Ethical Stakeholder Analysis

Ben Mitchell

Introduction

This article introduces a variant of Stakeholder Analysis as a method to help analyze the morality of a decision. It is particularly well suited to use with consequentialist ethical frameworks (e.g. utilitarianism, natural justice, etc.); other types of frameworks (e.g. deontology, virtue ethics) can be used, but it may be more challenging.

There are many different methodologies that can be used to help analyze the ethical impacts of a proposed decision. Like the choice of which ethical framework to use, different methodologies will have different strengths and weaknesses, so we shouldn't become overly fixated on a single method to the exclusion of all others. Nonetheless, I find this method to be a reasonable starting point for analyzing the impacts of technology on society.

The classic version of 'Stakeholder Analysis' is used in the business and project design communities; our version here borrows several key parts of the process, but ultimately we will use them differently since we have a different goal. After all, 'success' from a business perspective is not a guarantee of moral rectitude.

What is a Stakeholder?

For our purposes, a 'stakeholder' is any individual (or group of individuals) that may be impacted by the proposed system or design choice we are analyzing. Here, the 'impacts' we are primarily concerned with are those that are ethically relevant, but it is typically easier to begin by figuring out what the impacts will be, and then later determining the *value* of those impacts.

As an example, we might consider a proposal for a music player app with a recommender system (e.g. Spotify, Pandora, etc.). For such a system, the most obvious stakeholders will be those who create/sell the app, and those who use the app.

It is important to note that there are different ways we can break out these groups. In our example, perhaps we have an app with two types of users: those who pay a monthly subscription, and those who use the app for 'free' but have to listen to advertisements. Likewise, unless the app is created by a single individual, we may want to subdivide the 'owners' to distinguish between programmers, executives, support staff, etc.

The decision about when to treat a set of individuals as a 'group' and when to subdivide them should be based on whether the impacts on the members of the group will be impacted similarly. For example, if we expect that individuals of any gender will be impacted in the same way, it would make sense to group them all together. On the other hand, if we believe there may be differences in impacts to individuals of different genders, then it might be better to do the analysis for each gender identity group separately.

I find it helpful to use a tree-based hierarchical approach, starting at the top with something like 'people,' which can be split into 'creators' and 'users', which in turn can each be subdivided further. By doing as much analysis as we can at the higher levels of the tree, we can avoid replicating work across different groupings at lower levels.

Of course, this raises the issue of how finely grained our decomposition needs to be, and when we should stop. For instance, how many identity-based groups do we need? Do we divide based on age? Race? Gender? Wealth? Class? This is especially difficult when categories are socially constructed and have no clear, agreed upon boundaries. Additionally, there may not be a meaningful 'hierarchy'; e.g. we may want to consider both race and gender, but putting one 'above' the other in the tree seems somewhat arbitrary.

The guiding principle for making these choices should be, "what seems most relevant to our ethical analysis." Keep in mind that this entire process is just a tool to help you reason through ethical decision making, and make choices that facilitate that purpose.

Assessing Impacts

Once we have identified groups of stakeholders, we can start assessing the impacts that our proposed system or design choice might have. I find it helpful when thinking about impacts to consider three different types:

- **Direct Impacts** A *direct* impact is a change that is experienced by a stakeholder as a direct result of our choice. For example, a user of our music app discovering a new artist they like would be a direct impact.
- **Indirect Impacts** An *indirect* impact is a change that results from a chain of cause-and-effect that *begins* with a result of our choice, but may also be influenced by choices made by others. For example, if our music app tends to recommend shorter songs more than longer ones, this could result in artists creating shorter and shorter tracks. The app is not *directly* shortening the length of songs, but it is still clearly a causal factor.
- **Relational Impacts** A *relational* impact is one that, rather than being a unilateral change for one group, changes the *relationship* between groups. One of the most common types of relational impact is a change in power dynamics, where our decision results in a change in the balance of power between two (or more) groups. However, relational impacts are not restricted to *just* power dynamics; this label covers any impact that involves multiple groups in relation to each other. For example, our music app recommender system might impact the dynamics of the relationships between

artists, publishers, and consumers. It might also impact relationships between parents and children, resulting in less parental control (or knowledge) about what their kids are hearing.

Analyzing and Making Choices

Once we have a clearer idea about the impacts a proposed design decision might have, we can start thinking about whether we should implement it or not. At the simplest level, we can try to apply one of the classic consequentialist ethical frameworks to evaluate whether a specific proposal is justified. This requires assigning numeric value to the different impacts and groups, which exposes us to all the usual dangers of consequentialism, but it can still be useful.

Another way to use this tool is to consider a range of different design choices, and think about how to minimize harms or maximize goods. For example, if we have several different options for how to implement our music recommender system, we might find that some options may have indirect or relational impacts on artists that seem likely to negatively impact the entire music scene in the long run. It might be worth choosing an option with less short-term payoff for company or the consumers in exchange for a more stable business model that doesn't cannibalize its content creators.