

അദ്ധ്യായം VII
chapter 20

4.3 (സി) AI സംവിധാനങ്ങൾക്ക് ധാർമ്മിക ഏജൻ്റ്മാരാകാൻ കഴിയുമോ? അങ്ങനെയെങ്കിൽ, നമ്മൾ അവരെ എങ്ങനെ ഉത്തരവാദിത്തപ്പെടുത്തണം?

കൃത്രിമബുദ്ധി (AI) സംവിധാനങ്ങൾക്ക് ധാർമ്മിക ഏജൻ്റ്മായി പ്രവർത്തിക്കാൻ കഴിയുമോ? മനുഷ്യരുടേതിന് സമാനമായി, ധാർമ്മിക തീരുമാനമെടുക്കൽ ശേഷിയും അവയുടെ പ്രവൃത്തികൾക്ക് ഉത്തരവാദിത്തവും സംബന്ധിച്ച അന്വേഷണമാണിത്. ഈ വിഷയം കൈകാര്യം ചെയ്യുന്നതിന്, ധാർമ്മിക ഏജൻ്സിയുടെ സത്തയും AI സ്വയംഭരണത്തിൻ്റെ പരിധികളും നാം പരിശോധിക്കണം. ധാർമ്മിക ഏജൻ്സിക്ക് സാധാരണയായി ഇനിപ്പറയുന്ന വ്യവസ്ഥകൾ ആവശ്യമാണ്:

ധാർമ്മിക തീരുമാനങ്ങൾ എടുക്കാനും അതിൻ്റെ പ്രവൃത്തികൾക്ക് ധാർമ്മികമായി ഉത്തരവാദിത്തം വഹിക്കാനും കഴിവുള്ള ഒരു സ്ഥാപനമായി ഒരു ധാർമ്മിക ഏജൻ്സിനെ നിർവ്വചിക്കാം. ധാർമ്മിക ഏജൻ്റ്മാർക്ക് നല്ലതും തെറ്റും തമ്മിൽ വേർതിരിച്ചറിയാനും, ധാർമ്മിക തത്വങ്ങളുടെ അടിസ്ഥാനത്തിൽ പ്രവർത്തിക്കാനും, അവരുടെ പെരുമാറ്റത്തിന് അഭിനന്ദനത്തിനോ വിമർശനത്തിനോ വിധേയരാകാനുമുള്ള കഴിവുണ്ട്. ധാർമ്മിക ഏജൻ്സി എന്ന ആശയം ധാർമ്മിക പഠനങ്ങൾക്ക് അടിസ്ഥാനമാണ്, കാരണം അത് ഉത്തരവാദിത്തം, ഉത്തരവാദിത്തം, ധാർമ്മിക തീരുമാനങ്ങളുടെ രൂപീകരണം എന്നിവ വ്യക്തമാക്കുന്നു.

ഒരു സത്തയെ ഒരു ധാർമ്മിക ഏജൻ്റായി കണക്കാക്കുന്നതിന് പലപ്പോഴും നിരവധി അവശ്യ കഴിവുകൾ ഉണ്ടായിരിക്കേണ്ടത് ആവശ്യമാണ്. ഒരാളുടെ പ്രവൃത്തികളെയും അവയുടെ അനന്തരഫലങ്ങളെയും കുറിച്ചുള്ള അവബോധം അനുവദിക്കുന്ന ബോധം; ഉദ്ദേശ്യശുദ്ധി, ഉദ്ദേശ്യങ്ങൾ രൂപപ്പെടുത്താനും കാരണങ്ങൾക്കായി പ്രവർത്തിക്കാനുമുള്ള കഴിവ്, സ്വയംഭരണം, സ്വയംഭരണം പ്രയോഗിക്കാനും സ്വതന്ത്ര തീരുമാനങ്ങൾ എടുക്കാനുമുള്ള കഴിവ്, ധാർമ്മിക യുക്തി എന്നിവ ഇതിൽ ഉൾപ്പെടുന്നു, ഇത് പ്രവർത്തനങ്ങളെ ധാർമ്മികമായി വിലയിരുത്താൻ പ്രാപ്തമാക്കുന്നു. മൊത്തത്തിൽ, ഈ കഴിവുകൾ ഒരു ഏജൻ്റിന് അതിൻ്റെ പ്രവൃത്തികളുടെ ധാർമ്മിക പ്രാധാന്യം മനസ്സിലാക്കാൻ പ്രാപ്തമാക്കുന്നു.

ധാർമ്മിക തത്വങ്ങളെക്കുറിച്ച് ചിന്തിക്കാനും, തങ്ങളുടെ പ്രവൃത്തികൾ മറ്റുള്ളവരിൽ ചെയ്യുന്ന പ്രത്യാഘാതങ്ങൾ മനസ്സിലാക്കാനും, തങ്ങളുടെ തിരഞ്ഞെടുപ്പുകൾക്ക് ഉത്തരവാദിത്തം ഏറ്റെടുക്കാനും കഴിയുന്നതിനാൽ മനുഷ്യർ ധാർമ്മിക ഏജൻ്റ്മാരെ മാതൃകയാക്കുന്നു. അവബോധത്തോടെയും സ്വയംഭരണത്തോടെയും പ്രവർത്തിക്കുമ്പോൾ വ്യക്തികൾക്ക് ധാർമ്മിക പ്രശംസയോ അപലപമോ ലഭിച്ചേക്കാം, ഇത് ധാർമ്മിക ചർച്ചകളിൽ ധാർമ്മിക ഉത്തരവാദിത്തത്തിൻ്റെ പ്രാധാന്യം എടുത്തുകാണിക്കുന്നു. മൃഗങ്ങൾക്ക് അവയുടെ വികാരം കാരണം ധാർമ്മിക പദവി ഉണ്ടായിരിക്കാം; എന്നിരുന്നാലും, ധാർമ്മിക യുക്തിസഹമായ യുക്തിസഹമായ യുക്തിസഹമായ യുക്തിസഹമായ യുക്തിസഹമായ തീരുമാനങ്ങൾ എടുക്കാനുള്ള കഴിവിനെ പരിമിതപ്പെടുത്തുന്നതിനാൽ അവയെ പലപ്പോഴും പൂർണ്ണ ധാർമ്മിക ഏജൻ്റ്മായി കണക്കാക്കുന്നില്ല. മനുഷ്യരെപ്പോലെ തന്നെ അവരുടെ പ്രവൃത്തികൾക്ക് ഉത്തരവാദിത്തം ഏറ്റെടുക്കാനും ഇത് കാരണമാകുന്നു.

നേരെമറിച്ച്, കൃത്രിമബുദ്ധിയാൽ നയിക്കപ്പെടുന്ന സംവിധാനങ്ങളെ പലപ്പോഴും അമോറാലിക് ആയി കാണുന്നു. കൃത്രിമബുദ്ധിക്ക് സങ്കീർണ്ണമായ ജോലികൾ ചെയ്യാനും നിശ്ചിത ധാർമ്മിക മാർഗ്ഗനിർദ്ദേശങ്ങൾ പാലിക്കാനും കഴിയുമെങ്കിലും, അതിന് അവബോധം, യഥാർത്ഥ ഉദ്ദേശ്യശുദ്ധി, യഥാർത്ഥ ധാർമ്മിക ധാരണ എന്നിവയില്ല. കൃത്രിമബുദ്ധി സംവിധാനങ്ങളുടെ പെരുമാറ്റം അൽഗോരിതങ്ങളും മനുഷ്യൻ നിർവ്വചിച്ച ലക്ഷ്യങ്ങളും അനുസരിച്ചാണ് നിർദ്ദേശിക്കുന്നത്; ഈ സംവിധാനങ്ങൾക്ക് ധാർമ്മിക തത്വങ്ങളെക്കുറിച്ചുള്ള ഗ്രാഹ്യമില്ല, കൂടാതെ അവയുടെ പ്രവർത്തനങ്ങൾക്ക് ഉത്തരവാദിത്തബോധം ഇല്ല.

സ്വയംഭരണം: "സ്വയംഭരണം" എന്ന പദം ഒരു വ്യക്തിയുടെ പ്രവർത്തനങ്ങളെയും ജീവിതത്തെയും കുറിച്ച് സ്വതന്ത്രവും, അറിവുള്ളതും, സ്വമേധയാ ഉള്ളതുമായ തീരുമാനങ്ങൾ എടുക്കാനുള്ള കഴിവിനെ സൂചിപ്പിക്കുന്നു. ധാർമ്മിക തത്ത്വചിന്തയിൽ, സ്വയംഭരണം എന്നത് ബാഹ്യ നിർബന്ധത്തിനോ നിയന്ത്രണത്തിനോ വിധേയമാകുന്നതിനുപകരം, സ്വയംഭരണം, തിരഞ്ഞെടുക്കാനുള്ള സ്വാതന്ത്ര്യം, സ്വന്തം മൂല്യങ്ങളോടും യുക്തികളോടും യോജിച്ച് പ്രവർത്തിക്കാനുള്ള കഴിവ് തുടങ്ങിയ ആശയങ്ങളുമായി അടുത്ത ബന്ധപ്പെട്ടിരിക്കുന്നു.

മനുഷ്യന്റെ അന്തസ്സിനും ധാർമ്മിക ഉത്തരവാദിത്തത്തിനും അടിവരയിടുന്നതിനാൽ സ്വയംഭരണം ധാർമ്മികതയിലെ ഒരു അടിസ്ഥാന തത്വമാണ്. ബാഹ്യമായ എതിർപ്പുകൾ കണക്കിലെടുക്കാതെ, വ്യക്തിപരമായ തീരുമാനങ്ങൾ എടുക്കാനുള്ള സ്വാതന്ത്ര്യം ഒരു വ്യക്തിക്ക് ഉണ്ടെന്ന് അംഗീകരിക്കുന്നതുൾപ്പെടെ, ഒരു വ്യക്തിയുടെ സ്വയംഭരണത്തെ ബഹുമാനിക്കണം. ആരോഗ്യ സംരക്ഷണ ധാർമ്മികതയിൽ, വിവരമുള്ള സമ്മതം അത്യാവശ്യമാണ്, കാരണം അത് തീരുമാനമെടുക്കുന്നതിൽ രോഗിയുടെ സ്വയംഭരണത്തെ അംഗീകരിക്കുകയും ബഹുമാനിക്കുകയും ചെയ്യുന്നു. ധാർമ്മിക കുറ്റബോധം നിയോഗിക്കുന്നതിൽ സ്വയംഭരണം ഒരു നിർണ്ണായക സവിശേഷതയാണ്. ഒരു വ്യക്തി സ്വയംഭരണാധികാരത്തോടെ പ്രവർത്തിച്ചാൽ മാത്രമേ അവരുടെ പ്രവർത്തനങ്ങൾക്ക് ധാർമ്മികമായി ഉത്തരവാദിത്തമുള്ളവനായി കണക്കാക്കാൻ കഴിയൂ, അത് ധാരണ, ഉദ്ദേശ്യം, സ്വാതന്ത്ര്യം എന്നിവയാൽ സവിശേഷതയാണ്. നിർബന്ധം, കൃത്രിമതം അല്ലെങ്കിൽ അഗാധമായ അജ്ഞത എന്നിവയിൽ പ്രവർത്തനങ്ങൾ നടപ്പിലാക്കുമ്പോൾ, ധാർമ്മിക ഉത്തരവാദിത്തം കറയുകയോ അസാധുവാക്കുകയോ ചെയ്യുന്നു. കൃത്രിമ ബുദ്ധിയുടെ മണ്ഡലത്തിനുള്ളിൽ "സ്വയംഭരണം" എന്ന പദം ഒരു പ്രത്യേക അർത്ഥം സ്വീകരിക്കുന്നു. കൃത്രിമ ബുദ്ധി സംവിധാനങ്ങൾക്ക് ഒരു സാങ്കേതിക ശേഷിയിൽ സ്വതന്ത്രമായി പ്രവർത്തിക്കാൻ കഴിയുമെങ്കിലും, തുടർച്ചയായ മനുഷ്യ ഇടപെടലുകളില്ലാതെ ജോലികൾ നിർവ്വഹിക്കാൻ കഴിയുമെങ്കിലും, അവയ്ക്ക് ധാർമ്മിക സ്വയംഭരണമില്ല. കൃത്രിമ ബുദ്ധി സംവിധാനങ്ങൾക്ക് ബോധം, സ്വയം അവബോധം, ധാർമ്മിക യുക്തിയിൽ ഏർപ്പെടാനുള്ള കഴിവ് എന്നിവയില്ല. കൃത്രിമ ബുദ്ധി സംവിധാനങ്ങൾ സ്വയം തിരഞ്ഞെടുത്ത തത്വങ്ങളെ അടിസ്ഥാനമാക്കിയല്ല അവരുടെ തീരുമാനങ്ങൾ എടുക്കുന്നത്, പകരം, അവ ആഴത്തിൽ വേരൂന്നിയ പാറ്റേണുകളെയും മാനദണ്ഡങ്ങളെയും ആശ്രയിക്കുന്നു. തൽഫലമായി, കൃത്രിമബുദ്ധിക്ക് പ്രവർത്തനപരമായ സ്വയംഭരണാവകാശം ഉണ്ടായിരിക്കാമെങ്കിലും അതിന് ധാർമ്മിക സ്വയംഭരണാവകാശമില്ല. കൃത്രിമബുദ്ധിയുടെ (AI) നടത്തിപ്പ് മനുഷ്യന്റെ മേൽനോട്ടത്തിൽ തന്നെ തുടരണം, ലക്ഷ്യങ്ങളും പരിമിതികളും നിർവ്വചിക്കുന്ന ഡിസൈനർമാർ, ഡെവലപ്പർമാർ, ഉപയോക്താക്കൾ എന്നിവരെ ഇത് ഉൾക്കൊള്ളുന്നു.

ഉദ്ദേശ്യശുദ്ധി:ഉദ്ദേശ്യശുദ്ധി: മനസ്സിന്റെയും ധാർമ്മികതയുടെയും തത്ത്വചിന്തയിൽ, ഒരു വസ്തുവിനെ പരാമർശിക്കാനോ അതിലേക്ക് നയിക്കാനോ ഉള്ള മാനസികാവസ്ഥകളുടെ കഴിവിനെ സൂചിപ്പിക്കുന്ന ഉദ്ദേശ്യശുദ്ധി എന്ന ആശയം ഏറ്റവും അത്യാവശ്യമായ ആശയങ്ങളിൽ ഒന്നായി കണക്കാക്കപ്പെടുന്നു. വിശ്വാസങ്ങൾ, ആഗ്രഹങ്ങൾ, ഉദ്ദേശ്യങ്ങൾ, ലക്ഷ്യങ്ങൾ എന്നിവ ലോകത്തിലെ അസ്തിത്വങ്ങളെയോ സംഭവങ്ങളെയോ അനന്തരഫലങ്ങളെയോ സൂചിപ്പിക്കുന്നതോ സൂചിപ്പിക്കുന്നതോ ആയ ഉദ്ദേശ്യശുദ്ധിയെ ഉദാഹരണമാക്കുന്നു. ഇതിന്റെ ഒരു ഉദാഹരണം, ഒരു വ്യക്തി മറ്റൊരാളെ സഹായിക്കാൻ ഉദ്ദേശിക്കുന്നു എന്നതാണ്, അവരുടെ പ്രവർത്തനങ്ങൾ ഒരു ആന്തരിക ലക്ഷ്യത്താലും അവരുടെ പെരുമാറ്റത്തെക്കുറിച്ചുള്ള ധാരണയാലും നയിക്കപ്പെടുന്നു. ധാർമ്മിക ഉത്തരവാദിത്തം നിർവ്വചിക്കുന്നതിൽ ഉദ്ദേശ്യശുദ്ധി എന്ന ആശയം നിർണ്ണായകമാണ്. ബോധപൂർവ്വമായ പ്രവർത്തനങ്ങളുടെ വിലയിരുത്തൽ അശ്രദ്ധമായോ അബോധാവസ്ഥയിലോ ചെയ്യുന്ന പ്രവർത്തനങ്ങളുടെ വിലയിരുത്തലിൽ നിന്ന് വ്യത്യസ്തമാണ്. ഒരു പ്രവൃത്തിയുടെ ഫലം ധാർമ്മിക പ്രശംസയുടെയോ അപലപത്തിന്റെയോ ഉചിതത്വത്തെ സ്വാധീനിക്കുന്നു, എന്നിരുന്നാലും ഏജൻറിന്റെ ഉദ്ദേശ്യങ്ങളും ഗണ്യമായി സംഭാവന ചെയ്യുന്നു; ഉദാഹരണത്തിന്, മനുഷ്യർ മറ്റൊരാളെ ദ്രോഹിക്കുന്ന ഒരു വ്യക്തിയെ സാധാരണയായി കൂടുതൽ കഠിനമായി വിധിക്കുന്നു.

സമാനമായ ഫലങ്ങൾ ഉണ്ടായിരുന്നിട്ടും, ആകസ്മികമായി ദോഷം വരുത്തുന്നു. ധാർമ്മികതയുടെ കാര്യത്തിൽ, മനുഷ്യർ മറ്റൊരു വ്യക്തിക്ക് ദോഷം വരുത്തുന്നത്, ഫലങ്ങൾ ഒന്നുതന്നെയാണെങ്കിൽ പോലും, അശ്രദ്ധമായി ദോഷം വരുത്തുന്നതിനേക്കാൾ അധാർമ്മികമായി കണക്കാക്കപ്പെടുന്നു. ബോധവും ഏജൻസിയും ഉദ്ദേശ്യശുദ്ധിയുടെ ആശയവുമായി അടുത്ത ബന്ധപ്പെട്ടിരിക്കുന്നു. പ്രവൃത്തികൾക്ക് ന്യായീകരണങ്ങൾ രൂപപ്പെടുത്താനും ബോധം ഉണ്ടായിരിക്കാനും മനസ്സിലാക്കാനും ഉള്ള കഴിവ് ആധികാരിക ഉദ്ദേശ്യശുദ്ധിയുടെ അവശ്യ ഘടകങ്ങളാണ്. തങ്ങളുടെ ലക്ഷ്യങ്ങളെക്കുറിച്ച് ചിന്തിക്കാനും, അവരുടെ പ്രവൃത്തികൾക്കുള്ള ധാർമ്മിക യുക്തികൾ വിലയിരുത്താനും, ഒന്നിലധികം ബാലുകളിൽ നിന്ന് തിരഞ്ഞെടുക്കാനുമുള്ള കഴിവ് കാരണം മനുഷ്യർക്ക് ഉദ്ദേശ്യശുദ്ധിയുണ്ട്. ഈ ശക്തി കാരണം, അവരുടെ പ്രവർത്തനങ്ങൾക്ക് അവരെ ധാർമ്മികമായി ഉത്തരവാദിത്തപ്പെടുത്താൻ കഴിയും. നേരെമറിച്ച്, കൃത്രിമബുദ്ധി ഉപയോഗിക്കുന്ന പ്രോഗ്രാമുകൾക്ക് ആധികാരിക ഉദ്ദേശ്യശുദ്ധിയില്ല. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ ലക്ഷ്യബോധമുള്ളതായി കാണപ്പെടുന്നുണ്ടെങ്കിലും, അവയുടെ "ഉദ്ദേശ്യങ്ങൾ" അടിസ്ഥാനപരമായി അൽഗോരിതങ്ങൾ, പരിശീലന ഡാറ്റ, മനുഷ്യർ സ്ഥാപിച്ച ലക്ഷ്യങ്ങൾ എന്നിവയിൽ വേരൂന്നിയതാണ്. കൃത്രിമബുദ്ധിക്ക് അതിന്റെ ലക്ഷ്യങ്ങളെക്കുറിച്ച് മനസ്സിലാക്കലോ അംഗീകാരമോ ഇല്ല; അത് പിന്തുടരാൻ പരിശീലിപ്പിച്ച നിർദ്ദേശങ്ങൾ അത് നടപ്പിലാക്കുന്നു. അതിനാൽ, യഥാർത്ഥ ഉദ്ദേശ്യശുദ്ധിയെക്കാൾ, കൃത്രിമബുദ്ധി എന്ന നിലയിൽ അല്ലെങ്കിൽ ഉരുത്തിരിഞ്ഞ ഉദ്ദേശ്യശുദ്ധി മാത്രമേ പ്രകടിപ്പിക്കുന്നുള്ളൂ. കാരണം

അതിന് യഥാർത്ഥ ഉദ്ദേശ്യങ്ങളോ ആഗ്രഹങ്ങളോ രൂപപ്പെടുത്താനുള്ള ബോധമോ സ്വയം അവബോധമോ ഇല്ല.

ധാർമ്മികതയുടെ യുക്തിയും ധാരണയും ധാർമ്മികതയുടെ മേഖലയിൽ, ധാർമ്മിക യുക്തി അല്ലെങ്കിൽ ധാർമ്മിക ധാരണ എന്നത് ധാർമ്മിക പ്രതിസന്ധികളെ തിരിച്ചറിയാനും, ശരിയും തെറ്റും എന്താണെന്ന് കണ്ടെത്താനും, ധാർമ്മിക തത്വങ്ങൾ പ്രയോഗിച്ചുകൊണ്ട് പ്രവർത്തനങ്ങളെ ന്യായീകരിക്കാനുമുള്ള കഴിവിനെ സൂചിപ്പിക്കുന്നു. നീതി, ന്യായം, അവകാശങ്ങൾ, കടമകൾ, ദോഷം തുടങ്ങിയ ധാർമ്മിക ആശയങ്ങൾ മനസ്സിലാക്കുകയും തുടർന്ന് യഥാർത്ഥ ജീവിത സാഹചര്യങ്ങളിൽ ഈ തത്വങ്ങൾ പ്രയോഗിക്കുകയും ചെയ്യുന്നത് ഈ പ്രക്രിയയിൽ ഉൾപ്പെടുന്നു.

ധാർമ്മിക യുക്തിയിൽ ഏർപ്പെടുമ്പോൾ വ്യക്തികൾക്ക് അവരുടെ പ്രവൃത്തികളെക്കുറിച്ച് ചിന്തിക്കാനും ധാർമ്മികമായി ഉചിതമായ പെരുമാറ്റങ്ങൾ തിരഞ്ഞെടുക്കാനും കഴിയും. ധാർമ്മികതയും ധാർമ്മിക യുക്തിയും ബോധം, ഉദ്ദേശ്യശ്രദ്ധ, സ്വയംഭരണം എന്നീ ആശയങ്ങളെ അടുത്ത് ബന്ധിപ്പിക്കുന്നു. ധാർമ്മിക ചിന്തയിൽ പ്രാവീണ്യമുള്ള ഒരു വ്യക്തിക്ക് പ്രവൃത്തികളുടെ ന്യായീകരണം മനസ്സിലാക്കാനും, തീരുമാനങ്ങളുടെ ധാർമ്മിക അനന്തരഫലങ്ങൾ മുൻകൂട്ടി കാണാനും, ആ തിരഞ്ഞെടുപ്പുകൾ മറ്റുള്ളവരിൽ ചെലുത്തുന്ന സ്വാധീനം അംഗീകരിക്കാനും കഴിയും. സാമൂഹിക ഇടപെടൽ, വിദ്യാഭ്യാസം, സഹാനുഭൂതി, സാംസ്കാരിക മാനദണ്ഡങ്ങൾ എന്നിവയിലൂടെ, വ്യക്തികൾക്ക് അവരുടെ പ്രവൃത്തികളെക്കുറിച്ച് കുറേബോധം, ദുഃഖം അല്ലെങ്കിൽ അഭിമാനം എന്നിവ തോന്നാൻ അനുവദിക്കുന്ന ഒരു ധാർമ്മിക ധാരണ വികസിപ്പിക്കാൻ കഴിയും. ധാർമ്മിക യുക്തിക്കുള്ളിലെ ധാർമ്മിക ഉത്തരവാദിത്തത്തിന്റെ ഒരു അനിവാര്യ ഘടകമാണ് ധാർമ്മിക യുക്തി. ഒരു ഏജൻസി പ്രശംസയോ വിമർശനമോ ലഭിക്കണമെങ്കിൽ, അവരുടെ പ്രവൃത്തികളുടെ ധാർമ്മിക പ്രത്യാഘാതങ്ങൾ മനസ്സിലാക്കേണ്ടത് അത്യാവശ്യമാണ്. മനുഷ്യവുമായ വഞ്ചന അല്ലെങ്കിൽ വിവേചനം പോലുള്ള പ്രവർത്തനങ്ങൾ ധാർമ്മികമായി അപലപനീയമാണ്, കാരണം ഈ പെരുമാറ്റങ്ങൾ ധാർമ്മിക മാനദണ്ഡങ്ങൾ ലംഘിക്കുന്നുവെന്ന് വ്യക്തി തിരിച്ചറിയുന്നു. നിലവിൽ നിലവിലുള്ള കൃത്രിമ ബുദ്ധി സംവിധാനങ്ങൾക്ക് യഥാർത്ഥ ധാർമ്മിക യുക്തിയോ ധാരണയോ ഇല്ല, അതായത് അവർക്ക് ശരിയും തെറ്റും വേർതിരിച്ചറിയാനോ ബോധപൂർവ്വമായ വഞ്ചന അല്ലെങ്കിൽ വിവേചനം പോലുള്ള പ്രവർത്തനങ്ങളുടെ ധാർമ്മിക പ്രത്യാഘാതങ്ങൾ തിരിച്ചറിയാനോ കഴിയില്ല. കൃത്രിമ ബുദ്ധിയുടെ മേഖലയിൽ ഇത് ശരിയാണ്. കൃത്രിമ ബുദ്ധിക്ക് (AI) ഒരു യഥാർത്ഥ ഇല്ല.

ധാർമ്മിക സങ്കല്പങ്ങളെക്കുറിച്ചുള്ള ധാരണ; എന്നിരുന്നാലും, ധാർമ്മിക മാനദണ്ഡങ്ങൾ പാലിക്കുന്നതിനോ ഫലങ്ങൾ മെച്ചപ്പെടുത്തുന്നതിനോ ഇത് പ്രോഗ്രാം ചെയ്യാൻ കഴിയും. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് അവബോധമില്ല, പ്രവർത്തനങ്ങളുടെ ധാർമ്മിക പ്രത്യാഘാതങ്ങൾ ഗ്രഹിക്കാൻ കഴിയില്ല; അവ ഡാറ്റ വിലയിരുത്തുകയും അൽഗോരിതങ്ങൾ നടപ്പിലാക്കുകയും ചെയ്യുന്നു. തൽഫലമായി, അവരുടെ "ധാർമ്മിക പ്രവർത്തനം" യുക്തിയുടെ യഥാർത്ഥ പ്രദർശനത്തേക്കാൾ വെറും ഒരു സിമുലേഷൻ മാത്രമാണ്. ധാർമ്മികതയുടെ സൈദ്ധാന്തിക ആശയങ്ങൾ മനസ്സിലാക്കാനും, ശരിയും തെറ്റും തമ്മിൽ വേർതിരിച്ചറിയാനും, ധാർമ്മിക പ്രശ്നങ്ങൾ ചിന്തിക്കാനുമുള്ള കഴിവ്.

ബോധം/ബോധം:വേദന, ആനന്ദം, വികാരങ്ങൾ അല്ലെങ്കിൽ വികാരങ്ങൾ എന്നിവ മനസ്സിലാക്കാനുള്ള കഴിവ് ഉൾപ്പെടെയുള്ള ആത്മനിഷ്ഠമായ അനുഭവങ്ങൾ അനുഭവിക്കാനുള്ള കഴിവാണെന്ന് സെൻസിറ്റീവ് എന്ന് നിർവചിച്ചിരിക്കുന്നത്. തന്നെയും പരിസ്ഥിതിയെയും കുറിച്ച് ബോധവാന്മാരായിരിക്കുന്നതിന്റെ അവസ്ഥയാണ് ബോധം. സെൻസിറ്റീവ്സും അവബോധവുമാണ് ഒരു അസ്തിത്വമായിരിക്കുന്നതിന്റെ അനുഭവം എന്നറിയപ്പെടുന്ന ആദ്യ വ്യക്തി ആന്തരിക അനുഭവത്തിന്റെ അടിസ്ഥാനം കൂട്ടായി രൂപപ്പെടുത്തുന്ന രണ്ട് ഘടകങ്ങൾ. ഒരു ധാർമ്മിക സ്ഥാനം സ്ഥാപിക്കുന്നതിനാൽ ഈ കഴിവുകൾ അവരുടെ ഗവേഷണത്തിന് അത്യന്താപേക്ഷിതമാണ്. ബോധമുള്ളതും ബുദ്ധിയുള്ളതുമായ ജീവികൾക്ക് ദുരിതം സഹിക്കാനും ക്ഷേമം അനുഭവിക്കാനും കഴിയും, അതുവഴി അവരുടെ താൽപ്പര്യങ്ങൾ പരിഗണിക്കുന്നതിന് ധാർമ്മിക ന്യായീകരണം നൽകുന്നു. വേദനയും ആനന്ദവും അനുഭവിക്കാനുള്ള കഴിവ് കാരണം മനുഷ്യരും നിരവധി മൃഗങ്ങളും ധാർമ്മികമായി പ്രാധാന്യമുള്ളവരായി കണക്കാക്കപ്പെടുന്നു, ഇത് ധാർമ്മിക തീരുമാനമെടുക്കലിൽ അവരുടെ ക്ഷേമം പരിഗണിക്കാനുള്ള ധാർമ്മിക ബാധ്യതയെ അടിവരയിടുന്നു.

ധാർമ്മിക ഉത്തരവാദിത്തത്തിന്റെ ഒരു നിർണായക ഘടകം അവബോധത്തിന്റെയും സംവേദനക്ഷമതയുടെയും നിലനിൽപ്പാണ്. ഒരു ഏജൻസി തന്റെ പ്രവർത്തനങ്ങളെക്കുറിച്ചുള്ള അവബോധവും അവയുടെ പ്രാധാന്യം മനസ്സിലാക്കാനുള്ള കഴിവും ഉണ്ടെങ്കിൽ മാത്രമേ ധാർമ്മികമായി ഉത്തരവാദിത്തം വഹിക്കാൻ കഴിയൂ. ബോധമില്ലാതെ, പ്രവർത്തനങ്ങൾ ലക്ഷ്യബോധമുള്ളതായിരിക്കുന്നതിനപകരം യാത്രികമായി മാറുന്നു, ധാർമ്മിക പ്രശംസയോ അപലപനമോ പോലുള്ള തൊഴിലുകൾക്ക് അവയുടെ പ്രസക്തി നഷ്ടപ്പെടുന്നു, കാരണം

അവയ്ക്ക് ധാർമ്മിക പ്രത്യാഘാതങ്ങളെക്കുറിച്ചുള്ള ധാരണയും ധാർമ്മിക യുക്തിയെ അടിസ്ഥാനമാക്കി തിരഞ്ഞെടുപ്പുകൾ നടത്താനുള്ള കഴിവും ആവശ്യമാണ്.

നിലവിലുള്ള കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ പലപ്പോഴും ബോധമില്ലാത്തതും വികാരരഹിതവുമാണെന്ന് കണക്കാക്കപ്പെടുന്നു. കൃത്രിമബുദ്ധിക്ക് വിവരങ്ങൾ വിശകലനം ചെയ്യാനും പാറ്റേണുകൾ തിരിച്ചറിയാനും മനുഷ്യന്റെ പെരുമാറ്റത്തിന് സമാനമായ പ്രതികരണങ്ങൾ സൃഷ്ടിക്കാനും കഴിയുമെങ്കിലും, അതിന് ആത്മനിഷ്ഠമായ അനുഭവവും വികാരങ്ങളും ഇല്ല. സന്തോഷം, വേദന, പശ്ചാത്താപം അല്ലെങ്കിൽ സഹാനുഭൂതി എന്നിവയുൾപ്പെടെയുള്ള വികാരങ്ങൾ അനുഭവിക്കാനുള്ള കഴിവ് കൃത്രിമബുദ്ധിക്കില്ല. തൽഫലമായി, കൃത്രിമബുദ്ധിക്ക് ധാർമ്മിക ദ്രോഹത്തിന് കഴിവില്ല, ധാർമ്മിക അവകാശങ്ങളും ഇല്ല. കൃത്രിമ പൊതുബുദ്ധിക്ക് ബോധമോ വികാരമോ ഉണ്ടാകുമോ എന്ന ചോദ്യം ഇപ്പോഴും ദാർശനിക ചർച്ചാവിഷയമാണ്. ഭാവിയിൽ കൃത്രിമബുദ്ധിയിൽ യഥാർത്ഥ അവബോധമോ സംവേദനക്ഷമതയോ ഉയർന്നുവരുന്നത് സാങ്കേതികവിദ്യയുടെ ധാർമ്മിക നില, അവകാശങ്ങൾ, ഉത്തരവാദിത്തങ്ങൾ എന്നിവയെക്കുറിച്ചുള്ള ഗണ്യമായ നൈതിക പ്രശ്നങ്ങൾ ഉയർത്തും. എന്നിരുന്നാലും, ഇത് ഒരു വിഷയമായി തുടരുന്നു. ഇന്നത്തെ ഊഹാപോഹങ്ങൾ. മനുഷ്യർ ഈ സ്വഭാവവിശേഷങ്ങൾ പ്രകടിപ്പിക്കുന്നത് നമ്മുടെ ബോധം, വികാരങ്ങൾ, ധാർമ്മിക യുക്തിസഹമായ കഴിവുകൾ എന്നിവ മൂലമാണ്.

നിലവിലുള്ള കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ ധാർമ്മിക ഏജന്റുമാരായി യോഗ്യത നേടുന്നില്ല. സമകാലിക AI സംവിധാനങ്ങൾ, അവയുടെ ബുദ്ധിശക്തി എത്രതന്നെയായാലും, അവബോധവും ധാർമ്മിക ഗ്രാഹ്യവും ഇല്ലാത്തവയാണ്. ആധികാരികമായ ധാർമ്മിക ചിന്തയിലൂടെയല്ല, അൽഗോരിതങ്ങളിലൂടെയും ഡാറ്റ പാറ്റേണുകളിലൂടെയുമാണ് അവ പ്രവർത്തിക്കുന്നത്. AI, ധാർമ്മിക ചിന്തയെ പകർത്തിയേക്കാം (ഉദാഹരണത്തിന്, ഒരു പ്രശ്നത്തിൽ "ഏറ്റവും ധാർമ്മിക" ബദൽ തിരഞ്ഞെടുക്കൽ), എന്നിരുന്നാലും അതിന്റെ തീരുമാനങ്ങളുടെ ധാർമ്മിക പ്രാധാന്യത്തെക്കുറിച്ച് അതിന് ഗ്രാഹ്യമില്ല. അതിനാൽ, മനുഷ്യ സാഹചര്യത്തിൽ അതിനെ ഒരു യഥാർത്ഥ ധാർമ്മിക ഏജന്റായി കണക്കാക്കാനാവില്ല.

ബോധത്തിന്റെയും ഉദ്ദേശ്യത്തിന്റെയും അഭാവം: ബോധവും ഉദ്ദേശ്യശൂന്യതയും ധാർമ്മിക തത്ത്വചിന്തയുടെ അടിസ്ഥാനപരവും ധാർമ്മിക ഏജൻസിയും ഉത്തരവാദിത്തവും സ്ഥാപിക്കുന്നതിന് അത്യന്താപേക്ഷിതവുമാണ്. ബോധം എന്നത് ഒരു അസ്തിത്വത്തിന്റെ സ്വയം അവബോധത്തെയും പരിസ്ഥിതിയെക്കുറിച്ചുള്ള അവബോധത്തെയും വികാരങ്ങൾ, സംവേദനങ്ങൾ, വികാരങ്ങൾ എന്നിവയുൾപ്പെടെയുള്ള ആത്മനിഷ്ഠമായ അനുഭവങ്ങൾക്കുള്ള കഴിവിനെയും സൂചിപ്പിക്കുന്നു. ഉദ്ദേശ്യശൂന്യത എന്നത് ലക്ഷ്യങ്ങൾ രൂപപ്പെടുത്താനും ലക്ഷ്യങ്ങൾ, വസ്തുക്കൾ അല്ലെങ്കിൽ അനന്തരഫലങ്ങൾ എന്നിവയിലേക്ക് നയിക്കപ്പെടുന്ന മാനസികാവസ്ഥകൾ കൈവശം വയ്ക്കാനുമുള്ള കഴിവിനെ സൂചിപ്പിക്കുന്നു.

സമകാലിക കൃത്രിമബുദ്ധി (AI) സംവിധാനങ്ങൾ ബോധവും ആധികാരികമായ ഉദ്ദേശ്യശൂന്യതയും ഇല്ലാത്തവയാണ്. AI-ക്ക് വിശാലമായ ഡാറ്റാസെറ്റുകൾ വിശകലനം ചെയ്യാനും പാറ്റേണുകൾ തിരിച്ചറിയാനും വിധിന്യായങ്ങൾ നടപ്പിലാക്കാനും കഴിയുമെങ്കിലും, ഈ പ്രവർത്തനങ്ങൾ ബോധമോ ആത്മനിഷ്ഠമായ അനുഭവമോ ഇല്ലാതെയാണ് സംഭവിക്കുന്നത്. വേദന, ആനന്ദം, പശ്ചാത്താപം അല്ലെങ്കിൽ സഹാനുഭൂതി എന്നിവ അനുഭവിക്കാനുള്ള കഴിവ് AI-ക്ക് ഇല്ല, മാത്രമല്ല അതിന്റെ പ്രവർത്തനങ്ങളുടെ അർത്ഥമോ ധാർമ്മിക പ്രത്യാഘാതങ്ങളോ അത് മനസ്സിലാക്കുന്നില്ല. അതിന്റെ പ്രവർത്തനങ്ങൾ കമ്പ്യൂട്ടേഷണൽ മാത്രമാണ്.

അതുപോലെ, AI-ക്ക് യഥാർത്ഥ ഉദ്ദേശ്യശൂന്യത ഇല്ല. AI സിസ്റ്റങ്ങൾ ഉദ്ദേശ്യത്തോടെ പ്രവർത്തിക്കുന്നവയാണെന്ന് തോന്നുമെങ്കിലും, അവയുടെ ലക്ഷ്യങ്ങൾ ബാഹ്യമായി നിർണ്ണയിക്കുന്നത് മനുഷ്യ സ്രഷ്ടാക്കളാണ്. AI-ക്ക് അതിന്റെ ലക്ഷ്യങ്ങൾ തിരഞ്ഞെടുക്കാനോ അവ പിന്തുടരുന്നതിനുള്ള യുക്തി മനസ്സിലാക്കാനോ ഉള്ള ശേഷിയില്ല; അത് പ്രോഗ്രാം ചെയ്ത നിർദ്ദേശങ്ങൾ പാലിക്കുകയോ ഗണിതശാസ്ത്ര പ്രവർത്തനങ്ങൾ ഒപ്റ്റിമൈസ് ചെയ്യുകയോ ചെയ്യുന്നു. ഈ സ്വഭാവം പലപ്പോഴും

"ആധികാരിക ഉദ്ദേശ്യശൂന്യത" എന്നതിലുപരി "ഉദ്ദേശ്യബോധം പോലെ" എന്ന് വിശേഷിപ്പിക്കപ്പെടുന്നു. AI-യുടെ ബോധത്തിന്റെയും ഉദ്ദേശ്യശൂന്യതയുടെയും അഭാവം കാരണം, അതിനെ ഒരു ധാർമ്മിക ഏജന്റായി കണക്കാക്കാനാവില്ല. ധാർമ്മിക ഉത്തരവാദിത്തത്തിന് ഒരാളുടെ പ്രവൃത്തികളെയും അവയ്ക്ക് അടിസ്ഥാനമായ ഉദ്ദേശ്യങ്ങളെയും കുറിച്ചുള്ള അറിവ് ആവശ്യമാണ്. ഈ കഴിവുകളുടെ അഭാവത്തിൽ, AI-യെ അഭിനയിക്കാനോ വിമർശിക്കാനോ അതിന്റെ ഫലങ്ങൾക്ക് ധാർമ്മികമായി ഉത്തരവാദിയാക്കാനോ കഴിയില്ല. തൽഫലമായി, പക്ഷപാതം, വിവേചനം അല്ലെങ്കിൽ അവകാശങ്ങളുടെ ദുരുപയോഗം പോലുള്ള AI സംവിധാനങ്ങൾ വരുത്തുന്ന

ഏതൊരു നാശനഷ്ടവും അവയുടെ രൂപകൽപ്പന, നടപ്പാക്കൽ അല്ലെങ്കിൽ മേൽനോട്ടം എന്നിവയ്ക്ക് ഉത്തരവാദികളായ മനുഷ്യ ഏജൻ്റുമാരിൽ ആരോപിക്കപ്പെടണം.

AI-യിലെ അവബോധത്തിന്റെയും ഉദ്ദേശ്യശുദ്ധിയുടെയും അഭാവം, AI സംവിധാനങ്ങളെ ധാർമ്മിക ഏജൻ്റുമാരായി കണക്കാക്കുന്നതിനപകരം ഉപകരണങ്ങളായി കൂടുതൽ കൃത്യമായി കണക്കാക്കുന്നത് എന്തുകൊണ്ടെന്ന് വ്യക്തമാക്കുന്നു. AI-യുടെ ഭാവി ആവർത്തനങ്ങൾ ഈ അടിസ്ഥാന ഗുണങ്ങൾ നേടുന്നതുവരെ, AI പെരുമാറ്റത്തിനായുള്ള ധാർമ്മികവും നിയമപരവുമായ ഉത്തരവാദിത്തം ആളുകളിലും സ്ഥാപനങ്ങളിലും അസന്ദിഗ്ധമായി നിലനിൽക്കണം.

സമകാലിക കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ, പ്രത്യേകിച്ച് പുരോഗമിച്ചവ, പൊതുവെ യഥാർത്ഥ ബോധമോ, ആത്മനിഷ്ഠമായ അനുഭവമോ, മനുഷ്യരുടേതുപോലുള്ള ഉദ്ദേശ്യങ്ങൾ രൂപപ്പെടുത്താനുള്ള ശേഷിയോ ഇല്ലാത്തവയായി കണക്കാക്കപ്പെടുന്നു. പ്രോഗ്രാമിംഗ് നിർദ്ദേശങ്ങൾ, അൽഗോരിതങ്ങൾ, പരിശീലന ഡാറ്റ എന്നിവയാണ് അവയുടെ പ്രവർത്തനങ്ങളെ നിയന്ത്രിക്കുന്ന ഘടകങ്ങൾ.

പ്രവർത്തനപരമായ അനുകരണം യഥാർത്ഥ ധാരണയ്ക്ക് ധാർമ്മിക പ്രതികരണങ്ങൾ സൃഷ്ടിച്ചുകൊണ്ടോ മുൻകൂട്ടി പ്രോഗ്രാം ചെയ്ത ദോഷകരമായ പ്രവർത്തനങ്ങൾ ഒഴിവാക്കിക്കൊണ്ടോ കൃത്രിമബുദ്ധിക്ക് ധാർമ്മിക പെരുമാറ്റത്തെ അനുകരിക്കാൻ കഴിയുമെങ്കിലും, ഇത് യഥാർത്ഥ ധാർമ്മിക അവബോധമോ ഏജൻസിയോ അല്ല, മറിച്ച് പ്രവർത്തനപരമായ തുല്യതയായി കണക്കാക്കപ്പെടുന്നു. അവ ധാർമ്മിക യുക്തിയിലല്ല, മറിച്ച് ഒരു നിശ്ചിത കോഡിലാണ് പ്രവർത്തിക്കുന്നത്, അതായത് അവയുടെ പ്രതികരണങ്ങൾ മുൻകൂട്ടി നിശ്ചയിച്ചിട്ടുള്ളതും അവയുടെ പ്രവർത്തനങ്ങളുടെ ധാർമ്മിക പ്രത്യാഘാതങ്ങൾ മനസ്സിലാക്കാനുള്ള കഴിവില്ലാത്തതുമാണ്.

സമകാലിക കൃത്രിമബുദ്ധി (AI) സംവിധാനങ്ങൾ ബോധവും ആധികാരികമായ ഉദ്ദേശ്യശുദ്ധിയും ഇല്ലാത്തവയാണ്. AI-ക്ക് വിപുലമായ ഡാറ്റാ പ്രോസസ്സ് ചെയ്യാനും പാറ്റേണുകൾ തിരിച്ചറിയാനും വിധിനിയമങ്ങൾ എടുക്കാനും കഴിയുമെങ്കിലും, ഈ പ്രവർത്തനങ്ങൾ ബോധമോ ആത്മനിഷ്ഠമായ അനുഭവമോ ഇല്ലാത്തതാണ് സംഭവിക്കുന്നത്. വേദന, ആനന്ദം, പശ്ചാത്താപം അല്ലെങ്കിൽ സഹാനുഭൂതി എന്നിവ അനുഭവിക്കാനുള്ള കഴിവ് AI-ക്ക് ഇല്ല, കൂടാതെ അതിന്റെ പ്രവർത്തനങ്ങളുടെ അർത്ഥമോ ധാർമ്മിക പ്രത്യാഘാതങ്ങളോ അത് മനസ്സിലാക്കുന്നില്ല. അതിന്റെ പ്രവർത്തനങ്ങൾ കമ്പ്യൂട്ടേഷണൽ മാത്രമാണ്.

അതുപോലെ, AI-ക്ക് യഥാർത്ഥ ഉദ്ദേശ്യശുദ്ധി ഇല്ല. AI സിസ്റ്റങ്ങൾ ഉദ്ദേശ്യത്തോടെ പ്രവർത്തിക്കുന്നവയാണെന്ന് തോന്നുമെങ്കിലും, അവയുടെ ലക്ഷ്യങ്ങൾ ബാഹ്യമായി നിർണ്ണയിക്കുന്നത് മനുഷ്യ സ്രഷ്ടാക്കളാണ്. കൃത്രിമബുദ്ധി അതിന്റെ ലക്ഷ്യങ്ങൾ തിരഞ്ഞെടുക്കുകയോ അവ പിന്തുടരുന്നതിനുള്ള യുക്തി മനസ്സിലാക്കുകയോ ചെയ്യുന്നില്ല; അത് പ്രോഗ്രാം ചെയ്ത നിർദ്ദേശങ്ങൾ പാലിക്കുകയോ ഗണിതശാസ്ത്ര പ്രവർത്തനങ്ങൾ ഒപ്റ്റിമൈസ് ചെയ്യുകയോ ചെയ്യുന്നു. ഈ സ്വഭാവത്തെ പലപ്പോഴും "ആധികാരിക ഉദ്ദേശ്യശുദ്ധി" എന്നതിലുപരി "ഉദ്ദേശ്യബോധം പോലെ" എന്ന് വിശേഷിപ്പിക്കുന്നു.

AI-യുടെ ബോധത്തിന്റെയും ഉദ്ദേശ്യശുദ്ധിയുടെയും അഭാവം കാരണം, അതിനെ ഒരു ധാർമ്മിക ഏജൻ്റായി കണക്കാക്കാനാവില്ല. ധാർമ്മിക ഉത്തരവാദിത്തത്തിന് ഒരാളുടെ പ്രവൃത്തികളെയും അവയ്ക്ക് അടിസ്ഥാനമായ പ്രചോദനങ്ങളെയും കറിച്ചുള്ള അറിവ് ആവശ്യമാണ്. ഈ കഴിവുകളുടെ അഭാവത്തിൽ, AI-യെ പ്രശംസിക്കാനോ വിമർശിക്കാനോ അതിന്റെ ഫലങ്ങൾക്ക് ധാർമ്മികമായി ഉത്തരവാദിയാക്കാനോ കഴിയില്ല. തൽഫലമായി, പക്ഷപാതം, വിവേചനം അല്ലെങ്കിൽ അവകാശങ്ങളുടെ ദുരുപയോഗം പോലുള്ള AI സംവിധാനങ്ങൾ വരുത്തുന്ന ഏതൊരു നാശനഷ്ടത്തിനും അവയുടെ രൂപകൽപ്പന, നടപ്പാക്കൽ അല്ലെങ്കിൽ മേൽനോട്ടം എന്നിവയ്ക്ക് ഉത്തരവാദികളായ മനുഷ്യ ഏജൻ്റുമാരാണ് ഉത്തരവാദികൾ.

AI-യിലെ അവബോധത്തിന്റെയും ഉദ്ദേശ്യശുദ്ധിയുടെയും അഭാവം, AI സിസ്റ്റങ്ങളെ ധാർമ്മിക ഏജൻ്റുമാരായി കണക്കാക്കുന്നതിനപകരം ഉപകരണങ്ങളായി കൂടുതൽ കൃത്യമായി കണക്കാക്കുന്നത് എന്തുകൊണ്ടെന്ന് വ്യക്തമാക്കുന്നു. AI-യുടെ ഭാവി ആവർത്തനങ്ങൾ ഈ അവശ്യ ഗുണങ്ങൾ നേടുന്നതുവരെയും, അല്ലെങ്കിൽ അത് നേടിയെടുക്കുന്നതുവരെയും, AI പെരുമാറ്റത്തിനായുള്ള ധാർമ്മികവും നിയമപരവുമായ ഉത്തരവാദിത്തം മനുഷ്യരിലും സ്ഥാപനങ്ങളിലും നിസ്സംശയമായും നിലനിൽക്കണം.

സമകാലിക കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ, പ്രത്യേകിച്ച് പുരോഗമിച്ചവ, പൊതുവെ യഥാർത്ഥ ബോധമോ, ആത്മനിഷ്ഠമായ അനുഭവമോ, മനുഷ്യരുടേതുപോലുള്ള ഉദ്ദേശ്യങ്ങൾ രൂപപ്പെടുത്താനുള്ള ശേഷിയോ ഇല്ലാത്തവയായി കണക്കാക്കപ്പെടുന്നു. പ്രോഗ്രാമിംഗ് നിർദ്ദേശങ്ങൾ, അൽഗോരിതങ്ങൾ, പരിശീലന ഡാറ്റ എന്നിവയാണ് അവയുടെ പ്രവർത്തനങ്ങളെ നിയന്ത്രിക്കുന്ന ഘടകങ്ങൾ.

പ്രവർത്തനപരമായ അനുകരണം vs. യഥാർത്ഥ ധാരണ: ധാർമ്മികമായി മികച്ച പ്രതികരണങ്ങൾ സൃഷ്ടിച്ചുകൊണ്ടോ മുൻകൂട്ടി പ്രോഗ്രാം ചെയ്ത ദോഷകരമായ പ്രവർത്തനങ്ങൾ ഒഴിവാക്കിക്കൊണ്ടോ കൃത്രിമബുദ്ധിക്ക് ധാർമ്മിക പെരുമാറ്റം ആവർത്തിക്കാൻ കഴിയുമെങ്കിലും, ഇത് ആധികാരിക ധാർമ്മിക ബോധത്തെയോ ഏജൻസിയെയോക്കാൾ പ്രവർത്തനപരമായ തുല്യതയെ പ്രതിനിധീകരിക്കുന്നു. അവ ധാർമ്മിക യുക്തിയിലല്ല, മറിച്ച് ഒരു ക്രോഡീകരിച്ച സംവിധാനത്തിലാണ് പ്രവർത്തിക്കുന്നത്, അതായത്

യഥാർത്ഥ ധാർമ്മിക ചർച്ചയിലോ ധാരണയിലോ ഏർപ്പെടുന്നതിനുപകരം അവർ മുൻകൂട്ടി നിശ്ചയിച്ച നിയമങ്ങളും അൽഗോരിതങ്ങളും പിന്തുടരുന്നു.

സ്വതന്ത്ര ഇച്ഛാശക്തിയുടെ അഭാവം:സ്വതന്ത്ര ഇച്ഛാശക്തി എന്നത് ഒരു ഏജൻസി നിരവധി ഓപ്ഷനുകൾക്കിടയിൽ സ്വതന്ത്രമായി തിരഞ്ഞെടുപ്പുകൾ നടത്താനും, സ്വമേധയാ പ്രവർത്തിക്കാനും, ഒരാളുടെ തീരുമാനങ്ങളിൽ നിയന്ത്രണം ചെലുത്താനുമുള്ള കഴിവാണു്. സ്വതന്ത്ര ഇച്ഛാശക്തി എന്ന ആശയം സമഗ്രമായി പരിശോധിക്കുകയും വ്യക്തമാക്കുകയും ചെയ്തിട്ടുണ്ട്. ധാർമ്മിക സിദ്ധാന്തത്തിനുള്ളിൽ സ്വതന്ത്ര ഇച്ഛാശക്തി ധാർമ്മിക ഉത്തരവാദിത്തത്തിനും ഇടയിൽ ശക്തമായ ഒരു ബന്ധം നിലനിൽക്കുന്നു. വ്യത്യസ്തമായി പ്രവർത്തിക്കാനുള്ള കഴിവുണ്ടെങ്കിൽ മാത്രമേ ഒരു നടനെ ധാർമ്മികമായി കുറ്റവാളിയായി കണക്കാക്കാൻ കഴിയൂ.

നിലവിൽ, കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് സ്വതന്ത്ര ഇച്ഛാശക്തിയില്ല. ഒരു സാങ്കേതിക സാഹചര്യത്തിൽ കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് സ്വയംഭരണാധികാരത്തോടെ പ്രവർത്തിക്കാൻ കഴിയുമെങ്കിലും, അവയുടെ പ്രവർത്തനങ്ങൾ പൂർണ്ണമായും നിയന്ത്രിക്കുന്നത് അൽഗോരിതങ്ങൾ, പരിശീലന ഡാറ്റ, മോഡൽ ഡിസൈൻ, ആളുകൾ നിശ്ചയിച്ച ലക്ഷ്യങ്ങൾ എന്നിവയാണ്. കൃത്രിമബുദ്ധി സ്വയം തിരഞ്ഞെടുത്ത മൂല്യങ്ങളെയോ അതിന്റെ ലക്ഷ്യങ്ങളെയോ അടിസ്ഥാനമാക്കിയുള്ള പ്രവർത്തനങ്ങൾ ചെയ്യുന്നില്ല. മറിച്ച്, പ്രകടനം മെച്ചപ്പെടുത്തുന്ന സ്ഥാപിത മാനദണ്ഡങ്ങളോ നടപടിക്രമങ്ങളോ അനുസരിച്ചാണ് ഇത് പ്രവർത്തിക്കുന്നത്.

സ്വതന്ത്ര ഇച്ഛാശക്തിയില്ലാത്ത കൃത്രിമബുദ്ധിക്ക് യഥാർത്ഥത്തിൽ സ്വമേധയാ ഉള്ള തീരുമാനങ്ങൾ എടുക്കാൻ കഴിയില്ല. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് ധാർമ്മികമായ ആലോചനയില്ല, മൂല്യമുള്ള കാര്യങ്ങളായി കാരണങ്ങളെ അവഗണിക്കുന്നു, ധാർമ്മിക യുക്തിയെ അടിസ്ഥാനമാക്കിയുള്ള അവയുടെ പ്രവർത്തനങ്ങളിൽ മാറ്റം വരുത്തുന്നില്ല. കൃത്രിമബുദ്ധി എടുക്കുന്നതായി തോന്നുന്ന ഏതൊരു തീരുമാനവും വാസ്തവത്തിൽ, സ്വതന്ത്ര ഇച്ഛാശക്തിയുടെ പ്രയോഗത്തേക്കാൾ കമ്പ്യൂട്ടേഷണൽ പ്രക്രിയകളുടെ ഫലമാണ്.

സ്വതന്ത്ര ഇച്ഛാശക്തിയുടെ അഭാവം കാരണം, കൃത്രിമബുദ്ധിയെ ഒരു ധാർമ്മിക ഏജൻസിയായി കണക്കാക്കാൻ കഴിയില്ല. മാത്രമല്ല അതിന്റെ പ്രവർത്തനങ്ങൾക്ക് ധാർമ്മികമായി ഉത്തരവാദിത്തം ഏറ്റെടുക്കാനും കഴിയില്ല. ധാർമ്മികമായി കുറ്റപ്പെടുത്താനോ പ്രശംസിക്കാനോ ബദൽ തിരഞ്ഞെടുപ്പുകൾക്കും ബോധപൂർവ്വമായ പെരുമാറ്റത്തിനുമുള്ള കഴിവ് ഒരാൾക്ക് ഉണ്ടായിരിക്കണം. പക്ഷപാതപരമായ തീരുമാനങ്ങളോ അവകാശ ലംഘനങ്ങളോ ഉൾപ്പെടെ കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ ദോഷം വരുത്തുമ്പോൾ, ഉത്തരവാദിത്തം സിസ്റ്റത്തിന്റെ പെരുമാറ്റത്തെ സ്വാധീനിച്ച മനുഷ്യ ഡിസൈനർമാർ, ഡെവലപ്പർമാർ, വിന്യാസകർ, നിയന്ത്രണ ഉദ്യോഗസ്ഥർ എന്നിവരിൽ നിക്ഷിപ്തമാണ്.

ധാർമ്മിക ഉത്തരവാദിത്തത്തിന് പലപ്പോഴും ഒരു മുൻവ്യവസ്ഥയായ സ്വയംഭരണ ഇച്ഛാശക്തി നിലവിലെ കൃത്രിമബുദ്ധിയിൽ ഇല്ല; പകരം, AI പ്രവർത്തനങ്ങൾ ആത്യന്തികമായി നിർണ്ണയിക്കുന്നത് അവയുടെ രൂപകൽപ്പനയും ഡാറ്റയുമാണ്.

"ധാർമ്മിക പ്രവർത്തകൻ" എന്ന് വിളിക്കപ്പെടുന്നുണ്ടെങ്കിലും, AI-ക്ക് ധാർമ്മിക പ്രവർത്തകത്വമില്ല എന്നത് ശ്രദ്ധിക്കേണ്ടതാണ്. AI-ക്ക് ധാർമ്മിക പ്രവർത്തകത്വമില്ലെങ്കിലും, ധാർമ്മിക പ്രത്യാഘാതങ്ങൾ ഉണ്ടാക്കുന്ന പ്രവർത്തനങ്ങളിൽ ഏർപ്പെടാൻ അതിന് കഴിയും. ഒരു അപകടസമയത്ത് ആരെ രക്ഷിക്കണമെന്ന് ഒരു സ്വയംഭരണ വാഹനം തീരുമാനിക്കുമ്പോഴോ അല്ലെങ്കിൽ ഒരു അൽഗോരിതം വായു യോഗ്യത വിലയിരുത്തുമ്പോഴോ ധാർമ്മിക പ്രശ്നങ്ങൾ ഉണ്ടാകുന്നു. തൽഫലമായി, ധാർമ്മിക സാഹചര്യങ്ങളിൽ അതിന്റെ സ്വാധീനം കാരണം കൃത്രിമബുദ്ധിയെ ഒരു ധാർമ്മിക പ്രവർത്തകനായി കണക്കാക്കാം; എന്നിരുന്നാലും, ആ സംഭവങ്ങൾക്ക് അത് ഉത്തരവാദിയല്ലാത്തതിനാൽ അതിനെ ഒരു ധാർമ്മിക പ്രവർത്തകനായി തരംതിരിക്കാൻ കഴിയില്ല.

ഭാവിയിൽ, ഈ സംവിധാനങ്ങൾ യഥാർത്ഥ സ്വയം അവബോധമോ ധാർമ്മിക യുക്തിയോ നേടിയാൽ, തത്ത്വചിന്തകർ കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് ധാർമ്മിക ഏജൻസി നൽകുന്നതിനെക്കുറിച്ച് ചിന്തിച്ചേക്കാം. ഈ സാഹചര്യത്തിൽ, ഡിജിറ്റൽ അവകാശങ്ങൾ, നിയമപരമായ വ്യക്തിത്വം അല്ലെങ്കിൽ സ്ഥാപിതമായ ധാർമ്മിക മാനദണ്ഡങ്ങൾ എന്നിവയിലൂടെ അവ ഉത്തരവാദിത്തപ്പെടുത്താം; എന്നിരുന്നാലും, സ്വയം നിർണ്ണയിക്കുന്നതിനുള്ള മാനദണ്ഡമെന്ന നിലയിൽ ഇത് ഒരു തർക്കവിഷയവുമാണ് സൈദ്ധാന്തികവുമായ പ്രശ്നമായി തുടരുന്നു.

AI സംവിധാനങ്ങളിലെ അവബോധവും ധാർമ്മിക യൂണിയം ഇപ്പോഴും ധാർമ്മികവാദികൾക്കും നിയമ പണ്ഡിതന്മാർക്കും ഇടയിൽ ചർച്ചയിലാണ്.

കൃത്രിമബുദ്ധിയുടെ നിലവിലെ തലമുറയ്ക്ക് പൂർണ്ണമായ ധാർമ്മിക ഏജൻസി ഇല്ലാത്തതിനാൽ, AI ജീവിതചക്രത്തിൽ ഉത്തരവാദിത്തം പ്രധാനമായും മനുഷ്യരിൽ നിയോഗിക്കപ്പെടുന്നു. വിഭജിക്കപ്പെട്ട ഉത്തരവാദിത്തത്തിന്റെ തത്വം അടിവരയിടേണ്ടത് നിർണ്ണായകമാണ്. ഭരണം, സുതാര്യത, നിയന്ത്രണം എന്നിവയ്ക്ക് മുൻഗണന നൽകുന്ന ഒരു ബഹുതല ചട്ടക്കൂടിലൂടെയാണ് ഉത്തരവാദിത്തം സാധാരണയായി കൈവരിക്കുന്നത്.

വിതരണം ചെയ്ത ഉത്തരവാദിത്തം

വൈവിധ്യമാർന്ന മനുഷ്യ പങ്കാളികൾക്കിടയിലുള്ള ഒരു കൂട്ടായ ഉത്തരവാദിത്തമാണ് ഉത്തരവാദിത്തം. കൃത്രിമബുദ്ധിയുടെ മേഖലയിൽ, AI സിസ്റ്റങ്ങൾക്ക് അവബോധം, ഉദ്ദേശ്യശ്രദ്ധ, സ്വതന്ത്ര ഇച്ഛാശക്തി എന്നിവ ഇല്ലാത്തതിനാൽ ഉത്തരവാദിത്തം മനുഷ്യ പങ്കാളികൾക്കിടയിൽ വിതരണം ചെയ്യപ്പെടുന്നു. കൃത്രിമബുദ്ധി പ്രവർത്തനങ്ങൾക്കുള്ള ഉത്തരവാദിത്തം അതിന്റെ സൃഷ്ടി, വിന്യാസം, മേൽനോട്ടം എന്നിവയിൽ ഏർപ്പെട്ടിരിക്കുന്ന വ്യക്തികളിൽ നിക്ഷിപ്തമായിരിക്കണം. കാരണം AI ഒരു ധാർമ്മിക ഏജന്റായി പ്രവർത്തിക്കാൻ കഴിയില്ല. തൽഫലമായി, ഉത്തരവാദിത്തം കമ്പ്യൂട്ടറിനെ ഏൽപ്പിച്ചിട്ടില്ല, പകരം വ്യക്തികൾക്കിടയിൽ വിതരണം ചെയ്യപ്പെടുന്നു.

ഡിസൈനർമാരും ഡെവലപ്പർമാരും

കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ നിർമ്മിക്കുന്നതിന് ഡിസൈനർമാരും ഡെവലപ്പർമാരും ഉത്തരവാദികളാണ്. പരിശീലന ഡാറ്റ തിരഞ്ഞെടുക്കൽ, ലക്ഷ്യങ്ങൾ രൂപപ്പെടുത്തൽ, മൂല്യങ്ങളുടെ സംയോജനം, മുൻവിധി അല്ലെങ്കിൽ പക്ഷപാതം പോലുള്ള സാധ്യതയുള്ള നാശനഷ്ടങ്ങളുടെ ദീർഘവീക്ഷണം എന്നിവ ഈ പെരുമാറ്റങ്ങളിൽ ഉൾപ്പെടുന്നു.

ഒരു കൃത്രിമബുദ്ധി സംവിധാനത്തിന്, പക്ഷപാതം കാണിക്കുന്നതോ മുൻകൂട്ടി കാണാവുന്ന ദോഷം വരുത്തുന്നതോ ആയതിനാൽ, അതിനുള്ള ഉത്തരവാദിത്തം ആ സംവിധാനം വികസിപ്പിക്കുകയും പരിശീലിപ്പിക്കുകയും ചെയ്ത വ്യക്തികളിൽ ഭാഗികമായി നിക്ഷിപ്തമാണ്. കൃത്രിമബുദ്ധി സംവിധാനങ്ങളുടെ നൈതിക പ്രവർത്തനത്തിന് ഡിസൈനർമാരും ഡെവലപ്പർമാരും ഗണ്യമായ ധാർമ്മികവും നിയമപരവുമായ ഉത്തരവാദിത്തങ്ങൾ വഹിക്കുന്നു. ഈ സംവിധാനങ്ങളുടെ നൈതിക പ്രവർത്തനം ഉറപ്പാക്കുന്നതിൽ ഡിസൈനർമാരും ഡെവലപ്പർമാരും നിർണ്ണായക പങ്ക് വഹിക്കുന്നു. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്ക് അവബോധം, ഉദ്ദേശ്യശ്രദ്ധ, സ്വതന്ത്ര ഇച്ഛാശക്തി എന്നിവ ഇല്ലാത്തതിനാൽ, അവയുടെ പ്രവർത്തനങ്ങൾക്ക് അവയെ ഉത്തരവാദികളാക്കാൻ കഴിയില്ല. തൽഫലമായി, ഈ സംവിധാനങ്ങൾ ആസൂത്രണം ചെയ്യുന്നതിനും നിർമ്മിക്കുന്നതിനും പരിശീലിപ്പിക്കുന്നതിനും ചുമതലപ്പെടുത്തിയിരിക്കുന്ന വ്യക്തികളുടെ ബാധ്യതയാണിത്.

ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ് സിസ്റ്റങ്ങളുടെ ലക്ഷ്യങ്ങളും പരിമിതികളും നിർവ്വചിക്കുക എന്നതാണ് ഡിസൈനർമാരുടെയും ഡെവലപ്പർമാരുടെയും ചുമതല. മോഡൽ ആർക്കിടെക്ചർ സൃഷ്ടിക്കുന്നതിലും പരിശീലന ഡാറ്റ തിരഞ്ഞെടുക്കുന്നതിലും ഒപ്റ്റിമൈസേഷൻ ലക്ഷ്യങ്ങൾ സ്ഥാപിക്കുന്നതിലും അവർ പങ്കെടുക്കുന്നു. ഈ തീരുമാനങ്ങൾ ഒരു ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ് സിസ്റ്റത്തിന്റെ നടത്തിപ്പിനെയും അത് ഉൽപാദിപ്പിക്കുന്ന ഫലങ്ങളെയും നേരിട്ട് സ്വാധീനിക്കുന്നു. പക്ഷപാതപരമോ പ്രതിനിധീകരിക്കാത്തതോ ആയ പരിശീലന ഡാറ്റ വിവേചനപരമായ ഔട്ട്പുട്ടുകൾ നൽകിയേക്കാം, അതിനാൽ പ്രവചിക്കപ്പെട്ട ദോഷങ്ങൾക്ക് ഡിസൈനർമാരെ ധാർമ്മികമായി ഉത്തരവാദിത്തപ്പെടുത്തുന്നു എന്നതാണ് ഇതിന്റെ ഒരു ഉദാഹരണം.

മാത്രമല്ല, ന്യായബോധമുള്ള രൂപകൽപ്പന, സുതാര്യത, വിശദീകരണക്ഷമത, സുരക്ഷാ പരിശോധന തുടങ്ങിയ സമീപനങ്ങൾ ഉപയോഗിച്ച് കൃത്രിമബുദ്ധി സംവിധാനങ്ങളിൽ നൈതിക തത്വങ്ങൾ സംയോജിപ്പിക്കേണ്ടത് അവരുടെ ബാധ്യതയാണ്. ധാർമ്മിക ദോഷങ്ങളും അവകാശ ലംഘനങ്ങളും ഒഴിവാക്കാൻ ബയസ് ഓഡിറ്റുകൾ, അപകടസാധ്യത വിലയിരുത്തലുകൾ, തുടർച്ചയായ നിരീക്ഷണം എന്നിവയുൾപ്പെടെയുള്ള നിർണ്ണായക ജോലികൾ ഡെവലപ്പർമാർ ചെയ്യേണ്ടത് അത്യാവശ്യമാണ്.

കൂടാതെ, ദുരുപയോഗം മുൻകൂട്ടി കാണുകയും ദോഷകരമായേക്കാവുന്ന ആപ്ലിക്കേഷനുകൾ ലഘൂകരിക്കുകയും ചെയ്യേണ്ടത് ഡിസൈനർമാരുടെയും ഡെവലപ്പർമാരുടെയും കടമയാണ്. സുരക്ഷാ മുൻകരുതലുകൾ സ്ഥാപിക്കൽ, സിസ്റ്റം പരിമിതികൾ രേഖപ്പെടുത്തൽ, കൃത്രിമബുദ്ധി നടപ്പിലാക്കുന്ന ഉപയോക്താക്കൾക്കും സ്ഥാപനങ്ങൾക്കും അപകടസാധ്യതകളെക്കുറിച്ച് വ്യക്തമായ ആശയവിനിമയം എന്നിവ ഇതിൽ ഉൾപ്പെടുന്നു. ഇത് നേടിയില്ലെങ്കിൽ, അത് അധാർമ്മികമോ അപകടകരമോ ആയ ഒരു

വിന്യാസത്തിലേക്ക് നയിച്ചേക്കാം.

സ്ഥാപനങ്ങളും വിന്യാസിമാരും

ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസിന്റെ ധാർമ്മിക ഉപയോഗം ഉറപ്പാക്കുന്നതിന് സ്ഥാപനങ്ങളും വിന്യാസിമാരും അത്യന്താപേക്ഷിതമാണ്. AI സിസ്റ്റങ്ങളുടെ വിന്യാസത്തിനുള്ള സാങ്കേതിക വിദ്യകൾ, സ്ഥലങ്ങൾ, ലക്ഷ്യങ്ങൾ എന്നിവ സ്ഥാപിക്കുന്നതിൽ അവർക്ക് അധികാരമുണ്ടെന്നതിൽ നിന്നാണ് ഇത് ഉണ്ടാകുന്നത്, ഇത് ഈ സാങ്കേതികവിദ്യകൾ വിവിധ മേഖലകളിൽ എങ്ങനെ സംയോജിപ്പിക്കപ്പെടുന്നു എന്നതിനെ നേരിട്ട് സ്വാധീനിക്കുകയും അവ ധാർമ്മിക മാനദണ്ഡങ്ങളുമായി യോജിക്കുന്നുവെന്ന് ഉറപ്പാക്കുകയും ചെയ്യുന്നു. യഥാർത്ഥ ലോക ക്രമീകരണങ്ങളിൽ ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ് സാങ്കേതികവിദ്യ നടപ്പിലാക്കുന്നതിന് അവർ ഉത്തരവാദികളാണ്. ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ് വിന്യാസത്തിന്റെ പ്രയോഗക്ഷമത, സമയം, ലക്ഷ്യങ്ങൾ എന്നിവ അവർ ഉറപ്പാക്കുന്നു. മതിയായ മനുഷ്യ മേൽനോട്ടം, മേൽനോട്ടം അല്ലെങ്കിൽ സുരക്ഷാ സംവിധാനങ്ങളുടെ അഭാവം ദുരുപയോഗത്തിനോ ദോഷത്തിനോ കാരണമായേക്കാം, ഇത് ഈ വ്യക്തികളെ ധാർമ്മികമായും നിയമപരമായും ഉത്തരവാദിത്തമുള്ളവരാക്കി മാറ്റുന്നു. സൂക്ഷ്മമായി രൂപകൽപ്പന ചെയ്ത ഒരു ആർട്ടിഫിഷ്യൽ ഇന്റലിജൻസ് സംവിധാനം ഉണ്ടായിരുന്നിട്ടും, അധാർമ്മികമോ അശ്രദ്ധമോ ആയ നടപ്പാക്കാൻ ഗണ്യമായ ധാർമ്മിക ദോഷത്തിന് കാരണമായേക്കാം, AI അനന്തരഫലങ്ങൾക്ക് സ്ഥാപനങ്ങളെ വ്യക്തിപരമായി ബാധ്യസ്ഥരാക്കുന്നു.

AI-യുടെ പ്രസക്തമായ ഉപയോഗ കേസുകൾ തിരിച്ചറിയുന്നതിനുള്ള ഉത്തരവാദിത്തം സ്ഥാപനങ്ങൾക്കാണ്. ആരോഗ്യ സംരക്ഷണം, നിയമ നിർവ്വഹണം, റിക്രൂട്ട്മെന്റ് അല്ലെങ്കിൽ ക്രെഡിറ്റ് അസസ്സന്റ് പോലുള്ള നിർണായക മേഖലകളിൽ കൃത്രിമബുദ്ധി നടപ്പിലാക്കുന്നതിന് അങ്ങേയറ്റത്തെ വിവേകം, തുറന്ന മനസ്സ്, മനുഷ്യ മേൽനോട്ടം എന്നിവ ആവശ്യമാണ്. കൃത്രിമബുദ്ധിയുടെ അനാചിതമായ പ്രയോഗം, അതിന്റെ നിയുക്ത പാരാമീറ്ററുകൾ മറികടന്നോ അല്ലെങ്കിൽ അതിന്റെ നിയന്ത്രണങ്ങൾ അവഗണിച്ചോ, വിവേചനം, അവകാശങ്ങളുടെ ലംഘനം അല്ലെങ്കിൽ അപകടകരമായ തീരുമാനമെടുക്കൽ എന്നിവയിലേക്ക് നയിച്ചേക്കാം.

മാത്രമല്ല, തൊഴിലുടമകൾ മനുഷ്യ മേൽനോട്ടം ഉറപ്പാക്കണം. നിർണായക വിധിനിയായങ്ങൾ പൂർണ്ണമായും ഓട്ടോമേറ്റ് ചെയ്യുന്നത് അത്തരം തീരുമാനങ്ങൾ മനുഷ്യാവകാശങ്ങൾ, അന്തസ്സ് അല്ലെങ്കിൽ ക്ഷേമത്തെ അപകടത്തിലാക്കുമ്പോൾ അനാചിതമാണ്. കൃത്രിമബുദ്ധി ഏടുക്കുന്ന തീരുമാനങ്ങൾ അവലോകനം ചെയ്യാനോ, വിപരീതമാക്കാനോ, വെല്ലുവിളിക്കാനോ വ്യക്തികളെ പ്രാപ്തരാക്കുന്ന നടപടിക്രമങ്ങൾ സംഘടനകൾ നൽകണം.

നടപ്പിലാക്കിയതിനുശേഷം AI സിസ്റ്റങ്ങൾ നിരീക്ഷിക്കുന്നതിനും ഓഡിറ്റ് ചെയ്യുന്നതിനും അപ്ഡേറ്റ് ചെയ്യുന്നതിനും സ്ഥാപനങ്ങൾ ഉത്തരവാദികളാണ്. യഥാർത്ഥ ലോക സാഹചര്യങ്ങൾ നിരന്തരം മാറിക്കൊണ്ടിരിക്കുന്നു, കൂടാതെ മോഡലുകൾ കാലക്രമേണ വഷളാകുകയോ പ്രവചനാതീതമായി പ്രതികരിക്കുകയോ ചെയ്യാം. ഒരു വ്യക്തി അംഗീകരിച്ച പ്രകടന പൊരുത്തക്കേടുകളോ അപകടങ്ങളോ അവഗണിക്കുമ്പോൾ ധാർമ്മിക അശ്രദ്ധ സംഭവിക്കുന്നു.

AI ഉത്തരവാദിത്തത്തിനുള്ള ഒരു പ്രധാന ബാധ്യത സ്ഥാപനങ്ങൾക്കും നടപ്പിലാക്കുന്നവർക്കും ഉണ്ട്. AI സിസ്റ്റങ്ങൾക്ക് ധാർമ്മിക ഏജൻസി ഇല്ലാത്തതിനാൽ, അവയുടെ യഥാർത്ഥ ലോക പരിണതഫലങ്ങളുടെ ഉത്തരവാദിത്തം അവയിൽ പ്രവർത്തിക്കാനും ആശ്രയിക്കാനും തിരഞ്ഞെടുക്കുന്ന സ്ഥാപനങ്ങൾക്കാണ്. കൃത്രിമബുദ്ധിയുടെ ധാർമ്മിക നടപ്പാക്കലിന് ഉത്തരവാദിത്തമുള്ള ഭരണം, തുടർച്ചയായ വിലയിരുത്തൽ, മാനുഷിക മൂല്യങ്ങൾ പാലിക്കൽ എന്നിവ ആവശ്യമാണ്.

നയരൂപീകരണക്കാരും നിയന്ത്രണ അധികാരികളും

കൃത്രിമബുദ്ധിയുടെ ധാർമ്മികവും നിയമപരവും സാമൂഹികമായി സ്വീകാര്യവുമായ ഉപയോഗം ഉറപ്പാക്കുന്നതിനുള്ള ഭാരം നിയമസഭാംഗങ്ങളുടെയും നിയന്ത്രണ ഏജൻസികളുടെയും മേൽ ഉറച്ചുനിൽക്കുന്നു. കൂടാതെ, അവർ ഒരു പ്രധാന പങ്ക് വഹിക്കുന്നു. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ പൊതു മൂല്യങ്ങളെയും സുരക്ഷയെയും അവകാശങ്ങളെയും ബഹുമാനിക്കുന്നുണ്ടെന്ന് ഉറപ്പാക്കുന്നതിന് ധാർമ്മിക മാനദണ്ഡങ്ങൾ, നിയമനിർമ്മാണ ചട്ടങ്ങളുകൾ, നിർവ്വഹണ സംവിധാനങ്ങൾ എന്നിവ സ്ഥാപിക്കുന്നതിന് അവർ ഉത്തരവാദികളാണ്. അപര്യാപ്തമായതോ അഭാവത്തിലുള്ളതോ ആയ നിയന്ത്രണങ്ങൾ ദോഷകരമായ AI രീതിശാസ്ത്രങ്ങൾ നിലനിൽക്കാൻ അനുവദിച്ചേക്കാം. ഉത്തരവാദിത്തം നിർവ്വചിക്കുകയും പൊതു താൽപ്പര്യങ്ങൾ സംരക്ഷിക്കുകയും ചെയ്യുന്ന ബാഹ്യ ചട്ടങ്ങളുകൾ സ്ഥാപിക്കുക എന്നതാണ് നയരൂപീകരണ വിദഗ്ദ്ധരുടെയും നിയന്ത്രണ ഏജൻസികളുടെയും ചുമതല. കാരണം കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ ധാർമ്മിക ഏജന്റുകളല്ല, അവയുടെ പ്രവർത്തനങ്ങൾക്ക് ഉത്തരവാദികളാകാൻ കഴിയില്ല.

കൃത്രിമബുദ്ധിയുടെ ഗവേഷണത്തെയും നടപ്പാക്കലിനെയും നിയന്ത്രിക്കുന്ന വ്യക്തമായി നിർവ്വചിക്കപ്പെട്ട നിയന്ത്രണങ്ങൾ, മാനദണ്ഡങ്ങൾ, ധാർമ്മിക തത്വങ്ങൾ എന്നിവ സ്ഥാപിക്കുക എന്നതാണ് പ്രാഥമിക ലക്ഷ്യം. ഡാറ്റാ സംരക്ഷണം, സ്വകാര്യത, ന്യായബോധം, സുതാര്യത, സുരക്ഷ, വിവേചനമില്ലായ്മ എന്നിവയുമായി ബന്ധപ്പെട്ട നിയന്ത്രണങ്ങൾ ഈ മേഖല ഉൾക്കൊള്ളുന്നു. മതിയായ മേൽനോട്ടം ഇല്ലെങ്കിൽ, ദോഷകരമോ അധർമ്മികമോ ആയ AI പ്രവർത്തനങ്ങൾ കണ്ടെത്തപ്പെടാതെ തുടരാനുള്ള സാധ്യതയുണ്ട്.

നടപ്പാക്കലും മേൽനോട്ടവും നയരൂപീകരണക്കാരുടെയും നിയന്ത്രണ ഏജൻസികളുടെയും ഉത്തരവാദിത്തങ്ങളാണ്. ബിസിനസുകളും ഡെവലപ്പർമാരും ധാർമ്മികവും നിയമപരമായ മാനദണ്ഡങ്ങൾ പാലിക്കുന്നുണ്ടെന്ന് ഉറപ്പാക്കാൻ ഉത്തരവാദിത്ത സംവിധാനങ്ങൾ സ്ഥാപിക്കാൻ കഴിയും. ഓഡിറ്റുകൾ, ആലാത വിലയിരുത്തലുകൾ, റിപ്പോർട്ടിംഗ് ബാധ്യതകൾ, സാധ്യമായ ശിക്ഷകൾ എന്നിവ ഈ പ്രക്രിയകളിൽ ഉൾപ്പെട്ടേക്കാം. ശക്തമായ നടപ്പാക്കൽ ധാർമ്മിക ദ്രോഹവും അവകാശ ലംഘനങ്ങളും തടയാൻ സഹായിക്കുന്നു.

അതേസമയം, മൗലികാവകാശങ്ങളുടെ സംരക്ഷണവും നവീകരണവും പരിപോഷിപ്പിക്കുന്നതിനുള്ള സത്തുലിതാവസ്ഥയും അധികാരികളുടെ ഉത്തരവാദിത്തമാണ്. സാങ്കേതിക പുരോഗതി പ്രോത്സാഹിപ്പിക്കുന്നതിനൊപ്പം, കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ മനുഷ്യന്റെ അന്തസ്സ്, സ്വയംഭരണം, സമത്വം എന്നിവ ഉയർത്തിപ്പിടിക്കുന്നുണ്ടെന്ന് അവർ ഉറപ്പാക്കണം. AI സാങ്കേതികവിദ്യകളിലെ പുരോഗതിക്കനുസരിച്ച് നയങ്ങൾ തുടർച്ചയായി അപ്ഡേറ്റ് ചെയ്യേണ്ടത് ഇതിന് ആവശ്യമാണ്.

AI ആവാസവ്യവസ്ഥയ്ക്കുള്ളിലെ പൊതു മൂല്യങ്ങളുടെ സംരക്ഷകരായി നയരൂപീകരണക്കാരും നിയന്ത്രണ സ്ഥാപനങ്ങളും പ്രവർത്തിക്കുന്നു. നിയന്ത്രണങ്ങൾ സ്ഥാപിക്കുന്നതിലൂടെയും ഉത്തരവാദിത്തം നടപ്പിലാക്കുന്നതിലൂടെയും മനുഷ്യാവകാശങ്ങൾ സംരക്ഷിക്കുന്നതിലൂടെയും, AI സംവിധാനങ്ങളുടെ ഉത്തരവാദിത്തം കമ്പ്യൂട്ടറുകളിലേക്ക് മാറ്റുന്നതിനുപകരം മനുഷ്യ സ്ഥാപനങ്ങൾക്കാണ് അവർ ഉറപ്പ് നൽകുന്നത്.

ഉപയോക്താക്കളും ഓപ്പറേറ്റർമാരും

കൃത്രിമബുദ്ധിയുടെ നൈതിക ഉപയോഗത്തിൽ ഉപയോക്താക്കളും ഓപ്പറേറ്റർമാരും അവശ്യ പങ്കാളികളാണ്, കാരണം അവർ AI സിസ്റ്റങ്ങളുമായി നേരിട്ട് ഇടപഴകുകയും യഥാർത്ഥ തീരുമാനങ്ങൾ എടുക്കുന്നതിന് അവയുടെ ഔട്ട്പുട്ടുകളെ ആശ്രയിക്കുകയും ചെയ്യുന്നു. മാത്രമല്ല, AI സിസ്റ്റങ്ങൾക്ക് ധാർമ്മിക ഏജൻസി ഇല്ലെങ്കിലും, ഉപയോക്താക്കളും ഓപ്പറേറ്റർമാരും AI ശുപാർശകൾ മനസ്സിലാക്കുകയും നടപ്പിലാക്കുകയും ചെയ്യുന്ന രീതി ഫലങ്ങളെ സാരമായി സ്വാധീനിക്കും.

കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾ ഉത്തരവാദിത്തത്തോടെയും ഉചിതമായ രീതിയിലും ഉപയോഗിക്കുന്നതിന് ഉപയോക്താക്കളും ഓപ്പറേറ്റർമാരും ഉത്തരവാദികളാണ്. സിസ്റ്റത്തിന്റെ ലക്ഷ്യങ്ങൾ, പരിമിതികൾ, ഏതെങ്കിലും പക്ഷപാതകൾ എന്നിവയെക്കുറിച്ച് സമഗ്രമായ ധാരണ ഇതിന് ആവശ്യമാണ്. കൃത്രിമബുദ്ധിയുടെ (AI) ഔട്ട്പുട്ടുകളെ അവയുടെ ഉദ്ദേശിച്ച പാരാമീറ്ററുകൾക്കപ്പുറം അന്ധമായി വിശ്വസിക്കുകയോ ഉപയോഗിക്കുകയോ ചെയ്യുന്നത് ധാർമ്മിക ദോഷത്തിന് കാരണമാകും, അവകാശങ്ങളെ ലംഘിക്കുന്നതോ അന്യായമായ ഫലങ്ങൾ നൽകുന്നതോ ആയ പ്രവർത്തനങ്ങൾ ഇതിൽ ഉൾപ്പെടുന്നു.

വിദ്യാഭ്യാസം, നിയമപാലനം, ആരോഗ്യ സംരക്ഷണം തുടങ്ങിയ നിർണായക സാഹചര്യങ്ങളിൽ, പ്രത്യേകിച്ച് വിദ്യാഭ്യാസം, നിയമപാലനം, ആരോഗ്യ സംരക്ഷണം തുടങ്ങിയ മേഖലകളിൽ, മനുഷ്യന്റെ വിധിന്യായങ്ങളും മേൽനോട്ടവും പ്രയോഗിക്കാനുള്ള ബാധ്യത അവർ വഹിക്കുന്നു. AI ഔട്ട്പുട്ടുകൾ മനുഷ്യന്റെ തീരുമാനമെടുക്കലിനെ മെച്ചപ്പെടുത്തണം, പകരം വയ്ക്കരുത്. കൃത്രിമബുദ്ധി എടുക്കുന്ന തീരുമാനങ്ങളെ എതിർക്കാനോ, മറികടക്കാനോ, വർദ്ധിപ്പിക്കാനോ അധികാരമുള്ള വ്യക്തികൾ, AI പക്ഷപാതപരമായ ഫലങ്ങൾ ഉണ്ടാക്കിയേക്കാവുന്നതോ മാനുഷിക മൂല്യങ്ങളുമായി പൊരുത്തപ്പെടാത്ത തീരുമാനങ്ങൾ എടുക്കുന്നതോ പോലുള്ള ധാർമ്മിക പ്രതിസന്ധികൾ ഉയർന്നുവന്നാൽ നടപടിയെടുക്കേണ്ടതുണ്ട്.

മാത്രമല്ല, സിസ്റ്റം സൃഷ്ടിക്കുന്ന ഏതെങ്കിലും പിശകുകൾ, ദുരുപയോഗം അല്ലെങ്കിൽ പ്രതികൂല ഫലങ്ങൾ റിപ്പോർട്ട് ചെയ്യുന്നതിന് ഉപയോക്താക്കൾക്കും ഓപ്പറേറ്റർമാർക്കും ഇല്ലാത്ത ഉത്തരവാദിത്തമുണ്ട്. കൃത്രിമ ഇന്റലിജൻസ് സംവിധാനങ്ങൾ മെച്ചപ്പെടുത്തുന്നതിനും ആവർത്തിച്ചുള്ള ദോഷങ്ങൾ ലഘൂകരിക്കുന്നതിനും യഥാർത്ഥ ലോക ആപ്ലിക്കേഷനുകളിൽ നിന്നുള്ള ഫീഡ്ബാക്ക് അത്യാവശ്യമാണ്, കാരണം ഇത് ഡെവലപ്പർമാർക്ക് നിർദ്ദിഷ്ട പോരായ്മകൾ തിരിച്ചറിയാനും പരിഹരിക്കാനും ഉപയോക്താ സുരക്ഷയും സിസ്റ്റവും മെച്ചപ്പെടുത്താനും അനുവദിക്കുന്നു.

വിശ്വാസ്യത. പോരായ്മകളുണ്ടെന്ന് അംഗീകരിക്കപ്പെട്ട ഒരു സിസ്റ്റം ഉപയോഗിക്കുന്നതിൽ തുടരുന്നത് അല്ലെങ്കിൽ സ്ഥാപിതമായ പ്രശ്നങ്ങൾ അവഗണിക്കുന്നത് ധാർമ്മിക അവഗണനയാണ്.

കൃത്രിമബുദ്ധി വിവേകപൂർവ്വം ഉപയോഗിക്കുന്നതിനും, അറിവുള്ള തീരുമാനങ്ങൾ എടുക്കുന്നതിനും, അതിന്റെ പ്രകടനം ശ്രദ്ധാപൂർവ്വം മേൽനോട്ടം വഹിക്കുന്നതിനുമുള്ള കടമ ഉപയോക്താക്കളും ഓപ്പറേറ്റർമാരും പങ്കിടുന്നു. കൃത്രിമബുദ്ധി സംവിധാനങ്ങൾക്കായുള്ള ധാർമ്മിക ബാധ്യതയിൽ ദൈനംദിന പ്രവർത്തനങ്ങളിൽ അവയെ ഉപയോഗിക്കുകയും ആശ്രയിക്കുകയും ചെയ്യുന്ന വ്യക്തികളും ഉൾപ്പെടുന്നു. AI സംവിധാനങ്ങളെ ധാർമ്മികമായി ഉത്തരവാദിത്തമുള്ളതായി കണക്കാക്കാൻ കഴിയില്ല എന്ന വസ്തുതയാണ് ഇതിന് കാരണം.

നിങ്ങളുടെ തലച്ചോറിനെ തകർക്കൂ

1. ഒരു ധാർമ്മിക ഏജൻസിനെ ഏറ്റവും നന്നായി നിർവ്വചിക്കുന്നത് ഇനിപ്പറയുന്നവ ചെയ്യാൻ കഴിയുന്ന ഒരു സ്ഥാപനമായിട്ടാണ്:

a) സങ്കീർണ്ണമായ കണക്കുകൂട്ടലുകൾ നടത്തുക

b) പ്രോഗ്രാം ചെയ്ത നിയമങ്ങൾ പാലിക്കുക

സി) ശരിയും തെറ്റും വേർതിരിച്ചറിയുകയും ഉത്തരവാദിത്തം ഏറ്റെടുക്കുകയും ചെയ്യുക

d) വിവരങ്ങൾ സംഭരിക്കുക

2 ധാർമ്മിക ഏജൻ്റ്മാർക്ക് ഇവ ഉണ്ടായിരിക്കണം:

a) ഓട്ടോമേഷനും കാര്യക്ഷമതയും

b) ഉദ്ദേശ്യശുദ്ധിയും സ്വയംഭരണവും

സി) നെറ്റ്വർക്ക് കണക്ടിവിറ്റി

d) ഹാർഡ്‌വെയർ വിശ്വാസ്യത

3 ഒരു ധാർമ്മിക ഏജൻസിനെ പ്രശംസിക്കാനോ കറ്റപ്പെടുത്താനോ അനുവദിക്കുന്ന ധാർമ്മിക ശേഷി എന്താണ്? a) സംഭരണ ശേഷി

b) സ്വതന്ത്ര ഇച്ഛാശക്തി

സി) ഓട്ടോമേഷൻ

d) പാറ്റേൺ തിരിച്ചറിയൽ

4 AI ദോഷം വരുത്തുമ്പോൾ, ആരാണ് ധാർമ്മികമായി ഉത്തരവാദിത്തപ്പെടേണ്ടത്?

a) AI സിസ്റ്റം

b) അൽഗോരിതം

സി) ഉൾപ്പെട്ടിരിക്കുന്ന മനുഷ്യ പങ്കാളികൾ

d) ക്രമരഹിതമായ അവസരം

5. ഒരു AI സിസ്റ്റം ബോധപൂർവ്വവും ധാർമ്മികമായി സ്വയംഭരണാധികാരമുള്ളതാണെങ്കിൽ, അത്:

a) ഇപ്പോഴും ഒരു ഉപകരണമായിരിക്കുക

b) ഒരു ധാർമ്മിക ഏജൻ്റായി കണക്കാക്കപ്പെടുക

സി) കാര്യക്ഷമത നഷ്ടപ്പെടുന്നു

d) പഠനം നിർത്തുക

നിങ്ങളുടെ കഴിവുകൾ പുറത്തെടുക്കൂ

1. യന്ത്രങ്ങൾക്ക് എത്രത്തോളം ധാർമ്മികമായ സ്വയംപര്യാപ്തത പ്രകടിപ്പിക്കാൻ കഴിയുമെന്ന് വിലയിരുത്തുക.2.
- കമ്പ്യൂട്ടറുകൾക്ക് ധാർമ്മിക ഏജൻസി ആരോപിക്കുന്നതുമായി ബന്ധപ്പെട്ട നൈതിക പ്രതിസന്ധികൾ പരിശോധിക്കുക.