

**From:** Lawrence Krauss <[REDACTED]>

**To:** "jeffrey E." <jeevacation@gmail.com>

**Subject:** Fwd: Adverse Outcomes & Teams for Origins Project Scientific Workshop

**Date:** Tue, 21 Feb 2017 01:20:14 +0000

**Attachments:** AI\_Workshop\_-\_Adverse\_Outcomes\_&\_Teams.pdf; AI\_Workshop\_-\_Participant\_List\_022017.pdf

**Inline-Images:** DA866543-7401-4A5A-8E50-FD32E33A50EC.png; image001.png

---

fyi..

**Lawrence M. Krauss**

**Director, The Origins Project at ASU**

**Co-Director, Cosmology Initiative**

**Foundation Professor**

School of Earth & Space Exploration and Physics Department  
Arizona State University, P.O. Box 871404, Tempe, AZ 85287-1404  
Research Office: 480.965.6378, Assistant (Jessica): 480.965.9825  
Origins Office (Cynthia): 480.965.0053

[REDACTED]

[origins.asu.edu](http://origins.asu.edu) | [REDACTED] | [REDACTED]



Begin forwarded message:

**From:** Amelia Huggins [REDACTED]

**Subject:** Adverse Outcomes & Teams for Origins Project Scientific Workshop

**Date:** February 20, 2017 at 3:32:48 PM PST

**To:** Amelia Huggins [REDACTED]

**Cc:** Eric Horvitz <[REDACTED]>, Jaan Tallinn <[REDACTED]>, Lawrence Krauss <[REDACTED]>

Dear Participants,

We're looking forward to see you all on Friday in sunny Tempe, Arizona.

Thanks to all who submitted write-ups on proposed scenarios and adverse outcomes. We've worked to pool these and other inputs (including reflections from the organizers) into 6 areas that capture a class of potential adverse AI outcomes. We have included a particular adverse outcome per challenge area. We realize that some of the adverse outcome descriptions are more specific than others.

We have attached two documents. The first describes the set of 6 proposed adverse outcomes of AI that we will focus on at the meeting, along with team assignments and session chairs. The second document contains the confirmed participants and their contact information.

Based on our sense for interests and backgrounds, we have assigned a chair for each topic, and we have assigned each of you to teams for *two* adverse outcomes. We have divided each team into *red* and *blue* subgroups. We ask that folks assigned to the red subgroup work on specific ways to prepare ideas on increasing the cost, duration, or inability to address the outcome. The latter includes coming up with mechanisms or arguments that lower the probability of

disrupting the outcome. We ask folks assigned to the blue subgroup to think about the converse: What might be done to address and disrupt the adverse outcome from occurring, to make it less likely, or to reduce its cost should it occur. What might be done proactively in terms of mechanisms, policies, designs, best practices, or regulatory activities and treaties to address a concern—or to disrupt or lower the probability of the poor outcome? What evidence might come to the fore along the way so as to provide an alert or early warning of an impending challenge—so as to work to marshal resources, designs, or policies?

Please take a look at the topic descriptions and reflect in advance about the red and blue assignments. On Friday late morning and early afternoon we will sequentially assign 30 minute sessions for each adverse outcome team to meet for discussions and comments at the sessions on the full workshop day on Saturday and on Sunday morning. Groups can also meet Friday evening after dinner.

As noted above, each outcome discussion will be led by the topic chair. We will ask the chairs to share the overview presentations with the larger teams by Thursday evening if possible, so folks on the larger teams can reflect about their comments during the open discussion in the context of the specific content that the chairs will present, recognizing that they need not restrict their discussions to this content. Your plans can be refined when the groups meet.

As for the flow, after the chair presents the overview, we will first invite comments from the blue subgroup panel and then the red subgroup panel, before opening up the discussion to all. The timing for each outcome discussion will be a 10 minute chair presentation, 15 minute a piece for red and blue team panels, and 45 minutes for open discussion among all participants.

These details and materials are also available on our [participant website](#).

Thanks again for agreeing to participate in this experiment. We're looking forward to seeing you.

Eric, Jaan, Lawrence

---

**Amelia Huggins**

**Executive Director**

Origins Project at ASU, Arizona State University

P.O. Box 871902, Tempe, AZ 85287-1902

O: 480.727.7115 | C: [REDACTED] F: 480.965.0090

[REDACTED]  
[origins.asu.edu](https://origins.asu.edu) | [REDACTED] | [REDACTED]