



# مبادئ التزييف العميق

النسخة 1



**SDAIA**

الهيئة السعودية للبيانات  
والذكاء الاصطناعي  
Saudi Data & AI Authority

## جدول المحتويات

2	1. الملخص التنفيذي
3	2. مقدمة وبيان المشكلة
5	3. الغرض من الأدلة الاسترشادية ونطاقها
6	4. نظرة عامة على تقنيات التزييف العميق الخبيثة
6	■ 4.1 العمليات الاحتيالية
6	■ 4.2 التلاعب غير المصرح به
7	■ 4.3 التضليل والدعاية
7	5. إرشادات لمطوري تقنيات التزييف العميق
10	6. إرشادات لصنّاع محتوى التزييف العميق
11	7. إرشادات للجهات التنظيمية
13	8. إرشادات لعملاء التزييف العميق لاكتشاف التزييف العميق ومنع المخاطر
13	■ 8.1 تقييم الرسالة
13	■ 8.2 تحليل العناصر السمعية والبصرية
14	■ 8.3 مصادقة المحتوى باستخدام الأدوات
15	■ 8.4 الإبلاغ عن الحادث
17	■ 8.5 أفضل الممارسات لحماية النفس والآخرين
18	9. نظرة عامة على تقنيات التزييف العميق غير الخبيثة
18	■ 9.1 التسويق
19	■ 9.2 الترفيه
20	■ 9.3 تجارة التجزئة
20	■ 9.4 التعليم
21	■ 9.5 الرعاية الصحية
22	■ 9.6 الثقافة
23	10. الخلاصة
23	11. الملحق
23	11.1. الملحق 1: تعريفات
24	11.2. الملحق 2: التدريب وتنمية المهارات
26	11.3. الملحق 3: دراسات الحالة والأمثلة
28	11.4. الملحق 4: إطار عمل تقييم المخاطر لمطوري تقنيات التزييف العميق
30	11.5. الملحق 5: نموذج موافقة على استخدام البيانات الشخصية في صناعة محتوى التزييف العميق

## 1. الملخص التنفيذي

يوفر التقدم السريع في تقنيات الذكاء الاصطناعي، ولا سيما تقنيات التزييف العميق، فرصاً وتحديات كبيرة في مختلف القطاعات. والتزييف العميق هو وسائط اصطناعية واقعية للغاية تم إنشاؤها من خلال تقنيات التعلم العميق التي يمكنها التعامل مع وسائط الصوت أو الفيديو أو أي محتوى رقمي آخر بطرق يصعب تمييزها عن الواقع. وفي حين أن التزييف العميق يقدم فوائد محتملة في مجالات، مثل: التسويق، والترفيه، وتجارة التجزئة، والتعليم، والرعاية الصحية، والتطبيقات الثقافية، فإنه يؤدي إلى مخاطر جسيمة، بما في ذلك انتحال الهوية والتلاعب غير المصرح به ونشر المعلومات المضللة.

تحدد هذه الوثيقة أدلة استرشادية مفصلة لمعالجة الآثار المترتبة على تقنيات التزييف العميق والتخفيف من حدة مخاطرها، كما تحدد أساليب التزييف العميق وتدرس تأثيره المجتمعي، وتسلط الضوء على تطبيقاته الخبيثة وغير الخبيثة، وتؤكد هذه الأدلة الاسترشادية على أهمية الالتزام بالمبادئ الأخلاقية في تطوير واستخدام تقنيات التزييف العميق، بما في ذلك الخصوصية والشفافية والمساءلة والمنافع المجتمعية.

تتضمن التوصيات المقدمة للمطورين وصنّاع المحتوى بتقنية التزييف العميق تنفيذ ممارسات قوية لحماية البيانات، والحصول على الموافقة لاستخدام البيانات الشخصية، وضمان الشفافية من خلال التوثيق الواضح والتفسيرات الدقيقة لعمليات إنشاء محتوى التزييف العميق، ويتم تشجيع المطورين أيضاً على فرض رقابة بشرية وطرق مساءلة واضحة عن نتائج ومخرجات الذكاء الاصطناعي، وبالإضافة إلى ذلك تدعو الأدلة الاسترشادية هذه إلى توجيه تقنيات التزييف العميق نحو التطبيقات التي تقدم فوائد اجتماعية وبيئية متنوعة، فاتباع هذه الأدلة لا يساعد فقط على حماية الأفراد والمجتمع من الأضرار المحتملة، بل يساعد أيضاً على بناء الثقة والقبول لتقنيات الذكاء الاصطناعي.

يتم تشجيع مستخدمي تقنيات التزييف العميق على تنفيذ أفضل الممارسات لاكتشاف ومعالجة حالات التزييف العميق، ويتضمن ذلك التحقق من صحة المصادر، وفحص العناصر الصوتية والمرئية بحثاً عن التناقضات، واستخدام أدوات الذكاء الاصطناعي لتحديد علامات التلاعب، وتعد مبادرات التوعية والتعليم العامة ضرورية لمساعدة الأفراد على التعرف على أساليب التزييف العميق والاستجابة لها بشكل فعال، وتعتبر هذه الإجراءات ضرورية للحفاظ على سلامة المعلومات وحماية الأفراد من سوء استخدام تقنيات التزييف العميق.

وتؤكد الأدلة الاسترشادية أيضاً على أهمية التعلم المستمر وتنمية المهارات لإدارة تطبيقات الذكاء الاصطناعي بشكل فعال، ويوصى بأن يشارك الأفراد والمنظمات في ورش العمل ودورات التدريب العملي لاكتساب الخبرة اللازمة للتعامل مع تقنيات الذكاء الاصطناعي، ويمكن أن يؤدي تعزيز القدرات التنظيمية من خلال ورش العمل المستهدفة وممارسات التوظيف الاستراتيجية إلى تحسين مهارات الابتكار وحل المشكلات بشكل كبير، بما يساعد الجهات في التغلب على تعقيدات الذكاء الاصطناعي وتقنيات التزييف العميق، ويُعد التعليم المستمر أمراً حيوياً لضمان الحفاظ على مهارة الأفراد والمنظمات على حدٍ سواء في مواجهة التحديات الناشئة واغتنام الفرص الجديدة.

ومن خلال الالتزام بهذه الأدلة الاسترشادية الشاملة، يمكن لأصحاب المصلحة الاستفادة من الجوانب الإيجابية لتقنيات التزييف العميق مع تقليل المخاطر المرتبطة بها، ويضمن هذا النهج الاستخدام الأخلاقي والمسؤول لتقنيات التزييف العميق، ويشجع الابتكار، ويحافظ على ثقة الجمهور في تطورات الذكاء الاصطناعي.

## 2. مقدمة وبيان المشكلة

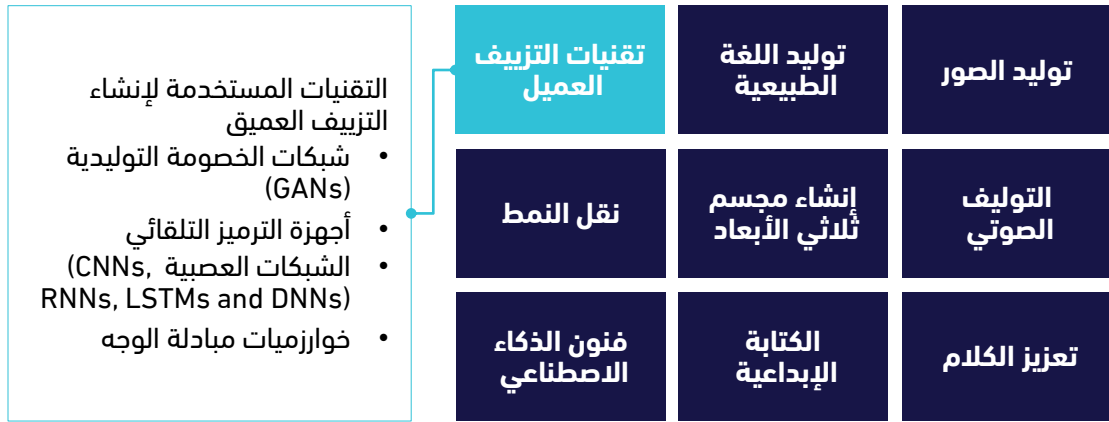
أصبحت تقنيات الذكاء الاصطناعي ذات أهمية مركزية في مختلف القطاعات بشكل متزايد، مما يؤثر بشكل عميق على الاقتصاد العالمي، وتسهل هذه التقنيات عدداً من التطبيقات، وتفتح مزيداً من الفرص في قطاعات متنوعة، مثل: الرعاية الصحية، والتمويل، والنقل، والترفيه. ومع ذلك فإن هذه التطورات تأتي أيضاً مصحوبة بمخاطر كبيرة، ويعد الذكاء الاصطناعي بشكل عام من الأدوات التي يمكن إساءة استخدامها، مما يؤدي إلى مشاكل كثيرة، مثل: التحيز، وانتهاكات الخصوصية، ويؤدي ظهور تقنيات التزييف العميق إلى تضخيم هذه المخاطر، حيث تعتبر تقنيات التزييف العميق وسائط اصطناعية شديدة الواقعية تم إنشاؤها باستخدام تقنيات التعلم العميق، التي تتلاعب بوسائط الصوت أو الفيديو أو أي محتوى رقمي آخر، مما يجعل من الصعب التمييز بين المواد الأصلية والمزيفة.

في حين أن تقنيات التزييف العميق لها تطبيقات مشروعة، فإنها تشكل مخاطر جسيمة، بما في ذلك انتشار المعلومات المضللة والاحتيال وتقويض الثقة في الوسائط والوسائل الإعلامية.

فمعالجة هذه القضايا أمر ضروري للحفاظ على سلامة المعلومات، وحماية سمعة الأفراد، وضمان السلامة العامة، إذ تؤدي صعوبة اكتشاف وتحديد أساليب التزييف العميق وتقنياته إلى تفاقم هذه المخاطر، مما يستلزم تطوير أطر سياسية وتنظيمية شاملة بالإضافة إلى اتخاذ تدابير مضادة فعّالة.

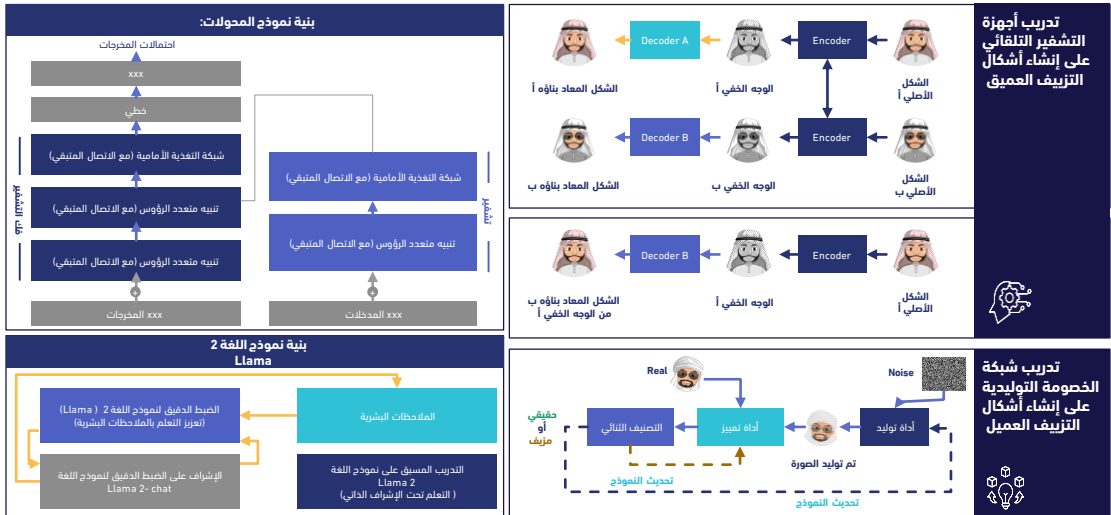
وكما هو موضح في الشكل (1)، تمثل تقنيات التزييف العميق إحدى حالات الاستخدام الكثيرة للذكاء ، والشبكات (GANs) الاصطناعي التوليدي التي تتضمن تقنيات متنوعة، مثل: شبكات الخصومة التوليدية ( العصبية) (وللتعرف على مزيد من المعلومات، راجع "الأدلة الاسترشادية للذكاء الاصطناعي التوليدي للجمهور العام").

### حالات استخدام الذكاء الاصطناعي التوليدي



الشكل 1: حالات استخدام الذكاء الاصطناعي التوليدي<sup>1</sup>

على عكس أدوات تحرير الصور والصوت التقليدية، تستخدم تقنية التزييف العميق خوارزميات معقدة لإنشاء محتوى معدل واقعي للغاية، وتسمح هذه التقنية بإنتاج وسائط مزيفة لا يمكن تمييزها تقريباً عن اللقطات الحقيقية، مما يخلق تحديات كبيرة للكشف والتحقق (انظر إلى الشكل رقم 2).



الشكل 2: عمليات إنشاء التزييف العميق <sup>2</sup>.

تسمح تقنية التزييف العميق، التي نشأت من شبكات الخصومة التوليدية (GANs) في عام 2014م، للذكاء الاصطناعي بإنتاج بيانات جديدة بناءً على الأنماط المدخلة، وفي البداية يتطلب إنشاء التزييف العميق بيانات وموارد حسابية كبيرة، مما جعلها في متناول المتخصصين المهرة في المقام الأول، ولكن التقدم في الذكاء الاصطناعي والبرمجيات منذ ذلك الحين جعل الوصول إلى تقنيات التزييف العميق أكثر سهولة وفعالية من حيث التكلفة وأسرع في الإنتاج، ومع استمرار تطور تقنية الذكاء الاصطناعي، من المتوقع أن تصبح تقنية التزييف العميق واقعية بشكل متزايد وأكثر صعوبة في اكتشافها، مما يسلب الضوء على الحاجة إلى التقدم المستمر في طرق الكسّف والتدابير التنظيمية لمعالجة المخاطر المرتبطة بها.

إن تقنيات التزييف العميق ليست خبيثة أو ضارة بطبيعتها؛ ويعتمد تأثيرها على نيتها وتطبيقها (انظر إلى الشكل رقم 3).

### تطبيقات التزييف العميق

	تقنيات التزييف العميق غير الخبيثة	تقنيات التزييف العميق الخبيثة	
...	<b>الدعاية والإعلان</b> تستخدم الشركات تقنية التزييف العميق لصياغة إعلانات مبتكرة وجدّابة	<b>المادة الصريحة</b> تستغل الجهات المهاجمة تقنيات التزييف العميق لإنشاء صور صريحة للمضايقة أو الابتزاز	<b>إنشاء الصور</b> استخدام الذكاء الاصطناعي لتعديل أو إنشاء صور واقعية للأشخاص أو المشاهد
...	<b>استنساخ الصوت</b> توظيف الأصوات المستنسخة للوصول إلى المعلومات الحصرية وتفويض المعاملات أو الأوامر	<b>تمكين التزييف العميق للصوت</b> إنشاء تعليقات صوتية ودلجة واقعية بمختلف اللغات دون الحاجة إلى الممثلين الأصليين	<b>توليد الصوت</b> إنشاء صوت اصطناعي يقلد صوت شخص لإنتاج خطاب مقنع ولكنه كاذب
...	<b>التجسس على الشركات</b> يمكن تقنية التزييف العميق محاكاة الاتصالات التي يجريها المدراء التنفيذيون لسرقة معلومات الشركة الحساسة	<b>التلاعب السياسي</b> باستخدام تقنية التزييف العميق، يمكن للأشخاص تغيير المقابلات والخطب لتشويه سمعة المرشحين	<b>توليد مقاطع الفيديو</b> يمكن لمقاطع الفيديو التي تم إنشاؤها بواسطة أدوات الذكاء الاصطناعي تغيير مظهر الأفراد لإنتاج محتوى مرئي واقعي ولكنه غير حقيقي

**أشكال أخرى من المحتوى الرقمي**  
اعتباراً من يونيو 2024، لا توجد إشارات محددة للتزييف العميق في أنواع أخرى من المحتوى الرقمي، ومع ذلك، قد تشمل التطورات المستقبلية المحتملة التزييف العميق الجيني، والتزييف العميق الخاص بحاسة الشم، والتزييف العميق القائم على الذوق

الشكل 3: تطبيقات تقنية التزييف العميق <sup>3</sup>.

إن تقنيات التزييف العميق الخبيثة عبارة عن وسائل يتم إنشاؤها بالذكاء الاصطناعي بهدف خداع الأفراد أو إيذائهم أو استغلالهم، وغالباً ما يتم توظيفهم لأغراض مثل: نشر معلومات كاذبة، أو التشهير، أو الاحتيال (انظر إلى الشكل رقم 4).

تقنيات التزييف العميق غير الخبيثة
<p><b>التعريف:</b> الوسائط التي تم إنشاؤها بواسطة أدوات الذكاء الاصطناعي لأغراض حميدة مثل الترفيه أو التعليم. عندما يتم ذلك دون نية الخداع أو التسبب في الأذى</p>
<p>وضع علامة واضحة على المحتوى للإشارة إلى أنه تم إنشاؤه أو تعديله صناعياً. وتضمن هذه الشفافية أن يدرك الجمهور أنهم يشاهدون وسائط تم التلاعب بها</p>
<p>الحصول على موافقة صريحة من الأفراد الذين تستخدم صوركهم في إنشاء صور مزيفة يساعد على منع إساءة استخدام الهويات</p>
<p>استخدم التزييف العميق لأغراض بناءة ومفيدة مثل التعليم والترفيه وإعادة بناء السردية التاريخية وتعزيز إمكانية الوصول.</p>
<p>الالتزام بالمبادئ التوجيهية الأخلاقية الشاملة لإنشاء واستخدام تقنيات التزييف العميق، وذلك بدعم المراقبة المستقلة، بما يضمن المساءلة ويحترم المعايير &gt;</p>

تقنيات التزييف العميق الخبيثة
<p><b>التعريف:</b> الوسائط التي تم إنشاؤها بواسطة أدوات الذكاء الاصطناعي بهدف خداع الأفراد أو إيداعهم أو استغلالهم، وغالباً ما تستخدم في نشر معلومات كاذبة أو التشهير أو الاحتيال</p>
<p>يمكن لمحتوى التزييف العميق الذي لا يكشف عن طبيعته الاصطناعية أن يشوه التصور العام ويسبب الضرر بسبب افتقاره إلى الشفافية.</p>
<p>استخدام صور الأفراد دون الحصول على موافقتهم بعد انتهاكاً لحقوقهم وخصوصيتهم. يمكن أن يؤدي الاستخدام غير المصرح به إلى سرقة الهوية والتشهير.</p>
<p>استخدام التزييف العميق لأغراض ضارة مثل نشر معلومات كاذبة أو الانخراط في التلاعب السياسي أو ارتكاب جريمة التشهير.</p>
<p>عدم الالتزام بالمبادئ التوجيهية واللوائح الأخلاقية في إنشاء واستخدام تقنيات التزييف العميق. يمكن أن يؤدي هذا إلى ضرر واسع النطاق بسبب الافتقار إلى المساءلة والرقابة</p>

<p><b>الشفافية والإفصاح</b></p>
<p><b>التفويض والموافقة</b></p>
<p><b>تقديم الطلبات</b></p>
<p><b>الامتثال</b></p>

الشكل 4: العوامل المميزة بين تقنيات التزييف العميق الخبيثة وغير الخبيثة 4

### 3. الغرض من الأدلة الاسترشادية ونطاقها

تقدم هذه الوثيقة شرحاً وافياً لتقنيات التزييف العميق وآثارها والتدابير الأساسية اللازمة للتخفيف من مخاطرها على مطوري التقنيات وصناع المحتوى والمستخدمين. وتتركز الأدلة الاسترشادية على:

- تحديد تقنيات التزييف العميق وتأثيرها على المجتمع.
- رفع مستوى الوعي بالتطبيقات الخبيثة وغير الخبيثة لتقنيات التزييف العميق.
- إرساء مبادئ أخلاقية للمطورين المشاركين في إنشاء تقنيات التزييف العميق: تم وضع مبادئ توجيهية للمطورين لتعزيز التطوير والاستخدام المسؤول لتقنيات التزييف العميق، مع التركيز على الشفافية و الحصول على الموافقات، واحترام خصوصية الأفراد
- تقديم إرشادات واضحة لصناع المحتوى الذين يستخدمون تقنيات التزييف العميق: تعليمات محددة لصناع المحتوى حول كيفية إنتاج وتوزيع محتوى التزييف العميق، والتأكد من التزامهم بالمعايير الأخلاقية والمتطلبات النظامية
- تقديم استراتيجيات تخفيف المخاطر وتمكين الجهات التنظيمية والجهات الحكومية.
- تقديم تدابير وقائية لتحديد ومكافحة تقنيات التزييف العميق.
- تشجيع التعليم والتوعية المستمرين لتعزيز القدرة على الصمود في مواجهة التهديدات المرتبطة بالتزييف العميق.

تنطبق هذه الوثيقة على ثلاث مجموعات أساسية متورطة في التزييف العميق التقنيات: المطورون، صناع المحتوى، ومستخدمون التزييف العميق.

## 4. نظرة عامة على تقنيات التزييف العميق الخبيثة

نظراً إلى أن تقنيات التزييف العميق أصبحت أكثر تقدماً ويمكن الوصول إليها، فقد تصاعدت المخاطر المرتبطة بها عالمياً، مما زاد من إمكانية تعرض الأفراد والمنظمات لمخاطرها، وتستهدف مخططات انتحال الهوية المتقدمة التي يتم تمكينها بتقنيات التزييف العميق الضحايا بشكل متزايد، مما يثير مخاوف كبيرة بشأن مستويات الأمان والخصوصية، ويسلط الارتفاع السريع في حالات انتحال الهوية والاحتيال المتعلقة بالتزييف العميق الضوء على الحاجة إلى التعاون الدولي لمعالجة هذه التهديدات المعقدة.

توضح الأقسام التالية فئات المخاطر الأساسية المرتبطة بتقنيات التزييف العميق الخبيثة:

### 4.1. العمليات الاحتيالية

تستخدم العمليات الاحتيالية تقنيات التزييف العميق لتقليد الأفراد الموثوق بهم بشكل مقنع من أجل الحصول على معلومات حساسة أو خداع الضحايا، ومن خلال تكرار الأصوات وتعبيرات الوجه والسلوكيات، تستغل العمليات الاحتيالية هذه الثقة الممنوحة للشخص المنتحل، وتتضمن السيناريوهات النموذجية مكالمات احتيالية أو مكالمات عبر الفيديو حيث يطلب المحتال الذي يتنكر في صورة شخصية تنفيذية أو ذات سلطة معروفة، تفاصيل سرية أو يسمح بإجراء معاملات مالية، إذ إن تعقيد العمليات الاحتيالية هذه يتحدى الإجراءات الأمنية التقليدية، ويسلط الضوء على الحاجة إلى تقنيات كشف متقدمة وزيادة التوعية لمنع الاحتيال.

مثال:

الاحتيال على الشركات المتعددة الجنسيات<sup>5</sup>: في حالة ملحوظة تم الإبلاغ عنها في منطقة آسيوية، تم خداع موظف في شركة متعددة الجنسيات لتحويل مبلغ كبير من المال إلى المحتالين الذين استخدموا تقنية التزييف العميق لانتحال شخصية أحد كبار المسؤولين التنفيذيين خلال مكالمة عبر الفيديو، وتوضح هذه الحادثة كيف يمكن استخدام تقنيات التزييف العميق لتسهيل عمليات الاحتيال المالي على نطاق واسع.

### 4.2. التلاعب غير المصرح به

يتضمن التلاعب غير المصرح به استخدام تقنية التزييف العميق لإنشاء محتوى صريح أو مساومة للأفراد دون إذنهم، وغالباً ما يكون ذلك لأغراض متنوعة، مثل: المضايقة، أو الابتزاز، أو الإضرار بالسمعة. ويعاني ضحايا مثل هذه الأفعال من ضائقة عاطفية شديدة، ومشاكل في الصحة العقلية الطويلة الأمد، وعلاقات شخصية متوترة، ونكسات مهنية، ويسلط هذا الانتهاك للخصوصية والثقة الضوء على الحاجة الملحة إلى حماية نظامية وتقنية قوية.

مثال:

فضيحة التزييف العميق للمشاهير<sup>6</sup>: قبل بضع سنوات، تم تداول صور ومقاطع فيديو صريحة للتزييف العميق لعدد من المشاهير عبر الإنترنت دون موافقتهم، وتسببت إساءة استخدام تقنية التزييف العميق في حدوث مشكلات عاطفية كبيرة وانتهاكات خطيرة لخصوصية الأفراد المتضررين.

5. شبكة قنوات CNN  
6. موقع Yahoo News

## 4.3. التزييل والدعاية

يمكن أن تنشر تقنيات التزييل العميق معلومات كاذبة، وتتلاعب بالتصور العام، وتؤثر على النتائج السياسية من خلال تصوير شخصيات سياسية تدلي ببيانات كاذبة أو تنخرط في سلوك غير أخلاقي، ويمكن استخدام هذا التكتيك لتقويض انتشار المعارضين، والتأثير على الرأي العام، وزعزعة استقرار المجتمعات، مما يؤدي إلى انتشار المعلومات الخاطئة، وتضاؤل مستوى الثقة في المؤسسات العامة، والاضطرابات الاجتماعية، وتمثل الطبيعة المقنعة للتزييل العميق تهديداً كبيراً للخطاب السياسي والاستقرار المجتمعي.

مثال:

**حادثة التزييل العميق لزيم سياسي<sup>7</sup>:** خلال الانتخابات الأخيرة، تم استخدام تقنيات التزييل العميق لإنشاء مقاطع فيديو مزيفة للمرشحين، ونشر معلومات مضللة وربما تغيير تصورات الناخبين، إذ صور أحد مقاطع الفيديو المزيفة أحد المرشحين البارزين وهو يُدلي بتصريحات مثيرة للجدل، مما أدى إلى انخفاض ملحوظ في شعبيته ومستوى دعمه، على الرغم من الجهود اللاذقة لدحض الفيديو وبيان زيفه.

## 5. إرشادات لمطوري تقنيات التزييل العميق

يؤكد هذا القسم على الدور الحاسم لأخلاقيات الذكاء الاصطناعي في تطوير تقنيات التزييل العميق.

ويتم تشجيع المطورين على التأكد من أن تقنيات التزييل العميق تتوافق مع المعايير الأخلاقية الموضوعية، وتحمي البيانات الشخصية، وتمنع سوء الاستخدام، ويُعد الامتثال للقوانين واحترام الأخلاق الاجتماعية والالتزام بالقيم الأساسية أمراً ضرورياً لا يمكن التفريط فيه، وينبغي للتقنية أن تسهل التحقق من صحة المحتوى وتمنع إساءة استخدام الوسائط الاصطناعية، وتعزز الممارسات المسؤولة والجديرة بالثقة.

ونظراً إلى التأثير المجتمعي العميق للتزييل العميق، فمن الأهمية بمكان توجيه تطورها نحو تطبيقات إيجابية وبناءة مع تخفيف المخاطر المرتبطة بها.

ويجب على المطورين الالتزام بالإرشادات التالية أدناه عند تطوير تقنيات التزييل العميق وإدارتها:

### 1. الامتثال التنظيمي

**1.1.** ضمان الالتزام بجميع قوانين خصوصية البيانات المحلية والدولية ذات الصلة، مثل: اللائحة العامة لحماية البيانات (GDPR)، ونظام حماية خصوصية المستخدم (CCPA)، ونظام حماية البيانات الشخصية (PDPL) في المملكة العربية السعودية، ونظام مكافحة الجرائم الإلكترونية، بالإضافة إلى دمج فحوصات الامتثال النظامي في بيئة التطوير الخاصة بك، باستخدام أدوات متنوعة، مثل: أطر السياسة كرمز (على سبيل المثال: Terraform و Cedar Policy Language).

(المراجع: المادة (3) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - قسم الامتثال، الصفحة (32))

### 2. خصوصية البيانات وحمايتها

**2.1.** تنفيذ تدابير قوية لحماية البيانات، بما في ذلك التشفير (على سبيل المثال: AES-256)<sup>8</sup> وضوابط الوصول (على سبيل المثال: المصادقة المتعددة العوامل)، لحماية البيانات الشخصية المستخدمة في التدريب وإنشاء أساليب التزييل العميق.

(المراجع: المادة (19) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2))

7. موقع Business Insider.

8. معيار التشفير المتقدم (AES) عبارة عن تشفير كتلة متماثل يستخدم طول مفتاح 256 بت لكل من تشفير وفك تشفير كتل الرسائل.



**2.2.** الالتزام بالبروتوكولات الصارمة لجمع البيانات، باستخدام الحد الأدنى من البيانات الضرورية للأغراض المقصودة فقط، وتطبيق إخفاء الهوية والأسماء المستعارة لتقليل مخاطر إعادة تحديد الهوية.

(المراجع: المواد (10) و(11) و(12) و(14) و(18) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2))

**2.3.** دمج تقنيات الحفاظ على الخصوصية في معالجة البيانات، وتقليل الاحتفاظ بالبيانات، وإنشاء دورات مراجعة آلية للتقييم المنتظم والحذف الآمن للبيانات غير الضرورية (على سبيل المثال: باستخدام محو التشفير)، يقلل هذا النهج من المخاطر المرتبطة بالانتهاكات والوصول غير المصرح به.

(المراجع: المادتان (4) و(28) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2))

**2.4.** دمج أنظمة إدارة الموافقة في أدوات الذكاء الاصطناعي الخاصة بك، مما يضمن الموافقة بشكل صحيح على البيانات المستخدمة للتدريب، وتسجيل وإدارة معاملات الموافقة بشكل آمن.

(المراجع: المواد (4) و(5) و(10) و(24) من نظام حماية البيانات الشخصية)

### 3. الشفافية والوضوح

**3.1.** توثيق كل مكون من مكونات نموذج الذكاء الاصطناعي الخاص بك، بما في ذلك مصادر البيانات وخطوات المعالجة المسبقة والبنية الخوارزمية وعملية اتخاذ القرار، والتأكد من أن هذه الوثائق قابلة للقراءة آلياً ومدمجة في مسار التكامل المستمر/ النشر المستمر (CI/CD) الخاص بك<sup>9</sup> للحصول على التحديثات المستمرة.

(المراجع: ومبادئ أخلاقيات الذكاء الاصطناعي - المبادئ (5) و(7)).

**3.2.** دمج ميزات القابلية للتفسير في نماذج الذكاء الاصطناعي باستخدام تقنيات، مثل: (LIME) (التفسيرات المحلية غير القابلة للتفسير)<sup>10</sup>، أو (SHAP) (التوسعات الإضافية لشابلي)<sup>11</sup> لتوضيح كيفية إنشاء المخرجات للمستخدمين وأصحاب المصلحة.

(المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (6)).

**3.3.** تطوير وتطبيق تقنيات العلامات المائية الرقمية غير المتطفلة (مثل: طرق إخفاء المعلومات) والتأكد من أن هذه العلامات مقاومة للهجمات الشائعة (مثل: إعادة الضغط، والقص، وتغيير المظهر).

(المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**3.4.** الحفاظ على الشفافية من خلال نشر تقارير منتظمة حول استخدام تقنيات التزييف العميق وتأثيرها، بما في ذلك التدابير المتخذة لمنع سوء الاستخدام.

(المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).

### 4. التحيز

**4.1.** تنفيذ عمليات متنوعة، مثل: أخذ العينات، وزيادة البيانات؛ لتحديد التحيزات في البيانات والقضاء عليها، والحفاظ على التوثيق الشامل لمصادر البيانات وخطوات الإعدادات وعمليات التحقق من الصحة.

(المراجع: المادة (14) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (1)).

9. سلسلة من الخطوات الآلية التي تساعد فرق البرمجيات في تقديم التعليمات البرمجية بسرعة أكبر وأكثر أماناً وموثوقية.

10. تعرف التقنية التي تقارب أي نموذج للتعليم الآلي للصندوق الأسود بنموذج محلي قابل للتفسير لشرح كل تنبؤ فردي باسم التفسيرات المحلية غير القابلة للتفسير. (LIME)

11. (SHAP) (التوسعات الإضافية لشابلي) هي طريقة نظرية للعبة تُستخدم لشرح مخرجات أي نموذج للتعليم الآلي من خلال إسناد مساهمة كل ميزة إلى التنبؤ.

## 5. المساءلة والمسؤولية

- 5.1.** دمج آليات نماذج الإنسان في الحلقة ( HITL) للرقابة المستمرة خلال المراحل الحرجة، مثل: التدريب، والتحقق، والنشر؛ لضمان اتخاذ القرارات الأخلاقية ومعالجة المخاطر المرتبطة بسلوك النموذج المستقل. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).
- 5.2.** تطوير وتنفيذ إطار حوكمة شامل يتضمن ضوابط الوصول المستندة إلى الأدوار (RBAC)، وأنظمة التحكم في الإصدار للنماذج، ومسارات التدقيق لجميع تكرارات النماذج. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).
- 5.3.** إنشاء نظام تقارير آلي يستفيد من تحليل السجلات والتعلم الآلي لاكتشاف الاستخدام غير المصرح به أو غير الأخلاقي لتقنية التزييف العميق والإبلاغ عنه.

## 6. المسؤولية الاجتماعية

- 6.1** توجيه تقنيات التزييف العميق نحو التطبيقات المفيدة اجتماعياً وبيئياً، مثل: التعليم، والترفيه، والرعاية الصحية. وإن إساءة استخدام هذه التقنيات أو تطبيقها بشكل غير أخلاقي قد يؤدي إلى عواقب نظامية. (المراجع: ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (4)، نظام مكافحة الجرائم الإلكترونية، المادة (3)).
- 6.2** تطوير خوارزميات الذكاء الاصطناعي (على سبيل المثال: (Amazon SageMaker، و(ELKI، و(PyOD))؛ لاكتشاف أنماط النشاط المشبوهة والإبلاغ عنها، مثل: النشر السريع للوسائط، أو الزيادات غير العادية في المشاركة، ويجب تنفيذ ميزات لتحديد وتقييد الحسابات والسلوكيات غير الحقيقية باستخدام تقنيات التعلم الآلي (على سبيل المثال: الكشف عن الحالات الشاذة، وتحليل السلوك)، حيث قد تؤدي الجرائم الإلكترونية إلى السجن أو الغرامات أو كليهما. (المراجع: نظام مكافحة الجرائم الإلكترونية المادة (3)).
- ولضمان أن تطوير وتطبيق تقنيات التزييف العميق يلتزم بأعلى المعايير الأخلاقية، يجب دمج إطار تقييم المخاطر في جميع المراحل، من التطوير إلى النشر.
- وتم تصميم هذا الإطار لمعالجة المخاطر المحتملة وفرض الالتزام بالمبادئ الأخلاقية الأساسية.
- ويرد في الملحق إطار تقييم المخاطر التفصيلي، بما في ذلك الأسئلة المحددة واستراتيجيات التخفيف من المخاطر لكل مرحلة.

## 6. إرشادات لصنّاع محتوى التزييف العميق

يجب على صنّاع المحتوى استخدام تقنية التزييف العميق بشكل مسؤول من خلال الالتزام بالمعايير الأخلاقية وحماية البيانات الشخصية ومنع سوء الاستخدام، ويُعد الامتثال للقوانين واحترام الأخلاق الاجتماعية والمواثيق مع القيم الأساسية أمراً ضرورياً.

ونظراً إلى التأثير المجتمعي الكبير للتزييف العميق، فمن الضروري توجيه صنّاع المحتوى نحو الاستخدامات الإيجابية والبناءة.

وتهدف المتطلبات التالية إلى الحفاظ على ثقة الجمهور، ومنع سوء الاستخدام، ومعالجة المخاطر، مثل: العمليات الاحتيالية، والتلاعب غير المصرح به، والتضليل، والدعاية.

### 1. الامتثال التنظيمي

**1.1.** التأكد من عدم استخدام خدمات التزييف العميق لإنتاج أو إعادة إنتاج أو نشر أو نقل محتوى محظور بموجب القوانين أو اللوائح الإدارية، ويجب أن يتوافق كل محتوى التزييف العميق مع حماية البيانات ذات الصلة والملكية الفكرية واللوائح الخاصة بالصناعة.

(المراجع: المادة (5) من نظام حماية البيانات الشخصية، والمادة (6) من نظام مكافحة الجرائم الإلكترونية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (6)).

### 2. خصوصية البيانات وحمايتها

**2.1.** تنفيذ بروتوكولات أمان قوية، بما في ذلك التشفير الشامل والمصادقة المتعددة العوامل (MFA) وضوابط الوصول الصارمة لحماية البيانات الشخصية المشاركة في صنّاع محتوى التزييف العميق، ويجب التأكد من أن أنظمة تخزين البيانات مرنة في مواجهة الانتهاكات والوصول غير المصرح به. (المراجع: المادة (19) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**2.2.** إعداد وتطوير خطة الاستجابة للحوادث لضمان سرعة معالجة وتخفيف أي انتهاكات للبيانات أو التهديدات الأمنية، مما يضمن الحد الأدنى من التأثير على سلامة البيانات الشخصية. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**2.3.** الحصول على موافقة كتابية صريحة من الأفراد أو الجهات قبل استخدام بياناتهم في صنّاع أو إنشاء محتوى تزييف عميق. (المراجع: المواد (4) و(7) و(11) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).

**2.4.** الاحتفاظ بسجلات الموافقة بشكل آمن، مع التأكد من تشفيرها ونسخها احتياطياً لمنع فقدانها أو التلاعب بها، ويجب أن تكون أنظمة إدارة الموافقة قابلة للتدقيق ومتوافقة مع قوانين حماية البيانات ذات الصلة. (المراجع: المواد (4) و(7) و(11) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**2.5.** تزويد أصحاب البيانات بمعلومات شاملة حول كيفية استخدام بياناتهم، بما في ذلك الأغراض المحددة والمخاطر المحتملة والسياق الأوسع لتطبيق المحتوى. (المراجع: المواد (4) و(7) و(11) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**2.6.** تخزين سجلات الموافقة بتنسيق آمن وسهل الوصول إليه، بما يضمن إمكانية استرجاعها بسهولة لإجراء عمليات التدقيق أو التحقق النظامي. (المراجع: المواد (4) و(7) و(11) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)). يمكن الاطلاع على نموذج الموافقة في الملحق.

**2.7.** التأكد من استخدام البيانات الشخصية فقط للأغراض المحددة في اتفاقية الموافقة، وأي توسيع للاستخدام بما يتجاوز الموافقة الأصلية يتطلب اتفاقية جديدة وصريحة. (المراجع: المواد (4) و(16) و(18) من نظام حماية البيانات الشخصية، ومبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (2)).

**2.8.** التعبير بوضوح عن الحق في سحب الموافقة في أي وقت، من خلال عملية مباشرة للأفراد لممارسة هذا الحق، والتأكد من عدم استخدام البيانات المسحوبة وحذفها بشكل آمن إذا طلب ذلك. (المراجع: المادة (5) من نظام حماية البيانات الشخصية).

### 3. الشفافية والوضوح

**3.1.** الاحتفاظ بسجلات تفصيلية لعملية إنشاء التزييف العميق، بما في ذلك مصادر البيانات والخوارزميات والأدوات والتقنيات المستخدمة، ويجب التأكد من أن هذه الوثائق خاضعة للتحكم في الإصدار وتتضمن الأساس المنطقي وراء الاختيارات المنهجية. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (6)).

**3.2.** استخدام الأنظمة الآلية لتطبيق العلامات المائية بشكل متسق ومنع الحذف العرضي، ويجب دمج هذه الأنظمة في سير عمل إنشاء المحتوى لتقليل الأخطاء البشرية. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (6)).

**3.3.** تنفيذ علامات مائية مرئية ومقاومة للتلاعب (على سبيل المثال: استخدام أساليب إخفاء المعلومات) على جميع محتويات التزييف العميق، مما يضمن إمكانية التعرف دائماً على الطبيعة الاصطناعية للمحتوى، بغض النظر عن النظام الأساسي أو الوسيط، ويجب أن تكون العلامات المائية متينة عبر التنسيقات وأنواع الوسائط المختلفة. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (6)).

**3.4.** استخدام أنظمة قوية للتحكم في الإصدار لجميع الوثائق لضمان تتبع جميع التغييرات وتسجيلها وإمكانية تتبعها بسهولة.

### 4. المساءلة والمسؤولية

**4.1.** تحمل مسؤولية المحتوى الذي تم إنشاؤه باستخدام تقنية التزييف العميق والتأكد من التزامه باللوائح والمعايير ذات الصلة. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).

**4.2.** وضع أدلة استرشادية داخلية صارمة تغطي المعايير الأخلاقية والنظامية والفنية لإنشاء محتوى التزييف العميق، ويجب ضمان تحديث هذه الأدلة الاسترشادية بانتظام لتعكس أفضل الممارسات ومعايير الصناعة. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).

**4.3.** توزيع محتوى التزييف العميق من خلال قنوات آمنة وخاضعة للرقابة؛ لمنع الوصول أو التغيير غير المصرح به، ويجب تنفيذ إدارة الحقوق الرقمية (DRM) وشبكات (CDN) الآمنة لحماية سلامة المحتوى، وقد تؤدي الانتهاكات إلى عقوبات، بما في ذلك الغرامات أو السجن. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (5)، والمادة (3) من نظام مكافحة الجرائم الإلكترونية).

**4.4.** دمج تقنية التجزئة المشفرة وتقنية سلسلة الكتل "البلوك تشين" لإنشاء سجلات ثابتة للمحتوى الأصلي، ويجب أن تتيح هذه الأدوات إمكانية التحقق من سلامة المحتوى، مما يضمن اكتشاف أي تعديلات وتتبعها حتى المصدر. (المراجع: مبادئ أخلاقيات الذكاء الاصطناعي - المبدأ (7)).

## 7. إرشادات للجهات التنظيمية

يُعد التنظيم الفعال لتقنيات التزييف العميق أمراً ضرورياً للتخفيف من المخاطر المتعلقة بالخصوصية والأمان، ويجب على الجهات التنظيمية تنفيذ مجموعة متنوعة من التدابير لمعالجة هذه المخاطر، مع التركيز أيضاً على تسخير الفوائد المحتملة للتزييف العميق. ويهدف هذا القسم إلى تحقيق التوازن بين التنظيم وتسهيل مزايا تقنية التزييف العميق.

### 1. تنظيم المنصات ومراقبتها

**1.1.** تفويض وسطاء الإنترنت ومقدمي خدمات الوسائط ومنصات التواصل الاجتماعي بتعطيل الوصول إلى المحتوى المزيف العميق الذي تم تحديده على أنه مُلفق أو مُضلل أو منع انتشاره. وينبغي أن يشمل ذلك إنشاء عملية شفافة للإبلاغ عن هذا المحتوى وإزالته.

**1.2.** إجراء مراقبة منتظمة لهذه المنصات لضمان الالتزام باللوائح وفرض عقوبات على عدم الامتثال للتوجيهات المتعلقة بالمحتوى الضار.

## 2. عمليات تقييم المخاطر والموافقة عليها

2.1 إجراء تقييمات شاملة لتقنيات التزييف العميق المستخدمة أو المعتمدة للاستخدام داخل الدولة، ويتضمن ذلك تقييم الخوارزميات ومصادر البيانات والآثار الأخلاقية المرتبطة بتطويرها.

2.2 إنشاء عملية موافقة رسمية على تقنيات التزييف العميق للتأكد من استيفائها للمعايير الأخلاقية والنظامية والتقنية المحددة مسبقاً قبل نشرها أو استخدامها تجارياً.

2.3 إنشاء أطر محددة لتقييم المخاطر مصممة خصيصاً للجهات الحكومية التي ستستخدم تقنيات التزييف العميق أو توافق على استعمالها.

2.4 تطوير وإنفاذ معايير لإنشاء وصناعة المحتوى، مثل تلك التي يقترحها التحالف من أجل مصدر المحتوى والأصالة (C2PA)، لضمان الشفافية والمساءلة في إنشاء وتوزيع وسائط التزييف العميق وغيرها من الوسائط التي يُنشأها الذكاء الاصطناعي.

## 3. العقوبات والجزاءات

3.1 التأكد من أن العقوبات والجزاءات الخاصة بإساءة استخدام تقنيات التزييف العميق تتناسب مع خطورة الجريمة التي يتم ارتكابها، مع الأخذ في الاعتبار عوامل مختلفة، مثل: النية، والتأثير، والتكرار.

3.2 وضع أحكام لتخفيف العقوبات أو الجزاءات على الاستخدامات العرضية أو البسيطة للتزييف العميق.

(على سبيل المثال: إذا كان الفيلم يعرض مشهداً قصيراً لحشد ما حيث تظهر نسخة رقمية طبق الأصل من شخص معروف بين عدد من الأشخاص الآخرين، يمكن اعتبار ذلك بمثابة تفاصيل بسيطة وعرضية بدلاً من كونها نقطة محورية للمشهد، وعلى هذا النحو، سيتم استبعاده من اعتبارات العقوبة).

## 4. الشفافية والمساءلة

4.1 تكليف بإعداد تقارير منتظمة (سنوياً على سبيل المثال) حول استخدام تقنيات التزييف العميق من قبل الجهات الحكومية، بما في ذلك تفاصيل حول الأغراض والنتائج وأي مشكلات تتم مواجهتها وذلك لتعزيز الشفافية والثقة العامة.

4.2 تنفيذ عمليات تدقيق وآليات إشراف منتظمة لضمان الاستخدام الأخلاقي لتقنيات التزييف العميق وفقاً للآطر التنظيمية، وينبغي إجراء عمليات التدقيق من قبل هيئات مستقلة لضمان الحياد والموضوعية.

4.3 تطوير والحفاظ على جرد حالة الاستخدام السنوي الذي يجمع البيانات من مختلف الوكالات الحكومية حول استخدام تقنيات التزييف العميق، بما في ذلك المحتوى الإجمالي الذي تم إنتاجه والاتجاهات بمرور الوقت.

## 5. التثقيف والتوعية العامة

5.1 تطوير وتنفيذ برامج تدريب إلزامية للموظفين الحكوميين والجهات التنظيمية تتناول الجوانب الأخلاقية والنظامية والتقنية لتقنيات التزييف العميق، وينبغي تحديث هذه البرامج بانتظام لتشمل آخر التطورات والتحديات.

5.2 إبرام الشراكات مع المنظمات المختلفة، مثل: مؤسسة (EndTAB) (إنهاء إساءة استخدام التقنية) للاستفادة من خبراتها في مجال التعليم والإبلاغ عن إساءة الاستخدام، ويمكن لمواردها أن تعزز برامج التدريب وحملات التوعية العامة، وتجهيز كل من الموظفين الحكوميين والجمهور للتعامل مع الانتهاكات التي تدعمها التقنيات والإبلاغ عنها، بما في ذلك التزييف العميق.

5.3 إطلاق حملات توعية عامة لتثقيف المواطنين حول تقنيات التزييف العميق وتطبيقاتها المحتملة والضمانات المعمول بها لمنع إساءة الاستخدام، وهذا من شأنه تعزيز الخطاب العام المستنير.

5.4. تنفيذ التقارير السنوية للبرامج التعليمية المتعلقة بالتزييف العميق لموظفي الحكومة، ويجب أن تتضمن هذه التقارير تفاصيل عدد الدورات التدريبية التي تم عقدها، وأعداد الحضور، والمحتوى المغطى، وفعالية البرنامج بشكل عام.

5.5. مطالبة جميع الموظفين الحكوميين المعيّنين بالمشاركة في التدريب السنوي أو ورش العمل التنشيطية حول التزييف العميق لضمان بقائهم على اطلاع بأحدث التطورات والتحديات في المجال.

## 8. إرشادات لعملاء التزييف العميق لاكتشاف التزييف العميق ومنع المخاطر

يمكن تحديد التزييف العميق بشكل فعّال من خلال نهج يتكون من ثلاث خطوات لفحص صحة المحتوى، وفيما يلي طريقة تفصيلية لاكتشاف حالات التزييف العميق المحتملة:

### 8.1. تقييم الرسالة:

1. **المصدر:** التأكد من مصداقية المصدر، مثل: الجهات الرسمية، أو الأفراد المعروفين، والتحقق من المعلومات مع المواقع الرسمية أو الاتصال بالمصدر مباشرة من خلال القنوات المعترف بها للتأكد من صحتها.
2. **السياق:** تحديد ما إذا كان المحتوى متوافقاً مع ما هو متوقع عادة من المصدر، على سبيل المثال: من غير المرجح أن يؤدي المسؤولون العموميون خطط الاستثمار أو يروجوا لعروض غير عادية.
3. **الهدف من المحتوى:** كن حذراً من الرسائل التي تحثك على اتخاذ إجراء فوري، مثل: النقر على الروابط، أو تنزيل التطبيقات، أو تقديم معلومات شخصية، وناقش أي رسائل مشبوهة مع الأصدقاء أو العائلة للتعرف على وجهة نظرهم.

### 8.2. تحليل العناصر السمعية والبصرية:

**حركات الوجه غير المنتظمة:** راقب تعبيرات الوجه غير الطبيعية أو المخرجة، خاصة حول الفم والعينين والحاجبين، إذ قد تواجه تقنيات التزييف العميق حركات خفية، مثل: التعبيرات الدقيقة، أو تشنجات العضلات الطفيفة، وعلى الرغم من تحسن تقنية التزييف العميق المتقدمة، إلا أن أوجه الشذوذ والمخالفات الصغيرة لا تزال تشير إلى التلاعب.

**تناقضات الإضاءة:** على الرغم من التقدم، قد لا تزال أساليب التزييف العميق تظهر تناقضات في الإضاءة، فابحث عن الظلال أو النقاط البارزة على الوجه التي لا تتوافق مع مصدر الضوء أو المناطق المحيطة، حيث يمكن أن تشير هذ التناقضات، خاصة في المشاهد المعقدة أو الحركية، إلى وجود حالة تلاعب.

**تغيرات لون الجلد:** في حين أن العرض المتقدم يقلل من التغيرات الملحوظة في لون البشرة، إلا أنه قد تظهر بعض حالات الشذوذ والمخالفات أثناء انتقالات الإضاءة أو الحركات السريعة، فانتبه للتغيرات المفاجئة في نسيج الجلد أو لونه التي لا تتطابق مع المشهد، لأنها قد تشير إلى التلاعب.

**الأنماط الواضحة:** تحاكي تقنية التزييف العميق الحديثة الوميض بشكل طبيعي أكثر، ولكن لا يزال من الممكن أن تكشف مشكلات حركة العين الدقيقة عن حركات مزيفة، فابحث عن حركات العين غير الطبيعية، مثل: العيون التي تبدو ثابتة بشكل مفرط، أو حركة الرمض ببطء أو بسرعة غير عادية، أو غير متوافقة مع السياق العاطفي للمشهد، وقد يكون عدم التزامن الدقيق بين الرمض والحالة العاطفية للموضوع بمثابة دليل أيضاً.

**مشكلات مزمنة حركة الشفاه:** لقد تطورت تقنية مزمنة حركة الشفاه، مما يجعل اكتشاف عدم التطابق أكثر صعوبة، ومع ذلك فإن التأخير الطفيف، أو حركات الفك غير الطبيعية، أو الشفاه التي لا تنغلق بشكل كامل أثناء إصدار أصوات معينة يمكن أن تكشف عن وجود تناقضات، فانتبه إلى تزامن حركات الشفاه مع الصوتيات المعقدة، حيث قد تكون الأخطاء أكثر وضوحاً

**الحالات الشاذة العالية الدقة:** مع تقدم تقنية التزييف العميق في الجودة المرئية، قد لا تظهر المخالفات إلا في ظل الفحص التفصيلي أو بدقة عالية، ولاكتشاف التشوهات الخفية في المقاطع المتقنة، يمكن تكبير اللقطات العالية الدقة، مع التركيز على حواف الوجه، حيث تلتقي بخط الشعر أو الأذنين أو خلفية المقطع أو الصورة، فابحث عن أي تشوهات خفية أو مزج غير طبيعي في هذه المناطق، حيث لا يزال بإمكانها الإشارة إلى وجود تقنية التزييف العميق.

**تحليل المشهد الحركي:** غالباً ما تعاني تقنية التزييف العميق من الحركات السريعة أو المشاهد المعقدة التي تتضمن مصادر إضاءة متعددة أو خلفيات معقدة، فقم بتحليل المشاهد من خلال الحركات السريعة أو التفاعلات مع البيئات المعقدة، حيث يمكن البحث عن عدم تناسق الحركة الضبابية أو المزج غير الطبيعي للموضوع مع الخلفية أثناء التسلسلات الحركية.

## 8.3. مصادقة المحتوى باستخدام الأدوات:

**أدوات الكشف:** استخدم أدوات الذكاء الاصطناعي المصممة لتحليل المحتوى بحثاً عن مؤشرات التلاعب، ويمكن لهذه الأدوات التدقيق في عناصر مختلفة من الوسائط لتحديد التناقضات التي قد تشير إلى استخدام تقنية التزييف العميق:

- **تحليل البيكسل:** تستلزم هذه الطريقة فحص التفاصيل على مستوى البيكسل لصورة أو مقطع فيديو للكشف عن الحالات الشاذة التي قد تشير إلى التلاعب، ويمكن أن يكشف هذا التحليل عن تناقضات في الإضاءة والظلال وتعبيرات الوجه، التي غالباً ما تشير إلى التزييف العميق.
- **تناقضات الحركة:** يمكن لبرنامج تحليل الحركة اكتشاف أوجه الشذوذ في حركة الأشياء أو الأشخاص داخل الفيديو، ونظراً إلى أن التزييف العميق غالباً ما يواجه صعوبة في تكرار الحركات الطبيعية بدقة، يمكن أن تساعد هذه الأدوات في تحديد أنماط الحركة غير الطبيعية.
- **المزامنة السمعية والبصرية:** تم تصميم أدوات معينة للتحقق مما إذا كان الصوت يتوافق مع المحتوى المرئي، وقد تواجه تقنية التزييف العميق صعوبة في مزامنة حركات الشفاه مع الكلمات المنطوقة، خاصة عندما يتم إنشاء الصوت أو تغييره بشكل مصطنع.
- **التحليل القائم على الذكاء الاصطناعي:** توفر الأدوات المفتوحة المصدر، مثل: ( Deepware Scanner ) و ( Sensity ) القائمة على الذكاء الاصطناعي طرقاً يمكن الوصول إليها لاكتشاف الوسائط التي تم التلاعب بها، وتستفيد هذه الأدوات من الذكاء الاصطناعي لفحص مقاطع الفيديو بحثاً عن مؤشرات شائعة للتزييف العميق، مثل: حركات الوجه غير الطبيعية، أو أدوات التلاعب.
- **الكشف عن المعلومات المضللة:** تقوم أدوات مثل المؤشر العالمي للتضليل بتعقب المعلومات المضللة عبر منصات مختلفة باستخدام معلومات استخباراتية مفتوحة المصدر ويمكن استخدامها أيضاً لتحديد التزييف العميق.

**مصدر المحتوى:** يُعد تحديد أصل المحتوى أمراً ضرورياً للتحقق من صحته، ومن خلال فحص البيانات الوصفية والعلامات المائية الرقمية المضمنة، يمكن للمرء تتبع المحتوى مرة أخرى وصولاً إلى مصدره، وفيما يلي بعض الطرق والأدوات المستخدمة لمعرفة مصدر المحتوى:

- **تحليل البيانات الوصفية:** توفر البيانات الوصفية تفاصيل حول الملف، بما في ذلك تاريخ ووقت إنشائه والجهاز المستخدم والبرامج المستخدمة، ويمكن أن يكشف تحليل هذه المعلومات عن تناقضات أو علامات تلاعب، فعلى سبيل المثال: يمكن أن تكون التناقضات بين تاريخ إنشاء الملف والحدث المزعم تمثيله بمثابة إشارة تحذير.

- **العلامة المائية الرقمية:** يتم تضمين العلامات المائية في المحتوى لتأكيد أصلها، وقد أسهمت التطورات في مجال التقنية في إتاحة الفرصة أمام إدراج علامات مائية رقمية غير ظاهرة بشكل واضح في الصور ومقاطع الفيديو تعود إلى صانع المحتوى الأصلي، وهذه العلامات المائية مقاومة لطرق التحرير الشائعة ويمكن استخدامها للتحقق مما إذا كان قد تم تغيير المحتوى من عدمه.
- **أدوات أصالة المحتوى:** توفر منصات، مثل: مبادرة (Adobe Content Authenticity Initiative) و (Truepic) حلولاً لضمان سلامة المحتوى الرقمي، وتقوم هذه الأدوات بتضمين بيانات اعتماد المحتوى عند نقطة الالتقاط، مما يسمح للمستخدمين بتتبع سجل الوسائط والتأكد من عدم العبث بها.
- **تقنية سلسلة الكتل "البلوك تشين":** يتم استكشاف تقنية سلسلة الكتل "البلوك تشين" كطريقة لتتبع المحتوى والمصادقة عليه، ومن خلال تخزين المحتوى في مستودع بيانات لامركزي، تقدم تقنية البلوك تشين سجلاً ثابتاً للأصل المحتوى وأي تعديلات خضع لها، وتتيح هذه التقنية إنشاء سجل شفاف ويمكنه تتبع المحتوى، مما يزيد من صعوبة تقديم محتوى التزييف العميق على أنه محتوى أصلي.
- **الاستفادة من أدوات ومعايير الصناعة:** مع تزايد انتشار التزييف العميق، يقوم قادة الصناعة بإنشاء واعتماد معايير لضمان صحة المحتوى، وغالباً ما تتضمن هذه المعايير أدوات ومنصات مصممة لاكتشاف التلاعب والتحقق من صحة المحتوى.

ومن الأمثلة على ذلك:

- **معرف المحتوى في (YouTube):** تم تصميم نظام معرف المحتوى (Content ID) في موقع (YouTube) لتحديد المواد المحمية بحقوق الطبع والنشر التي يتم تحميلها إلى المنصة، وعلى الرغم من أنه ليس مخصصاً على وجه التحديد لاكتشاف حالات التزييف العميق، إلا أنه يمكن تكييف النظام لاكتشاف المحتوى الذي تم التلاعب به من خلال تحليل مسارات الصوت والفيديو.
- **أصل المشروع في برنامج (Adobe):** يهدف أصل المشروع في برنامج (Adobe) إلى إرفاق أصل يمكن التحقق منه بالوسائط، مما يمكن المشاهدين من الوثوق بمصدر المحتوى.

## 8.4. الإبلاغ عن الحادث:

عندما يتأثر شخص ما بالتزييف العميق، يمكن اتخاذ خطوات محددة لمعالجة الموقف وحماية سمعته، وفيما يلي نهج منظم لضحايا حوادث التزييف العميق.

فمن الأهمية بمكان الاعتراف بأن هذه الأدلة الاسترشادية تعمل ضمن الإطار النظامية الأوسع للمملكة العربية السعودية، حيث توفر القوانين واللوائح الإدارية القائمة أحكاماً أكثر تحديداً، ولها الأسبقية على التوجيهات والأدلة الاسترشادية المقدمة هنا؛ فيجب أن يكون المرجع الأساسي لمعالجة حوادث التزييف العميق دائماً هو القوانين واللوائح الشاملة للمملكة.

- **توثيق الأدلة:** يجب على الضحايا حفظ نسخ من محتوى التزييف العميق على الفور، بما في ذلك الروابط ولقطات الشاشة والبيانات الوصفية، ويُنصح بإنشاء ملف شامل يتضمن جميع الأدلة ذات الصلة، مثل: التواريخ، والأوقات، والمنصات، التي تم فيها نشر حالات التزييف العميق، وسيكون هذا التوثيق الشامل حاسماً للإجراءات النظامية أو إبلاغ السلطات.



• **التعرف على أنواع الاستخدام الخبيث:** من الأهمية بمكان إجراء تقييم دقيق لاستخدام تقنية التزييف العميق ما إذا كان يتم استخدامه للاحتيال، أو التشهير، أو الابتزاز، أو النشر غير المصرح به للبيانات الشخصية، ويجب على الضحايا تصنيف وسائط التزييف العميق بناءً على الاستخدام الخبيث والضرر المقصود والاستعداد للإبلاغ عنه وفقاً لذلك، بما يضمن توثيق كل جانب من جوانب إساءة الاستخدام بدقة لدعم أي إجراءات نظامية. وللمساعدة في تحديد نوع معين من وسائط التزييف العميق، يمكن للضحايا الرجوع إلى موارد مختلفة، مثل: كتاب حالات التلاعب بالوسائط الذي يقدم أمثلة مفصلة وأبحاثاً حول أشكال مختلفة من التلاعب بالوسائط. وفقاً للمادة (36) من نظام حماية البيانات الشخصية، سيتم التحقيق في الانتهاكات، مع إصدار تحذيرات أو فرض غرامات بناءً على طبيعة الانتهاك وشدته وتأثيره. بالإضافة إلى ذلك تحظر المادة (24) من نظام حماية البيانات الشخصية الكشف عن البيانات الشخصية دون موافقة.

• **إبلاغ المنصة عن الحادث:** يجب إبلاغ المنصة عن حالات التزييف العميق على الفور المنشورة عليها تلك الحالات (على سبيل المثال: فيسبوك وإكس (تويتر سابقاً))، وطلب إزالتها، وعند تقديم البلاغ، يجب على الضحية الرجوع إلى شروط خدمة المنصة، وخاصة البنود التي تحظر انتهاكات خصوصية المستخدم ونشر المحتوى الضار، وتطبق معظم منصات التواصل الاجتماعي سياسات صارمة ضد المحتوى الذي ينتهك حقوق الخصوصية، ويمكن أن يؤدي الاستشهاد بهذه الشروط المحددة إلى تعزيز قضية اتخاذ إجراءات سريعة، بالإضافة إلى ذلك يمكن للضحايا طلب الدعم من منظمات، مثل: (Cyber Smile Foundation) المتخصصة في معالجة جميع أشكال الإساءة عبر الإنترنت، بما في ذلك الحوادث التي تنطوي على التزييف العميق.

• **إبلاغ السلطات المحلية بالحادث:** بعد تقديم التقرير إلى المنصة، يجب على الضحايا إخطار السلطات المحلية لضمان اتخاذ مزيد من الإجراءات، ويمكن القيام بذلك عن طريق إعادة إرسال المحتوى الضار إلى (330330) أو تقديم شكوى عبر تطبيق كلنا أمن، الذي يربط المستخدمين مباشرة بجهات إنفاذ النظام، ويجب على الضحايا أيضاً الإبلاغ عن الحادث إلى وحدة الجرائم الإلكترونية في وزارة الداخلية من خلال البوابة الإلكترونية أو عن طريق زيارة مركز الشرطة، وبالنسبة للحالات التي تنطوي على احتيال مالي أو سرقة هوية، من الضروري إبلاغ الخدمة الإلكترونية للبنك المركزي السعودي وإبلاغ البنك لمنع حدوث أضرار إضافية والشروع في اتخاذ تدابير وقائية. (بموجب المادة (15) من نظام مكافحة الجرائم الإلكترونية، فإن هيئة التحقيق والادعاء العام مسؤولة عن التحقيق في الجرائم الإلكترونية. بالإضافة إلى ذلك تنص المادة (14) من نظام مكافحة الجرائم الإلكترونية على أنه يجب على هيئة الاتصالات والفضاء والتقنية (CST) تقديم المساعدة والدعم الفني للأجهزة الأمنية المختصة أثناء التحقيقات في الجرائم الإلكترونية).

• **طلب المشورة النظامية:** من المهم التشاور مع محام متخصص في الحقوق الرقمية وقوانين حماية البيانات، ويجب على الضحايا تعيين ممثل نظامي (محام) لتولي تقديم المشورة في هذه العملية، وتزويدهم بجميع الأدلة والحجج النظامية ذات الصلة للتحضير للإجراءات المدنية والجنائية.

• **إشراك الطب الشرعي الرقمي:** ينصح بإشراك خبراء الطب الشرعي الرقمي لتتبع أصول التزييف العميق واستخدام أدوات الكشف عن الذكاء الاصطناعي، ويوصى أيضاً بالتعاقد مع شركة أبحاث جنائية رقمية ذات سمعة طيبة لتحليل محتوى التزييف العميق، بما يضمن توثيق النتائج بدقة ويمكن استخدامها بفاعلية في السياقات النظامية.

• **تعزيز الأمن الشخصي:** يجب على الضحايا اعتماد تدابير أمنية معززة عبر جميع المنصات الرقمية والنظر في الاشتراك في خدمات المراقبة للحماية من مزيد من إساءة استخدام بياناتهم الشخصية، وتشمل الخطوات الرئيسية تعزيز إعدادات الخصوصية على جميع الحسابات عبر الإنترنت، وتمكين المصادقة الثنائية، والمراقبة المنتظمة للحوادث الإضافية.

• **الرجوع إلى القوانين الشاملة:** تقدم هذه الوثيقة إرشادات حول تطوير واستخدام وتأثير التزييف العميق، ومع ذلك يجب أن تكون الأسبقية للقوانين واللوائح الإدارية المعمول بها، ولذلك من الأهمية بمكان استخدام نظام حماية البيانات الشخصية نظام مكافحة الجرائم الإلكترونية والنظام الجنائي لمعالجة حالات التزييف العميق، ويجب على الضحايا التعاون مع محاميهم لصياغة استراتيجية نظامية شاملة تتضمن هذه القوانين لاستكشاف جميع السبل النظامية المتاحة في إطار الاستجابة لحالات التزييف العميق.

## 8.5 أفضل الممارسات لحماية النفس والآخريين:

للتخفيف من المخاطر المرتبطة بالتزييف العميق، يجب على الأفراد والمؤسسات تنفيذ أفضل الممارسات التالية:

- **الحد من التعرض للبيانات الشخصية عبر الإنترنت:** التقليل من كمية البيانات والصور الشخصية التي تتم مشاركتها على وسائل التواصل الاجتماعي والمنصات العامة الأخرى لتقليل المواد الخام المتاحة لإنشاء التزييف العميق، وتجنب نشر صور عالية الدقة ومقاطع فيديو شخصية يمكن استغلالها لأغراض ضارة.
- **تثقيف الجمهور حول مخاطر التزييف العميق:** زيادة الوعي حول أساليب التزييف العميق ومخاطرها المحتملة، وتثقيف الأشخاص حول كيفية التعرف على المحتوى المريب والاستجابة له، ويمكن أن تتضمن حملات التوعية العامة عقد ورش عمل تطوير محتوى تثقيفي عبر الإنترنت وتعاوناً مع وسائل الإعلام لنشر المعلومات حول التزييف العميق.
- **التعرف على خصائص التزييف العميق:** بالنسبة للأفراد يعد فهم خصائص التزييف العميق أمراً ضرورياً، ويمكن أن يساعد التعرف على العلامات الشائعة لمحتوى التزييف العميق، مثل: حركات الوجه غير الطبيعية، وعدم انتظام الإضاءة والظلال، وعدم الاتساق في الصوت، في التشكيك في صحة هذه الوسائط، ويمكن أن يؤدي البقاء على اطلاع من خلال المصادر الموثوقة واستخدام الأدوات المتاحة لتحليل المحتوى والتحقق منه إلى التخفيف بشكل كبير من تأثير تهديدات التزييف العميق.
- **تنفيذ عمليات تحقق قوية:** بالنسبة للمؤسسات لا سيما تلك العاملة في مجال الإعلام والاتصالات، إذ يعد إنشاء عمليات تحقق صارمة أمراً بالغ الأهمية لضمان صحة المحتوى، ويتضمن ذلك استخدام أدوات الكشف عن الذكاء الاصطناعي جنباً إلى جنب مع طرق التحقق اليدوية التي يجربها محترفون مدربون، وقد تتضمن عمليات التحقق ما يلي:
  - استخدام أدوات الذكاء الاصطناعي والتحقق من البيانات الوصفية: استخدم أدوات الذكاء الاصطناعي المتقدمة لتحليل مقاطع الفيديو والصور بحثاً عن علامات التلاعب، بالإضافة إلى ذلك افحص البيانات الوصفية للمحتوى الرقمي التي توفر تفاصيل حول تواريخ الإنشاء ومحفوظات التحرير والبرامج المستخدمة، مما يساعد في التحقق مما إذا كان المحتوى قد تم تغييره.
  - التحقق اليدوي والاحترافي:
    - خدمات التحقق من الحقائق: استخدم الخدمات الاحترافية لتدقيق الحقائق للتحقق من صحة المحتوى المشبوه.
    - تحليل الخبراء: يمكن إشراك خبراء الطب الشرعي الرقمي لفحص المحتوى يدوياً بحثاً عن علامات التلاعب بالتزييف العميق، مثل: ملامح الوجه غير المنتظمة، والحركات غير الطبيعية، والتناقضات بين الصوت والمرئيات.
    - مصادر الإسناد الترافقي والعلامة المائية: مقارنة المحتوى المشبوه بمصادر موثوقة ورسمية، وإذا ظهر مقطع فيديو أو مقطع صوتي مشكوك في صحته، فتتحقق منه مقابل الإصدارات الأصلية المعروفة أو اتصل بالمصدر مباشرة للتأكد من صحته، وحماية المحتوى الأصلي باستخدام العلامات المائية والتوقيعات الرقمية إذ تساعد هذه التقنيات على ضمان بقاء الوسائط دون تغيير منذ إنشائها.
    - تعزيز تنمية المهارات: تشجيع التعلم المستمر وتنمية المهارات في مجال الذكاء الاصطناعي والأمن السيبراني، ويمكن أن يؤدي تقديم برامج تدريبية وورش عمل للموظفين والجمهور إلى تحسين قدراتهم على اكتشاف تهديدات التزييف العميق والاستجابة لها بفاعلية.

## 9. نظرة عامة على تقنيات التزييف العميق غير الخبيثة

في حين أن تقنية التزييف العميق ترتبط في كثير من الأحيان بالاستخدامات الخبيثة والضارة، إلا أنها تحمل أيضاً إمكانات تحويلية عبر مختلف القطاعات عند تطبيقها بشكل أخلاقي ومسؤول، ويوضح الشكل (5) ستة مجالات رئيسية يمكن أن توفر فيها التطبيقات غير الخبيثة للتزييف العميق فرصاً في ستة قطاعات أساسية، هي: **التسويق، والترفيه، وتجارة التجزئة، والتعليم، والرعاية الصحية، والثقافة.**

المثال	وصف الغرض	القطاع
يمكن لوزارة التجارة بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي دعم الشركات السعودية في إنشاء مؤثرين افتراضيين على وسائل التواصل الاجتماعي، مما يتيح تسويق شخصي على نطاق واسع	دعم الشركات الصغيرة والمتوسطة السعودية في إنشاء إعلانات مخصصة وجذابة، وتعزيز الجهود التسويقية.	التسويق
يمكن للهيئة العامة للترفيه بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي الاستفادة من تقنية التزييف العميق لإنشاء مضيفين افتراضيين للفعاليات والبرامج التلفزيونية، مما يعزز مشاركة المشاهدين	تعزيز صناعة الترفيه في المملكة العربية السعودية من خلال تسهيل إنشاء شخصيات رقمية وخفض تكاليف الإنتاج	الترفيه
يمكن لوزارة التجارة بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي دعم تجار التجزئة على تنفيذ ميزات التجربة الافتراضية للملابس والإكسسوارات، مما يسمح للعملاء بمشاهدة شكل المنتجات عليهم في الوقت الفعلي.	تنفيذ برنامج تحول قطاع التجزئة في المملكة العربية السعودية من خلال تقديم التجارب الافتراضية، مما يعزز مشاركة العملاء، وتوفير تجارب تسوق مخصصة	تجارة التجزئة
يمكن لوزارة التعليم استخدام مدرسين افتراضيين في المناطق النائية، وذلك لتوفير دروس متخصصة وتفاعلية للطلاب.	تعزيز التجربة التعليمية في المملكة العربية السعودية من خلال تقديم محتوى تعليمي تفاعلي وشخصي، مما يساهم في تقديم تجربة تعليم أكثر جاذبية وسهولة في الوصول إليها.	التعليم
يمكن لوزارة الصحة بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي دمج تقنية التزييف العميق في كليات الطب لتزويد الطلاب بمحاكاة تدريبية واقعية ومتقدمة	إحداث تغيير في التدريب الطبي وتعليم المرضى في المملكة العربية السعودية من خلال تقديم عمليات محاكاة واقعية وقدرات إعادة بناء الصوت	الرعاية الصحية
يمكن لوزارة الثقافة بالتعاون مع الهيئة السعودية للبيانات والذكاء الاصطناعي لإعادة تمثيل أحداث تاريخية محورية في التاريخ السعودي، مثل توحيد المملكة أو حياة الملك عبد العزيز آل سعود - رحمه الله	الحفاظ على التراث الثقافي السعودي وإحيائه من خلال إعادة تمثيل الأحداث التاريخية الواقعية والحفاظ على اللهجات القديمة	الثقافة

الشكل 5: أمثلة غير خبيثة للتزييف العميق في القطاعات.

فيما يلي سوف نستكشف أمثلة نموذجية حول كيفية استفادة هذه القطاعات من التطبيقات المبتكرة والقيمة للتزييف العميق.

### 9.1. التسويق:

**إعلانات مخصصة:** تسمح تقنية التزييف العميق بإنشاء إعلانات مخصصة للغاية تضم المشاهير، بشرط أن يكونوا قد أعطوا موافقة صريحة على القيام بذلك، ويمكن لهذه الطريقة تعزيز معدلات المشاركة والتحويل من خلال صدى أعمق مع الجماهير المستهدفة.

**المحتوى الحركي:** يمكن للمسوقين الاستفادة من هذه التقنية لإنتاج محتوى ترويجي واقعي موجه نحو جمهور معين بتكلفة أقل، وهذا يسهل التوليد والإنتاج السريع للمواد التسويقية المتنوعة التي يمكن أن تتكيف مع تفضيلات الجمهور.

**المؤثرون الافتراضيون:** يمكن لتقنية التزييف العميق أيضاً إنشاء مؤثرين افتراضيين يروجون للمنتجات والخدمات بشكل فعال، وتوفر هذه الشخصيات التي تم إنشاؤها أدوات الذكاء الاصطناعي طرقياً مبتكرة للتفاعل مع الجماهير، وتقديم رسائل متسقة للعلامة التجارية دون الخضوع للقيود التي يفرضها المؤثرون البشريون.

وفيما يلي قائمة بأهم المزايا والتحديات في إطار ضمان الاستخدام الأخلاقي لتقنيات التزييف العميق في مجال التسويق.

#### المزايا:

- ضمان الحصول على موافقة واضحة وموثقة من الأفراد الذين سيتم استخدام صورههم أو أصواتهم في المواد التسويقية.
- استخدام تقنية التزييف العميق لجعل الأفراد يؤيدون المنتجات أو الخدمات التي وافقوا صراحة على الترويج لها.

#### التحديات:

- تجنب استخدام تقنية التزييف العميق بطرق قد ينظر إليها على أنها تلاعبية أو غير أخلاقية، مثل: تعديل أحداث، أو بيانات العالم الحقيقي، لتناسب السرد التسويقي.
- تجنب استخدام شكل الفرد دون إذنه، لأن هذا يمكن أن يؤدي إلى انتهاكات خطيرة للخصوصية وعواقب نظامية.

## 9.2. الترفيه:

**دعم الإنتاج** يمكن استخدام تقنية التزييف العميق لإنشاء أو تعديل المشاهد في الأفلام، مما يلغي الحاجة إلى إعادة التصوير المكلفة أو المؤثرات الخاصة، ويمكن أن يقلل ذلك بشكل كبير من وقت الإنتاج والنفقات مع تقديم مقاطع مرئية عالية الجودة.

**إعداد مقاطع البودكاست:** يمكن لصناعة الترفيه تسخير تقنيات التزييف العميق لإنتاج مقاطع بودكاست منخفضة التكلفة تتميز بأصوات رواة القصص البارزين أو المشاهير، ويمكن أن يجذب هذا النهج جمهوراً أكبر ويوسع نطاق وصول مقاطع البودكاست.

**تفاعل المعجبين:** يمكن للمشاهير استخدام تقنيات التزييف العميق للتواصل مع معجبيهم من خلال الرسائل أو التفاعلات المخصصة، وتعزيز ولاء المعجبين وخلق تجارب لا تنسى.

وفيما يلي قائمة بأهم المزايا والتحديات في إطار توجيه الاستخدام الأخلاقي لتقنيات التزييف العميق في قطاع الترفيه.

#### المزايا:

- إمكانية استخدام تقنية التزييف العميق لتعزيز سرد القصص في الأفلام والوسائط، بما يسمح بالتكامل السلس للتأثيرات المرئية وتصوير الممثلين.
- استخدام تقنيات التزييف العميق لتسهيل تفاعلات المعجبين الإيجابية التي لا تنسى، مثل: الرسائل المخصصة من المشاهير، أو التجارب الافتراضية التي تعزز ولاء المعجبين، بشرط حصولك على موافقة الشخصية العامة.

#### التحديات:

- تجنب استخدام تقنيات التزييف العميق لإنتاج محتوى يمكن اعتباره ضاراً، حتى لو لم يكن المقصود منه في البداية أن يكون كذلك.

## 9.3. تجارة التجزئة:

**التجارب الافتراضية:** تتيح تقنية التزييف العميق للعملاء تجربة الملابس والإكسسوارات والمكياج افتراضياً، مما يعزز بشكل كبير تجربة التسوق عبر الإنترنت، ويمكن لهذه الميزة التفاعلية خفض معدلات الإرجاع وتعزيز رضا العملاء من خلال تقديم معاينة واقعية للمنتجات.

**عروض المنتجات:** يمكن للشركات استخدام تقنيات التزييف العميق لتقديم عروض المنتجات للمستثمرين عن بُعد، وصياغة عروض تقديمية جذابة تعمل على تحسين فرص تأمين التمويل، ويمكن لهذا النهج تكرار التفاعلات وجهاً لوجه، مما يجعل العروض أكثر إقناعاً.

وفيما يلي قائمة بأمثلة على المزايا والتحديات في إطار ضمان الاستخدام الأخلاقي لتقنيات التزييف العميق في قطاع تجارة التجزئة.

### المزايا:

- الاستفادة من تقنيات التزييف العميق لإنشاء تجارب افتراضية واقعية للملابس والإكسسوارات والمكياج، مما يعزز تجربة التسوق عبر الإنترنت ويقلل معدلات الإرجاع.
- استخدام تقنيات التزييف العميق في عروض المنتجات للمستثمرين لصياغة عروض تقديمية جذابة ومقنعة، خاصة عندما يكون الاتصال عن بُعد.

### التحديات:

- تجنب استخدام التزييف العميق لتقديم المنتجات بطريقة مضللة تشوه مظهرها أو وظائفها أو جودتها، لأن ذلك قد يؤدي إلى استياء العملاء.
- الامتناع عن استخدام تقنيات التزييف العميق لإنشاء إعلانات أو تجارب مخصصة متطفلة بشكل مفرط قد تجعل العملاء يشعرون بعدم الارتياح أو الانتهاك.

## 9.4. التعليم

**التعلم التفاعلي:** يمكن استخدام تقنية التزييف العميق لتطوير محتوى تعليمي جذاب وتفاعلي مع مدرسين افتراضيين مخصصين، وهذا يمكن أن يجعل تجارب التعلم أكثر إثارة ومصممة خصيصاً لتلبية احتياجات الطلاب الفردية.

**الفصول الافتراضية:** يمكن لتقنية التزييف العميق تحسين مستوى التعليم عبر الإنترنت من خلال إنشاء صور رمزية واقعية للمعلمين والطلاب، مما يعزز بيئة تعليمية أكثر تفاعلية وجاذبية.

**التدريب عن بُعد:** يمكن للمدرسين الافتراضيين المدعومين بتقنيات التزييف العميق إجراء جلسات تدريبية عن بُعد، مما يساعد على تقليل التكاليف والتحديات اللوجستية المرتبطة بطرق التدريب التقليدية.

وفيما يلي قائمة بأمثلة على المزايا والتحديات في إطار ضمان الاستخدام المسؤول لتقنيات التزييف العميق في قطاع التعليم.

## المزايا:

- استخدام تقنيات التزييف العميق لإنشاء مواد تعليمية تفاعلية وغامرة، بما يعزّز المشاركة والفاعلية للطلاب.
- استخدام المدربين الافتراضيين الذين يعملون بتقنية التزييف العميق لتوسيع نطاق الوصول إلى خدمات التدريب والتعليم عن بُعد، خاصة للأفراد في المناطق النائية أو المحرومة.
- التأكد من أن المحتوى التعليمي الذي تم إنشاؤه باستخدام تقنية التزييف العميق دقيق ومناسب للأعراف الثقافية، مما يعزّز الفهم الشامل للموضوع.

## التحديات:

- تجنب استخدام التزييف العميق لإنتاج محتوى تعليمي قد ينشر معلومات خاطئة أو مضللة، لأن ذلك قد يكون له آثار ضارة على المتعلمين.
- الامتناع عن الإفراط في الاعتماد على تقنيات التزييف العميق كبديل للمعلمين أو المدربين الحقيقيين، لأن التفاعل والتوجيه الشخصي أمران حيويان ومهمان للغاية في التعلم الفاعل.

## 9.5. الرعاية الصحية

**التدريب الطبي:** يمكن استخدام تقنية التزييف العميق لإنشاء صور طبية واقعية للمرضى، مما يعزّز تدريب الأطباء على تقنيات التشخيص والعلاج، ويمكن لهذا التقدم أن يرفع جودة التعليم الطبي ورعاية المرضى.

**إعادة بناء الصوت:** بالنسبة للمرضى الذين فقدوا قدرتهم على التحدث بسبب المرض أو الحوادث، يمكن لتقنية التزييف العميق تجديد أصواتهم الأصلية، مما يحسّن حياتهم بشكل كبير.

**الحملات الصحية:** يمكن للمنظمات الصحية الاستفادة من تقنيات التزييف العميق لتطوير حملات توعية صحية واقعية وجذابة، وتوصيل الرسائل المهمة بشكلٍ فعّال ودعم مبادرات الصحة العامة.

فيما يلي قائمة بأمثلة على المزايا والتحديات في إطار ضمان الاستخدام الآمن والأخلاقي لتقنيات التزييف العميق في قطاع الرعاية الصحية.

## المزايا:

- استخدام تقنيات التزييف العميق لإنتاج محاكاة طبية واقعية للتدريب، ومساعدة العاملين في القطاع الطبي في تعزيز مهاراتهم ومعرفتهم.
- استخدام تقنيات التزييف العميق لصياغة حملات توعية صحية جذابة وواقعية تصل بشكلٍ فعّال إلى جماهير متنوعة وتدعم مبادرات الصحة العامة.

## التحديات:

- تجنب استخدام تقنيات التزييف العميق لإنشاء محتوى طبي أو نوائح قد تضلل المرضى أو مقدمي الرعاية الصحية، لأن ذلك قد يتسبب في أضرار كبيرة.
- الامتناع عن استخدام تقنيات التزييف العميق لمحاكاة نتائج المرضى أو تطور المرض بطريقة يمكن أن تغرس أملاً كاذباً أو خوفاً لا مبرر له لدى المرضى أو عائلاتهم.

## 9.6. الثقافة

**إعادة تمثيل الأحداث التاريخية:** يمكن لتقنيات التزييف العميق إنشاء عمليات إعادة تمثيل واقعية للأحداث التاريخية، مما يجعل التراث الثقافي جذاباً وتعليمياً، ويمكن لهذا التطبيق إحياء التاريخ وجعله في متناول جمهور أوسع.

**الحفاظ على اللغة واللهجة:** يمكن أن تساعد تقنيات التزييف العميق في الحفاظ على اللغات واللهجات المهددة بالانقراض وإحيائها من خلال توليد أصوات واقعية، وهذا يدعم الحفاظ على الثقافة والتنوع اللغوي.

وفيما يلي قائمة بأمثلة على المزايا والتحديات في إطار ضمان الاستخدام المسؤول لتقنيات التزييف العميق في الحفاظ على الثقافة:

### المزايا:

- استخدام تقنية التزييف العميق لإعادة إنشاء الفعاليات والتقاليد الثقافية والحفاظ عليها، بما يجعلها في متناول جماهير أوسع والأجيال القادمة.
- تطبيق تقنيات التزييف العميق لإحياء وتعليم اللغات واللهجات المهددة بالانقراض، ودعم الحفاظ على التنوع الثقافي والتراث.

### التحديات:

- تجنب استخدام تقنيات التزييف العميق لتصوير الممارسات أو الشخصيات أو الفعاليات الثقافية بشكل غير دقيق، ويمكن أن يؤدي التحريف إلى الاستيلاء الثقافي وسوء الفهم وعدم الاحترام.
- عدم استغلال التراث الثقافي لأغراض تجارية دون الحصول على إذن مناسب واحترام المجتمع المصدر، والتأكد من أن استخدام تقنيات التزييف العميق يفيد المجتمع الثقافي المعني.

إن تسخير قوة تقنيات التزييف العميق لأغراض إيجابية وأخلاقية يمكن أن يؤدي إلى تقدم ملحوظ وتحسين التجارب في مختلف القطاعات، ومن خلال إعطاء الأولوية للتعليم المستمر وتنمية المهارات، وضمان الاستعداد التنظيمي، والتركيز على التطبيقات الأخلاقية، يمكن تحقيق الإمكانيات التحويلية لتقنيات التزييف العميق بالكامل مع تقليل المخاطر المرتبطة بها،

وفيما يلي عرض الأفكار الرئيسية الثلاث:

- **التعلم المستمر وتنمية المهارات:** يُعد التعليم المستمر والتدريب العملي ضروريين لمواكبة تطورات الذكاء الاصطناعي ومنهجيته واعتباره الأخلاقية، بما يضمن الإدارة الفعّالة لتطبيقات الذكاء الاصطناعي.
- **الاستعداد التنظيمي:** يؤدي بناء الخبرة الداخلية في مجال الذكاء الاصطناعي من خلال ورش العمل المستهدفة وممارسات التوظيف الاستراتيجية إلى تعزيز قدرات الابتكار وحل المشكلات، مما يمكن المؤسسات من التعامل ببراعة مع تقنيات الذكاء الاصطناعي والتزييف العميق.
- **التطبيقات الأخلاقية والإيجابية:** يمكن أن يؤدي التركيز على الاستخدامات غير الخبيثة في قطاعات، مثل: التسويق، والترفيه، وتجارة التجزئة، والتعليم، والرعاية الصحية، والثقافة، إلى تحقيق تقدم كبير وإثراء التجارب مع إدارة المخاطر المرتبطة بها بشكل مسؤول.

## 10. الخلاصة

توفر تقنيات التزييف العميق مخاطر كبيرة وفرصاً وأعداء في مختلف القطاعات، وعند استخدامها بطريقة أخلاقية ومسؤولة يمكن أن تؤدي تقنيات التزييف العميق إلى تحقيق تقدم ملحوظ في مجالات التسويق والترفيه وتجارة التجزئة والتعليم والرعاية الصحية والتطبيقات الثقافية، مما يعزز تجارب المستخدم ويرفع مستوى الابتكار، وتسلب الأدلة الاسترشادية المقدمة الضوء على أهمية التعلم المستمر وتطوير المهارات، وتحث المؤسسات على تنمية قدرات الذكاء الاصطناعي الداخلية من خلال ورش عمل مخصصة وممارسات توظيف استراتيجية لإدارة تقنيات التزييف العميق بشكل فعال.

تمثل الاعتبارات الأخلاقية عاملاً حاسماً في نشر تقنيات التزييف العميق، ويجب على المطورين وصناع المحتوى إعطاء الأولوية للخصوصية والشفافية والمساءلة والفوائد الاجتماعية للتخفيف من المخاطر، ويمكن أن يساعد تنفيذ تدابير قوية لحماية البيانات، وتأمين الموافقات، والحفاظ على الشفافية في بناء الثقة ومنع إساءة الاستخدام، بالإضافة إلى ذلك يجب على مطوري تقنيات التزييف العميق التأكد من أن أدواتهم تحتوي على آليات لتتبع صحة المحتوى والتحقق منه لمكافحة إساءة الاستخدام، ويجب على المستخدمين المشاركة في عمليات تحقق صارمة والبقاء على اطلاع بخصائص التزييف العميق وطرق الكشف عنه لحماية أنفسهم والآخرين.

يُعد الوعي العام والتعليم أمرين حيويين لمعالجة الأضرار المحتملة الناجمة عن تقنيات التزييف العميق، ويمكن أن تساعد حملات التوعية والمبادرات التعليمية الأفراد على التعرف على التزييف العميق والاستجابة له بشكل أفضل، مما يُقلل مخاطر التلاعب والمعلومات المضللة، ومن خلال الالتزام بهذه التوصيات يمكن لأصحاب المصلحة ضمان استخدام تقنيات التزييف العميق بطريقة مفيدة وأخلاقية، مما يحافظ على ثقة الجمهور ويعزز الابتكار المسؤول، ومن خلال الجهود التعاونية والالتزام بالممارسات الأخلاقية، يمكن تسخير الإمكانيات التحويلية لهذه الأدوات لدفع التغيير الإيجابي عبر مجالات متعددة.

## 11. الملحق

### 11.1 الملحق 1: تعريفات

**التزييف العميق:** وسائط اصطناعية شديدة الواقعية تم إنشاؤها باستخدام تقنيات التعلم العميق التي تتلاعب بالصوت أو الفيديو أو أي محتوى رقمي آخر لتغيير شكل الأشخاص أو الأشياء أو الأحداث أو أفعالهم أو مظهرهم بشكل مقنع، مما يجعل من الصعب التمييز بين المحتوى الحقيقي والمزيف.

**تقنيات التزييف العميق الخبيثة:** أساليب تزييف عميق يتم إنشاؤها بقصد خداع الأفراد أو إيذائهم أو استغلالهم، وغالباً ما تُستخدم لنشر معلومات كاذبة أو تشهير أو أنشطة احتيالية.

**تقنيات التزييف العميق غير الخبيثة:** أساليب التزييف العميق التي يتم استخدامها لأغراض إيجابية وبناءة، دون نية الخداع أو التسبب في ضرر، وتشمل الأمثلة التطبيقات في قطاعات التسويق والترفيه وتجارة التجزئة والتعليم والرعاية الصحية والثقافة.

**مطور تقنيات التزييف العميق:** فرد أو منظمة مسؤولة عن إنشاء التقنيات الأساسية وخوارزميات الذكاء الاصطناعي التي تمكن من توليد أساليب التزييف العميق، مع التركيز على تطوير الأدوات والأنظمة لإنشاء الوسائط الاصطناعية.

**صانع محتوى التزييف العميق:** فرد أو كيان يستخدم تقنيات التزييف العميق لإنتاج محتوى رقمي، مثل: الصور، أو مقاطع الفيديو، أو الأصوات لأغراض مختلفة، بما في ذلك إنشاء المحتوى للشركات أو الأغراض التعليمية أو البحثية.

**مستخدم التزييف العميق:** أي فرد أو منظمة تتفاعل مع تقنيات التزييف العميق أو تشاهدها أو تتأثر بها، بمن في ذلك الجمهور العام والموظفون والشركات والمؤسسات.



## 11.2. الملحق 2: التدريب وتنمية المهارات

في مجال الذكاء الاصطناعي وتقنيات التزييف العميق السريع التطور، يُعد التعلم المستمر أمراً حيوياً للبقاء على اطلاع بالتطورات والمنهجيات والاعتبارات الأخلاقية الجديدة، ويساعد التعلم المستمر الأفراد والمنظمات على التكيف مع التقدم التقني، وتعزيز مهاراتهم، والحفاظ على ميزة تنافسية.

بالنسبة للمستخدمين النهائيين، يضمن التعلم المستمر قدرتهم على تطوير تطبيقات الذكاء الاصطناعي وتنفيذها وإدارتها بشكل فعال.

وتسهم ورش العمل بدور حاسم في تزويد المستخدمين النهائيين بالمهارات والمعرفة اللازمة لمكافحة الاستخدامات الضيئة لهذه التقنية.

تغطي العناصر الرئيسية لورش العمل الجوانب التالية:

- 1. فهم تقنية التزييف العميق:** تقديم نظرة شاملة حول كيفية إنشاء حالات التزييف العميق، بما في ذلك خوارزميات وتقنيات الذكاء الاصطناعي الأساسية المعنية.
  - 2. آليات الكشف:** تعليم الأساليب المتقدمة لتحديد التزييف العميق، مثل: تحليل التناقضات في الفيديو والصوت، باستخدام أدوات الكشف القائمة على الذكاء الاصطناعي، والتعرف على علامات الوسائط الاصطناعية.
  - 3. الاعتبارات الأخلاقية والنظامية:** استكشاف الآثار الأخلاقية والأطر النظامية المتعلقة باستخدام التزييف العميق لضمان ممارسات مسؤولة وممتثلة للقانون.
  - 4. التدريب العملي:** تقديم جلسات عملية حيث يمكن للمستخدمين النهائيين العمل مع أدوات الكشف ومجموعات البيانات لاكتساب خبرة عملية في تحديد التهديدات العميقة والتخفيف منها.
- ولضمان بقاء المستخدمين النهائيين مطلعين على أحدث التطورات وأفضل الممارسات، يُوصى بإجراء التدريب وورش العمل على فترات منتظمة، ويتضمن الجدول المقترح ما يلي:
- ورش العمل النصف السنوية: استضافة ورش عمل كل ستة أشهر لتقديم تحديثات حول التطورات الأخيرة في تقنيات التزييف العميق وآليات الكشف.
  - الدورات التنشيطية السنوية: تقديم دورات تنشيطية سنوية متعمقة لتعزيز المفاهيم الأساسية ومراجعة الاعتبارات الأخلاقية والنظامية الجديدة وضمان الكفاءة المستمرة في أدوات الكشف.
  - الجلسات الخاصة: تنظيم جلسات بعد التقدم التقني الكبير أو التغييرات في الأطر التنظيمية لإبقاء المشاركين على اطلاع واستعداد للتكيف مع التحديات الناشئة.

إن الالتزام بجدول التدريب هذا سيساعد المستخدمين النهائيين على الحفاظ على مستوى عالٍ من الاستعداد والفعالية في إدارة التهديدات التي تشكلها تقنية التزييف العميق، بالإضافة إلى ذلك يمكن اعتماد هذه المبادئ التوجيهية التدريبية الموجهة من قبل الجهات الحكومية والمطورين والجمهور العام لتعزيز التعليم الفعال والقدرة على مواجهة إساءة استخدام تقنية التزييف العميق.

### 11.2.1. الجهات الحكومية:

تهدف هذه الأدلة الاسترشادية إلى تزويد المسؤولين الحكوميين بالمعرفة والأدوات الأساسية لفهم تهديدات التزييف العميق وتحديدها ومعالجتها، وبالتالي الحفاظ على سلامة الاتصالات العامة والأمن.

ويجب أن يركز التدريب على تقنية التزييف العميق، وطرق الكشف، والأطر النظامية، والآثار السياسية، ودراسات الحالة ذات الصلة، ويجب على المسؤولين اكتساب فهم شامل لإنشاء التزييف العميق، والقضايا الأخلاقية والنظامية المرتبطة به، وتأثيرها المحتمل على الأمن الوطني وثقة الجمهور.

ويُنصح بأن تشارك الجهات الحكومية في ورش عمل ربع سنوية لمواكبة أحدث التطورات، وعلاوة على ذلك ينبغي إجراء جلسات تدريب عملي نصف سنوية باستخدام أدوات الكشف المتقدمة لتقديم الخبرة العملية وتحسين قدرتها على مواجهة تهديدات التزييف العميق بشكلٍ فعّال.

### 11.2.2. المطورون:

تم تصميم هذا القسم لمساعدة المطورين على اكتساب فهم شامل لتقنية التزييف العميق، مما يضمن قدرتهم على إنشاء واكتشاف التزييف العميق بشكلٍ مسؤول ودمج أدوات الكشف الفاعلة في طول البرامج الخاصة بهم.

ويجب أن يركز التدريب على التفاصيل الفنية لإنشاء واكتشاف التزييف العميق، وممارسات تطوير الذكاء الاصطناعي الأخلاقية، وتكامل أدوات الكشف داخل البرامج، ويجب أن تتناول الجلسات المتقدمة خوارزميات الذكاء الاصطناعي، وتقنيات التعلم الآلي، وأن تتضمن تدريبات عملية على البرمجة.

ويتم تشجيع المطورين على المشاركة في الندوات التقنية الشهرية عبر الإنترنت للبقاء على اطلاع بأحدث التطورات في مجال الذكاء الاصطناعي وتقنيات التزييف العميق، بالإضافة إلى ذلك يجب أن تشمل ورش العمل المكثفة النصف السنوية على جلسات تفسير عملية، وخوارزميات الذكاء الاصطناعي المتقدمة، وتمارين حل المشكلات لضمان إعداد المطورين جيداً لمواجهة تحديات التزييف العميق بفعالية.

### 11.2.3. الجمهور العام والمنظمات:

تهدف هذه الأدلة الاسترشادية إلى تعزيز الوعي بين مؤسسات القطاعين العام والخاص حول مخاطر التزييف العميق وتزويدهم بالمعرفة والأدوات اللازمة للتعرف على المحتوى الضار والحماية منه.

ويجب أن يركز التعليم العام على إعلام الأفراد بوجود مخاطر التزييف العميق، كما يجب أن يغطي التدريب طرق الكشف الأساسية، وأهمية التحقق من المعلومات، واستراتيجيات حماية البيانات الشخصية عبر الإنترنت، وبعد تقديم النصائح العملية بشأن تحديد التزييف العميق المشتبه به والإبلاغ عنه أمراً بالغ الأهمية.

ينبغي تنظيم حملات توعية عامة نصف سنوية للتأكيد على مخاطر وعلامات التزييف العميق، ويجب أن تقدم الندوات الإعلامية الشهرية عبر الإنترنت تقنيات الكشف الأساسية ونصائح لحماية البيانات الشخصية، ويجب أن تقدم ورش العمل المجتمعية الربع السنوية جلسات تفاعلية حول كيفية التعرف على التزييف العميق والرد عليه.

ومن خلال تصميم محتوى التدريب وتكراره بما يتوافق مع الاحتياجات المحددة للجهات الحكومية والمطورين والجمهور العام، يمكن لأصحاب المصلحة معالجة التحديات التي تفرضها تقنية التزييف العميق بشكلٍ فعّال، ويُعد التعليم المستمر والمشاركة الاستباقية أمراً ضرورياً للحفاظ على المرونة ضد إساءة استخدام الذكاء الاصطناعي وتقنيات التزييف العميق.

## 11.3. الملحق 3: دراسات الحالة والأمثلة

### 11.3.1. السيناريوهات الضارة للتزييف العميق

**مثال على العمليات الاحتيالية<sup>12</sup>: احتيال للخطف في جنوب آسيا**

في إحدى دول جنوب آسيا، استخدم أحد المحتالين تقنية استنساخ الصوت المتقدمة المعتمدة على الذكاء الاصطناعي لتقليد صوت الطفل، وقام المحتال بالاتصال بالوالدي الطفل، مدعياً كذباً أن طفلهم قد تم اختطافه وطالب بفدية، وفي حالة الذعر، أرسل والدا الطفل مبلغاً كبيراً من المال عبر منصة الدفع الرقمية.

أدت عملية الاحتيال هذه إلى حدوث ضائقة عاطفية كبيرة للوالدين، وخسارة مالية للعائلة، وزيادة الوعي العام بالتهديدات التي تشكلها تقنية التزييف العميق، وتسبب استخدام المحتال غير الأخلاقي للتقنية المتقدمة للتلاعب بالأفراد وخداعهم في حدوث ضرر كبير، يسلط هذا الحادث الضوء على الحاجة الماسة إلى تعزيز المعرفة الرقمية وأهمية التحقق من الاتصالات غير المتوقعة للحماية من مثل عمليات الاحتيال المعقدة هذه.

وقد أطلقت السلطات المحلية تحقيقاً لتتبع عملية الدفع الرقمي وتحديد هوية المتصل، في حين تم إصدار تحذيرات عامة بشأن عمليات الاحتيال التي تعتمد على الذكاء الاصطناعي، ونصح المواطنين باستخدام كلمات رمزية خاصة بالأسرة للتحقق، بالإضافة إلى ذلك شاركت شركات التقنية في فهم الأدوات المستخدمة في عملية الاحتيال ووضع استراتيجيات للوقاية.

**مثال على التلاعب غير المصرح به<sup>13</sup>: فضيحة التزييف العميق غير المصرح به:**

في عام 2023م، عبّرت شخصية سياسية أمريكية بارزة علناً عن معاناتها من صورة اصطناعية تم تداولها لها على إحدى منصات التواصل الاجتماعي دون إذنها تسببت الصورة المزيفة التي أُنبتت صدمات الماضي في صدمة كبيرة وأذى نفسي، ووصفت التزييف العميق بأنه يحمل "نفس النية المتمثلة في الإذلال العنيف".

ولم يتوقف ضرر الحادث على حدوث ضائقة عاطفية كبيرة للضحية، بل أدى أيضاً إلى تقويض ثقة الجمهور، وحفز اتخاذ إجراءات نظامية محتملة ضد المبدعين، وزيادة الوعي العام والاهتمام التشريعي بتنظيم تقنيات التزييف العميق، وإنشاء وتوزيع منتجات التزييف العميق بشكل غير مصرح به يقوض المعايير الأخلاقية من خلال انتهاك الاستقلالية الشخصية والخصوصية مع التسبب في ضرر عاطفي، وتؤكد هذه القضية على الحاجة الملحة إلى وضع مبادئ توجيهية أخلاقية وأطر نظامية قوية لتنظيم إنشاء الوسائط الاصطناعية ونشرها، وضمان حماية الأفراد والحفاظ على ثقة المجتمع.

أدى هذا الحدث إلى طرح مناقشات حول ضرورة وضع لوائح أكثر صرامة لمكافحة إساءة استخدام تقنيات التزييف العميق، بما في ذلك صياغة قوانين جديدة لمنع محتوى التزييف العميق الضار، واقترح المشرعون الفيدراليون ومشروع الولايات مشاريع قوانين مختلفة تستهدف التعامل مع حالات التزييف العميق، خاصة في السياقات السياسية، بهدف حماية الأفراد من انتحال الشخصية والاحتيال، وتعكس هذه الجهود التشريعية مبادرة أوسع لمعالجة التهديد المتطور المتمثل في المعلومات الخاطئة والمضللة التي يتركها الذكاء الاصطناعي وحماية الشخصيات العامة والجمهور العام.

## مثال على التزليل والدعاية<sup>14</sup>: فضيحة مقطع صوتي مزيف:

انتشر على نطاق واسع عبر مواقع التواصل الاجتماعي مقطع صوتي مزيف ومغبرك، يظهر صوت زعيم سياسي من حزب ليبرالي في إحدى الدول الأوروبية، يشير هذا المقطع كذباً إلى أن السياسي كان يشتري الأصوات من مجموعة أقلية.

لم يزل المقطع الصوتي المزيف الرأي العام ويؤثر على القرارات الحاسمة فحسب، بل شوه أيضاً سمعة المشاركين فيه، وقد أثار دعوات إلى إجراء إصلاحات تنظيمية وتحسين أدوات الكشف عن التزييف العميق، فإثناء ونشر التسجيلات الصوتية المزيفة ينتهك المعايير الأخلاقية من خلال نشر معلومات مضللة وخداع الجمهور وتقويض الثقة، ويسلط هذا الحادث الضوء على الحاجة الملحة إلى تعزيز المعرفة الرقمية وأدوات التحقق الفاعلة للحماية من إساءة استخدام الوسائط الاصطناعية، وبالتالي الحفاظ على ثقة المجتمع ونزاهته.

رداً على ذلك كانت هناك دعوات إلى أنظمة وسياسات أكثر صرامة لمكافحة التزييف العميق والأشكال الأخرى من الوسائط الاصطناعية، وبدأت الهيئات التشريعية بالنظر في قوانين جديدة لمنع إنشاء وتوزيع مثل هذا المحتوى المخادع، ونظراً إلى أن البلاد كانت تفتقر إلى تشريعات محددة للتزييف العميق في ذلك الوقت، تم تطبيق قوانين التشهير والجرائم الإلكترونية الحالية، وأدى الحادث إلى تسريع الجهود الرامية إلى تطوير وتنفيذ تشريعات شاملة تهدف إلى معالجة التزييف العميق والوسائط الاصطناعية لحماية الشخصيات العامة والعملية الديمقراطية من الهجمات المماثلة في المستقبل.

## 11.3.2 ■ السيناريوهات المفيدة للتزييف العميق

### مثال تطبيقي في قطاع التسويق<sup>15</sup> : حملة إعلانات بتقنيات التزييف العميق

استخدمت إحدى منصات توصيل الطعام الشهيرة تقنية التزييف العميق لإعلان جذاب وذو صلة بالسياق، وظهر في الإعلان ممثل معروف بدا أنه يطلب الأطعمة المفضلة له، وكان الجانب المبتكر هو أن الإعلان تم تعديله وفقاً للموقع حسب نظام تحديد المواقع العالمي الخاص بالمشاهد، مما يجعل المحتوى أكثر ملاءمة محلياً، وأدى التكامل في تجسيد صورة الممثل مع المحتوى المخصص إلى جعل هذه الحملة الإعلانية تجذب انتباه المشاهدين على مستوى محلي وقريب من واقعهم.

إن استخدام تقنيات التزييف العميق في الحملة الإعلانية أدى إلى تخصيص فَعَالٍ من حيث التكلفة، وزيادة تفاعل المشاهدين، وارتفاع معدلات التحويل، مما أسهم بشكل كبير في زيادة إيرادات المنصة والولاء للعلامة التجارية من خلال استراتيجيات الإعلان المبتكرة، وتعتبر الحملة مثالاً على الاستخدام الأخلاقي لتقنيات التزييف العميق في مجال الإعلان، حيث تتسم بالموافقة والشفافية والإفصاح، ويظهر هذا المثال إمكانيات تقنيات التزييف العميق كأداة قوية للإعلانات المخصصة والحركية، مُبرزاً أهمية الحفاظ على المعايير الأخلاقية، مثل: الموافقة، والإفصاح، والشفافية؛ لتعزيز الثقة والمصداقية في التسويق الرقمي.

### مثال تطبيقي في قطاع الترفيه<sup>16</sup>: تقنية تصغير العمر في الأفلام

في فيلم من إخراج مخرج مشهور من بطولة ممثل بارز في الدور الرئيسي كقاتل محترف، تطلب الفيلم مشاهد استحضار ذكريات فائتة وأحداث ماضية حيث يحتاج الممثلون إلى الظهور بشكل أصغر سناً، ولتحقيق ذلك استخدم صانعو الأفلام تقنية التزييف العميق لإزالة علامات الشيخوخة من الممثل الرئيسي والممثلين الآخرين، وتعديل مظهرهم رقمياً لمنحهم مظهراً أكثر شباباً.

أدى تطبيق تقنية التزييف العميق في الفيلم إلى توفير كبير في التكاليف وكفاءة الإنتاج، حيث أتاحت التقنية تأثيرات إزالة علامات الشيخوخة الواقعية العالية الجودة. كما أثار هذا الابتكار ضجة إعلامية كبيرة، مما أدى إلى تعزيز أداء شبكات التذاكر ومشاركة الجمهور.

كان استخدام تقنية التزييف العميق بعد الحصول على الموافقات اللازمة، مما يدل على استخدام أخلاقي ومفيد للتقنيات في التعبير الفني لإثراء سرد القصة وتعزيز تجربة المشاهدين، ويوضح الفيلم كيف يمكن استخدام تقنية التزييف العميق بشكل خلاق ومسؤول في صناعة الترفيه، مما يؤكد على إمكانات التقدم التقني العالية لرفع مستوى المشاريع الفنية مع الحفاظ على المعايير الأخلاقية.

14. مكتب الصحافة الاستقصائية

15. موقع ديجيتال تريندس

16. سي بي إس نيوز

## • مثال تطبيقي في قطاع الرعاية الصحية<sup>17</sup>: إعادة بناء الصوت لمرضى التصلب الجانبي الضموري

أظهرت إعادة بناء الصوت باستخدام تقنية التزييف العميق نتائج واعدة للأفراد المصابين بالتصلب الجانبي الضموري (ALS)، وهو مرض تنكس عصبي يضعف الخلايا العصبية في الدماغ والحبل الشوكي، مما يؤدي إلى فقدان القدرة على الحركة والكلام.

من خلال استخدام تقنية التزييف العميق لإعادة بناء الصوت، شهدت حياة مرضى التصلب الجانبي الضموري تحسناً كبيراً، وقد مكنتهم هذه التقنية من استعادة قدرتهم على التواصل، ومواصلة أنشطتهم المهنية، والحفاظ على العلاقات الشخصية، مع إطلاق مبادرات جديدة لدعم الآخرين الذين يعانون من ظروف مماثلة، إذ يسلط هذا الاستخدام لتقنية التزييف العميق الضوء على تطبيق مفيد وأخلاقي، مما يعزز بشكل كبير تحسن حياة مرضى التصلب الجانبي الضموري من خلال استعادة قدراتهم على التواصل والحفاظ على جودة حياتهم، ويظهر تطبيقه في السياقات الطبية إمكانات التقنيات في تحسين حياة الأشخاص الذين يعانون من تحديات تواصل حادة، مع الحفاظ على المعايير الأخلاقية.

## 11.4. الملحق 4: إطار عمل تقييم المخاطر لمطوري تقنيات التزييف العميق

### 11.4.1. المرحلة الأولى: التخطيط والتصميم

المرحلة	السؤال	المبدأ	تخفيف حدة المخاطر
PD.1	هل قمت بتصميم المستوى المناسب من الرقابة البشرية على نظام الذكاء الاصطناعي وحالة الاستخدام؟	المساءلة والمسؤولية	التأكد من أن الرقابة البشرية جزء لا يتجزأ من جميع المراحل الحاسمة لمنع إساءة الاستخدام.
PD.2	هل يمنع تصميم نظام الذكاء الاصطناعي الخاص بك الثقة المفرطة في نظام الذكاء الاصطناعي بآليات التدخل البشري الضرورية؟	المساءلة والمسؤولية	تنفيذ نقاط الأمان والتدخل البشري لتصحيح أي أحكام خاطئة بشأن الذكاء الاصطناعي.
PD.3	هل قمت بتحديد عمليات الرقابة البشرية باستخدام مؤشرات الأداء الرئيسية المناسبة وإسناد المسؤولية إلى الأطراف ذات الصلة؟	المساءلة والمسؤولية	مراجعة مؤشرات الأداء الرئيسية وتحديثها بانتظام، مما يضمن المساءلة الواضحة.
PD.4	هل قمت بتصميم استراتيجية التشغيل والحوكمة لتعطيل النظام أو التدخل فيه عندما لا يعمل بالطريقة المقصودة؟	المساءلة والمسؤولية	تطوير واختبار بروتوكولات الاستجابة لفشل النظام أو الانحرافات.
PD.5	هل أخذت في الاعتبار متطلبات المسؤولية وحماية أصحاب البيانات؟	المساءلة والمسؤولية	الامتثال لقوانين حماية البيانات وتأمين الموافقات اللازمة.
PD.6	هل قمت بتعيين الحدود القصوى لمؤشرات الأداء الرئيسية وهل وضعت إجراءات حوكمة أو إجراءات مستقلة لتفعيل خطط بديلة / خطط للتراجع؟	المساءلة والمسؤولية	إنشاء ومراقبة مؤشرات الحد الأدنى لتفعيل خطط الطوارئ.
PD.7	هل قمت بتوفير التدريب والتعليم للمساعدة في تطوير ممارسات المساءلة؟	المساءلة والمسؤولية	إجراء دورات تدريبية منتظمة حول المعايير الأخلاقية والمساءلة.

14. مكتب الصحافة الاستقصائية

15. موقع ديجيتال تريندس

16. سي بي إس نيوز

## 11.4.2 . المرحلة الثانية: تحضير بيانات الإدخال

المرحلة	السؤال	المبدأ	تخفيف حدة المخاطر
PID.1	تنفيذ عمليات فحص آلية وبدوية للامتثال لخصوصية البيانات.	الخصوصية والأمن	هل هناك آلية قائمة تحدد المشكلات المتعلقة بخصوصية البيانات أو حمايتها في عملية جمع البيانات ومعالجتها؟
PID.2	تصنيف البيانات بشكل مناسب وتقييد الوصول إلى البيانات الحساسة.	الخصوصية والأمن	هل تمت مراجعة البيانات من حيث النطاق والتصنيف؟
PID.3	هل تمت مراجعة البيانات للتحقق مما إذا كانت البيانات الشخصية واضحة ضمن مجموعة البيانات؟ هل هناك آلية راسخة تسمح لنموذج الذكاء الاصطناعي بالتدريب دون استخدام البيانات الشخصية أو الحساسة؟	الخصوصية والأمن	استخدام تقنيات إخفاء الهوية والأسماء المستعارة حيثما أمكن ذلك.

## 11.4.3 . المرحلة الثالثة: البناء والتحقق من الصحة

المرحلة	السؤال	المبدأ	تخفيف حدة المخاطر
BV.1	هل تم اختبار سلوك النظام في مواجهة المواقف والبيئات غير المتوقعة؟ هل هناك خطة احتياطية محددة إذا واجه نموذج الذكاء الاصطناعي هجمات معادية أو مواقف أخرى غير متوقعة؟	الموثوقية والسلامة	إجراء اختبارات شاملة في ظل سيناريوهات متنوعة وإنشاء آليات احتياطية قوية.
BV.2	هل هناك عمليات معينة تحدد الإجراءات اللازمة لوصف التدابير التي يجب اتخاذها عند فشل نظام الذكاء الاصطناعي في سياقات مختلفة؟	الموثوقية والسلامة	تطوير بروتوكولات واضحة واختبارها بانتظام.
BV.3	هل توجد آلية ثابتة للاتصال لضمان موثوقية النظام للمستخدمين النهائيين؟	الشفافية والوضوح	الحفاظ على قنوات اتصال مفتوحة وتقديم تحديثات منتظمة حول أداء النظام.
BV.4	هل هناك تعريفات واضحة ومفهومة تشرح نتائج نظام التزييف العميق؟	الشفافية والوضوح	التأكد من أن عملية التزييف العميق موثوقة ومفهومة للمستخدمين.

## 11.4.3 . المرحلة الرابعة: النشر والمراقبة

المرحلة	السؤال	المبدأ	تخفيف حدة المخاطر
DM.1	هل قام الفريق بتقييم نقاط الضعف في نظام الذكاء الاصطناعي أمام الهجمات المحتملة أو الكشف عن البيانات الحساسة أو انتهاك السرية؟	الخصوصية والأمن	إجراء عمليات تدقيق أمنية منتظمة وتقييم نقاط الضعف.
DM.2	هل توجد آليات لقياس ما إذا كان النظام ينتج قدرًا غير مقبول من النتائج غير الدقيقة؟	المساءلة والمسؤولية	تنفيذ تدابير المراقبة المستمرة ومراقبة الجودة.
DM.3	هل الأشخاص الذين يصلون إلى البيانات مؤهلون بالكفاءات اللازمة لفهم تفاصيل متطلبات حماية البيانات؟	الخصوصية والأمن	التأكد من حصول الموظفين على التدريب المناسب وشهادة الاعتماد الملائمة في مجال حماية البيانات.

## 11.5. الملحق 5: نموذج موافقة على استخدام البيانات الشخصية في صناعة محتوى التزييف العميق

### 11.5.1. الغرض من الموافقة

يوضح نموذج الموافقة هذا كيفية استخدام بياناتك الشخصية، بما في ذلك الصور و/أو الفيديو و/أو الصوت، لإنشاء محتوى وسائط اصطناعي (تزييف عميق)، وسيتم استخدام المحتوى الذي تم إنشاؤه في أغراض محددة، [تحديد الغرض، على سبيل المثال: مقاطع الفيديو التعليمية، وحملات التوعية العامة، والأبحاث وما إلى ذلك]، يُرجى مراجعة هذه الوثيقة بعناية قبل تقديم موافقتك عليها..

### 11.5.2. البيانات التي سيتم استخدامها

[إدراج أنواع بيانات محددة: على سبيل المثال: الصور الفوتوغرافية، ومقاطع الفيديو، والتسجيلات الصوتية، وما إلى ذلك].

### 11.5.3. نطاق ومدة الاستخدام

سيتم استخدام البيانات المقدمة فقط لغرض [غرض محدد].

وسيتم استخدام المحتوى الذي تم إنشاؤه من [تاريخ البدء] إلى [تاريخ الانتهاء] أو بشكل مستمر حتى يتم حذفه.

وسيتم تخزين البيانات بشكل آمن ولا يمكن الوصول إليها إلا من قبل الموظفين المفوضين.

### 11.5.4. الحق في سحب الموافقة

لديك الحق في سحب موافقتك في أي وقت، ولسحب الموافقة يُرجى الاتصال بـ [معلومات الاتصال].

عند الانسحاب لن يتم استخدام بياناتك بعد ذلك في إنشاء محتوى مزيف عميق في المستقبل.

### 11.5.5. المخاطر المحتملة

على الرغم من أنه سيتم اتخاذ جميع التدابير المعقولة لحماية بياناتك، فقد تكون هناك مخاطر كامنة في استخدام البيانات الشخصية للوسائط الاصطناعية، تشمل هذه المخاطر [تحديد المخاطر المحتملة، على سبيل المثال، سوء الاستخدام من قبل أطراف غير مُصرح لهم، واحتمال إساءة تفسير محتوى التزييف العميق، وما إلى ذلك].

### 11.5.6. تأكيد الموافقة

من خلال التوقيع أدناه، فإنك تؤكد أنك قد قرأت وفهمت المعلومات المقدمة في نموذج الموافقة هذا.

أنت توافق على استخدام بياناتك الشخصية على النحو المبين أعلاه.

[مربع اختيار] أوافق على استخدام بياناتي الشخصية لإنشاء محتوى مزيف عميق كما هو موضح.

[مربع اختيار] أقر بأنه تم إبلاغي بحقي في سحب الموافقة في أي وقت.

اسم مالك البيانات: \_\_\_\_\_

توقيع مالك البيانات: \_\_\_\_\_

التاريخ: \_\_\_\_\_

معلومات الاتصال للانسحاب أو الاستفسارات:

- [اسم جهة الاتصال]
- [عنوان البريد الإلكتروني]
- [رقم الهاتف]





**SDAIA**

الهيئة السعودية للبيانات  
والذكاء الاصطناعي  
Saudi Data & AI Authority