



Deepfakes sind Videos oder Bilder, in denen meist Gesichter ausgetauscht werden, um Betrachter zu täuschen.

Aktionsplan gegen Fälschungen

Die Bundesregierung hat einen Aktionsplan gegen Deepfake-Manipulationen präsentiert. Unter Leitung des Innenministeriums wurde der Aktionsplan interministeriell erarbeitet und abgestimmt.

Deepfakes werden als Überbegriff für verschiedene Formen audiovisueller Fälschungen verwendet. Deepfakes sind Videos oder Bilder, in denen meist Gesichter ausgetauscht werden und beispielsweise Stimmen, Mimik und Lippenbewegungen imitiert oder übernommen werden. Technisch veränderte Videos, Bilder oder Texte können unangenehme oder gar existenzbedrohende Folgen für Einzelpersonen und Unternehmen haben. Auch der gesellschaftliche Zusammenhalt und das Funktionieren staatlicher Institutionen können durch die Verbreitung von Deepfake-Videos gefährdet werden, die beispielsweise Politikerinnen und Politiker bei unerwünschten Äußerungen oder Handlungen zeigen.

Die Herstellung von Deepfake-Videos wird durch immer leistungsstärkere Smartphones und PCs erleichtert. Umso wichtiger ist es, rechtzeitig die zahlreichen Fälschungsmöglichkeiten anzugehen, die sich durch täuschend echt wirkende Deepfake-Manipulationen ergeben.

Der Nationale Aktionsplan gegen Deepfakes wurde am 25. Mai 2022 im Ministerrat vorgestellt. Unter Leitung

des Bundesministeriums für Inneres wurde der Aktionsplan im Vorfeld interministeriell erarbeitet und abgestimmt. Die Arbeitsgruppe deckte die rechtlich relevanten Tätigkeitsfelder, die Bild- und Tonmanipulationen betreffen, ebenso ab wie schützenswerte Bereiche, etwa Kunstfreiheit und politische Satire.

Gefährdungslage. Im Abschnitt zur Gefährdungslage listet der Nationale Aktionsplan verschiedene Bedrohungsszenarien inklusive geltender Rechtslage auf. Das Phänomen der Deepfakes ist hochriskant, weil die erstellten Videos und Tonaufnahmen täuschend echt wirken. Eine wesentliche Gegenmaßnahme ist die Entwicklung automatisierter Tools zur Kennzeichnung manipulierter Bilder und Videos.

„Es kann jede Einzelne und jeden Einzelnen treffen“, sagte Innenminister Mag. Gerhard Karner nach der Regierungssitzung. Auch Unternehmen, Medien, die gesamte Gesellschaft oder auch „ganze Staaten“ könnten betroffen sein, wenn mit derartigen Methoden gearbeitet werde. Deepfakes und Fake News hätten die „Destabilisierung unserer demokratischen Ordnung“ zur Folge, sagte der Innenminister. Die

Notwendigkeit, dagegen vorzugehen, hätten insbesondere die letzten zwei Jahre gezeigt – nämlich die „Querdenker“ im Rahmen der Corona-Pandemie und die Verbreitung von diversen Verschwörungstheorien oder auch Falschinformation im Rahmen des russischen Angriffskriegs auf die Ukraine.

Tool. Das AIT (*Austrian Institute of Technology*) arbeitet im Sicherheitsforschungsprogramm KIRAS am Projekt „Defalsif-AI“ an der Entwicklung eines solchen Werkzeugs. Defalsif-AI steht für „Detektion von Falschinformation mittels Artificial Intelligence“. Die geplante Software bewertet eingespeiste Videos und Bilder mithilfe von künstlicher Intelligenz und liefert die Wahrscheinlichkeit für Fälschungen am Bild bzw. am Video.

Das AIT legt einen Schwerpunkt auf die Nachvollziehbarkeit der Ergebnisse. Die Bewertung, ob Inhalte letztlich „fake“ sind, obliegt den Nutzerinnen und Nutzern. Der Prototyp, der im Zuge einer Präsentation auch BMI-Vertretern gezeigt wurde, zeigt vielversprechende Ergebnisse und wird künftig einen Mehrwert bei der automatisierten Erkennung von Deepfakes liefern.

Leonhard Kunz