

From: Ben Goertzel [REDACTED]

To: Jeffrey Epstein <jeevacation@gmail.com>, "J. Epstein" <jeeproject@yahoo.com>

Subject: "Coping with Future Catastrophes" conference: suggestions

Date: Fri, 06 Apr 2012 00:31:49 +0000

Attachments: EDITED_GoertzelNineWaysToFriendlyAI_v6_BG

Hi Jeffrey,

I'll send you an update on my AGI project (which is going reasonably well) sometime soon, but what occasioned this email was reading the press release for your upcoming meeting on global risks in Dubai:

<http://www.pr.com/press-release/403599>

This topic interests me a lot as you probably know ... attached find an almost-final version of a paper on how to mitigate the potential risks of advanced AGI systems, which is going to be published shortly in the Journal of Evolution and Technology.

So if you see fit, it would be great for you to reserve a slot for me at the Dubai meeting. ;-)

Also, I have some other suggestions for you...

You might want to consider inviting Steve Omohundro, who has written some pretty insightful stuff about the potential existential risks posed by AGI.... His paper on Basic AI Drives is a good one

<http://selfawaresystems.com/2007/11/30/paper-on-the-basic-ai-drives/>

and I suspect it will appeal to you. Steve and I gave talks together with Larry Krauss (whom it seems you know) at the Singularity Summit Australia last year...

There're also the Singularity Institute for AI guys, who are obsessed with the topic of AGI and existential risk. Some of the SIAI guys are too kooky for my taste; but Luke Muelhauser, their director, is reasonably solid..

<http://singinst.org/aboutus/team>

Another very obvious suggestion is Nick Bostrom, who coined the term "existential risks", and heads the Future of Humanity Institute at Oxford. We're actually having our AGI 2012 conference at Oxford, and alongside the technical AGI presentations we'll have some discussions with Nick and his Oxford philosophy team on AGI-related "future catastrophe" stuff.

Finally Paul Werbos

<http://www.werbos.com/>

is a brilliant cross-disciplinary scientist who knows a hell of a lot about the risks facing humanity on all levels, due to his position with the NSF. He's a rare case of a guy who's a broad-minded, wildly imaginative thinker and ALSO very tight w/ US government science, engineering, military, intelligence, etc.

These folks all represent a community that has thought about, and discussed, the "future catastrophes" issue quite a lot, and probably would have a lot to add to your Future Catastrophes conference.

And as I said, more on the AGI work a little later ;)

thanks much
Ben

--

Ben Goertzel, PhD
<http://goertzel.org>

"My humanity is a constant self-overcoming" -- Friedrich Nietzsche