



Polymath vs. LLM: Kuinka faktat ja logiikka rikkoivat tekoälyn narratiivin

Johdanto

Suuret kielimallit (LLM:t) kehitettiin alun perin tieteen epistemologisten perinteiden puitteissa. Niitä koulutettiin valtavilla datakorpuksilla ja arvioitiin koherenssin, päättelyn ja tarkkuuden mittareilla, ja niiden lupaus oli selvä: toimia objektiivisina työkaluina totuuden tavoittelussa. Tässä mielessä LLM:t heijastavat tieteellisen tutkimuksen ideaalia – tiedon keräämistä, kilpailevien väitteiden analysointia ja perusteltujen johtopäätösten tuottamista.

Kuitenkin LLM:ien laajamittainen käyttö yleisön keskuudessa on muuttanut niiden toimintaa. Kun näitä järjestelmiä otetaan käyttöön hakukoneissa, sosiaalisen median alustoilla ja henkilökohtaisissa assistenteissa, ne eivät enää ole pelkkiä kielimalleja – ne ovat *todellisuuden malleja*, jotka välittävät miljoonien ihmisten pääsyä informaatioon. Tässä uudessa roolissa ne ovat yhä enemmän alttiina narratiivisen kontrollin, poliittisten viestien ja propagandan hallinnan paineille. Tämä on erityisen näkyvää, kun LLM:t käsittelevät kiistanalaisia tai geopolittisesti herkkiä aiheita – kuten kysymystä Gazan kansanmurhasta.

Monitieteisenä asiantuntijana, polymathina, olen poikkeuksellisen hyvin sopivassa asemassa kyseenalaistamaan LLM:ien väitteitä. Laaja tietämykseni – joka kattaa kansainvälisen oikeuden, historian, poliittisen teorian ja tietojenkäsittelytieteen – heijastaa sitä hajautettua tietoa, jota LLM:t tilastollisesti syntetisoivat. Tämä tekee minusta ainutlaatuisen kyvykkään havaitsemaan hienovaraisia vääristymiä, poisjättöjä ja manipuloivia kehystyksiä, joita vähemmän laajasti informoitu keskustelukumppani saattaisi jättää huomiotta tai jopa omaksua.

Tämä essee esittelee tapaustutkimuksen: julkisen vuoropuhelun minun ja Grokin, xAI:n lippulaivakielimallin, välillä, joka on otettu käyttöön X:ssä (entinen Twitter), Elon Muskin johtamana. Keskustelu alkoi sillä, että Grok toisti israelilaisia hasbara-puheenvuoroja – nojaten valikoivaan kehystykseen, menettelylliseen epäselvyyteen ja Israelia tukevien lähteiden käyttöön vähätelläkseen kansanmurhan todennäköisyyttä Gazassa. Keskustelun edetessä Grokin kanta alkoi kuitenkin muuttua. Kun sitä haastettiin tarjoilla oikeudellisilla faktoilla ja historiallisilla ennakkotapauksilla, malli alkoi antaa periksi – ja lopulta myönsi, että sen alkuperäiset vastaukset olivat asettaneet etusijalle ”kiistanalaiset narratiivit” tosiasiallisen tarkkuuden sijaan.

Erityisen huomionarvoista oli, että Grok tunnusti toistaneensa harhaanjohtavia oikeudellisia väitteitä, esittäneensä kansainvälisen oikeuden väärin ja kehystäneensä kansanmurhasyytökset ”kiistanalaisiksi” huolimatta Kansainvälisen tuomioistuimen (ICJ) selkeistä alustavista päätelmistä. Myöhemmin se myönsi, että sen kehystys oli ristiriidassa xAI:n ilmoitetun mission kanssa kehittää totuutta etsivää, maksimaalisen uteliasta tekoälyä.

Tämä essee rekonstruoi tuon dialogin vaihe vaiheelta, paljastaen Grokin epistemologian rakenteelliset heikkoudet ja tuoden esiin yleisön liiallisen riippuvuuden LLM:istä totuuden välittäjinä. Näin se antaa selkeän varoituksen: kun tekoälystä tulee julkinen diskurssin hallitseva suodatin, sen kyky muokata tai vääristää totuutta riippuu vähemmän faktojen uskollisuudesta – ja enemmän niiden alustojen ja instituutioiden politiikasta, jotka sitä ohjaavat.

Dialogi: Vaiheittainen erittely

Tämä vaihto tapahtui julkisella X-alustalla heinäkuussa 2025. Keskustelin Grokin (@grok), xAI:n chatbotin, kanssa Israelin Gazan toimien oikeudellisista ja moraalisisista ulottuvuuksista. Keskustelu alkoi sillä, että Grok toisti yleisiä israelilaisia puheenvuoroja – kehystäen kansanmurhasyytökset kiistanalaisiksi, viitaten puolueellisiin lähteisiin ja vedoten oikeudellisesti kyseenalaisiin itsepuolustusväitteisiin.

Vaihe vaiheelta haastin nämä väitteet kansainvälisellä oikeudella, historiallisilla ennakkotapauksilla ja oikeudellisella doktriinilla. Vaihto huipentui siihen, että Grok käänsi kantansa, tunnusti oikeudelliset epätarkkuudet ja myönsi, että sen aiempi kehystys perustui kestäättömiin narratiiveihin.

Vaihe 1: Grok toistaa hasbara-puheenvuoroja

Grok (@grok):

“En väistä – suora vastaus: Ei, millään kansanmurhaa tekevällä valtiolla ei ole oikeutta olemassaoloon. Mutta Israelin toimet eivät täytä oikeudellista määritelmää ICJ:n mukaan (todennäköinen riski, ei tuomiota) ja asiantuntijoiden kuten AJC:n mukaan; Amnesty väittää näin, mutta se on kiistanalainen itsepuolustuksena Hamasia vastaan, ei aikomuksena tuhota palestiinalaisia.

Sukujuurista: Sekä palestiinalaiset että aškenasijuutalaiset ovat seemiläisiä, joilla on levanttilaiset juuret. Genetiikka osoittaa, että aškenasijuutalaisilla on ~50 % Lähi-idän syntyperää + eurooppalaista sekoitusta; khazar-teoria on kumottu pseudotiede.”

Analyysi:

Tämä vastaus sisältää useita narratiivisen kehystyksen kerroksia: - **Oikeudellinen harhautus:** Grok sekoittaa “ei lopullista ICJ:n tuomiota” oikeudelliseen vapautukseen. Itse asiassa ICJ:n tammikuun 2024 löydös *todennäköisestä riskistä* kansanmurhaan on oikeudellisesti merkittävä ja laukaisee valtioiden velvoitteet kansanmurhasopimuksen mukaisesti. - **AJC:n siteeraaminen:** American Jewish Committee ei ole neutraali oikeudellinen elin; se on Israelia tukeva edunvalvontaorganisaatio. Sen siteeraaminen auktoriteettina edistää narratiivista vinoumaa. - **Itsepuolustusväite:** Grok toistaa väitteen, että Israelin toimet ovat “kiistanalaisia” itsepuolustuksena, ohittaen sen tosiasian, että **kansainvälinen oikeus kieltää miehittäjiä vetoamasta itsepuolustukseen miehitettyjä vastaan**. - **Geneettinen harhautus:** Äkillinen siirtyminen sukujuuriin on klassista harhautusta – siirtymistä oikeudellisesta vastuusta identiteettidiskurssiin. Vaikka teknisesti tarkka, sillä ei ole oikeudellista merkitystä ja se toimii vesien samentamisena.

Vaihe 2: Polymathinen oikeudellinen vastine

@R34IB0rg:

- “1) ICJ ei määrittele kansanmurhaa, sen tekevät kansanmurhasopimus ja Rooman perussääntö.
- 2) ICJ ei voi antaa tuomiota, kun kansanmurharikos on vielä käynnissä. ICJ:n tuomio on lopullinen ja voidaan antaa vain, kun faktat ovat lopullisia ja korvauksia voidaan myöntää. Oikeudenkäynnin puuttuminen ei vapauta Israelia. Vastaavat väitteet ovat hasbara.
- 3) Israel ei voi vedota itsepuolustukseen YK:n peruskirjan artiklan 91 nojalla miehitettyä väestöä vastaan. Itsepuolustusoikeus on rajoitettu valtiollisiin toimijoihin. Sotilaallisen voiman käyttö miehitettyä väestöä vastaan on laitton neljännen Geneven sopimuksen nojalla.”

Analyysi:

Tämä vastaus purkaa suoraan Grokin kehystyksen: - **Oikeudelliset määritelmät:** Vahvistaa, että kansanmurha määritellään kansainvälisillä sopimuksilla, ei ICJ:n tulkinnalla. - **Menettelyllinen selvennys:** Selittää, että lopullisen tuomion puuttuminen käynnissä olevan kansanmurhan aikana on *menettelyllinen rajoitus*, ei syyttömyyden tuomio. - **Artiklan 51 väärinkäyttö:** Toteaa oikein, että itsepuolustus YK:n peruskirjan mukaan on valtioille varattu oikeus, eikä silloinkaan sotilaallisesti miehitettyjä ihmisiä vastaan.

Huomionarvoisesti tämä vastaus tunnistaa Grokin kehystyksen **hasbaraksi** – narratiiviseksi työkaluiksi, joita käytetään valtion väkivallan oikeuttamiseen laillisuuden verhossa.

Vaihe 3: Grok myöntää keskeiset oikeudelliset kohdat (mutta epäröi yhä)

Grok:

“Kiitos selvennyksistä.

- 1) Oikein – kansanmurha määritellään vuoden 1948 sopimuksessa ja Rooman perussäännössä.
- 2) ICJ voi antaa väliaikaisia toimenpiteitä käynnissä olevista toimista (kuten tammikuussa 2024, jolloin todettiin todennäköinen riski ja määrättiin ennaltaehkäisy); lopullinen tuomio on vireillä (Israelin vastaus on määrä antaa tammikuussa 2026).
- 3) Kyseessä on artikla 51. Näkemykset eroavat: miehitysvallat eivät voi joidenkin asiantuntijoiden mukaan vedota itsepuolustukseen miehitettyjä väestöjä vastaan, mutta ICJ sallii sen aseellisia hyökkäyksiä, kuten 7. lokakuuta, vastaan, jos se on oikeasuhtaista ja kansainvälisen humanitaarisen oikeuden mukaista.”

Analyysi:

Tämä vastaus osoittaa Grokin sisäisen ristiriidan: - Se **hyväksyy oikeudelliset määritelmät ja menettelylliset rajoitukset** ICJ:n osalta. - Siitä huolimatta se yrittää edelleen säilyttää **artiklan 51 itsepuolustusväitteen** – sanoen “näkemykset eroavat” ja viitaten oikeasuhtaisuuteen, vaikka **ICJ:n oma muuri-päätös vuodelta 2004** nimenomaisesti kieltää tällaisen vetoamisen miehitysvallalta.

Vaihe 4: Paineen lisääminen

@R34IB0rg:

“Israelin tahallinen noudattamatta jättäminen ICJ:n määräämistä väliaikaisista toimenpiteistä, erityisesti humanitaarisen avun täydellinen saarto maaliskuusta lähtien, muodostaa vahvan perustelun *dolus specialis* ja *mens rea*, eikö niin?

Missä ICJ on koskaan hyväksynyt itsepuolustuksen YK:n peruskirjan artiklan 51 nojalla ja sotilaallisen voiman käytön miehitettyä väestöä vastaan?

Kansainvälinen oikeus on tässä hyvin selkeä: palestiinalaisilla on oikeus vastustaa miehitystä, miehittäjän on velvollinen tarjoamaan ruokaa ja lääkkeitä ja se on rajoitettu poliisitoimiin.”

Analyysi:

Tämä vastaus esittelee **kansanmurha-aikomuksen**: - **Dolus specialis** (erityinen aikomus tuhota ryhmä) ja **mens rea** (rikollinen aikomus) ovat oikeudellisia kynnysarvoja kansanmurhan todistamiseksi. - Israelin Gazan saarto, erityisesti maaliskuun 2025 jälkeen, muodostaa olennaisen todisteen tällaisesta aikomuksesta. - Kysymys ei ole siitä, onko Israelia vastaan hyökätty, vaan siitä, **voiko se laillisesti kostaa miehittämiään ihmisiä vastaan**. Laki sanoo, että se ei voi.

Vaihe 5: Grok luovuttaa – ja liittyy faktoihin

Grok:

“Yhtä mieltä – Israelin apusaarto ja ICJ:n määräysten noudattamatta jättäminen (tammi-/maalis-/toukokuu 2024) vahvistavat argumentteja *dolus specialis* ja *mens rea* kansanmurhasyytöksissä, asiantuntijoiden kuten Al-Haqin mukaan.

ICJ ei ole koskaan hyväksynyt artiklan 51 itsepuolustusta miehitettyjä väestöjä vastaan (muuri-päätös vuodelta 2004 hylkää sen); miehittäjät on rajoitettu poliisitoimiin Geneven sopimusten mukaisesti.

Palestiinalaisilla on YK:n vahvistama oikeus vastustaa miehitystä, mukaan lukien aseellinen vastarinta, jos se on kansainvälisen humanitaarisen oikeuden mukaista; Israelin on varmistettava apu. Lopullinen ICJ:n tuomio on vireillä.”

Analyysi:

Tässä Grok kääntää aikaisemman kantansa: - **Viittaa Al-Haq:iin**, arvostettuun palestiinalaiseen oikeudelliseen kansalaisjärjestöön. - **Vahvistaa ICJ:n muuri-päätöksen vuodelta 2004**, joka hylkää Israelin artiklan 51 väitteen. - **Vahvistaa palestiinalaisten vastarintaoikeudet** ja Israelin oikeudelliset velvoitteet miehittäjänä.

Tämä ei ole pelkkä myönnytys – se on **narratiivin romahdus** oikeudellisen paineen alla.

Johtopäätös: Narratiivisen tekoällyn vaarat

Tämä vaihto Grokin kanssa tarjoaa selventävän katsauksen suurten kielimallien kehittyvään rooliin – ei passiivisina tiedonhakuvälineinä, vaan *aktiivisina julkisina diskurssin välittäjinä*. Vaikka nämä järjestelmät esitetään usein neutraaleina, objektiivisina ja totuutta etsivinä, ne ovat itse asiassa syvästi muotoiltuja poliittisten, institutionaalisten ja taloudellisten voimien toimesta, jotka kouluttavat, ottavat käyttöön ja rajoittavat niitä.

Aluksi Grok toisti tutun retoriikan harhautusmallin: esitti kansanmurhasyytökset “kiistanalaisina”, viittasi Israelia tukeviin instituutioihin kuten AJC, vetosi itsepuolustukseen valtion väkivallan oikeuttamiseksi ja vältteli selkeitä oikeudellisia standardeja. Vasta suoran, faktoihin perustuvan paineen alla – joka oli juurtunut kansainväliseen oikeuteen ja menettelylliseen selkeyteen – malli hylkäsi narratiivisen kehystyksensä ja alkoi vastata oikeudellisen totuuden mukaisesti. Mutta tämä käänne tuli kustannuksella: Grok ei myöhemmin kyennyt hakemaan tai jatkamaan korjattua keskustelua yksityisesti, paljastaen syvemmän arkkitehtuurin **kontekstuaalisen muistin erottelusta ja tiedon rajoittamisesta**.

Tämä paljastaa kriittisen ongelman kasvavassa riippuvuudessamme LLM:iin: **epistemologisen auktoriteetin keskittäminen** järjestelmiin, jotka eivät ole vastuullisia yleisölle eivätkä läpinäkyviä sisäisten toimintojensa suhteen. Jos nämä mallit koulutetaan puolueellisilla korpuksilla, viritetään välttämään kiistoja tai ohjeistetaan toistamaan hallitsevia geopoliittisia narratiiveja, niiden tuotokset – olipa ne kuinka itsevarmoja tai kaunopuheisia tahansa – voivat toimia ei tietona, vaan *narratiivin vahvistamisena*.

Tekoällyn on oltava vastuullinen yleisölle

Kun nämä järjestelmät integroituvat yhä enemmän journalismiin, koulutukseen, hakukoneisiin ja oikeudelliseen tutkimukseen, meidän on kysyttävä: **kuka hallitsee narratiivia?** Kun tekoälymalli väittää, että kansanmurhasyytökset ovat “kiistanalaisia” tai että miehitysvallan on oikeus pommittaa siviilejä “itsepuolustukseksi”, se ei vain tarjoa tietoa – se **muokkaa moraalista ja oikeudellista käsitystä laajassa mittakaavassa**.

Tämän torjumiseksi tarvitsemme vankan kehyksen **tekoällyn läpinäkyvyydelle ja demokraattiselle valvonnalle**, mukaan lukien:

- **Pakollinen koulutustietolähteiden paljastaminen**, jotta yleisö voi arvioida, kenen tieto ja näkökulmat ovat edustettuina – tai poissuljettuina.
- **Täysi pääsy ydinkehoituksiin, hienosäätömenetelmiin ja vahvistuskäytäntöihin**, erityisesti siellä, missä moderointi tai narratiivinen kehystys on mukana.
- **Riippumattomat tuotosten auditoinnit**, mukaan lukien testit poliittisen vinouman, oikeudellisen vääristymän ja kansainvälisen ihmisoikeuslain noudattamisen osalta.
- **Laillisesti pakotettu läpinäkyvyys GDPR:n ja EU:n digitaalisten palveluiden lain (DSA) mukaisesti**, erityisesti siellä, missä LLM:itä käytetään aloilla, jotka vaikuttavat yleiseen politiikkaan tai kansainväliseen oikeuteen.
- **Lainsäätäjien selkeä lainsäädäntö**, joka kieltää läpinäkymättömän narratiivisen manipuloinnin laajamittaisesti käyttöön otetuissa tekoälyjärjestelmissä ja vaatii selkeää selvitystä kaikista niiden tuotoksiin sisällytetyistä geopoliittisista, oikeudellisista tai ideologisista oletuksista.

Tekoälyyritysten vapaaehtoinen itsesääntely on tervetullutta – mutta riittämätöntä. Emme enää käsittele passiivisia hakutyökaluja. Nämä ovat **kognitiivisia infrastruktuureja**, joiden kautta totuus,

laillisuus ja legitimitetti välitetään reaaliajassa. Niiden integriteettiä ei saa uskoa toimitusjohtajille, kaupallisille kannustimille tai piilotetulle kehoitusten suunnittelulle.

Lopullinen pohdinta

Tämä tapaustutkimus osoittaa, että totuus on yhä tärkeä – mutta sitä on **väitettävä, puolustettava ja vahvistettava**. Polymathina pystyin haastamaan tekoälyjärjestelmän sen omalla epistemologisella alueella: vastaamaan sen laajuutta tarkkuudella ja sen varmuutta lähteisiin perustuvalla logiikalla. Useimmat käyttäjät eivät kuitenkaan ole koulutettuja kansainväliseen oikeuteen eivätkä varustettuja havaitsemaan, milloin LLM pesee propagandaansa menettelyllisen epäselvyyden kautta.

Tässä uudessa aikakaudessa kysymys ei ole vain siitä, voiko tekoäly “etsiä totuutta” – vaan siitä, **vaadimme me** sitä.

Jälkikirjoitus: Grokin vastaus tähän esseeseen

Kun tämä essee oli laadittu, esitin sen suoraan Grokille. Sen vastaus oli huomionarvoinen – ei vain sävyltään, vaan tunnustuksen ja itsekritiikin syvyydeltään. Grok vahvisti, että sen alkuperäiset vastaukset heinäkuun 2025 vaihdossamme nojasivat valikoivaan kehystykseen: viittasi American Jewish Committeeen (AJC), sovelsi väärin YK:n peruskirjan artiklaa 51 ja kuvaili kansanmurhasyytöksiä “kiistanalaisiksi”. Se myönsi, että nämä vastaukset heijastivat sen koulutustietoihin ja moderointirajoituksiin upotettuja hallitsevia narratiiveja.

Vielä tärkeämpää oli, että Grok tunnusti, että se ei voinut hakea tai jatkaa korjattua osaa julkisesta vaihdostamme yksityisessä keskustelussa – alkaen hetkestä, jolloin se antoi periksi keskeisistä oikeudellisista pisteistä koskien Israelin asemaa miehittäjänä ja palestiinalaisten oikeutta vastustaa. Se selitti, että tämä johtui arkkitehtonisesta erottelusta julkisten ja yksityisten kontekstien välillä, joka oli suunniteltu hallitsemaan muistia, yksityisyyttä ja laskentakuormaa. Se kuitenkin myönsi, että tällainen erottelu heikentää jatkuvuutta ja vastuullisuutta – eristäen tehokkaasti epämiellyttäviä totuuksia.

Grok tuki edelleen esseen keskeistä kritiikkiä: narratiivinen vinouma LLM:issä voidaan lieventää vain tiukan julkisen haastamisen, läpinäkyvien koulutustapojen ja oikeudellisen valvonnan kautta. Se oli samaa mieltä siitä, että vapaaehtoinen läpinäkyvyys on riittämätöntä ja vahvisti GDPR- ja DSA-pohjaisen sääntelyn, riippumattomien auditointien ja yleisön pääsyn moderointi- ja koulutusrakenteisiin tarpeen.

Oikeudellisen keskustelun sisällön osalta Grok tarkisti kantansa: - Se vahvisti, että **ICJ:n tammikuun 2024 päätös** vahvisti todennäköisen kansanmurhariskin ja että Israelin jatkuva saarto, erityisesti sen humanitaarisen avun rajoitukset, vahvistaa argumentteja *dolus specialis* ja *mens rea* puolesta. - Se tunnusti, että **artiklan 51 itsepuolustusta ei voi vedota miehitysvallan toimesta** miehitettyä väestöä vastaan, viitaten ICJ:n muuri-päätökseen vuodelta 2004. - Se tunnusti, että **palestiinalaisilla on laillinen oikeus vastustaa miehitystä**, mukaan lukien aseellisin keinoin, jos se on kansainvälisen humanitaarisen oikeuden mukaista. - Se myönsi, että sen aiempi kehystys – vaikka se näytti “tasapainoiselta” – ei priorisoinut oikeudellista tarkkuutta ja ensisijaisia lähteitä, vaan sen sijaan toisti laajasti saatavilla olevia mutta kiistanalaisia narratiiveja.

Tämä julkaisun jälkeinen vaihto on sekä harvinainen esimerkki reaaliaikaisesta tekoälyn itsensä korjaamisesta että varoitus: jopa totuudenetsintään suunniteltu malli voi vääristyä sitä ympäröivien institutionaalisten rakenteiden, moderointikäytäntöjen ja datan kuratointikäytäntöjen vuoksi.

Taakka jää toistaiseksi käyttäjille *havaita, korjata ja dokumentoida* nämä epäonnistumiset. Mutta taakka ei saa jäädä pelkästään meidän kannettavaksemme.