una de las personas más influyentes en el desarrollo de las tecnologías en las que se basa La Inteligencia artificial abandona Google para divulgar sobre sus peligros palantir lanza su plataforma de Inteligencia artificial aip que ofrece un chatbot al más puro estilo chat gpt para tomar acciones bélicas en el campo de batalla ya de vuelta después de una pequeña pausa involuntaria ya tienes un nuevo episodio de tierra de hackers comenzamos Hola hola y bienvenidos a tierra de hackers tu noticiero de ciberseguridad hecho podcast publicamos este episodio el 22 de mayo de 2023 es el episodio número 94 yo soy Martín vigo lo estaba leyendo o sea está leyendo mi nombre y está aquí mi compañero fiel al otro lado del aparato Don Alexis porrosol Alexis qué tal bienvenido de vuelta muy bien Martín gracias gracias Sí eso de que ya no sabes ni Cuál es tu nombre porque ya te piensas que eres tierra de hackers no porque tanto decimos más somos tierra de hacker que nuestros nombres pero sí nada un gusto estar de nuevo de vuelta con nuestros oyentes disculpad la pausa pero volvemos Con más ganas que nunca y lo de siempre Gracias por estar con nosotros en todos estos episodios Incluso en esta parada técnica que hemos tenido no solo escuchándonos sino también dejándonos vuestras sugerencias comentarios todo el feedback que nos dejáis para seguir mejorando episodio tras episodio Y de nuevo gracias y para aquellos que se hayan Unido hace poco estamos en todas las redes sociales más populares para hacer un resumen Rápido como tierra de hackers o arroba tierra de hackers también en todas las plataformas de podcast que si no lo estáis ahora mismo y por favor a suscribiros para que os llegue la notificación cuando este nuevo episodio salga calentito del horno y también tenemos el servidor de discord en el que podéis acceder vía tierra de hackers.com/di discord y lo de siempre del final de la intro agradecer vuestro apoyo a la pregunta del episodio que publicamos en Twitter como en cada episodio y la del anterior fue la siguiente has usado tiktok alguna vez una pregunta simple sin trampa Así que directo al grano tenemos un 52% con No jamás creo que me ha sorprendido bastante Un 30% con sí una cuenta 10% si con cuenta de otro está está también me parecía interesante y un 8% aún no Pero supongo el pero es me lo estoy pensando o bueno si aún si a más gente interesante que me atraiga Pues igual me hago una no pero ahí queda qué bueno sí sí que me sorprende a mí también que haya alguna mitad de oyentes que no hayan usado tiktok A lo mejor eso solo habla de la media de edad de nuestros oyentes yo por mi parte como dices tú también disculpas que estuvimos dos dos semanas sin sin publicar debido a que por un tema personal mío pues me ha ocupado literalmente las 24 horas del día Aparte de estar en la conferencia o sintomáticos me fue imposible pero aquí seguimos creo que en los tres años de vida que tenemos Es la primera vez que no publicamos ahí en dos fines seguidos Así que Que bueno que lo dicho disculpas pero nos Mola mucho que nos escribáis preguntando Oye qué pasa Qué pasa tierra de hackers mencionar que en asintomáticos estuve dando una charla y fue como siempre digo increíble reunirme con tantos y tantos oyentes quería mencionar especialmente uno que en la ronda de preguntas Solo hubo tiempo para una y Alexis me preguntó Bueno me hice una pregunta referente a la charla que hice pero hizo dos preguntas según ahí la segunda era que cómo se deletreaba proliferando así delante de 500 personas me partí de risa buenísimo y le dije Oye te mencionaré en el podcast tío que me ha molado mucho pero sí fue fue muy divertido y luego pues eso hablando con muchísima gente que escucha el podcast que le Mola el feedback es brutal y nada aquí dos preguntas deletreaste proliferando o profilerando o no dijiste aproveché me reí y ni me atreví O sea me re venga contesto a la otra pregunta no vaya a ser que me equivoque pero viste que ya me sale prolifera muy bien bueno y antes de empezar Ya con la noticia para no liarnos demasiado darle las gracias a nuestros mecenas de patreon que tenemos dos estas esta semana nuevos dos nuevos César auris y Diana Gracias Muchísimas gracias por uniros a nuestro grupo de patreons a nuestros oyentes favoritos Y además que hemos perdido nuestro editor que se ha movido a le han salido otros

proyectos que lo ocupan todo el tiempo y no puede seguir con nosotros Así que estamos a la búsqueda de editor que lo puso por ahí por discord y nada vuestro apoyo en patreon nos ayuda un montón para para poder pagarlo Así que muchísimas gracias y por supuesto también tierra que eso no podría existir sin nuestros sponsors como hombre branding una empresa formada por especialistas en varios ámbitos profesionales que se enfocan la reputación online a múltiples niveles han ayudado desde personas como tú y como yo querido oyente hasta famosos a llevar a juicio casos de ciberacoso mitigar situaciones donde la reputación de empresas estaba siendo dañada o incluso a borrar la huella digital que dejamos online no Solo han decidido Apoyar el podcast sino que si le contáis que venís de parte de tierra de hackers tendréis un descuento especial en sus servicios si necesitáis algún tipo de ayuda con vuestra identidad digital om branding es lo que estáis buscando Así que visita o branding.es o nbrand y n g punto es y también queremos darle las gracias a brawler Pro por ayudarnos en el podcast brawler pro es la herramienta más completa de seguridad en la WS y empresas de todos los tamaños grandes pequeñas y medianas se apoyan diariamente en brawler pro para que sus equipos puedan confiar en su modelo de seguridad de aws puedes probar bra hoy mismo y de manera totalmente gratuita Por qué Porque obtendrás paneles y gráficas con información concisa y accionable con todo lujo de detalles sobre la madurez de tu modelo de seguridad y una visión completa de tu infraestructura en todas las regiones de aws y además tendrás todos los resultados en apenas unos minutos empieza a usar brawler pro y beneficiarte de sus resultados visitando tierra de hackers.com barra brawler proprowl e rpro Muy bien pues vamos al lío que ha sido una intro un poquito larga pero que quería explicar el Por qué habíamos tardado en publicar esta vez volvemos al término del momento el palabro la expresión el concepto no que ilusiona a unos y bueno asusta a otros Inteligencia artificial hace cosa de un mes hablamos de como un grupo de expertos publicó y firmó una carta gritando al cielo no sus temores sobre el rápido avance de esta tecnología os contaba como en los últimos seis meses hemos visto tanta Innovación que han incluso hasta creado un Buscador de inteligencias artificiales una manera de Buscar todos los productos que están saliendo que hacen nuestra vida más fácil y tienen su fundamento pues bueno en la Inteligencia artificial traductores de texto impecables generadores de páginas web profesionales con solo pedirles lo que lo que necesitas generación de imágenes hiperrealistas cronado de Voces generación de contenido multimedia etcétera etcétera es innegable que estamos viviendo un punto de inflexión en la informática y yo me atrevería a decir que al menos una época que quedará grabada en los libros de historia como cuando llegó internet seguramente la Inteligencia artificial ya no es solo un tema académico una prueba de concepto una herramienta al alcance solo de los que pueden pagar a la Inteligencia artificial al igual que internet y los dispositivos móviles ha pasado de ser una tecnología al alcance de muy pocos a ser una herramienta más del día a día y ya de hecho acaba de Open eyes sacar la aplicación de chat gpt para para el móvil para iOS así con ella puedes tener Chan gpt desde el móvil y es súper útil y hoy volvemos a hablar de Inteligencia artificial porque Jeffrey Hilton considerado uno de las personas más influyentes en cuanto a Inteligencia artificial se refiere ha abandonado Google para en parte dedicarse como divulgador a alertar de los peligros de la Inteligencia artificial Jeffrey Hilton es una eminencia académica que como estudiante en los años 70 ya empezó merodear con la idea de lo que hoy de lo que soy la Inteligencia artificial y que gracias a él bueno conocemos como redes neuronales un sistema matemático de aprendizaje en base al análisis de datos curiosamente en esa época no muchos creían en la viabilidad de este sistema Pero está claro que estaban equivocados en los años 80 ya y de hecho con Jeffrey como doctorando de la Universidad de carnegie malone decidió abandonar dicha universidad Porque todos los fondos de investigación de Inteligencia artificial venían del pentágono en los años 80 y eso a él le

incomodaba de aquella y bueno a día de hoy está totalmente en contra del uso de la Inteligencia artificial en el campo de batalla lo que él denomina la creación de soldados robots ya en 2012 hace unos bueno 10 años no Jeffrey y dos de sus mejores estudiantes desarrollaron una red neuronal que era capaz de analizar miles de fotos y enseñarse a sí mismo es decir la propia Inteligencia artificial aprendía tenía la capacidad de aprender a identificar objetos como flores perros y coches insisto al igual que aprende un bebé viendo muchas veces gatos y perros Pues lo mismo la Inteligencia artificial fue tal el éxito que Google compró la empresa que montó en torno a esta tecnología por 44 millones de dólares o sea el Jeffrey este montó un montón son presillas que viró Google y dijo ven para acá cuánto quieres de hecho en 2018 tanto él como sus dos estudiantes que mencionabas recibieron el premio turing que es el más prestigioso en el mundo de la informática lo que sería como el equivalente al Premio Nobel en esta disciplina No pues bien joffrey se quedó en Google y uno de sus estudiantes aceptó un puesto como el jefe de investigación en Open ella hay la empresa detrás del archifamoso charge gpt y que está basado precisamente las redes neuronales desarrolladas por ellos con los jugadores en sus posiciones Google opening Y por supuesto otras empresas empezaron a trabajar seriamente en el desarrollo de redes neuronales aprendiendo de grandes cantidades de texto en ese momento nuestro querido protagonista de esta historia todavía pensaba que no Sería posible desarrollar una tecnología más capaz que el cerebro humano y que estos sistemas serían muy útiles Pero según peligro real Pero esto cambió el año pasado cuando Google y opening empezaron a entrenar sus modelos con muchísimos más datos es como si tienes un bebé que casi no le enseñas nada Y pasas de repente a constantemente estar enseñándole cosas no pues ese bebé va a aprender más si bien todos no en todos los aspectos Geoffrey empezó a pensar que la Inteligencia artificial estaba adelantándose a las capacidades de nuestro cerebro humano me quedo con esta frase que dijo en una entrevista al New York Times para justificar el cambio de postura el pasar de bueno las redes neuronales que yo he inventado Y que ahora se utilizan en Inteligencia artificial Nunca van a pasar la capacidad humana a la postura que tiene ahora de ostras ojo cuidado Pues dijo esto Mira cómo era la tecnología hace tan solo cinco años y mira cómo es Ahora ten en cuenta Esa diferencia y propágala al futuro asusta ese asusta por cierto no es mío es parte de sus declaraciones y aprovecho recordar que este es el tío que inventó los fundamentos de todo esto Ese es el tío que está que está diciendo esta frase uno de los problemas a los que apunta Jeffrey es que en su momento Google era un poco quien dictaminaba la dirección y las pautas del desarrollo tecnológico de la Inteligencia artificial pero apunta que desde que Open ella y otros competidores han empezado a trabajar en lo mismo Google se ha visto en un aprieto ya que su modelo de negocio las búsquedas en internet de repente corre peligro porque ahora ya le puedes preguntar directamente a la Inteligencia artificial a echar gpt no tienes que hacerle una pregunta Google y buscar entre los links a ver dónde está la respuesta por tanto Google ha dejado de lado el camino seguro y se ha puesto a investigar y publicar avances sin los controles rigurosos que solía aplicar antes de las publicaciones antes de que tuviese esta competencia Esta es una de las razones por las que josffrey ha abandonado la empresa de hecho su temor principal en torno a qué puede suceder si las tecnologías de Inteligencia artificial siguen avanzando sobre todo a este ritmo coincide de hecho con la que yo tengo en mente hace mucho tiempo fake news como dice él las fotos vídeos y texto que ya nos inundan con falsedades llegará un punto en que ya no tendremos la capacidad de determinar Qué es verdad y que no es como decía coincido de pleno yo también creo que eso será el riesgo más inmediato en un mundo donde tenemos a gente creyendo en conspiraciones como el terraplanismo las elecciones amañadas y Trump sigue siendo supuestamente presidente antivacunas ritos satánicos llevados a cabo por famosos y de más sin sentidos Cómo creéis que

va a afectar que las pruebas no entre comillas distribuidas por internet para para que la gente se crea estas conspiraciones sean realmente imposibles de distinguir de la realidad que se puedan crear vídeos de Obama junto a Jeffrey epstein imágenes de la NASA falseando la luisaje audios de declaraciones de Putin hablando de la implantación del nuevo orden mundial etcétera etcétera con entre comillas evidencias no fácilmente atribuibles a un bulo ya tenemos a miles y miles de personas creyendo las cosas más locas cómo será cómo será cuando no podamos demostrar que el material fotográfico de vídeo de audio sea imposible de demostrar que fue generado por Inteligencia artificial sesgo de confirmación aumentará exponencialmente y desde luego las escenas de Navidad pasarán a ser todavía más divertidas hablando con el cuñado de turno no pero bueno fuera bromas los fake news ya han causado estragos en nuestra sociedad la Inteligencia artificial facilitará aún más esa tarea comenté que los que los fake news es el temor más inmediato de Jeffrey pero realmente no es el único otro que destaca es el uso militar de la Inteligencia artificial concretamente le preocupa la creación de robots soldados como decía antes y como lo oís el auténtico escenario Terminator skynet no y es que de hecho leyéndome su viendo su entrevista es que Define tal cual skynet comenta que ya permitimos a la Inteligencia artificial escribir código de manera Autónoma Solo nos falta Añadir que lo ejecute también de manera Autónoma y sin supervisión con esto de hecho ya tenemos literalmente el plot de la peli de Terminator y acabó su argumento diciendo lo siguiente y cito textualmente la idea de que esta tecnología podría acabar siendo más inteligente que los humanos pues había gente que lo creía pero otros tantos incluidos yo creíamos que era exagerado que como mucho sería en 50-60 años evidentemente ya no opino lo mismo a día de hoy eso es lo que dice se le pone a uno los pelos de punta Comenta en una entrevista otra que me encontré algo interesante también dice que a diferencia esto es muy bueno a diferencia de las armas nucleares que es relativamente sencillo saber si una nación está intentando avanzar un programa nuclear debido a que requiere de material instalaciones muy grandes y hacerlo en secreto es muy complicado por lo que cualquier pues programa de espionaje de otra nación lo descubriría eso no sucede con la Inteligencia artificial por mucho que el resto del mundo lo usemos de manera responsable una nación o gobierno hostil podría estar trabajando en avances o aplicaciones de la Inteligencia artificial para uso bélico sin que fuera posible que el resto del mundo lo supiera hace unas semanas os traía la famosa carta firmada por muchos expertos avisando de los peligros de la Inteligencia artificial y ahora os traigo a uno de Los Pioneros en el tema diciendo lo mismo pero si esto no es suficiente habemos nueva carta pública En este caso de la asociación por el avance de la Inteligencia artificial y tiene el mismo tono de la carta que os hablé hace un par de episodios un tono un poco menos catastrófico pero el fondo es el mismo os Leo un párrafo un párrafo solo de la carta que se titula trabajando juntos en el futuro del avance de la Inteligencia artificial y que por supuesto os dejo las notas del episodio el párrafo dice así en los últimos dos años la Inteligencia artificial ha revelado la estructura de cientos de miles de proteínas y se está utilizando para mejorar la calidad de la atención en hospitales realizar predicciones detalladas del clima guiar el desarrollo de nuevos materiales y proporcionar a los ingenieros ideas que estimulen su creatividad creemos que la Inteligencia artificial tendrá un impacto cada vez mayor en la atención médica el clima la educación la ingeniería y muchos otros Campos Al mismo tiempo somos conscientes de las limitaciones y Preocupaciones en torno a los avances en Inteligencia artificial incluyendo la posibilidad de que los sistemas de Inteligencia artificial cometan errores ofrezcan recomendaciones sesgadas amanece amenace nuestra privacidad en poder en actores malintencionados con nuevas herramientas y tenga un impacto en el empleo y termino el extracto de esta carta en el tema del empleo porque hace dos semanas ya hemos tenido una prueba de ello ya sabéis que yo cuando preparo las noticias bueno y Alexis también

nos gusta no quedarnos solo en la noticia en sí Sino indagar bastante más y Buscar noticias relacionadas pues como decía acaba de publicarse IBM declarando públicamente que deja de rellenar 7.800 vacantes porque son puestos de trabajo que en unos años saben que podrá llevar a cabo la Inteligencia artificial y todo esto Solo en los próximos cinco años Esa es la estimación de del ceo de IBM pero si todavía ha querido oyente no es suficiente para temer a la Inteligencia artificial una carta firmada por expertos otra carta firmada por una asociación específica de expertos en Inteligencia artificial los avisos del inventor de los fundamentos de los que se basa La Inteligencia artificial ficial o que por ejemplo IBM ya esté anunciando cambios en la estrategia de contratación os dejo con las declaraciones del ceo de Open gpt frente al congreso de los diputados de los Estados Unidos hace unos días el audio está en inglés pero os digo Solo dos frases que dice el ceo de oppenhei mi mayor temor es que nosotros la tecnología la industria cause un daño muy significativo al mundo si esta tecnología sale mal saldrá muy mal como os digo Siempre Cómo se te queda el cuerpo Pues mira Dímelo contestando a la pregunta de este episodio A dónde nos lleva la Inteligencia artificial y os doy Solo dos opciones a un mundo mejor o a un mundo peor y para contestarlo ya sabéis os vais a Twitter arroba tierra de hackers y a ver qué nos contáis Y por supuesto os dejo la comparecencia entera en las notas del episodio del ceo de Open que está muy bien hasta aquí Hemos llegado así que recordad contestar ahí a la pregunta del episodio A dónde nos lleva al futuro o al pasado pues eso es una una manera bastante poética de ponerlo porque el futuro lo interpretamos como mejor el pasado Es como volver a la Edad de Piedra no que siempre se dice que la tercera guerra mundial va a ser pedradas Bueno sí también sí cibernética primero y luego pedrada porque ya habremos destrozado todo ya ya no queda de hecho había había visto una imagen por Twitter que era muy guapa que lo representaba como el típico la típica imagen de los días tíos Cavernícolas con una lanza y la lanza el pico estaba hecho de una placa base sabes como recortada está muy bien Es como volvemos a la edad Cavernícola pero con los restos todavía y de los ordenadores y de movidas tipo Mad Max o alguna de esas futuristas hablando de a mí que siempre me gusta comentar alguna serie que otra estoy viendo una que se llama sailo o silo en español en Apple TV y un poquito es de que unas no sé unas 10.000 personas se despiertan un día y están dentro de una especie de silo bajo tierra y no saben cómo han llegado hasta ahí y unos fundadores según los llaman han creado eso ahí y Bueno hay situaciones en las que se encuentran le llaman en reliquias Oh mira esta reliquia Qué es o déjame que voy a averiguar Qué es una chica que siempre está tocando objetos antiguos pero en este caso era una una videocámara de estas antiguas de hace 10 años de grabar y lo llama reliquia no sabía ni qué era o sea vamos en el futuro Igual cuando después de la tercera guerra mundial vamos a estar así en plan qué es esto Esto hace mil años que no lo veo porque ya no se producen o bueno lo que sea anyway muy buena noticia Martín lo de la lo de aplicación de opening lo he visto justo esta mañana antes de que habláramos y todavía no más tiempo ni de usarla pero algo que incluye también es que incluye lo del como lo llaman ellos whisper O sea que le puedes hablar te reconoce lo que le hablas y lo traduce a texto y se lo envía como un chrometa a la ella y te responde No si te responde por voz o no pero al menos te responde así no tienes que escribir lo que que es más fácil y más rápido y sobre además Pues la verdad es que no sé si comentar mucho más al respecto porque la mía mi noticia también va de Inteligencia artificial Así que estamos dejando a los oyentes fritos contentos con todo esto con todos estos avances todo todo para adelante para adelante mis valientes así que no sé si quieres comento la mierda también y luego si quieres hacemos una un comentario combinado un debatillo Vale pues antes de que pases a tu noticia solo darle las gracias a mónad una empresa que comparte los mismos valores que tierra de hackers hacer la seguridad más accesible y transparente nosotros a través de un podcast y monat con una herramienta de

gestión y visualización de telemetría y datos de seguridad fundada en Valley y que está buscando ingenieros con experiencia en ciberseguridad para ayudarles a construir y hacer realidad su misión contrata que en todo el mundo y en remoto así que ya sabéis echarle un vistazo a su web monat. com y mandarles vuestro currículum a tierra de hackers @monat.commonade.com Pues nada seguimos con la siguiente noticia que va de palantir De hecho hace Estaba mirando en nuestros archivos y la única vez que hemos mencionado esta empresa según parece es fue en el episodio 6 en junio de 2020 casi hace tres años Bueno ahí en esa ocasión en ese episodio comentábamos el tema de el análisis de datos del Sistema Nacional de salinidad británico y cómo lo estaban utilizando bueno para abusar un poco de los pacientes y similares pero en este caso como he dicho la noticia va de palantir y temas de Inteligencia artificial hace un par de semanas esta empresa que es la compañía del multimillonario Peter fiel lanzó palantir artificial intelligence platform a la que me voy a referir de ahora en adelante como aip una plataforma software diseñada para utilizar los modelos grandes de lenguaje en inglés large Language models como gpt 3.5 o gpt 4 o flan t5 o todos estos y similares en las redes privadas de sus clientes en las redes privadas de sus clientes que para lo que no lo sepáis la mayoría de los clientes de palantir son gobiernos y sobre todo los departamentos militares o las los grupos militares básicamente como dato curioso comentar que una palantir es un objeto ficticio del mundo del señor de los anillos de tolkien que es una piedra Vidente de color negro que parece vidrio o cristal que se utiliza como una forma de comunicación casi como un sistema de videoconferencia algo interesante bueno en esta plataforma no se utiliza tanto para comunicarse como Vais a entender Ahora cuando le explique más detalle pero bueno es interesante el nombre Supongo que lo cogieron Peter fiel cogió el nombre de palantir para su empresa como algo que puede ver todo lo que sucede en todo el mundo en uno de los vídeos de lanzamiento de la plataforma aip para Antier demostró como un ejército podría usar esta plataforma para luchar en la guerra en el vídeo el operador usa un chatbot estilo chat gpt para ordenar el reconocimiento de drones generar varios planes de ataque y organizar la interferencia o Jamie de las comunicaciones enemigas en el escenario de palantir un operador militar responsable de monitorear la actividad dentro de Europa del este recibe una alerta de la plataforma de que un enemigo está acumulando equipo militar cerca de fuerzas amigas o fuerzas aliadas luego el operador le pide al chat Bot que le muestre más detalles obtiene un poco más de información y luego le pide a la Inteligencia artificial que adivine Cuáles podrían ser las unidades enemigas el operador pregunta qué unidades enemigas hay en la región y aprovecha la Inteligencia artificial para construir una probable formación de unidades opuestas una estrategia Cuál es básicamente lo que está preguntando el operador es cuál puede ser el siguiente paso de nuestros enemigos después de obtener la mejor suposición de la Inteligencia artificial sobre lo que está sucediendo el operador le pide a la misma que tome mejores fotografías porque Supongo que la zona supongamos que esa zona no se tienen fotografías actuales probablemente tenga fotografías de hace tiempo pero las queremos de ahora no en tiempo real de saber qué está pasando Incluso si puede ser vídeo mejor pues lo que hace es lanzar un dron reaper mq9 de reconocimiento para tomar fotos y vídeo y descubre que sorpresa hay un tanque t80 un vehículo ruso cerca de las fuerzas aliadas luego el operador pregunta al chatbot que hacer al respecto y lo que hace el operador es usar la plataforma para generar tres posibles planes de acción para contrarrestar a este equipo enemigo todo esto sale en el vídeo no me lo estoy inventando es una situación hipotética no pero lo podéis ver lo vamos a poner en las notas del episodio para que lo reviséis acto seguido bueno como digo la plataforma le proporciona tres posibles planes de acción que incluyen la primera es atacar el tanque con un f-16 que es un avión de combate la segunda opción es utilizar artillería de largo alcance y la tercera opción es utilizar misiles havelin o javeling que son misiles antitanque

después de revisar estas tres opciones utiliza la plataforma a IP para enviarlas a la cadena de Mando es decir hace clic y las tres opciones que le han parecido bien al operador si no se entiende que Podría preguntarle del nuevo a la Inteligencia artificial Oye estas tres no me gustan cambia la primera o cambia las todas y vuelve a darme tres opciones bueno en este caso le parecieron bien y las envió para arriba digamos según el vídeo La Inteligencia artificial también determina si las tropas cercanas tienen suficientes jabalinas por ejemplo en la opción 3 que le dice utiliza estos misiles antitanque de estas jabalíns Pues también determina si esa opción es viable o no y la cantidad de misiles que se tienen en un Rango cercano para poder realizar ese ataque la plataforma de Inteligencia artificial está diseñada para que sus usuarios obtengan una ilusión de seguridad y control de sus activos militares y bueno Así ha sido como lo presentan el punto opuesto digamos a la empresa no porque te da un poco de ilusión de seguridad de control pero como voy a comentar en breve como todo el sistema software puede tener fallos según palantir en su nota de prensa comentaba que los lms y los algoritmos deben controlarse en este contexto altamente regulado y sensible para garantizar que se usen de manera legal y ética esto lo comenta porque desde un principio cuando ha publicado esta plataforma dice que todo lo que está haciendo respecto a contraofensiva militar utilizando Inteligencia artificial se basa en la ética y la moral según palantir hayp podría implementar estos sistemas en redes clasificadas y dispositivos en el borde táctico a lo que se refiere Con esto del borde táctico es que implementaría desplegaría esta Inteligencia artificial dentro de digamos aviones de guerra tanques e incluso sistemas cibernéticos de soldados como armas o sistemas de visión de realidad aumentada de hecho lo que ha mencionado brevemente Martín sobre Geoffrey que se oponía un poquito darle superpoderes a los soldados No pues en todos esos sistemas digitales también se podría incluir esta Inteligencia artificial afirma que podrá analizar datos clasificados y en tiempo real de manera responsable legal y ética de nuevo y según el vídeo los usuarios tendrán control sobre lo que cada lm y modelo de foundation model de Inteligencia artificial respaldado por palantir por todo lo que puede hacer es decir y de hecho esto se ve en el vídeo sales se puede definir lo que el operador puede hacer y lo que no puede hacer A través de la Inteligencia artificial y también se genera un registro digital seguro de las operaciones se ve el nombre de la persona que está interactuando con la Inteligencia artificial y Qué medidas toma O sea que hay un registro digital que se puede auditar luego para porque en Casos de guerra siempre puede haber acciones inesperadas y esto Bueno pues lo añaden en la plataforma para para que se aseguren de que todas las acciones que se tomen sean de forma responsable legal y ética y sobre todo para temas legales y regulatorios también porque hay que si se pide evidencia pues se necesitan esos blogs esos registros el vídeo personalmente es bastante interesante o espectacular es como un videojuego estrategia tipo starcraft o League of Legends que es un poquito extremo no la comparación Pero se ve un poquito así como es vídeo en tiempo real Pero tú puedes decirle que te muestre las diferentes opciones y se ven las diferentes los diferentes activos de militares en el campo de batalla o tanques o lo que fuera y te pone las líneas Hacia dónde se va a mover O no así que es bastante Interesante pero el símil más directo sería la película World games no O como Martín ha mencionado antes skynet la compañía planea comenzar a ofrecer la primera versión de IP a clientes seleccionados como ellos dicen en las próximas semanas o sea que la plataforma ya la tiene bastante la tiene lista para para venderla básicamente como he dicho antes los principales clientes de la empresa son el gobierno estadounidense y los gobiernos de otros países aliados Pero obviamente los que se van a beneficiar de esta plataforma es son los grupos militares de cada uno de estos países palantir de hecho ha estado promocionando sus ofertas de Inteligencia artificial durante un tiempo pero en su conferencia de resultados de este mes ha usado el término de Inteligencia artificial y sus sinónimos unas 50 veces o sea está

bastante invertido en la Inteligencia artificial y desde el 2019 palantir ya le ha puesto el ojo a la Inteligencia artificial para usarla en sus servicios comerciales de defensa en vídeos en su sitio web la compañía expuso como el producto podría usarse en un entorno comercial por ejemplo recuperándose de un incidente en la cadena de suministro también para temas de fraude para temas del lavado de dinero y para temas militares como el que estoy comentando en esta noticia sugiriendo estrategias para hacer frente a las amenazas de guerra y Bueno aquí tenemos un poquito la polémica Incluso el ceo de palantir lo reconoce y pidió precaución sobre el uso de Inteligencia artificial en temas de guerra en una carta a sus accionistas escribiendo literalmente la máquina debe permanecer permanecer subordinada a su Creador palantir ha enfatizado que sus herramientas de Inteligencia artificial en realidad no harán nada sin la supervisión humana eso le deja uno bastante satisfecho bastante tranquilo aún así algunos investigadores de Inteligencia artificial han mostrado su preocupación sobre utilizarla en contextos militares y opinan que esta tecnología debe mantenerse alejada de estos entornos sin embargo otros expertos argumentan algo interesante y dicen que el mayor peligro es dejar que un rival geopolítico descubra primero la Inteligencia artificial y la utilice antes que tú en el campo de batalla algo que te deja pensar no en plan deberíamos usar o no pero si la utilizan antes nuestros enemigos que nosotros Entonces qué al final Al fin y al cabo la voy a tener que usar Igualmente Porque si es las usan estoy en desventaja si no la uso Esto me recuerda un poquito como muchos estudiantes últimamente no que la polémica esta que sale en sobre todo en los Colegios o universidades que hay estudiantes que utilizan chat gpt pero los profesores se niegan un poco pero es que casi es imposible Bueno hay formas no de analizar un texto creado con Inteligencia artificial y ver que está creado por Inteligencia artificial Pero hay formas de inspirarse con la Inteligencia artificial y yo creo que si alguien no la está utilizando pues está en desventaja con todos los demás es como decir ahora que vamos vamos a no utilizar Google porque es hacer trampas me parece algo que no debería pensarse de esta forma todo se puede utilizar con una connotación ofensiva incluso Google puede sacar temas de digamos para atacar a otras personas a partir de ahí así que bueno no no entrar más en detalle pero un comentario que dejo Ahí lanzó el aire y quería comentar que en resumen básicamente esta plataforma vamos a pensar un poco ya ya tenemos desde hace tiempo el hecho de que se pueden controlar drones de forma remota para atacar para matar a enemigos Y esto no sé la primera si lo piensas de forma objetiva es algo frío y separado de la realidad Porque ya no estás en el campo de batalla matando a alguien haces un clic y esto ya lo hemos comentado en otro episodio si no me equivoco Pero eso haces un clic y puedes matar a alguien mientras Estás tomándote una coca-cola en casa o lo que sea y no sientes lo mismo no sientes el impacto que si estás delante de la otra persona o del grupo de personas y les disparas con lo que sea no esta plataforma me parece Incluso un nivel más elevado un nivel de abstracción mayor incluso porque estás chateando con un Bot y le dices Oye por qué no disparas a estas personas y ya está lo hace y se sigues tomando la coca-cola y ya está todo la vida continua a para para el operador que está detrás de este de esta plataforma y no siente vamos no siente es es cada vez la guerra se hace más más digital y más fría digamos Y de nuevo se hay que recalcar un poco los problemas que hay en todos estas plataformas de Inteligencia artificial uno de los primeros problemas es el tema de las alucinaciones que ya se ha comentado sobre todo los que quieren concienciar a internet a todo el mundo en sí de del uso de estas plataformas que no hay que tomárselas siempre al pie de la letra Porque pueden tener errores y en el escenario de la plataforma aip de palantir como ha dicho el sí o no existe un humano en el sistema para evitar tales acciones incorrectas a partir de alucinaciones del sistema no Aunque a partir del vídeo que se ve de la demostración de la plataforma el operador parece hacer poco más que aceptar lo que sugiere lo que le sugiere la plataforma el vídeo demostración tampoco

Explica qué pasos están tomando para evitar que los lms en los que se basa el sistema alucinen y proporcionen respuestas incorrectas Y en base a esta información pues el operador decida desplegar una acción ofensiva que impacta a fuerzas aliadas de forma incorrecta en lugar de hacerlo contra las fuerzas enemigas otro tema es qué pasa con ataques como prompt injection que podrían permitir a un operador malicioso saltarse la salvaguardas diseñadas en la plataforma por ejemplo he comentado anteriormente que cuando la plataforma le proporciona las tres posibles opciones de ataque pues estas las el operador las envía a su mando de control a sus superiores para que aprueben o no Entonces sí sí de alguna forma un operador malicioso pudiera saltarse esta aprobación de niveles superiores para ciertas acciones sobre todo las más fatales de lanzamiento de misiles utilizando ataques estos de pronto injection o inyección de instrucciones No pues eso es un riesgo también que habría que barajarlo en todo todo este tema en el episodio 83 el 83 para aquellos que no estéis Al Día de lo que es el tema del promping Ahí comenté Bueno un poco de qué va y algunos ejemplos que ya se han visto y Esto fue hace algunos meses Así que el tema seguro que sigue para más así que si queréis ir a escuchar ese episodio otro tema importante también es la capacidad de interacción de aip o estas plataformas de Inteligencia artificial con sistemas externos por ejemplo el desplegar un dron de reconocimiento o de ataque en el vídeo se ve que el usuario le pide al IP las opciones para obtener imágenes de la zona objetivo a lo que el aip contesta con dos opciones podemos desplegar un dron mq9 que está por aquí cercano o podemos utilizar un satélite para hacer fotografías y vídeo que incluso tiene una mayor resolución Pero va a tardar unos 39 minutos porque está en órbita y hasta 39 minutos no va a llegar a cubrir esta zona Entonces el operador finalmente decide desplegar el dron porque lo quiere ya no quiere esperarse 39 minutos y lo hace haciendo clic directamente en la plataforma de aipe esto significa que esta plataforma tiene acceso directo a interactuar con sistemas externos como el lanzamiento de drones en este caso para fotografiar la zona pero muy probablemente también para lanzar acciones ofensivas como lanzar misiles esta funcionalidad tendría que revisarse con detalle para evitar que atacantes puedan saltarse la autorización del operador a través del uso de datos de entrada maliciosos y con esto me refiero a temas por ejemplo seguro que lo habéis visto este meme que circula por internet no que hay gente que se ha puesto en la matrícula en lugar de una matrícula normal se pone una comilla y bueno una sintaxis no para eliminar una tabla y porque hay muchos sistemas de parkings o incluso los coches de policía tienen hacen fotos a las matrículas y utilizan el reconocimiento óptico de caracteres para Bueno pues para Añadir una nueva multa al sistema no con la información de la matrícula pero la información de la matrícula en este caso contiene una inyección contra la base de datos que hace que no se incluya la matrícula en la multa e incluso que haga daños como borrar tablas otro ejemplo no sé si la habéis visto es el del cómic este de la madre que está en x kcbs este sitio web que hacen cómics enfocado es un poco más en la informática con un tono de humor y hay una madre que recibe una llamada colegio de su hijo y le dicen Oiga usted realmente ha llamado a su hijo Robert comilla cierro paréntesis punto y coma droptable estudiantes punto y coma guión guión y la madre dice Sí sí claro que sí Ese lo llamamos pequeño Bobby tablas bueno total un ejemplo gracioso también para que podáis ver es un ejemplo gracioso pero en el campo de batalla si alguien utiliza temas similares cuando estos drones intentan reconocer por ejemplo un tanque no sé a través de algún tipo de etiqueta o texto o incluso forma si se puede interpretar que ese tanque de hecho no es enemigo sino que es Aliado pues Oye no se le va a atacar o incluso puede hacer que se le engañe que lo ataque al enemigo en lugar de que ataque al Aliado en lugar del enemigo En definitiva el tema es que la Inteligencia artificial es la nueva tecnología de moda y Está avanzando de forma muy rápida como ha dicho Martín sobre todo el tema de las cartas que han emitido estos líderes de Inteligencia artificial se está

desplegando está evolucionando casi Sin Control sobre todo en temas de seguridad y estabilidad sin tener en cuenta estas consecuencias en en la vida real de estos diseños digamos lo igual un poco inseguros quería comentar que blockchain ya ha pasado el segundo plano ya veis que casi no no comentamos muchas noticias de blockchain porque realmente según Todavía siguen habiendo un montón de de robos en plataformas de blockchain pero es que realmente ahora la Inteligencia artificial con este con esta conexión que tiene ahora casi con el mundo Real Como en este ejemplo que puede controlar drones y similares tenemos que concienciarnos y tenemos que incluso nos iba a decir salir a la calle pero tenemos que conocerlo conocerla bien y bueno pedir a estas empresas que la desarrollan que lo hagan de forma segura y de hecho comentar el tema este también que ahora a partir de o sea el tema de la inteligencia de la Inteligencia artificial es algo tan calentito a día de hoy que si creas una empresa y dices que estás haciendo algo de Inteligencia artificial O lms puedes recibir muy fácilmente inversiones de empresas de capital de riesgo porque es lo que les gusta ahora y es donde están queriendo poner el dinero ahora Así que si hay alguien interesado que no sabe qué hacer con su tiempo libre que se dedica a esto porque hay mucho dinero involucrado en el tema a todo esto palantir no es nuevo en el mercado de defensa y ya lleva años vendiendo sus servicios de vigilancia nacional al servicio de inmigración y control de aduanas de Estados Unidos y como digo ya hace tres años comentábamos también que se ofrecía servicios Bueno a los sistemas de sanidad de Inglaterra y temas similares recientemente y con esto Cierro la noticia se ha celebrado la conferencia del uso responsable de la Inteligencia artificial en el ejército en Holanda con 50 países asistentes entre los que se encuentran Estados Unidos y china me ha parecido interesante que china estuviera ahí como asistente pero a ver si Martina adivinas al país al que no se invitó a asistir Rusia o Corea del Norte Rusia en el clavo estaba clarísimo Supongo por temas de la guerra no Y tal Pero bueno también no sé sí porque si fuese por desconfianza antes no invitarían el objetivo de la conferencia yo De hecho pensaba que era un poco más para para para definir limitaciones en el uso de la Inteligencia artificial en la guerra pero esto se ve como algo Bastante lejos Supongo que quieren ir pasito a pasito y de hecho el objetivo de esta conferencia en concreto Es que la mayoría de las delegaciones los países que han asistido respalden una declaración de principios Así que es algo Bastante básico Pero bueno al menos en plan no sé es como las leyes de la robótica no en plan un robot no va a matar a su Creador pues Supongo que quieren definir temas similares y eso me trae recuerdos Es como si fuera el convenio de Ginebra no lo del tema de no utilizar bombas racimo y ciertas leyes en la guerra que puede parecer contradictorio pero sí que efectivamente hay ciertas cosas que se consideren pues crímenes contra la humanidad y cosas así claro quieren aplicar lo mismo los convenios de Ginebra pero para la Inteligencia artificial Qué bueno sí no sé si está pasada hace poquito no ha encontrado el documento en el que se ha firmado por por la mayoría de los países pero algo así entiendo yo aunque dice la noticia que no han llegado a tanto a limitar el uso en plan no vamos a hacer esto así que es eso que tú comentas sería lo ideal pero no sé si han llegado a tanto Aunque igual en la siguiente el ceo de palantir fue uno de los asistentes y dijo que un principio que el apoya para empresas como la suya es que deben poder explicar y verificar cómo se ha utilizado la tecnología la Inteligencia artificial en este tipo sobre todo en los temas de guerra dio el ejemplo de una decisión asistida por Inteligencia artificial de atacar a soldados enemigos cerca de una escuela u hospital comentaba que se necesita una arquitectura que permita la transparencia en las fuentes de datos que fuentes de datos se usaron y cuáles fueron las fuentes estas de entrada y obviamente que se registren todas las acciones que se llevaron a cabo Esto está muy en línea con como ha comentado Martín anteriormente en el episodio 89 en el que trajo el tema de la Carta abierta firmada por expertos y líderes de empresas tecnológicas sobre pausar el desarrollo de

tecnologías de este tipo aunque no sé si en ese campo en esa carta Comenta en el tema más cercano a temas bélicos Pero bueno en cualquier caso casos similares como el que Martín ha dicho esto se puede aplicar también a temas de fraude Se podrían las inteligencias divididas podría alucinar y e identificar como culpable a una persona inocente no en un robo de un billón de dólares Así que hay que tener en cuenta también todos estos problemas que pueden ocasionarse por temas por todos los temas que mencionan anteriormente alucinación inyección de instrucciones o proms y e interacción de Inteligencia artificial con sistemas externos que esto hasta hace poquito no se daba pero ahora últimamente chat gpt por ejemplo tiene el tema de los plugins no que puedes instalarlos en tu entorno y por ejemplo utilizar Open Table y le puedes decir Oye me puedes hacer una reserva en un restaurante vegetariano cerca de yo que sé el Times Square porque donde voy a estar en el viernes a las 8 de la tarde y eso charge gpt se conecta a través de un plugin que ha diseñado Open Table y hace esta reserva o te ofrece al menos la posibilidad de hacer esta reserva si hay disponibilidad o no con esto así que ahí vemos que el tema de interactuar con sistemas externos en este caso el ejemplo que doy desde Open Table pero en este caso de guerra de palantir aip sería interactuar con sistemas de guerra como drones como misiles antiaéreos antitaques o similares Pues sí que sí que hemos metido cañita la Inteligencia artificial palantir que sí que tenemos un episodio especial para ellos pero vamos ha salido en el podcast hemos hablado un montón sobre ellos porque vamos siempre van de la mano de los abusos a la privacidad y aparte es una empresa que desarrolla tecnología muy hollywoodiense como decimos siempre y estaba Claro que se iban a subir al carro de la Inteligencia artificial como y que vamos a seguir cubriendo temas de estos en los próximos episodios seguro según van saliendo nuevas noticias Sí totalmente lo gracioso que algunos dicen Oh no sí palantir Peter ciel Sí este Esta persona es muy ética obviamente lo dicen con ironía por todo toda la trayectoria que tiene que tú acabas de mencionar Así que es gracioso que diga no no esta plataforma es muy moral muy ética y muy responsable y la gente Sí sí sí que qué más Qué más me quieres engañar señor si es como poner a China no a marcarnos las pautas sobre la ética la moral los Derechos Humanos la democracia no tendría ningún sentido Pero bueno sí pues hasta aquí ha llegado el episodio después de dos semanitas en completo y absoluto silencio esperemos que os haya gustado muy centrados en Inteligencia artificial señal de que tierra hacker siempre cubre la actualidad veremos que cómo va evolucionando esto pero la verdad es que lo de la Inteligencia artificial cada vez levanta más cejas y miedos Así que esperemos que sigáis con nosotros recordad suscribiros en podéis visitarnos en tierra de hackers.com dejarnos comentarios en vuestras plataformas de podcasting o reviews que nos ayuda un montonazo y espero que nos escuchemos en el siguiente y por cierto ahora que lo quería comentar antes pero se me olvidó cuando comenté el tema de las leyes de la robótica que son de Isaac asimov también otra serie que estoy viendo que recomienda también muy interesante es la serie de la fundación que es una una de las de las sagas de Isaac asimov que es el que definió las tres leyes de la robótica y la verdad que me parece muy buena la serie en Apple TV de nuevo Así que la recomiendo para ver relacionado todo el tema de la Inteligencia artificial sale una robot que tiene miles de años ahí que esté bueno ya la veis si queréis tienes que hacerte un podcast de series tío pues lo dicho Adiós adiós hasta luego si te ha gustado este episodio y quieres ayudarnos a seguir con el podcast compártelo con tus amigos y compañeros con tu apoyo podremos atraer y despertar el interés por la ciberseguridad de mucha más gente Acuérdate de dejarnos un comentario y una valoración donde nos estés escuchando también puedes seguirnos en Twitter Instagram y Facebook te esperamos en el próximo episodio de tierra de hackers