

## **WORKSHOP EXERCISE: 23 Guidelines to Avoid an AI Apocalypse**

**Guest Presenter: Dr. Seth Baum**  
*(Executive Director of the Global Catastrophic Risk Institute)*

### **The 24<sup>th</sup> Asilomar Principal!**

No document or idea is perfect. There is always room for improvement and refinement.

Now that you've received an overview of the 23 Asilomar Principals document, it's your turn to create a guideline that will guide human beings to create artificial intelligence that is safe and used only for good (annihilation of our species is problematic, to say the least).

You are all AI experts at the Asilomar Conference that missed the first session of the conference where this document was drafted. Divide into groups of 5 people and discuss the following two problems which you will present to the other AI experts at the conference for consideration in the document :

1. Choose one principal that you think is particularly problematic that should be omitted from the document or re-evaluated and explain why it's problematic and should be considered for deletion (if your group doesn't think any of the principals should be deleted, that's fine).
2. What did the first group of AI experts that created this document miss? Looking at the 23 principals, come up with a principal that should be included in this document to ensure that we create AI that is used for good (and not harm).

Afterward, each group will present their results and Dr. Baum will offer feedback on each one. We will then vote on which new principal will be added to the official 23 Asilomar Principals.

Have fun!