Rights. *Journal of Human Rights, 9*(1), 81–99.

Metz, T. (2015). An African theory of social justice. In C. Biosen, & M. Murray (Eds.), *Distributive justice debates in political and social thought: perspectives on finding a fair share*. New York: Routledge.

Metz, T. (2017). Ancillary care obligations in light of an African bioethic: from entrustment to communion. *Theoretical Medicine and Bioethics, 38*(2), 111–126. https://doi.org/10.1007/s11017-017-9404-1

Metz, T., & Gaie, J. B. R. (2010). The African ethic of Ubuntu/Botho: implications for research on morality. *Journal of Moral Education*. https://doi.org/10.1080/03057240.2010.497609

Michael, K., & Miller, K. W. (2013). Big data: New opportunities and new challenges [guest editors' introduction]. *Computer, 46*(6), 22–24.

Milton, C. L. (2017). The ethics of big data and nursing science. *Nursing Science Quarterly, 30*(4), 300–302.

Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics, 22*(2), 303–341.

Muxe Nkondo, G. (2007). Ubuntu as public policy in South Africa: A conceptual framework. [Jour]. *International Journal of African Renaissance Studies - Multi-, Inter- and Transdisciplinarity, 2*(1), 88–100, doi:https://doi.org/10.1080/18186870701384202.

Nussbaum, B. (2003). Ubuntu: Reflections of a South African on our common humanity. *Reflections, 4*(4), 21–26.

O'Sullivan, J. W., Muntinga, T., Grigg, S., & Ioannidis, J. P. A. (2018). Prevalence and outcomes of incidental imaging findings: umbrella review. *BMJ, 361*, k2387. https://doi.org/10.1136/bmj.k2387

Richards, N. M., & King, J. H. (2014). Big data ethics. *Wake Forest l. Rev., 49*, 393.

Robinson, M. (1998). The universal declaration of human rights: Hope and history. *Health and Human Rights, 3*(2), 27–29.

Rüping, S. (2015). Big data in medicine and healthcare. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz, 58*(8), 794–798.

Saqr, M. (2017). Big data and the emerging ethical challenges. *International Journal of Health Sciences, 11*(4), 1.

Tavernaro-Haidarian, L. (2019). Decolonization and development: Reimagining key concepts in education. *Research in Education, 103*(1), 19–33. https://doi.org/10.1177/0034523719839750

The hippocratic oath (1998). *BMJ, 317*, 1110. https://doi.org/10.1136/bmj.317.7166.1110b

Thesmar, D., Sraer, D., Pinheiro, L., Dadson, N., Veliche, R., & Greenberg, P. (2019). Combining the power of artificial intelligence with the richness of healthcare claims data: Opportunities and challenges. *PharmacoEconomics, 37*(6), 745–752. https://doi.org/10.1007/s40273-019-00777-6

Vayena, E., Salathé, M., Madoff, L. C., & Brownstein, J. S. (2015). Ethical challenges of big data in public health. *PLoS Computational Biology, 11*(2), e1003904. https://doi.org/10.1371/journal.pcbi.1003904

Wielki, J. (2015). The social and ethical challenges connected with the big data phenomenon. *Polish Journal of Management Studies, 11*(2), 192–202.