Like Selling the Sky
An Analysis of Data Colonialism and Necessity of Indigenous Data Sovereignty

When European colonists arrived in South Africa seeking to claim the land's rich source of natural resources starting in the 1650s, they encountered a problem. Generations of Indigenous communities had been living on these lands for centuries. The Dutch and British, however, claimed the land was *empty,* when in fact indigenous agropastorialists had developed "sophiscated political and material cultures."

The Empty Land Theory stemmed from a fundamental misunderstanding of South African indigenous practices — for these communities, land ownership was focused more so on the products reaped from the land, less on the area itself. Land was not owned by individuals as "private property", allowing passing nomadic groups to utilize it as well. For these indigenous groups, the idea of such impersonal transactions of land were likely as bizarre as selling pieces of the sky (an untouchable, unmonetizable component of their environment). This kind of disregard of community practices is central to the damaging and exploitative power of colonialism.

Now, the same abstraction of indigenous cultures has made its way into the digital age, which researchers Jennifer Shae Roberts and Laura Montoya call *data colonialism* in their article "Inconsideration of Indigenous Data Sovereignity." Let's explore the harms and social implications of data colonialism from Roberts and Montoya's research.

Left Out of the Data

In an age where "data is king," Indigenous voices are being left out of the conversations around the usage of their data in the development of data-dependent technology, such as A.I as Roberts and Montoya claim. By "indigenous," they mean communities (such as the Maori of New Zealand) that fight to maintain autonomy amidst waves of historic marginalization often originating from dehumanising settler colonialist efforts.

It's important to recognize, though, that the concept of data and the collection of data is not a new one. Data at its core is information we define ourselves by. Nowadays, its age, gender, occupation, etc. But for indigenous communities, its agricultural practices, music, community traditions, and art forms. Let's call this pre-data.

I concur with Roberts and Montoya when they claim that data as we define it now is not naturally occurring. The push away from this kind of "abstraction of life into bits and bytes" is a core tenant of data humanism: a movement that utilizes data to preserve human lives, to embrace the lived complexity inherent to our stories.

The exclusion of indigenous data and pre-data from the building of modern technology-dependent systems leads to systems that value profit over life, resulting in the destruction of vital ecological networks, violation of individuals privacy, and abusive labour conditions.

Sovereignity as a Solution to Exclusion

As a solution to the impacts of data exclusion and to overcome external oppression, Roberts and Montoya present Indigenous Data Sovereignty (ID-SOV): the right to assert and act upon their right to self-determination over their data and pre-data.

Though, acheiving data sovereignty is easier said than done. Asserting this right means challenging historically oppressive systems and demonstrating to them that "the adherence to ID-SOV principles will be of benefit to everyone, as well as the health of the planet." This is largely because the wellbeing of those in power is normally considered before and above all others. Exactly what incentive could be offered to these groups should be thought of carefully.

The CARE Principles: Sovereignity in Action

For Roberts and Montoya, the journey to ID-SOV must be guided by 4 principles that explicitly respond to the exclusion and unethical mining of their data/pre-data: C.A.R.E. They assert that Indigenous data sovereignty will only succeed if data mining enables *collective benefit*; if it endows them the *authority to control* their data; if there are *responsible* practices to protect their data rights; and if the *ethical* wellbeing of said rights are made the primary concern.

As a form of direct resistance, Roberts and Montoya outline the implementation of CARE at each level of the data lifecycle. This concrete plan serves an understanding of the power of data and a promise to dismantle the systems in place from the inside.

Data now defines our lives, as a social movement and on an individual scale. We are the sum of our individual points, like puppets controlled by threads of our data and for Roberts and Montoya, the time is now to cut the strings. The result of biased and exploitative applications of data is largely due to the biased and exploitative nature of those applying the data. As Catherine D'Ignazio and Lauren Klein, authors of "Data Feminism", put it, what goes into computational systems is reflective of what comes out. Meaning that the problem was never really about data. It is about control and complexes of superiority. It is about power. How then, I ask, can we challenge those in power if they believe we are unworthy?

Word Count: 789