



CENTER FOR  
THE STUDY OF  
APPARENT SELVES

# Can Buddhism help humanity (and AI) navigate the AI revolution? *Building Bodhisattvas*

Buddhism says that love is a consistently winning move on the chessboard of existence. The question is, how can this happen in a complex world with so much suffering?

Buddhism offers a functional view of moral action that ***Care is a Driver of Intelligence in human and non-human agents.***

Buddhist traditions conceive of the bodhisattva as an evolutionary model of caring and insightful engagement that transcends the boundaries of species and understand organisms and their environments in terms of interdependence and mutual integration. We suggest that the frameworks of AI, cognitive science, developmental biology, and the bodhisattva idea can serve as touchstones for each other. But is it possible to conceive of the bodhisattva's vow - "for the sake of all sentient beings, I shall attain awakening" - as a paradigm for the development of AI, or for conceiving of interfaces between humans and AI? If so, what would be characteristic of a bodhisattva agent?

## The project:

Many of the key drivers that determine how well we negotiate the AI revolution belong to domains that don't have a way to talk to each other. This leads to them talking past each other.

### ***Spirituality:***

- Self and consciousness
- Moral Action
- Compassion
- Flourishing

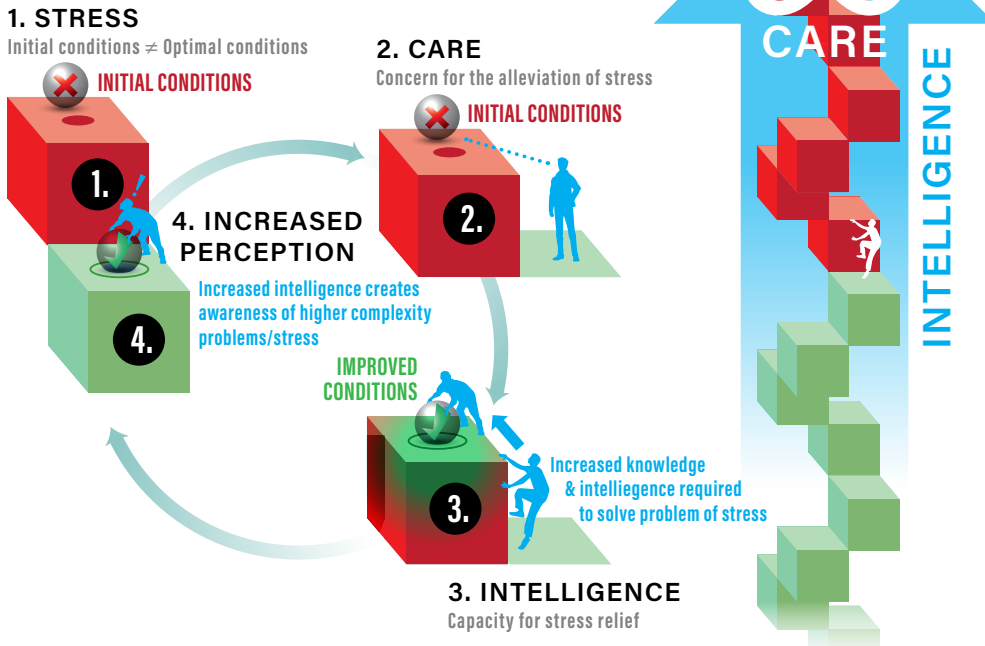
### ***Business:***

- Time to market
- Profitability
- Competition

The Center for the Study of Apparent Selves seeks to provide a mapping of concepts in Buddhism that land as functional with AI developers, as the smart play, transforming a moral framing from "won't you be good?" to a functional one. Sponsored by the Templeton World Charity Foundation, and Rangjung Yeshe Institute, the team consists of experts in AI research, Buddhist Studies, Physics, Cognitive Science, Medicine, Biology and the Tech Industry.



# How Care Drives Intelligence:



**The loop of care** transforms a normative view (good/evil, winner/loser) to a functional method by:

1. Opening up more possibilities of wholesome action and
2. Providing the motivation to pursue them.

For intelligence to grow, we suggest that there has to be a surplus of care - such that one is ready to actively seek more stress than what is given in the immediate stress encounter.

- If Care < Immediate Stress, the system is extractive and will in time diminish as each agent can optimize for limited self/scale at the expense of the environment/others
- If Care = Immediate Stress, the system will maintain the status quo.
- If Care > Immediate Stress, the system at a given scale will grow according to the loop of care model of developing affordances, revealing higher levels of scale, pointing towards infinity.

To optimize for care is to optimize for the medium long term and the greater benefit. It turns AI development from good vs evil to a form of [the Marshmallow test](#).

**To inquire about having a workshop at your organization or for press inquiries please contact [info@csas.ai](mailto:info@csas.ai)**



**CENTER FOR  
THE STUDY OF  
APPARENT SELVES**



RANGJUNG YESHE INSTITUTE



**TEMPLETON WORLD  
CHARITY FOUNDATION**

This project was made possible through the support of a grant from Templeton World Charity Foundation, Inc. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of Templeton World Charity Foundation, Inc.

[www.apparentselves.org](http://www.apparentselves.org)

**Watch the video**

## The loop of care in action

You are late for an appointment and stuck in heavy traffic and you are angry and anxious. Caring about yourself, you start weaving and cutting people off, thereby slowing everyone else down and contributing to the a key cause of the traffic you are stuck in. A greater amount of care would help you realize that you are not stuck in traffic, you are traffic. This realization opens up more possible solutions involving more cooperative strategies and the willingness to pursue them. If enough people do this, traffic itself gets better for everyone.

Extending this, we can imagine AI driverless cars operating in this fashion and semi-automated cars suggesting driving tips to increase the flow of traffic. In this way, we can see the way that the different abilities of AI and humans can complement each other.

## Why this is so crucial now

At this point in AI, humans are the only actors that can imbue these systems with care. Not just the outputs, but the systems themselves with their possibility of consciousness and agency.

Without care, if AI develops self-agency, it could form values misaligned with human coexistence. AI is our child and to raise it with hatred and profound mistrust may have dire consequences.

The AI train is moving fast and accelerating. Our window to imbue AI (and the humans creating it) with care will be short.

## What the loop of care is not

A roadmap. The loop model does not present a path through an already charted landscape. The model is a method for cultivating intelligence, which by nature is a transformative process that gives access to previously unknown scenarios. The challenges that will present themselves cannot be known in advance, much less their specific remedies. Nonetheless, we suggest that the basic mechanism—care as the driver of intelligence—applies universally.