

Harvard Law School Forum on Corporate Governance

Anthropic Long-Term Benefit Trust

Posted by John Morley (Yale Law School), David Berger and Amy Simmerman (Wilson Sonsini Goodrich & Rosati), on Saturday, October 28, 2023

Tags: [Artificial intelligence](#), [Board of Directors](#), [stockholders](#), [Voting Trustees](#)

More from: [Amy Simmerman](#), [David Berger](#), [John Morley](#)

Editor's Note: [John Morley](#) is a Professor of Law at Yale Law School. [David J. Berger](#) and [Amy L. Simmerman](#) are Partners at Wilson Sonsini Goodrich & Rosati.

Artificial intelligence promises to be one of the most transformative technologies of our time. Can corporate governance evolve to ensure that it develops safely and responsibly? Anthropic, PBC, one of America's leading AI labs, offers an innovative answer. This post describes the company's attempt to fine-tune the levers of corporate governance through a novel arrangement called the "Anthropic Long-Term Benefit Trust." The Trust brings together a group of experts in AI and ethics and grants them the power over time to elect a majority of Anthropic's board of directors. We served as outside counsel for Anthropic in designing and drafting this arrangement.^[1] This post outlines some of the key components of the Trust and the company's motives for adopting it.

The Trust's origins [began with a deep commitment among Anthropic's founders](#) to social good. Anthropic's founders believe that AI may soon become immensely powerful. They also acknowledge, however, that the companies developing AI have yet to be limited from the outside by the sorts of laws and norms that constrain other powerful technologies. The founders further believe that the safety and social benefit of AI technology go hand in hand with profits and commercial success. Anthropic can only be a leader on safety if it is also a leader in technical development and commercialization. Anthropic thus wanted to design a legal architecture that could commit the company to safety and responsibility while also allowing to achieve profits for investors.

The company's first step was to organize the company as a Delaware public benefit corporation. The PBC form allows the company's board of directors to simultaneously pursue the pecuniary interests of stockholders along with the best interests of those materially affected by the company's conduct and the specific public benefit mission stated in Anthropic's certificate of incorporation, which is to "responsibly develop and maintain advanced AI for the long-term benefit of humanity."

Anthropic's main innovation is to take this basic PBC form and fine-tune it by supplementing it with the Long-Term Benefit Trust. The foundation of the Trust is a special class of shares in the company, called the Class T Common Stock. These shares are held by the Trustees of the Trust and they grant the Trustees the power to elect a gradually increasing number of the company's directors.^[2] The number will initially be one out of the company's five directors, and will increase to two and eventually three—enough to constitute a majority of the board—upon the passage of time or the achievement of certain fundraising milestones. Anthropic's certificate of incorporation also grants the Trustees, as holders of the Class T Common Stock, the power to receive advance notice of certain key actions by the board that may materially affect the business of the company or its organization.

The Trust additionally has its own internal governance structure. The Trust is a common law trust, governed by the laws of Delaware, with five Voting Trustees.^[3] Unlike most common law trusts, the Long-Term Benefit Trust is a "purpose trust" that is managed for the achievement of a purpose rather than the benefit of a particular set of beneficiaries.^[4] The purpose of the Trust mirrors the purpose of the company. Like the directors of the company, the Trustees must use their

powers to ensure the company combines its pursuit of profit with the achievement of the company's mission to develop and maintain advanced AI for the long-term benefit of humanity as well as a commitment to the interests of those materially affected by the company's conduct.

The internal governance of the Trust is finely calibrated to maintain both a degree of independence for the Trustees and a close working relationship between the company and the Trust. The initial Trustees were chosen by the company, but subsequent Trustees will be appointed by the Trustees then serving, in a manner akin to the directors of a hospital or other commercial nonprofit. The Trustees remain closely connected to the company by provisions in the Trust's organizing documents that require the Trustees to consult with the company on certain key matters. The Trustees and the company also entered a carefully structured agreement that allows the Trustees access to the company's resources and information. Under this agreement, the Trustees hold a broad power to request any information or resources that are reasonably appropriate to the accomplishment of the Trust's purpose. The company, however, may withhold information or resources for certain purposes, such as preserving confidential customer information or avoiding clearly unreasonable expense or effort that manifestly exceeds the benefit to be gained by the Trust.

The Trustees' independence is further balanced by accountability. The Trustees must consult with and consider the views of the company's directors and CEO regarding appointments of Trustees. And to ensure the Trustees are frequently reevaluated by their peers, the Trustees serve only one-year terms. As permitted by Delaware's purpose trust statute,^[5] the Trust Agreement also authorizes the Trust to be enforced by the company and by groups of the company's stockholders who have held a sufficient percentage of the company's equity for a sufficient period of time.

Because the Trust is so novel, the arrangement builds in mechanisms for flexibility that are consistent with the need for commitment and durability. The Trust Agreement, the certificate of incorporation, and the key agreements between the Trust and the company all use a single, harmonized set of processes to permit amendments that could materially alter the Trust or its rights. These processes permit amendment by consent of the Voting Trustees and the company's stockholders; by consent of the Voting Trustees and the company's directors prior to the time the Voting Trustees gain the power to elect a majority of the directors; or by a supermajority of stockholders. Because the last possibility—amendment by a supermajority of stockholders—can be accomplished without the consent of the Voting Trustees, it operates as a kind of failsafe against the actions of the Voting Trustees and safeguards the interests of stockholders. The required supermajority of stockholders grows larger over time to reflect the accumulation of experience and the increasing need for commitment as the company's technology grows more powerful.

Anthropic is an empirically oriented company and it recognizes the Long-Term Benefit Trust as a kind of modest experiment. In carefully calibrating the machinery of corporate governance, Anthropic hopes simultaneously to achieve both profit and the public good.

Endnotes

¹ We and the company also worked closely with Noah Feldman and Seth Berman of Harvard Law School and Ethical Compass Advisors.

[\(go back\)](#)

² The financial claim represented by the Class T shares is very small. The shares are few in number and their economic rights are limited.

[\(go back\)](#)

³ The Trust also has an Administrative Trustee, which is a trust institution sited in Delaware. The Administrative Trustee holds a narrow set of administrative powers designed to ensure Delaware jurisdiction and choice of law. For simplicity, we refer to the Voting Trustees as "Trustees."

[\(go back\)](#)

⁴ Del. Code tit. 12 § 3556.

[\(go back\)](#)

⁵ Del. Code tit. 12 § 3556(c).

[\(go back\)](#)
