**Course 4 Video 26 Deepfake e copyright nell'IA**

***[Image on screen] Title Text visible "Deepfake e copyright nell'IA "***

Voice over: Deepfake e copyright nell'IA . Immaginate di scorrere il vostro cellulare e di imbattervi in un video di una persona famosa che dice qualcosa di controverso. Siete sorpresi, ma poi vi ricordate che potrebbe trattarsi di un cosiddetto deepfake.

***[Image on screen] A mobile phone displaying a video with captions.***

Un deepfake è un contenuto falso che è stato alterato o appositamente creato. I deepfake sono di solito contenuti audio o video e l’IA è stata utilizzata nella loro creazione.

I deepfake utilizzano tecniche avanzate di IA per sostituire la voce, l'immagine - o entrambe - di una persona reale con riproduzioni artificiali sorprendentemente simili. Questa tecnologia ha reso sempre più difficile discernere se qualcosa che si vede o si sente su Internet sia reale.

I deepfake sono spesso utilizzati per diffondere disinformazione e possono essere impiegati in truffe, manipolazioni elettorali, attacchi di social engineering, o ingegneria sociale, e altri tipi di frode.

***[Image on screen] A person in a mask with money, files, and an ID card.***

Riconoscendo l'urgente necessità di combattere i deepfake, il settore tecnologico sta agendo, comprese le aziende che si occupano della creazione di modelli di IA e di servizi ai consumatori come Microsoft.

Questi sforzi si sono evoluti nel tempo, passando dall'implementazione di funzioni anticontraffazione all'avanzamento delle tecnologie di protezione digitale.

Ecco cosa sta facendo il settore tecnologico.
Costruire un ambiente sicuro: misure di sicurezza sono state messe in atto per garantire che tutto funzioni senza intoppi e in modo sicuro. Ciò include controlli costanti, blocco dei comportamenti scorretti e azioni rapide contro coloro che abusano del sistema.

***[Image on screen] AI in a circle with badges for each topic.***

Credenziali dei contenuti: per combattere i video, le immagini o gli audio falsi, ai contenuti creati dall'IA vengono aggiunti marchi o icone speciali. Questo aiuta a determinare la provenienza del contenuto e la sua storia.

Mantenere i servizi sicuri: ci si adopera per individuare e rimuovere i contenuti dannosi e fuorvianti dalle piattaforme online. Ciò contribuisce a far sì che lo spazio online sia sicuro e rispettoso per tutti.

Lavorare insieme: poiché la collaborazione è fondamentale per raggiungere obiettivi condivisi, quando gli attori del settore tecnologico, le organizzazioni dedicate al benessere sociale e gli enti governativi si uniscono, possono contribuire collettivamente a creare un ambiente online più sicuro.

Questo approccio di gruppo può portare a soluzioni innovative e a maggiori garanzie per tutti nello spazio digitale.

***[Image on screen] AI in a circle with icons circling it.***

Aggiornare le leggi per affrontare le nuove sfide: con l'emergere di nuove sfide, si cerca di sviluppare nuove leggi e iniziative per proteggere le persone dall'abuso dell’IA.

Educare il pubblico: è importante che tutti siano ben informati. Sono in corso iniziative per mettere le persone in grado di distinguere i contenuti autentici da quelli falsi. Ciò include lo sviluppo di nuovi strumenti e programmi educativi per il pubblico. Queste strategie mirano a rendere le cose più trasparenti e ad aiutare la società a contrastare meglio i deepfake.

Se da un lato i deepfake pongono problemi etici e di sicurezza a causa del loro potenziale uso improprio per diffondere disinformazione e falsificazione di identità, dall'altro la tecnologia IA sta facendo passi da gigante nell'individuare i deepfake con grande precisione.

***[Image on screen] AI in a circle with real and deepfake videos that are checked.***

Questo ci porta a un altro argomento significativo: l'aumento e la consapevolezza dei contenuti generati dall'intelligenza artificiale.

Il copyright, o diritto d'autore, è un concetto legale che garantisce ai creatori e agli autori di opere originali diritti esclusivi sull’uso e sulla distribuzione, assicurando loro il riconoscimento e il beneficio economico delle proprie creazioni.

Ma cosa succede quando il creatore è un servizio o uno strumento di IA?

Sapendo che il copyright è una questione importante da affrontare per i contenuti generati dall’IA, sono state create iniziative come “Microsoft Copilot Copyright Commitment” per estendere il supporto esistente per la sicurezza della proprietà intellettuale ai servizi commerciali Copilot.

***[Image on screen] The Copilot logo.***

Questo impegno riguarda la potenziale responsabilità per violazione della proprietà intellettuale che potrebbe derivare dall'uso dei risultati dei Copilot di Microsoft e di Azure OpenAI Service.

Per creare ulteriore fiducia nei contenuti generati dall'IA, Microsoft ha sviluppato le credenziali dei contenuti. Questa funzione utilizza metodi crittografici per aggiungere una filigrana digitale invisibile a tutte le immagini generate dall'IA in Bing, comprese l'ora e la data di creazione originaria.

***[Image on screen] An image with an AI flag.***

Questo aiuta a determinare la provenienza dei contenuti e fornisce a chi li crea e li condivide strumenti migliori per decidere di cosa fidarsi. Questo passo è fondamentale per assicurarci di utilizzare la tecnologia IA in modo sicuro e responsabile.

***[Image on screen] A person connected with images and AI, with a tick in a thought bubble in the center.***

È importante rimanere informati. Ricordate che non tutto ciò che vedete o sentite può essere verità; quindi, quando si tratta di contenuti generati dall'IA è importante conoscere i vostri diritti e le misure in atto per proteggerli.

**ITALIAN VERSION**

**Corso 4 Video 26 Deepfake e copyright nell'IA**

***[Immagine sullo schermo] Testo del titolo visibile "Deepfake e copyright nell'IA "***

Voice over: Deepfake e copyright nell'IA . Immaginate di scorrere il vostro cellulare e di imbattervi in un video di una persona famosa che dice qualcosa di controverso. Siete sorpresi, ma poi vi ricordate che potrebbe trattarsi di un cosiddetto deepfake.

***[Immagine sullo schermo] Un telefono cellulare che visualizza un video con una didascalia.***

Un deepfake è un contenuto falso che è stato alterato o appositamente creato. I deepfake sono di solito contenuti audio o video e l’IA è stata utilizzata nella loro creazione.

I deepfake utilizzano tecniche avanzate di IA per sostituire la voce, l'immagine - o entrambe - di una persona reale con riproduzioni artificiali sorprendentemente simili. Questa tecnologia ha reso sempre più difficile discernere se qualcosa che si vede o si sente su Internet sia reale.

I deepfake sono spesso utilizzati per diffondere disinformazione e possono essere impiegati in truffe, manipolazioni elettorali, attacchi di social engineering, o ingegneria sociale, e altri tipi di frode.

***[Immagine sullo schermo] Una persona mascherata con denaro, incartamenti e una carta d'identità.***

Riconoscendo l'urgente necessità di combattere i deepfake, il settore tecnologico sta agendo, comprese le aziende che si occupano della creazione di modelli di IA e di servizi ai consumatori come Microsoft.

Questi sforzi si sono evoluti nel tempo, passando dall'implementazione di funzioni anticontraffazione all'avanzamento delle tecnologie di protezione digitale.

Ecco cosa sta facendo il settore tecnologico.
Costruire un ambiente sicuro: misure di sicurezza sono state messe in atto per garantire che tutto funzioni senza intoppi e in modo sicuro. Ciò include controlli costanti, blocco dei comportamenti scorretti e azioni rapide contro coloro che abusano del sistema.

***[Immagine sullo schermo] IA in un cerchio con schede per ogni argomento.***

Credenziali dei contenuti: per combattere i video, le immagini o gli audio falsi, ai contenuti creati dall'IA vengono aggiunti marchi o icone speciali. Questo aiuta a determinare la provenienza del contenuto e la sua storia.

Mantenere i servizi sicuri: ci si adopera per individuare e rimuovere i contenuti dannosi e fuorvianti dalle piattaforme online. Ciò contribuisce a far sì che lo spazio online sia sicuro e rispettoso per tutti.

Lavorare insieme: poiché la collaborazione è fondamentale per raggiungere obiettivi condivisi, quando gli attori del settore tecnologico, le organizzazioni dedicate al benessere sociale e gli enti governativi si uniscono, possono contribuire collettivamente a creare un ambiente online più sicuro.

Questo approccio di gruppo può portare a soluzioni innovative e a maggiori garanzie per tutti nello spazio digitale.

***[Immagine sullo schermo] IA in un cerchio con icone che lo circondano.***

Aggiornare le leggi per affrontare le nuove sfide: con l'emergere di nuove sfide, si cerca di sviluppare nuove leggi e iniziative per proteggere le persone dall'abuso dell’IA.

Educare il pubblico: è importante che tutti siano ben informati. Sono in corso iniziative per mettere le persone in grado di distinguere i contenuti autentici da quelli falsi. Ciò include lo sviluppo di nuovi strumenti e programmi educativi per il pubblico. Queste strategie mirano a rendere le cose più trasparenti e ad aiutare la società a contrastare meglio i deepfake.

Se da un lato i deepfake pongono problemi etici e di sicurezza a causa del loro potenziale uso improprio per diffondere disinformazione e falsificazione di identità, dall'altro la tecnologia IA sta facendo passi da gigante nell'individuare i deepfake con grande precisione.

***[Immagine sullo schermo] IA in un cerchio con video reali e deepfake che vengono verificati.***

Questo ci porta a un altro argomento significativo: l'aumento e la consapevolezza dei contenuti generati dall'intelligenza artificiale.

Il copyright, o diritto d'autore, è un concetto legale che garantisce ai creatori e agli autori di opere originali diritti esclusivi sull’uso e sulla distribuzione, assicurando loro il riconoscimento e il beneficio economico delle proprie creazioni.

Ma cosa succede quando il creatore è un servizio o uno strumento di IA?

Sapendo che il copyright è una questione importante da affrontare per i contenuti generati dall’IA, sono state create iniziative come “Microsoft Copilot Copyright Commitment” per estendere il supporto esistente per la sicurezza della proprietà intellettuale ai servizi commerciali Copilot.

***[Immagine sullo schermo] Il logo Copilot***

Questo impegno riguarda la potenziale responsabilità per violazione della proprietà intellettuale che potrebbe derivare dall'uso dei risultati dei Copilot di Microsoft e di Azure OpenAI Service.

Per creare ulteriore fiducia nei contenuti generati dall'IA, Microsoft ha sviluppato le credenziali dei contenuti. Questa funzione utilizza metodi crittografici per aggiungere una filigrana digitale invisibile a tutte le immagini generate dall'IA in Bing, comprese l'ora e la data di creazione originaria.

***[Immagine sullo schermo] Un'immagine di un contrassegno IA***

Questo aiuta a determinare la provenienza dei contenuti e fornisce a chi li crea e li condivide strumenti migliori per decidere di cosa fidarsi. Questo passo è fondamentale per assicurarci di utilizzare la tecnologia IA in modo sicuro e responsabile.

***[Immagine sullo schermo] Una persona connessa con immagini e IA, con un segno di spunta in un fumetto al centro.***

È importante rimanere informati. Ricordate che non tutto ciò che vedete o sentite può essere verità; quindi, quando si tratta di contenuti generati dall'IA è importante conoscere i vostri diritti e le misure in atto per proteggerli.