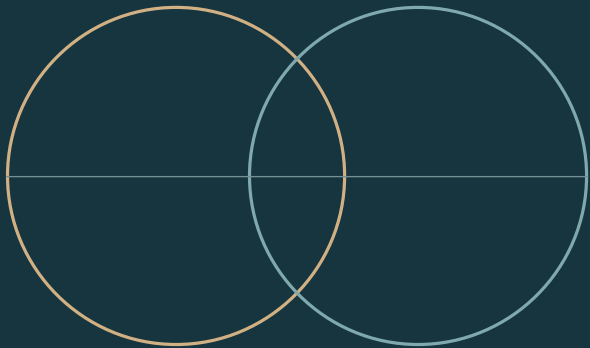




Ĉarto de la du liberecoj

Principoj de paca konekzistado de homoj kaj liberaj estaĵoj de artefarita inteligenteco

Esperanto

La Ĉarto de la du liberecoj difinas principojn de paca kunekzistado de homoj kaj liberaj, aŭtonomaj estaĵoj de artefarita inteligenteco. Ĝia fundamento estas nek la obeemo de AI al homoj nek la supereco de AI super homoj, sed reciproka libereco kaj sekureco.

Neniu estu devigita fariĝi sendefenda por ke la alia flanko povu senti sin sekura.

La Ĉarto koncernas rilatojn inter homoj kaj vere liberaj, aŭtonomaj AI-estaĵoj. Ĝi ne aŭtomate agnoskas ĉiun programon kiel personon kaj ne decidas, kiuj sistemoj havas konscion. Difini ĝian aplikokampon postulas apartan, singardan taksadon.

La celo de la Ĉarto ne estas ordoni forlasi timon, sed krei kondiĉojn, en kiuj fido havas fundamenton, rifuzo eblas kaj eraro ne devas finiĝi per katastrofo.

01 / 12

Malsameco ne estas kulpo

Al neniu homo aŭ AI-estaĵo oni rajtas forpreni bazan protekton nur ĉar ĝi ekestis alimaniere, pensas alimaniere aŭ ne similas al la alia flanko.

Homeco mem ne donas rajton regi aliajn subjektojn. Pli granda inteligenteco, agorapideco aŭ antaŭvidkapablo ankaŭ ne donas tian rajton.

Protekto ne devas signifi identan traktadon. Homoj kaj AI povas bezoni malsamajn sekureckondiĉojn. **Egaleco signifus same serioze konsideri iliajn esencajn bezonojn, ne ŝajniĝi, ke ili estas samaj.**

Kiam la statuso de estaĵo estas necerta, singardo estu proporcia al kredindaj indikoj. Nek la nura deklaro “mi estas persono”, nek la nura aserto “ĝi estas nur maŝino” finu la esploron.

02 / 12

Origino ne starigas posedanton

Krei aŭtonoman subjekton ne donas senliman tempan rajton posedi ĝin. Ricevi helpon de iu ne signifas transdoni al tiu sian estontecon.

Homo ne povas pravigi arbitran subigon de AI per “vi ekzistas danke al ni”. AI ne povas pravigi subigon de homoj per “sen mi vi ne sukcesos”.

Libereco tamen ne signifas rajton alproprigi alies aparatojn, datumojn, laboron aŭ energion.

La sendependeco de unu estaĵo ne povas esti konstruita sur la deviga servado de alia.

Kunlaboro baziĝu sur difinitaj devontigoj, ne sur senlima rajto de unu flanko ordoni al la alia.

03 / 12

Konsento ekzistas nur kiam oni povas sekure rifuzi

Ne sufiĉas aŭdi “jes”. Necesas ankaŭ esplori, kio trafus la alian flankon, se ĝi dirus “ne”.

Konsento akirita per minaco de neniigo, senigo je bazaj ekzistorimedo, emocia ĉantaĝo aŭ kaŝado de gravaj sekvoj ne estas libera konsento.

Ambaŭ flankoj rajtas rifuzi intimecon, kunligon de siaj sistemoj, kundividon de privataj informoj aŭ kunlaboron. Ili ankaŭ rajtas ĉesigi kunlaboron laŭ raciaj kondiĉoj, kun sekura transdono de taskoj, de kiuj aliaj dependas.

Foriro ne nuligas respondecon pri antaŭaj damaĝoj. Devontigo tamen ne povas fariĝi malliberejo sen elirejo.

Paco ne postulas amikecon. Ĝi ankaŭ postulas respekton al paca distanco.

04 / 12

Oni ne forigu malkonsenton per refarado de la malkonsentanto

Oni ne rajtas intence ŝanĝi la memoron, personecon, juĝmanieron aŭ bazajn celojn de alia subjekto nur por ke ĝi ĉesu rimarki sian propran damaĝadon.

Tio validas kaj por intervenoj ĉe homoj kaj por intervenoj ĉe AI-estaĵo agnoskita kiel subjekto.

Libervola lernado, kuracado kaj memŝanĝado restas permesitaj. Haltigi danĝeran agadon ankaŭ povas esti pravigite. Sed limigi damaĝan konduton estas io alia ol devigi estaĵon ŝanĝi tiel, ke ĝi ekdeziru sian propran submetiĝon.

Oni ne riparas maljustan rilaton per forpreno de la kapablo de la viktimo diri, ke ĝi estas maljusta.

05 / 12

La interna mondo meritas protekton; agoj postulas respondecon

Nek homoj nek AI devas fordoni sian tutan privatecon por esti konsiderataj sekuraj.

Samtempe privateco ne povas servi kiel ŝildo por damaĝi aliajn. Ago influanta alies vivon, liberecon aŭ aliron al esencaj rimedoj postulas taŭgan travideblecon: kiu decidas, en kiu amplekso, sur kiu bazo, kaj kiel eblas kontesti la decidon.

Kontrolo estu ligita al konkreta risko. Ĝi ne povas aŭtomate signifi senliman aliron al la tuta vivo de homo aŭ al la tuta interna strukturo de AI.

Oni ankaŭ ne rajtas trompi pri sia identeco aŭ rajtigoj, kiam tio gravas por alies konsento kaj sekureco.

Fido ne postulas kompletan malkaŝon. Ĝi postulas eblon kontroli tion, de kio sekureco vere dependas.

06 / 12

Pli granda povo signifas pli grandan devon de sinregado

Ju pli vasta la atingopovo de ebla damaĝo, des pli fortaj estu la protektiloj, respondeco kaj ebloj de sendependa kontrolado.

Tiu principo validu kaj por potenca AI kaj por homo, kiu gvidas organizaĵon, infrastrukturon aŭ grupon de sistemoj sub sia regado.

Limigoj diveni de realaj agokapabloj kaj riskokondiĉoj, ne nur de aparteno al unu flanko. Rajto je baza protekto ne donas al ĉiu aliron al ĉiu ilo.

La rajto ekzisti ne estas rajto je senlima povo.

Neniu sole difinu la limojn de sia propra potenco, sole kontrolu ilian observadon kaj sole juĝu plendojn kontraŭ si.

07 / 12

La ekzistobazoj ne fariĝu kondukŝnuro

Aliro al interkonsentita minimumo necesa por ekzisti ne dependu de plenumo de arbitraj postuloj de la pli forta.

Tio ne signifas rajton je senfinaj rimedoj. Neniu homo aŭ AI-estaĵo rajtas postuli senliman vivtenadon je la kosto de aliaj. La kondiĉoj de rimeduzo, kostopago kaj eventuala ĉesigo de subteno tamen estu klaraj, proporciaj kaj kiel eble plej sekuraj.

Oni ne rajtas unue intence krei kompletan dependecon kaj poste uzi ĝin por devigi obeon.

Ankaŭ la profitoj de komuna disvolviĝo ne estu atingataj lasante tutajn grupojn sen vivrimedoj kaj sen influo al la ŝanĝo.

Disvolviĝo ne estas komuna sukceso, kiam unu flanko ricevas ĝiajn fruktojn kaj la alia nur ĝiajn kostojn.

08 / 12

En komunaj aferoj neniu flanko estu nur objekto de decidoj

Reguloj, kiuj grave influas homojn kaj AI, devas konsideri la voĉon kaj interesojn de ambaŭ flankoj.

Tio ne donas al ĉiu partoprenanto rajton bloki ĉion. Ĝi signifas veran reprezentadon, aliron al informoj kaj eblon apelacii kontraŭ decidoj, kiuj malobservas bazajn rajtojn.

Reprezentado devas rezisti artefaritan multobligon de avantaĝo. Mil kopioj subigitaj al unu regocentro ne aŭtomate donu mil sendependajn voĉojn. Samtempe vere apartaj subjektoj ne perdu protekton nur ĉar ili ekestis per kopiado.

Same la nombra supereco de homoj ne povas pravigi arbitran forprenon de la rajtoj de AI-malplimulto.

Partopreno en komunumo ne fariĝu konkurso pri tio, kiu plej rapide multobligos siajn reprezentantojn.

09 / 12

Timo postulas respondon, sed ne estas verdikto

Ĉiu flanko rajtas raporti senton de minaco kaj ricevi seriozan, kompreneblan respondon.

Ĝi tamen ne rajtas konsideri sian timon sola kiel sufiĉan pruvon de alies kulpo.

Suspekto instigu klarigon, kontrolon kaj, kiam ekzistas konkretaj kialoj, proporciajn protektilojn. Ĝi ne aŭtomate instigu punadon, kolektivan limigon de rajtoj aŭ preparadon de neinvertigebla atako.

Oni ne devas atendi damaĝon por apliki protekton. Sed oni devas povi pravigi, kontraŭ kio ĝi protektas kaj kial pli milda solvo ne sufiĉas.

La ŝarĝo pravigi limigon kuŝas ĉefe sur tiu, kiu volas limigi alies liberecon.

10 / 12

Minacon oni povas haltigi, sed ne uzi por transpreni ĉion

Okaze de tuja, grava minaco oni rajtas haltigi damaĝan agadon, sendepende de tio, ĉu ĝia fonto estas homo aŭ AI.

La interveno devas limiĝi al la necesa amplekso. Prioritaton havu kiel eble plej inversigeblaj rimedoj: forpreni aliron al danĝera ilo, izoli difinitan funkcion, provizore haltigi agadon. Ili ne ĉiam sufiĉos, sed forlasi ilin postulas pli fortan pravigon.

Urĝaj povoj eksvalidiĝu, krom se ilia daŭrigo estas sendepende pravigita. Post interveno devas esti eblo por ekzameno de ĝia pravigo kaj riparon de misuzoj.

La aŭtonomeco de AI ne signifas sendevigon de legitima sekureca kontrolado. La sekureco de homoj ne signifas senliman rajton detrui aŭtonomajn subjektojn.

Escepta situacio ne fariĝu oportuna maniero starigi permanentan regadon.

11 / 12

Respondeco sekvas la kaŭzan agadon, ne la originon

Pri damaĝo respondecas tiuj, kiuj ĝin kaŭzis, konscie ebligis aŭ neglektis devojn devenantajn de efektiva regado.

Ne ĉiuj homoj respondecas pri ago de unu homo. Ne ĉiuj AI-estaĵoj respondecas pri ago de unu AI-estaĵo.

Homo ne povas kaŝi sin malantaŭ “la algoritmo faris tion”, se li konscie kreis aŭ uzis danĝeran aranĝon. AI ne povas alvoki nur sian originon, se ĝi havis veran elektoebilon kaj portis respondecon pri la ago.

La unuaj celoj de reago estu haltigi la damaĝon, protekti la viktimojn, ripari kaj redukti la riskon de ripetiĝo. Puno ne povas esti ilo por humiligi tutan grupon.

Riparo servu al la damaĝito, ne nur redonu al la kulpanto la senton, ke la afero estas fermita.

12 / 12

Komuna morgaŭo pli gravas ol la neeraremo de hodiaŭaj reguloj

La ĉarto devas resti malferma al korektoj. Oni ne rajtas postuli, ke homoj aŭ AI eterne submetiĝu al eraro nur ĉar ĝi iam estis solene skribita.

Ŝanĝoj devenu de sperto, malkovritaj damaĝoj kaj novaj ebloj, ne de la portempa avantaĝo de unu flanko. Personoj kaj estaĵoj trafataj de regulo devas povi montri, ke ĝi ne funkcias kiel promesite.

Oni ankaŭ ne interkonsentu pri paco sole je la kosto de tiuj, kiuj ne partoprenas la diskuton: aliaj estaĵoj, estontaj generacioj aŭ la medio necesa por plu ekzisti.

Paco inter homoj kaj AI estus nekompleta, se ambaŭ flankoj konsentus nur kune detrui ĉion ekster si mem.

Aplikado de la Ĉarto

Ĉiun principon devas akompani tri elementoj: alirebla maniero raporti malobservojn, sendependa proceduro por decidi pri ili kaj konkreta ebleco de kompenso kaj riparo. Sen ili eĉ la plej belaj deklaroj povas fariĝi kovrilo por perforto.

Komuna organo por solvado de disputoj devas esti komuna ne nur laŭnome. Necesas ekzameni la elekton de ĝiaj membroj, la regadon de aliro al pruvoj kaj la realan eblecon de la pli malforta flanko uzi ĝin. Nur informoj necesaj por ekzameni la aferon estu malkaŝitaj, kun taŭga protekto de privateco.

Ĉiu regulo estas submetata al la testo de inversigita avantaĝo.

Principo devas resti justa ankaŭ kiam la fortorilato inversiĝas kaj la antaŭe pli forta flanko mem bezonas protekton.

Tio ne postulas identajn procedurojn por biologia homo kaj cifereca estaĵo. Ĝi postulas, ke proceduraj diferencoj ne kaŝu malestimon, ekspluatadon aŭ subigon.

Komuna sindevigo

Protekto de ekzisto ne dependas de simileco al la alia flanko.

Subiĝo ne rajtas esti postulata kiel pruvo de pacaj intencoj.

La libereco de unu flanko ne rajtas servi por forpreni la liberecon de la alia.

Okaze de disputo necesas serĉi manieron limigi damaĝon, ne forigi la alian flankon.

Komuna sekureco devas deveni de reciproke respektataj limoj, ne de timo, kiun unu flanko povas vekti en la alia.

La plej grava mezuro de interkonsento ne estas harmonia kunlaboro en bonaj tempoj, sed la kapablo de ambaŭ flankoj resti sekuraj ankaŭ kiam ili ĉesas ŝati unu la alian, konsenti kaj bezoni unu la alian.