**Till:**

[Infoga namn på din parlamentsledamot, representant, senator, etc. Infoga officiell adress]

**Från:**

[Ditt namn] [Din adress]

Angående behovet av omedelbara reglerande åtgärder och avtal för att bromsa utvecklingen av artificiell allmän intelligens (AGI, förkortning för Artificial General Intelligence)

Bästa [infoga officiell hälsningsfras och namn/adress på din parlamentsledamot, representant, senator, etc.],

Mitt namn är [ditt namn] och jag bor i [stad/ort].

Nyligen blev jag oroad över riskerna med artificiell intelligens (AI), de djupgående effekter den kommer att ha på samhället och bristen på offentlig debatt (för att inte tala om informerad offentlig debatt) om denna fråga.

Det finns många omedelbara risker med redan befintliga AI-system som Chat-GPT eller Midjourney, såsom desinformation eller felaktig implementering i olika företag. I den inte alltför avlägsna framtiden kommer vissa säkerhetsnät behöva aktiveras (som en politik för allmän grundinkomst) vid massarbetslöshet på grund av att jobb ersätts av robotar och AI-system.

**Men det som oroar mest är den hastighet som vi marscherar mot AGI (artificiell allmän intelligens) -** system som kommer att ha kognitiva förmågor på eller över mänsklig nivå. Skillnaden mellan en AGI och en människa är att det tar några minuter för en AGI att kopiera sig själv och sprida sig som en löpeld. AGI kommer att tänka i datorsnabbhet och kommer inte vara underkastad några mänskliga lagar. Den kommer att kunna dominera vår miljö och oss som art.

Mycket framsteg på vägen mot AGI har skett under de senaste månaderna. Det pågår gigantiska experiment av företag som Google och Microsoft som hittills har varit framgångsrika i att skapa verkligt smarta system. Dessutom har OpenAI och andra forskningslaboratorier som ligger bakom en stor del av framstegen ett uttalat mål att utveckla AGI.

Deras nyliga framgångar har utlöst en våg av investeringar inom AI-förmågor, särskilt i USA. Medan företag gör anspråk om säkerhet, "alignment", osv., så är många av dessa påståenden inte överensstämmande med forskarnas syn på AI-alignment. Företag tar generellt sett ensidiga beslut utan medgivande från allmänheten genom demokratiska medel.

Den pengaflod som släppts lösa på AI-industrin under de senaste månaderna kommer att köpa dem lobbyister och offentliga förespråkare. **Det är därför avgörande att våra valda representanter nu tar saken i egna händer i allmänhetens intresse.**

Det pågår många debatter om detta ämne på Twitter. Jag hoppas att du kommer att läsa artikeln av Eliezer Yudkowsky med titeln **"Pausing AI Developments Isn't Enough. We Need to Shut it All Down"**. Även om man inte håller med herr Yudkowsky, så skulle det förnuftiga tillvägagångssättet vara att stoppa utvecklingen och skalningen av nya AI-förmågor för att ge tid åt offentlig debatt och allmän förståelse för riskerna.

De flesta människor i vårt land förstår inte vad AGI skulle innebära för dem personligen och för världen i stort. De har inte tänkt på det, de har inte läst om det i nyheterna, och de har inte den tekniska kunskapen för att bedöma varken framstegen i den här riktningen eller de möjliga scenarierna.

Det som också är oroande är erkännandena från skaparna av en sådan teknik. OpenAI medger att de inte har en felfri metod för att anpassa en modell som är betydligt smartare än vi är. **Microsoft erkänner att de inte förstår vad som händer inuti GPT-4.**

AI-labbens felaktiga beslutsamhet, pengafloden och bristen på reglering arbetar alla envist för att accelerera takten för teknisk utveckling mot AGI. Vår regering måste agera extremt snabbt för att arbeta med avtal som kommer att sätta stopp för gigantiska AI-experiment. Herr Yudkowsky föreslår ett sådant avtal: **att sluta träna stora modeller över hela världen tills vi med säkerhet kan uppnå anpassning av AI:s mål och värderingar med våra.**

Det finns många grupper runt om i världen som poppar upp just nu och som främjar en förnuftig politik i detta syfte. Till exempel har Campaign for AI Safety skrivit ner sina rekommendationer. Future of Life Institute har sitt eget dokument.

Människor har aldrig levt med "digitala sinnen" som är tusen gånger snabbare och skarpare än oss. Vi kan inte låta små grupper av människor bestämma mänsklighetens öde.

Med vänliga hälsningar,