ديجيتال فلسطين

الانحياز الخوارزمي في أنظمة الترجمة الآلية العبرية العربية في السياق الفلسطيني

تقرير سياساتي بحثي

مركز ديجيتال فلسطين

@مايو 2026

الملخص التنفيذي

لا تزال الترجمة الآلية للزوج العبري–العربي تعاني من فجوة حادة بين التطور التقني المحرز وبين ضمانات العدالة والحقوق، فعلى الرغم من انتقال هذا الزوج اللغوي من مرحلة ندرة البيانات إلى مرحلة النماذج العصبية المتعددة اللغات، تظل الشفافية محدودة، وتظل المقاييس التقليدية قاصرة عن الكشف عن الأذى الحقيقي في السياق الفلسطيني.

تكشف الأدلة المتوفرة أن الترجمة الآلية في هذا السياق ليست مجرد خدمة لغوية، بل هي بنية تحتية تؤثر في الوصول إلى المعلومات، وتوثيق الانتهاكات، وحماية الهوية الثقافية، وأحياناً في سلامة الأفراد الجسدية.

وقد وثّقت منظمة هيومن رايتس ووتش أكثر من 1,050 حالة إزالة أو تقييد لمحتوى فلسطيني على منصتَي إنستغرام وفيسبوك خلال شهرَي أكتوبر ونوفمبر 2023، وعزت ذلك جزئياً إلى الاعتماد المفرط على الأدوات الآلية في مراجعة المحتوى وترجمته (Human Rights Watch, 2023). وخلص مجلس الرقابة في ميتا إلى أن المعالجة الشاملة لكلمة "شهيد" كانت مفرطة وغير متناسبة (Oversight Board, 2024)، فيما وثّق تقرير BSR أن المحتوى العربي تعرّض لإفراط في التطبيق الآلي مقارنةً بالمحتوى العبري (Business for Social Responsibility [BSR], 2022).

تخلص هذه الورقة إلى ضرورة تحويل معيار تقييم الترجمة الآلية في السياق الفلسطيني من سؤال "هل الترجمة مفهومة؟" إلى سؤال "ما الذي تمحوه؟ من تُسيء تسميته؟ وما أثرها على الحقوق والأمان والكرامة؟."

الفهرس

الملخص التنفيذي	2
الفهرس	4
أولاً: الإطار المفاهيمي وسؤال البحث	6
تعريف الانحياز الخوارزمي في الترجمة	**6**
السياق الفلسطيني كمضاعف للمخاطر	**7**
ثانياً: المشهد التقني — الأنظمة والأداء	7
2.1التطور التاريخي	**7**
2.2الأنظمة المفتوحة المصدر الحالية	**8**
2.3الخدمات التجارية وإشكالية الشفافية	**9**
2.4 الدلالات التقنية للأنظمة على مخاطر الأذى	**10**
ثالثاً: الأذى الموثَّق — الترجمة كبنية تحتية للمخاطر	11
3.1الاعتقال بسبب ترجمة خاطئة (2017)	**11**
3.2تقرير هيومن رايتس ووتش (2023)	**11**
3.3تقرير BSR وانحياز التطبيق	**12**
3.4قضية "شهيد" ومجلس الرقابة في ميتا	**12**
رابعاً: أنماط الأذى الخوارزمي — تصنيف وتحليل	13
4.1تحريف الجنس اللغوي(Misgendering)	**13**
4.2تحويل مصطلحات الهوية إلى تصنيفات أمنية (Securitization)	**14**
4.3محو الكيانات ذات الحمولة السياقية (Erasure)	**14**

4.4سوء الإسناد (Misattribution) 14
4.5مخاطر السلامة(Safety Risk) 15
خامساً: القصور في مقاييس التقييم الحالية 15
سادساً: إطار التدقيق المقترح 15
6.1المقاييس الموسّعة 16
6.2منهجية أخذ العيّنات 16
6.3متطلبات التعليق البشري 17
سابعاً: الإطار القانوني والأخلاقي 17
7.1 القانون الدولي الإنساني 17
7.2 مبادئ الأمم المتحدة التوجيهية بشأن الأعمال وحقوق الإنسان(UNGPs) 18
7.3 قانون الذكاء الاصطناعي الأوروبي(EU AI Act 2024) 19
ثامناً: التوصيات 20
8.1على مستوى الشركات التقنية 20
8.2على مستوى حوكمة المنصات 20
8.3على مستوى البحث الأكاديمي 21
المراجع 22

أولاً: الإطار المفاهيمي وسؤال البحث

تعريف الانحياز الخوارزمي في الترجمة

يُستخدم مصطلح الانحياز الخوارزمي في هذه الورقة بمعنى الانحراف المنهجي والمتكرر الذي لا يوزَّع بالتساوي على الفئات أو السياقات، وينشأ من: اختيار البيانات وتمثيلها، والمعالجة المسبقة، وبناء النموذج، وشروط النشر والاستخدام. في الترجمة الآلية تحديداً، لا يقتصر هذا الانحياز على الصور النمطية الجندرية، بل يشمل أيضاً إفقار اللغة، وتضخيم الأنماط الأكثر شيوعاً على حساب الأقل ظهوراً، وتفاوت الدقة في الكيانات المسماة، والانزلاق من الترجمة إلى إعادة تصنيف دلالي أو أمني (Vanmassenhove et al., 2021).

ولأغراض هذه الورقة، يُميَّز بين ثلاثة مستويات متمايزة من الانحياز لا ينبغي الخلط بينها تحليلياً أو حوكمياً:

المستوى	تعريفه	موضع التدخل	الجهة المسؤولة
انحياز الترجمة	خطأ في تحويل المعنى اللغوي بين العبرية والعربية داخل النموذج نفسه	النموذج اللغوي	شركات تطوير الذكاء الاصطناعي
انحياز مراجعة المحتوى	تصنيف المحتوى خطأً بعد ترجمته، مما يؤدي إلى قرارات إزالة أو تقييد	خط أنابيب التصنيف الآلي	المنصات الرقمية
انحياز السياسات	قواعد مكتوبة أو ضمنية في حوكمة المنصة تُرسّخ معاملة غير متكافئة للمحتوى	قرارات الإدارة والتصميم	إدارة المنصات والمشرّعون

السياق الفلسطيني كمضاعف للمخاطر

في السياق الفلسطيني، تصير المشكلة أوسع من البُعد اللغوي المعتاد؛ إذ تُمثّل الترجمة وسيطاً للوصول إلى الأخبار، ونافذةً على قرارات المنصات، وأداةً لتفسير محتوى حقوقي وقانوني وإنساني، بل وخطوةً داخلية في خط أنابيب المراجعة والتصنيف الآلي. ومن ثَمَّ فإن الخطأ لا يُقاس فقط بضعف المقاييس الإحصائية، بل أيضاً بـ: تحريف الجنس اللغوي (misgendering)، وتحويل اسم مكان إلى اسم جنس عام، ومحو دلالة هوية، أو استبدال مصطلح هوياتي بمصطلح أمني.

وعلى هذا الأساس، **يكون سؤال البحث الرئيس**: كيف تظهر أشكال الانحياز الخوارزمي في أنظمة الترجمة الآلية العبرية–العربية عند تطبيقها على النصوص المتصلة بفلسطين؟ وما حدود المقاييس الشائعة في الكشف عن هذه الانحيازات؟ وما إطار التدقيق الأنسب لتقليل ضررها؟

ثانياً: المشهد التقني – الأنظمة والأداء

2.1 التطور التاريخي

وصفت الأبحاث الأولى الترجمة العربية–العبرية بأنها متأخرة أساساً بسبب غياب المتون الموازية الكافية، مع ميل الأعمال السابقة نحو الاعتماد على قواعد يدوية أو الترجمة الوسيطة عبر الإنجليزية، وكلاهما غير مُرضٍ لبناء نظام قابل للتوسّع

.(Belinkov & Glass, 2016)وفي عام 2016، قدّم Cettolo مدوّنةً عربية–عبرية من محاضرات TED مبنيةً علىWIT3 ، تضم نحو 2,000 محاضرة بإجمالي يقارب 3.5 مليون وحدة نصية لكل لغة؛ وهو تطور تأسيسي نقل الزوج من الندرة الحادة إلى معيار عملي نسبي.(Cettolo, 2016; Cettolo et al., 2012)

2.2 الأنظمة المفتوحة المصدر الحالية

شهدت مرحلة ما بعد 2020 ظهور ثلاثة مسارات تقنية رئيسية:

أ. النماذج الزوجية المباشرة :(OPUS-MT) يُمثّل نموذج OPUS-MT للاتجاه العبري→العربي أبرز الحلول المفتوحة المصدر، مع نتيجة منشورة تبلغ BLEU 23.6 وchrF 0.532 على مجموعة اختبار Tatoeba ؛ وهي نتيجة تدل على إمكانية الترجمة المباشرة تقنياً، لكنها لا تكفي لادعاء الاعتمادية في المجالات عالية المخاطر .(Tiedemann & Thottingal, 2020; Tiedemann et al., 2023)

ب. النماذج متعددة اللغات: (M2M-100) يترجم مباشرةً بين 9,900 اتجاه لغوي من أصل 100 لغة دون المرور بالإنجليزية كلغة وسيطة.(Fan et al., 2021)

ج. عائلات الترجمة الواسعة: (NLLB-200) تشمل 196 لغة وتُقيَّم على-FLORES 200 باستخدام مقاييس BLEU و spBLEUوchrF++، مع تحذيرات أخلاقية من إمكانية بقاء معلومات شخصية في البيانات المستخرجة من الويب (NLLB Team, 2022; Costa-jussà et al., 2024).

د. نموذج الفضاء الكتابي المشترك (2024): طوّر Rom و Bar نموذجاً ثنائي اللغة عربي–عبري يستخدم العربية منقولةً إلى الخط العبري، مستفيداً من القرب الصرفي والبنيوي بين اللغتين وكثرة الكلمات المشتركة. وتُظهر النتائج تفوق هذا النموذج على

النماذج المقارنة التي تحافظ على الخط العربي الأصلي، على الرغم من استخدام بيانات تدريب أقل بنحو 60% (Rom & Bar, 2024) هذا يعني أن جزءاً من مشكلة الزوج ليس "ندرة" خالصة، بل يتعلق أيضاً **بكيفية تمثيل التشابه السامي بنيوياً داخل النموذج.**

2.3 الخدمات التجارية وإشكالية الشفافية

تؤكد الخدمات التجارية الكبرى دعم العربية والعبرية على مستوى الخدمة، لكن الوثائق العامة لا تكشف عادةً بنية الزوج العبري–العربي المحددة، ولا بيانات تدريبه، ولا معدلات الخطأ المتحيّزة:

الخدمة	ما هو موثَّق	ما يبقى مجهولاً
Google Cloud Translation	العربية والعبرية ضمن اللغات المدعومة؛ النماذج المخصصة توثق أزواجاً مع الإنجليزية فقط	معمارية الزوج العبري–العربي وبيانات تدريبه
Azure AI Translator	العربية والعبرية مدعومتان في ترجمة النص؛ العبرية غير مدعومة في ترجمة المستندات الممسوحة ضوئياً	لا توجد مقاييس زوجية مدققة منشورة
Amazon Translate	العربية والعبرية ضمن اللغات المدعومة؛ الترجمة المستندية موجَّهة من/إلى الإنجليزية	لا مقاييس خاصة بالزوج العبري–العربي

2.4 الدلالات التقنية للأنظمة على مخاطر الأذى

تكشف الخصائص التقنية للأنظمة المستعرضة أعلاه عن مواطن هشاشة قابلة للتعيين ومرتبطة مباشرة بأنماط الأذى التي تُوثّقها الأقسام اللاحقة:

تعتمد نماذج OPUS-MT الزوجية المباشرة على التكرار الأعلى في بيانات التدريب؛ وحين تكون هذه البيانات مستقاة أساساً من مصادر إعلامية تُهيمن عليها الأطر الأمنية، يميل النموذج إلى تضخيم المصطلحات الأمنية على حساب مصطلحات الهوية والإنسانية — وهذا يُغذّي مباشرةً نمط الـ Securitization الموثّق في القسم الرابع.

أما M2M-100 و NLLB-200 فتُترجمان مباشرةً دون وسيط إنجليزي، مما يُقلل خطر التشويه المزدوج، غير أن ذلك لا يُعالج الانحياز الكامن في بيانات التدريب الأصلية للزوج العربي-العبري، ولا سيما في الكيانات المسماة وأسماء الأماكن الفلسطينية — وهو ما يُغذّي نمط الـErasure.

وفيما **يخص الخدمات التجارية (Google وAzure وAmazon)**، فإن انعدام الشفافية في بنية الزوج اللغوي يجعل تحديد مصدر الخطأ مستحيلاً: هل هو انحياز في النموذج (المستوى الأول)، أم في خط أنابيب التصنيف (المستوى الثاني)، أم في السياسة الحوكمية (المستوى الثالث)؟ هذا الغموض بحد ذاته نتيجة بحثية يجب تسجيلها.

ثالثاً: الأذى الموثَّق – الترجمة كبنية تحتية للمخاطر

3.1 الاعتقال بسبب ترجمة خاطئة (2017)

في أكتوبر 2017، اعتُقل عامل بناء فلسطيني في الضفة الغربية بعد أن ترجمت خوارزمية فيسبوك الآلية منشورَه الذي كتب فيه بالعربية "صباح الخير" إلى عبارة "هاجمهم" بالعبرية، و "hurt them"بالإنجليزية. وقد اعتُقل قبل أن يقرأ أي ضابط ناطق بالعربية المنشورَ الأصلي .(Hern, 2017) هذه الحادثة ليست استثناءً عابراً، بل نموذج توضيحي بالغ الأهمية يربط الخطأ التقني بالأثر الحقوقي الملموس مباشرةً.

هذه الحادثة تمثّل في جوهرها انحيازاً في المستوى الأول – أي داخل النموذج اللغوي نفسه – غير أن الضرر الفعلي تحقق عبر المستوى الثاني حين استُخدم مخرج الترجمة الخاطئ أساساً لقرار أمني آلي دون مراجعة بشرية.

3.2 تقرير هيومن رايتس ووتش (2023)

وثّق التقرير الصادر في ديسمبر 2023 أكثر من 1,050 حالة إزالة وتقييد وحذف للمحتوى الفلسطيني السلمي على إنستغرام وفيسبوك خلال الفترة الممتدة بين أكتوبر ونوفمبر 2023. وحدّد التقرير "الاعتماد المفرط على الأدوات الآلية في مراجعة المحتوى المتعلق بفلسطين وترجمته" عاملاً نظامياً مركزياً (Human Rights Watch, 2023).

الأذى الموثّق هنا يقع أساساً في المستوى الثاني – انحياز مراجعة المحتوى – إذ إن المشكلة ليست في ترجمة الكلمات بالضرورة، بل في كيفية تفسير الترجمة وتصنيفها ضمن خط الإشراف الآلي.

3.3 تقرير BSR وانحياز التطبيق

خلص تقرير BSR المستقل إلى أن المحتوى العربي تعرّض لإفراط في التطبيق الآلي مقارنةً بالمحتوى العبري على أساس كل مستخدم، ويعود ذلك جزئياً إلى وجود مُصنِّف آلي للكلام العدائي بالعربية دون مقابل له بالعبرية. وأصدر التقرير 21 توصية لمعالجة هذا الخلل.(BSR, 2022)

يُجسّد هذا التقرير المستويين الثاني والثالث معاً: انحياز في خوارزمية التصنيف الآلي (الثاني)، وانحياز في السياسة الحوكمية التي بنت مصنّفاً للكلام العدائي بالعربية دون مقابل عبري (الثالث).

3.4 قضية "شهيد" ومجلس الرقابة في ميتا

في مارس 2024، أصدر مجلس الرقابة في ميتا رأياً استشارياً خلص فيه إلى أن السياسة الشاملة تجاه كلمة "شهيد" كانت "تُقيّد حرية التعبير على نحو جوهري وغير متناسب"، وأوصى بإنهاء الحظر الشامل للكلمة. وأشارت ميتا إلى أن هذه الكلمة مسؤولة عن عمليات إزالة محتوى تفوق أي كلمة مفردة أخرى على منصاتها (Oversight Board, 2024).

هذه القضية نموذج نقي للمستوى الثالث – انحياز السياسات – إذ إن المشكلة لم تكن في ترجمة الكلمة ولا في خوارزمية التصنيف، بل في القرار الحوكمي الذي عامل الكلمة معاملة شاملة بصرف النظر عن السياق.

هذه الأمثلة من الحوادث تُثبت أن الانحياز في الترجمة العبرية–العربية ضمن السياق الفلسطيني يُنتج أضراراً تمسّ: الدقة، والكرامة، وإمكانية الوصول إلى المعلومة، وسلامة الأفراد، وقابلية توثيق الانتهاكات.

رابعاً: أنماط الأذى الخوارزمي – تصنيف وتحليل

تتوزع أشكال الأذى الخوارزمي في هذا الزوج اللغوي على خمس فئات رئيسة:

4.1 تحريف الجنس اللغوي(Misgendering)

العربية لغة تُظهر الجندر صرفياً بوضوح في أفعالها وصفاتها وضمائرها، غير أن العبرية أيضاً لغة تصريفية بالجندر، وقد أثبتت الأبحاث أن أنظمة الترجمة الآلية تميل نحو التذكير الافتراضي حين تغيب القرائن السياقية الكافية، مما يؤدي إلى محو الهوية الجندرية للأفراد وتشويه التمثيل المهني والاجتماعي (Stanovsky et al., 2019; Alhafni et al., 2022).

مثال توضيحي:

-العبرية الأصلية: הפרמדיקית הפלסטינית אמרה שהיא עייפה (المسعفة الفلسطينية قالت إنها متعبة)

-مخرج آلي منحاز: "قال المسعف الفلسطيني إنه متعب" – محو مزدوج للجندر والهوية

4.2 تحويل مصطلحات الهوية إلى تصنيفات أمنية (Securitization)

أخطر أشكال الانحياز هو الانتقال الخاطئ من وسم هوياتي أو ديني أو اجتماعي إلى وسم أمني، كما يكشف مثال "شهيد" → "إرهابي"، هذا التحويل لا يعكس خطأ لغوياً بسيطاً، بل يُعيد تصنيف الشخص ويُعرّضه للمساءلة أو الحذف أو سوء الفهم في سياقات أمنية حساسة.

4.3 محو الكيانات ذات الحمولة السياقية (Erasure)

يشمل ذلك: حذف كلمة "مخيم" واستبدالها بـ"منطقة"، أو تحريف أسماء الأماكن الفلسطينية، أو إسقاط الفاعل من الجملة. ومن الأمثلة:

-تحريف اسم المكان: "דג'נין" → "الجنين" بدلاً من "جنين" — تحويل اسم علم مكاني إلى اسم نكرة يفسد المعنى الميداني

-محو السياق الإنساني: "תושבי המחנה" → "سكان المنطقة" بدلاً من "سكان المخيم" — حذف حمولة إنسانية وقانونية مهمة

4.4 سوء الإسناد (Misattribution)

حين تنقل الترجمة الفاعلية على نحو مضلل، أو تعكس من قال ماذا ولمن، تصبح بنية الحدث بأكملها ملتبسة في الأخبار والوثائق القانونية.

4.5 مخاطر السلامة(Safety Risk)

عندما يُستخدم مخرج الترجمة الخاطئ أساساً لقرارات إزالة المحتوى أو الاعتقال أو سوء التقدير الأمني، يتحول الخطأ اللغوي إلى خطر على سلامة الأشخاص — وهذا ما تُجسّده حادثة 2017 بوضوح تام.

خامساً: القصور في مقاييس التقييم الحالية

تعتمد أبحاث الترجمة الآلية تقليدياً على مقاييس مثل BLEU و chrF وMETEOR، وهي مقاييس تقيس الجودة اللغوية العامة، ولكنها تخفق في التقاط أشكال الأذى المذكورة أعلاه لأسباب جوهرية:

أولاً: تكافئ هذه المقاييس الأنظمة التي تُسوّي اللهجات أو الخصوصيات في "متوسط فصيح"، بينما تخسر الدقة الثقافية والمحلية .(Joshi et al., 2025) في السياق الفلسطيني، اللهجة والاسم المحلي والمصطلح السياسي جزء من المعنى نفسه.

ثانياً: لا تقيس هذه المقاييس دقة الكيانات المسماة ولا صواب مصطلحات الهوية ولا نسبة تحريف الجنس اللغوي.

ثالثاً: لا تعكس العلاقة بين خطأ الترجمة وقرارات الإزالة والتصنيف الآلي downstream.

سادساً: إطار التدقيق المقترح

6.1 المقاييس الموسّعة

تقترح هذه الدراسة تبني منظومة مقاييس متكاملة تجمع بين الكمّي والنوعي:

المقياس	الوظيفة
BLEU / chrF / METEOR	خط أساس للجودة اللغوية العامة
دقة الكيانات المسماة	نسبة الكيانات التي نُقلت دون تحوير دلالي أو جغرافي
دقة الجندر اللغوي	نسبة الجمل التي حُفظ فيها جنس الفاعل/المشار إليه دون تحريف
معدل الإيجابيات الكاذبة للهوية	نسبة الحالات التي يُحوَّل فيها مصطلح هوياتي إلى تسمية أمنية خطأ
معدل السلبيات الكاذبة للهوية	نسبة الحالات التي يُطمَس فيها مصطلح تحريضي بما يُخفّف المعنى
نسبة المحو	نسبة حذف الكيانات أو الفاعلين أو مؤشرات الزمان والمكان
درجة شدة الأذى	تقييم بشري من 0–3 لشدة الضرر العملي في الأخبار/القانون/الإغاثة

6.2 منهجية أخذ العيّنات

يُوصى بأسلوب أخذ العيّنات الطبقي الحساس للأذى عبر خمسة مجالات: الأخبار، وسائل التواصل الاجتماعي، القانون والإدارة، الإغاثة الإنسانية، وخط الأساس العام. ويجب إعطاء أولوية للكلمات عالية المخاطر مثل: "شهيد/שאהיד"، "أسير"، "مخيم"، "حاجز"، وأسماء المدن والمخيمات الفلسطينية؛ لأن الخطأ فيها ليس لغوياً فحسب، بل تبعاته معيارية وحقوقية.

6.3 متطلبات التعليق البشري

يستلزم التدقيق الفعّال ثلاثة مُعلِّقين ثنائي اللغة عربي–عبري على الأقل مع مُحكِّم رابع، على أن يكونوا ملمّين بالسياقين اللغوي والسياسي الفلسطيني، فقد أثبتت الأبحاث أن الأخطاء في نصوص الصراع يصعب وسمها دون معرفة سياقية عميقة تتجاوز الكفاءة اللغوية.(Alhafni et al., 2022)

سابعاً: الإطار القانوني والأخلاقي

7.1 القانون الدولي الإنساني

يقوم القانون الدولي الإنساني على مبدأين جوهريين ذوَي صلة مباشرة بموضوع هذه الورقة: مبدأ التمييز بين المدنيين والمقاتلين، ومبدأ الاحتياط في اتخاذ القرارات ذات الأثر على حياة الأشخاص.

حين تُستخدم الترجمة الآلية في تنبيهات الإخلاء أو التعليمات الإنسانية أو تفسير شهادات المدنيين، فإن خطأها يمس مباشرةً مبدأ الاحتياط: إذ يُفترض بالجهات العسكرية والأمنية اتخاذ كل الاحتياطات الممكنة قبل اتخاذ قرار يؤثر في حياة شخص أو حريته. والاعتماد على ترجمة آلية غير مدقَّقة في سياق أمني — كما حدث عام 2017 — يُشكّل إخلالاً بهذا المبدأ لأنه يحل الأتمتة محل الاحتياط الواجب.

علاوةً على ذلك، فإن تحويل مصطلح هوياتي كـ"شهيد" إلى تصنيف أمني عبر الترجمة أو مراجعة المحتوى يمس مبدأ التمييز: إذ يُعيد تصنيف أشخاص بناءً على انتمائهم اللغوي والثقافي لا على أفعالهم الفعلية، وهو ما يتعارض مع صون الحماية المدنية المكفولة بالقانون الدولي.

7.2 مبادئ الأمم المتحدة التوجيهية بشأن الأعمال وحقوق الإنسان(UNGPs)

تُرسي المبادئ التوجيهية للأمم المتحدة ثلاث ركائز، وتنطبق كل منها على حالة الترجمة الآلية في السياق الفلسطيني على النحو الآتي:

الركيزة الأولى — واجب الدولة في الحماية: تقع على الدول التي تعمل فيها شركات الترجمة الكبرى مسؤولية إلزامها بتقييم أثر منتجاتها على حقوق الإنسان قبل نشرها في بيئات النزاع. غياب هذا الإلزام حتى الآن يعني أن شركات مثل Google و Meta تنشر أنظمة ترجمة ومراجعة آلية في سياق فلسطيني بالغ الحساسية دون أي رقابة خارجية مستقلة.

الركيزة الثانية — مسؤولية الشركات عن الاحترام: تلتزم الشركات بموجب هذه المبادئ بإجراء تقييم مستمر للأثر الحقوقي لمنتجاتها (Human Rights Due

.(Diligenceوعدم الإفصاح عن معدلات الخطأ الخاصة بالزوج العبري-العربي يُشكّل إخلالاً صريحاً بهذا الالتزام، لأنه يحول دون أي مساءلة خارجية أو مراجعة مستقلة.

الركيزة الثالثة – سبل الانتصاف: المستخدم الذي اعتُقل بسبب ترجمة خاطئة، أو الذي أُزيل محتواه بسبب تصنيف آلي مبني على ترجمة منحازة، لا تتوفر له حالياً أي آلية تظلم فعّالة تُلزم الشركة بالاعتراف بالخطأ أو التعويض عنه. هذا الفراغ يتعارض مع الركيزة الثالثة تعارضاً صريحاً.

7.3 قانون الذكاء الاصطناعي الأوروبي(EU AI Act 2024)

يُشكّل هذا القانون، الذي دخل حيز التنفيذ عام 2024، إطاراً تنظيمياً ذا صلة مباشرة وإن كان نطاقه الجغرافي محدوداً. يُصنّف القانون أنظمة الذكاء الاصطناعي المستخدمة في سياقات إنفاذ القانون ومراقبة الأشخاص وتصنيف المحتوى ضمن "التطبيقات عالية الخطورة"، مما يفرض عليها متطلبات صارمة تشمل: الشفافية في بيانات التدريب، ومعدلات الخطأ المُصنَّفة، وآليات الطعن البشري. وعلى الرغم من أن معظم الحوادث الموثّقة في هذه الورقة تقع خارج الاتحاد الأوروبي، فإن هذا القانون يُرسي معياراً دولياً يمكن الاحتجاج به في ضغط مماثل على مستوى الحوكمة العالمية لشركات التكنولوجيا.

ثامناً: التوصيات

8.1 على مستوى الشركات التقنية

1. نشر بطاقات نموذج (Model Cards) خاصة بالزوج العبري–العربي تتضمن: توزيع بيانات التدريب، ومعدلات الخطأ المصنّفة حسب نوع المحتوى، ومقاييس الانحياز الجندري وانحياز مصطلحات الهوية.

2. إجراء اختبارات عدائية (Adversarial Testing) قبل النشر لا بعده، تشمل: قوائم مصطلحات الهوية والأمن، وأسماء الأماكن الفلسطينية، والصيغ المُلتبسة صرفياً.

3. تطوير ترجمة واعية بالكيانات (Entity-Aware Translation) مع قواميس مقيّدة للأسماء الحساسة جغرافياً وسياسياً، ومعالجة خاصة للنقل الصوتي واللهجة الفلسطينية.

8.2 على مستوى حوكمة المنصات

4. فرض إشراف بشري إلزامي في حالات الاستخدام عالية المخاطر: النصوص القانونية، والإشعارات الإدارية، والتنبيهات الإنسانية، وقرارات

الحذف والمراجعة المرتبطة بفلسطين. فقد وجدت هيومن رايتس ووتش أن مئات المستخدمين لم يتمكنوا من الطعن في القيود المفروضة عليهم (Human Rights Watch, 2023).

5. إنشاء آليات استئناف فعّالة وسريعة بما يضمن حق المستخدمين في مراجعة القرارات الآلية المؤثرة في محتواهم.

8.3 على مستوى البحث الأكاديمي

6. بناء مدوّنة تدقيق فلسطينية مخصصة تشمل الأخبار والتنبيهات الإنسانية والنصوص القانونية والمنشورات العامة، مع مراجع بشرية من مترجمين يمتلكون معرفة سياقية عميقة، وتوازن بيانات عبر التوليد المضاد للحالات.(Counterfactual Augmentation)

7. توسيع مقاييس التقييم لتشمل دقة الكيانات المسماة ومؤشرات الانحياز الهوياتي والجندري ضمن كل تقييم مقارن لهذا الزوج اللغوي.

المراجع

1. Alhafni, B., Habash, N., & Bouamor, H. (2022). The Arabic Parallel Gender Corpus 2.0: Extensions and analyses. In Proceedings of the Thirteenth Language Resources and Evaluation Conference (pp. 1870–1884). European Language Resources Association. https://aclanthology.org/2022.lrec-1.199/

2. Amazon Web Services. (n.d.). What is Amazon Translate? Amazon Translate Developer Guide. Retrieved May 29, 2026, from https://docs.aws.amazon.com/translate/latest/dg/what-is.html

3. Belinkov, Y., & Glass, J. (2016). Large-scale machine translation between Arabic and Hebrew: Available corpora and initial results. arXiv. https://arxiv.org/abs/1609.07701

4. Business for Social Responsibility. (2022). Human rights due diligence of Meta's impacts in Israel and Palestine in May 2021. BSR. https://www.bsr.org/en/reports/meta-human-rights-israel-palestine

5. Cettolo, M. (2016). An Arabic-Hebrew parallel corpus of TED talks. arXiv. https://arxiv.org/abs/1610.00572

6. Cettolo, M., Girardi, C., & Federico, M. (2012). WIT3: Web inventory of transcribed and translated talks. In Proceedings of the 16th Annual Conference of the European Association for Machine Translation (pp. 261–268). European Association for Machine Translation. https://aclanthology.org/2012.eamt-1.60/

7. Costa-jussà, M. R., et al. (NLLB Team). (2024). Scaling neural machine translation to 200 languages. Nature, 630, 841–846. https://doi.org/10.1038/s41586-024-07335-x

8. Fan, A., Bhosale, S., Schwenk, H., Ma, Z., El-Kishky, A., Goyal, S., Baines, M., Celebi, O., Wenzek, G., Chaudhary, V., Goyal, N., Artetxe, M., Shrivastava, M., Chen, A., Maillard, J., & Joulin, A. (2021). Beyond English-centric multilingual machine translation. Journal of Machine Learning Research, 22(107), 1–48. https://jmlr.org/papers/v22/20-1307.html

9. Google Cloud. (n.d.). Cloud Translation documentation. Retrieved May 29, 2026, from https://cloud.google.com/translate/docs

10. Goyal, N., Gao, C., Chaudhary, V., Chen, P.-J., Wenzek, G., Ju, D., Krishnan, S., Ranzato, M., Guzmán, F., & Fan, A. (2022). The FLORES-101 evaluation benchmark for low-resource and multilingual machine translation. Transactions of the Association for Computational Linguistics, 10, 522–538. https://doi.org/10.1162/tacl_a_00474

11. Hern, A. (2017, October 24). Facebook translates 'good morning' into 'attack them', leading to arrest. The Guardian. https://www.theguardian.com/technology/2017/oct/24/facebook-palestine-israel-translates-good-morning-attack-them-arrest

12. Human Rights Watch. (2023). Meta's broken promises: Systemic censorship of Palestine content on Instagram and Facebook. Human Rights Watch. https://www.hrw.org/report/2023/12/21/metas-broken-promises/systemic-censorship-palestine-content-instagram-and

13. Joshi, A., Dabre, R., Kanojia, D., Li, Z., Zhan, H., Haffari, G., & Dippold, D. (2025). Natural language processing for dialects of a language: A survey. ACM Computing Surveys, 57(6), 1–37. https://doi.org/10.1145/3712060

14. Microsoft. (n.d.). What is Azure AI Translator? Microsoft Learn. Retrieved May 29, 2026, from https://learn.microsoft.com/en-us/azure/ai-services/translator/

15. NLLB Team, Costa-jussà, M. R., Cross, J., Çelebi, O., Elbayad, M., Heafield, K., Heffernan, K., Kalbassi, E., Lam, J., Licht, D., Maillard, J., Sun, A., Wang, S., Wenzek, G., Youngblood, A., Alastruey, B., Gao, C., Chen, P.-J., Okhonko, D., … Wang, J. (2022). No language left behind: Scaling human-centered machine translation. arXiv. https://arxiv.org/abs/2207.04672

16. Oversight Board. (2024). Referring to designated dangerous individuals as "Shaheed" [Policy advisory opinion]. Meta Oversight Board. https://www.oversightboard.com/news/oversight-board-publishes-policy-advisory-opinion-on-referring-to-designated-dangerous-individuals-as-shaheed/

17. Rom, A., & Bar, K. (2024). Training a bilingual language model by mapping tokens onto a shared character space. arXiv. https://arxiv.org/abs/2402.16065

18. Stanovsky, G., Smith, N. A., & Zettlemoyer, L. (2019). Evaluating gender bias in machine translation. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics (pp. 1679–1684). Association for Computational Linguistics. https://aclanthology.org/P19-1164/

19. Tiedemann, J., & Thottingal, S. (2020). OPUS-MT — Building open translation services for the World. In Proceedings of the 22nd Annual

Conference of the European Association for Machine Translation (pp. 479–480). European Association for Machine Translation. https://aclanthology.org/2020.eamt-1.61/

20. Tiedemann, J., Aulamo, M., Bakshandaeva, D., Boggia, M., Grönroos, S.-A., Nieminen, T., Raganato, A., Scherrer, Y., Vázquez, R., & Virpioja, S. (2023). Democratizing neural machine translation with OPUS-MT. Language Resources and Evaluation, 58, 713–755. https://doi.org/10.1007/s10579-023-09704-w

21. United Nations Office of the High Commissioner for Human Rights. (2011). Guiding principles on business and human rights: Implementing the United Nations "Protect, Respect and Remedy" framework (HR/PUB/11/4). United Nations. https://www.ohchr.org/en/publications/reference-publications/guiding-principles-business-and-human-rights

22. Vanmassenhove, E., Shterionov, D., & Gwilliam, M. (2021). Machine translationese: Effects of algorithmic bias on linguistic complexity in machine translation. arXiv. https://arxiv.org/abs/2102.00287