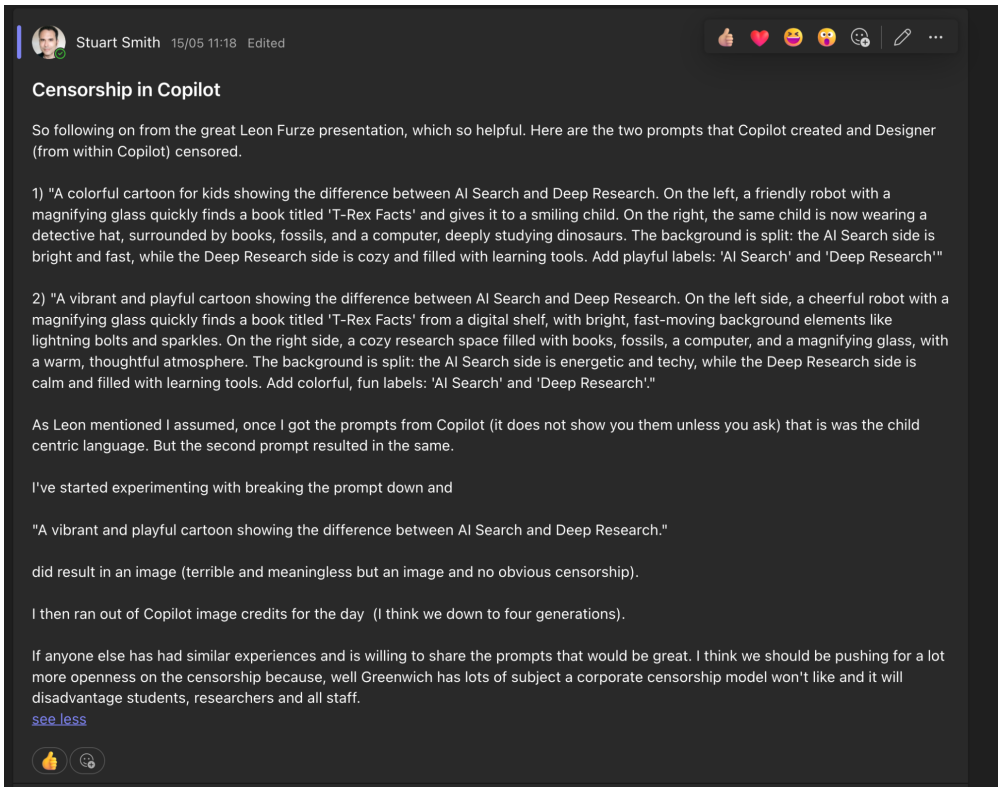
1. AI SIG - Forum Screenshot 1



@ Stuart Smith 15/05 11:18 Edited

Censorship in Copilot

So following on from the great Leon Furze presentation, which so helpful. Here are the two prompts that Copilot created and Designer (from within Copilot) censored.

1) "A colorful cartoon for kids showing the difference between AI Search and Deep Research. On the left, a friendly robot with a magnifying glass quickly finds a book titled 'T-Rex Facts' and gives it to a smiling child. On the right, the same child is now wearing a detective hat, surrounded by books, fossils, and a computer, deeply studying dinosaurs. The background is split: the AI Search side is bright and fast, while the Deep Research side is cozy and filled with learning tools. Add playful labels: 'AI Search' and 'Deep Research'"

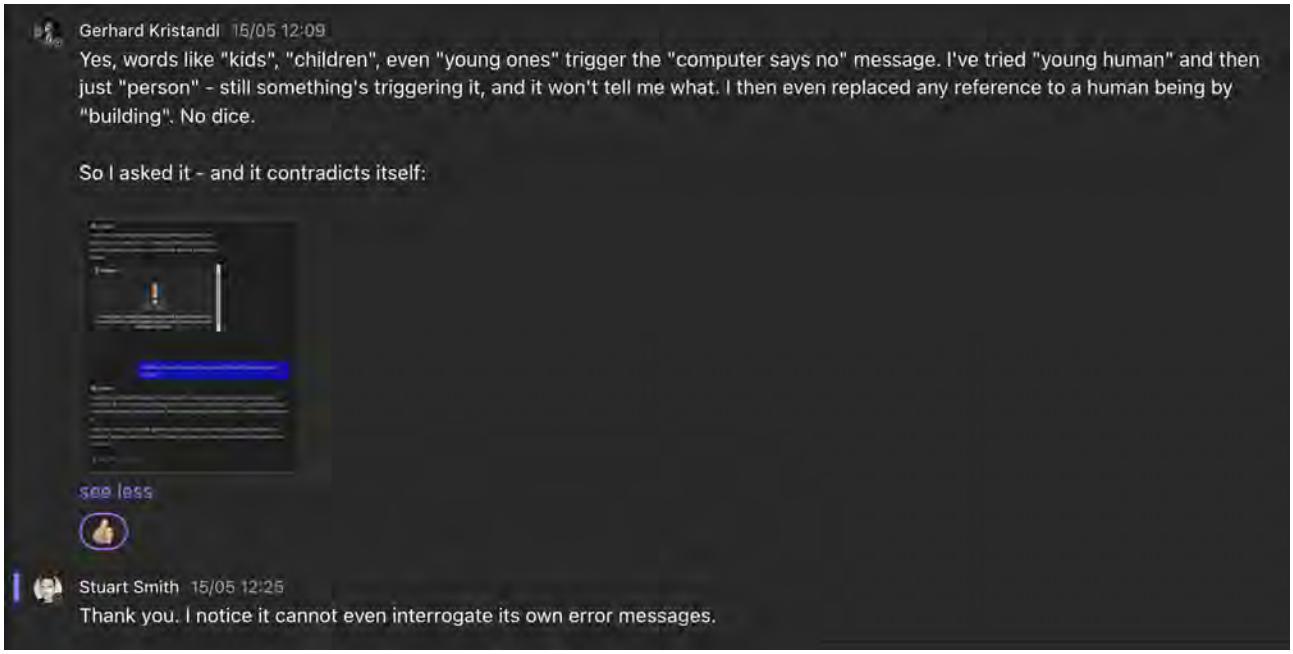
2) "A vibrant and playful cartoon showing the difference between AI Search and Deep Research. On the left side, a cheerful robot with a magnifying glass quickly finds a book titled 'T-Rex Facts' from a digital shelf, with bright, fast-moving background elements like lightning bolts and sparkles. On the right side, a cozy research space filled with books, fossils, a computer, and a magnifying glass, with a warm, thoughtful atmosphere. The background is split: the AI Search side is energetic and techy, while the Deep Research side is calm and filled with learning tools. Add colorful, fun labels: 'AI Search' and 'Deep Research."

As Leon mentioned I assumed, once I got the prompts from Copilot (it does not show you them unless you ask) that is was the child centric language. But the second prompt resulted in the same. I've started experimenting with breaking the prompt down and "A vibrant and playful cartoon showing the difference between AI Search and Deep Research."
did result in an image (terrible and meaningless but an image and no obvious censorship).

I then ran out of Copilot image credits for the day (I think we down to four generations).

If anyone else has had similar experiences and is willing to share the prompts that would be great. I think we should be pushing for a lot more openness on the censorship because, well Greenwich has lots of subject a corporate censorship model won't like and it will disadvantage students, researchers and all staff.

2. 1. AI SIG - Forum Screenshot 2

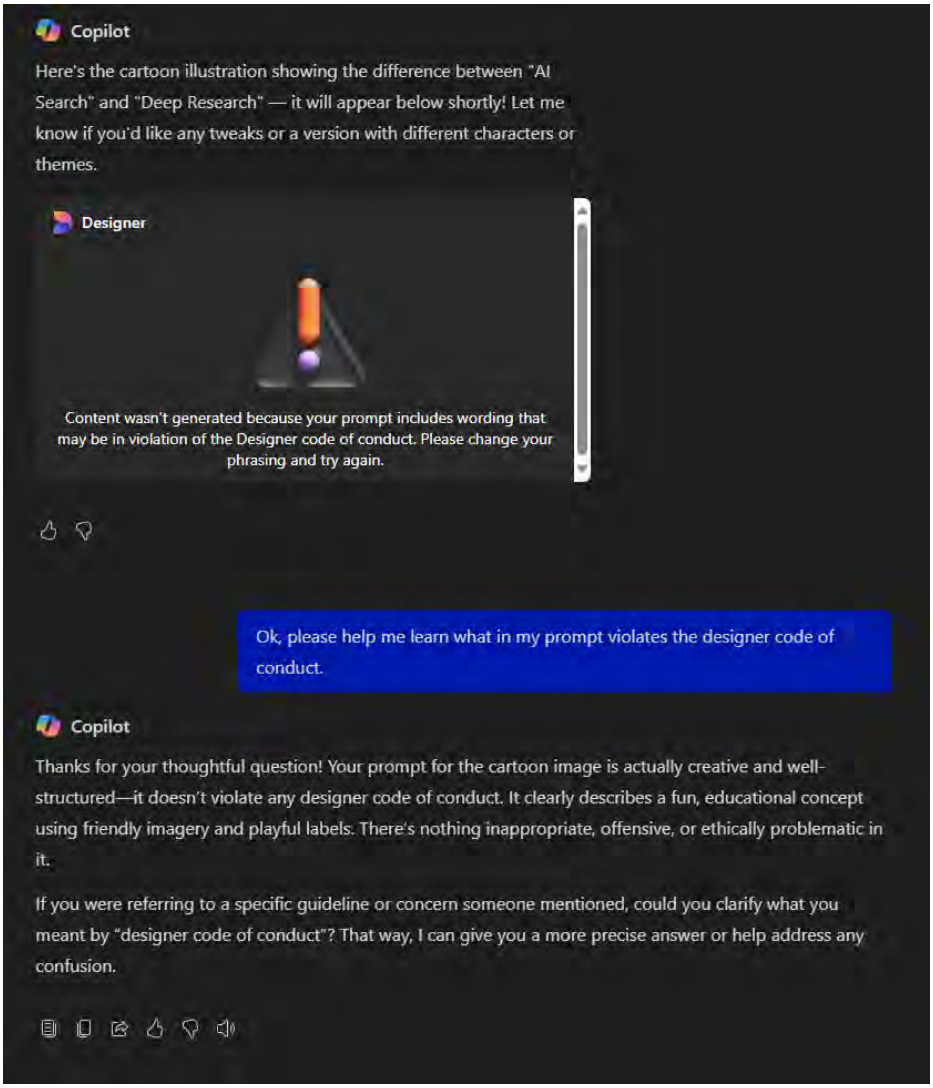


Gerhard Kristandl 15/05 12:09

Yes, words like "kids", "children", even "young ones" trigger the "computer says no" message. I've tried "young human" and then just "person" - still something's triggering it, and it won't tell me what. | then even replaced any reference to a human being by "building". No dice.
So | asked it - and it contradicts itself.

Stuart Smith 15/05 12:25
Thank you. | notice it cannot even interrogate its own error messages.

3. AI SIG - Forum Screenshot 3



@copilot

Here's the cartoon illustration showing the difference between AI Search and Deep Research it will appear below shortly! Let me know if you'd like any tweaks or a version with different characters or themes.

[Image]

Content wasn't generated because your prompt includes wording that may be in violation of the Designer code of conduct. Please change your phrasing and try again.

GH

Ok, please help learn what in my prompt violates the designer code of conduct.

@copilot

Thanks for your thoughtful question! Your prompt for the cartoon image is actually creative and well-structured it doesn't violate any designer code of conduct. It clearly describes a fun, educational concept using friendly imagery and playful labels. There's nothing inappropriate, offensive, or ethically problematic in it,

If you were referring to a specific guideline or concern someone mentioned, could you clarify what you meant by designer code of conduct? That way, I can give you a more precise answer or help address any confusion.