الرقم الدولي المعياري للمؤتمر (معتمد ومصنف دوليًا) ISBN 978 - 9953 - 0 - 2970 - 2

# المؤتمر الدولي الحادي عشر للغة العربية

22 - 24 أكتوبر 2025م الموافق 30 ربيع الآخر - 2 جمادى الأولى 1447هـ

دبي - الإمارات العربية المتحدة

## الهيئات العربية والدولية أعضاء المجلس الدولي للغة العربية

**د. محمد أحمد الشهراني**

# مستقبل الأمن اللغوي العربي في عصر الذكاء الاصطناعي: مقاربة لسانية اجتماعية نقدية للهيمنة الرقمية، وتحديات الهوية اللغوية

**الملخص:**

يشكل التقدم المتسارع للذكاء الاصطناعي (AI) لحظة محورية للأمن اللغوي العربي، مما يعيد تشكيل المشهد اللساني الاجتماعي بعمق ومستقبل اللغة العربية في عالم يزداد رقمنة. تتبنى هذه الورقة مقاربة لسانية اجتماعية نقدية لبحث التداعيات متعددة الأوجه للذكاء الاصطناعي على الأمن اللغوي العربي، لمعالجة فجوة ملحوظة في الأدبيات الحالية التي غالبًا ما تتجاهل الديناميكيات البنيوية العميقة وعلاقات القوة المؤثرة. تستقصي الدراسة بشكل خاص التحديات المتصاعدة للهيمنة الرقمية، وتبحث في كيفية قيام تطوير ونشر الذكاء الاصطناعي بتعزيز الاختلالات اللغوية العالمية، وتكريس التحيزات الخوارزمية، والمساهمة في أشكال جديدة من الاستعمار الرقمي الذي يهدد السيادة المعلوماتية والثقافية العربية. بالتزامن مع ذلك، تتعمق الدراسة في التحديات الهوياتية العميقة التي يطرحها الذكاء الاصطناعي، بما في ذلك تأثيره على المعايير اللغوية، والتعبير الثقافي، وتشكيل الهويات العربية الفردية والجماعية في الفضاء الرقمي. وبينما تقيّم هذه التهديدات بشكل نقدي، تستكشف الدراسة أيضًا الفرص الكامنة التي يوفرها الذكاء الاصطناعي لتعزيز الحيوية الرقمية للغة العربية. من خلال تحليل مقارن لأطر السياسات والمبادرات الاستراتيجية في دولة المملكة العربية السعودية والإمارات العربية المتحدة العربية المتحدة (كممثل رئيس لمنطقة الخليج العربي)، والتوجهات الأوروبية البارزة، ومنطقة شمال إفريقيا، تقيّم هذه الورقة نماذج متنوعة لحوكمة التقاطع بين الذكاء الاصطناعي واللغة، بهدف استخلاص رؤى عملية ذات صلة بمختلف أنحاء العالم العربي. في نهاية المطاف، تجادل الدراسة بضرورة التطوير العاجل لاستراتيجيات سياسات عربية شاملة واستباقية يمكنها التعامل مع تعقيدات عصر الذكاء الاصطناعي، وتخفيف المخاطر، والاستفادة استراتيجياً من الذكاء الاصطناعي لحماية وتعزيز الأمن اللغوي العربي في المستقبل.

**الكلمات المفتاحية:**

الأمن اللغوي العربي؛ الذكاء الاصطناعي (AI)؛ اللسانيات الاجتماعية؛ الهيمنة الرقمية؛ السياسة اللغوية؛ الهوية الثقافية

**المقدمة:**

يشهد العصر الراهن تحولاً تكنولوجياً جذرياً يتمحور حول الذكاء الاصطناعي (AI)، الذي تتجاوز تداعياته إعادة هيكلة النظم الاقتصادية والصناعية لتمس جوهر الممارسات اللغوية والتفاعلات الاجتماعية (Brynjolfsson & McAfee, 2014; Schwab, 2017). في هذا السياق، تبرز اللغة الإنسانية، باعتبارها الأداة المحورية للفكر والتواصل وبناء الهوية الاجتماعية، ونتاج تطور معرفي واجتماعي معقد، كميدان استراتيجي يتأثر ويتفاعل بعمق مع التطورات المتسارعة للذكاء الاصطناعي. وتكتسب اللغة العربية، بما تحمله من إرث حضاري ودور مركزي في تشكيل الهوية الفردية والجماعية لملايين الناطقين بها (المسدي، 2014؛ العزير، 2023)، أهمية خاصة في ظل التقدم المذهل في معالجة اللغات الطبيعية (NLP) وتطور نماذج اللغة الكبيرة (LLMs) القادرة على فهم وتوليد النصوص بمستويات قد تحاكي الطلاقة البشرية (Brown et al., 2020).

من منظور اللسانيات الاجتماعية وعلم اللغة الجنائي، يتجاوز مفهوم "الأمن اللغوي" مجرد الحفاظ على البنية الصرفية والنحوية للغة، ليشمل حماية وظائفها الاجتماعية، وحيويتها في سياقات الاستخدام المتنوعة، وضمان سلامة وموثوقية الخطاب اللغوي في الفضاء الرقمي. إن قدرة المجتمعات الناطقة بالعربية على استخدام لغتها بفعالية وثقة، والتعبير عن هوياتها المتعددة، والمشاركة في إنتاج المعرفة، والحفاظ على التماسك الاجتماعي، وحماية الفضاء التواصلي من التزييف والتضليل، تمثل أبعاداً جوهرية للأمن اللغوي (التركي، 2001؛ زاهد، 2000). في هذا الإطار، تطرح تقنيات الذكاء الاصطناعي تحديات

وفرصاً غير مسبوقة؛ فبينما قد تسهم في تطوير أدوات تعليمية أو ترجمية، فإنها تثير قضايا معقدة تتعلق بتغير المعايير اللغوية الاجتماعية، وتأثيرات التحيز الخوارزمي على تمثيل الجماعات اللغوية، وإمكانية استخدامها في إنتاج محتوى مضلل أو منسوب زوراً، وهي قضايا تقع في صميم اهتمامات اللسانيات الاجتماعية التطبيقية وعلم اللغة الجنائي (Coulthard & Johnson, 2007; Olsson, 2008).

**الفجوات البحثية وأهمية الدراسة:**

على الرغم من الاهتمام الأكاديمي المتنامي بتقاطع الذكاء الاصطناعي واللغة العربية (Habash, 2010; Zaghouani, 2021)، تظل هناك فجوات بحثية واضحة تستدعي المعالجة:

1. **ندرة التحليلات اللسانية الاجتماعية النقدية:** يميل الخطاب السائد إلى التركيز على الجوانب التقنية أو الفرص الاقتصادية أو المخاطر الأخلاقية العامة (القايدي، 2023؛ عطية وآخرون، 2019)، مع إغفال لافت للتحليل اللساني الاجتماعي المعمق الذي يتناول كيفية تأثير الذكاء الاصطناعي على أنماط الاستخدام اللغوي الفعلي، والتنوع اللهجي، والمعايير الاجتماعية للغة، وديناميكيات القوة والهيمنة اللغوية في الفضاء الرقمي العربي (Couldry & Mejias, 2019).
2. **غياب منظور علم اللغة الجنائي:** تكاد تنعدم الدراسات التي تتناول، من منظور علم اللغة الجنائي، التحديات الناشئة عن قدرة الذكاء الاصطناعي على توليد نصوص عربية يصعب تمييزها عن النصوص البشرية. ويشمل ذلك قضايا تحديد المؤلف الحقيقي (Authorship Attribution)، وانتشار المعلومات المضللة (Disinformation)، وإمكانية استخدام النصوص المولدة آليًا في أنشطة احتيالية أو ضارة، وتأثير ذلك على الثقة في المحتوى الرقمي العربي.

3. **قصور في ربط التحديات بالسياسات اللغوية:** غالبًا ما تفشل الدراسات القائمة في تجاوز المستوى الوصفي أو التقني لتقديم رؤى عملية لصياغة سياسات لغوية وثقافية عربية استباقية وفعالة، قادرة على التعامل مع الأبعاد الاجتماعية والقانونية لهذه التقنيات، وضمان "أمن لغوي" شامل (المسدي، 2014).

من رحم هذه الفجوات، تستمد هذه الدراسة أهميتها البالغة وضرورتها الملحة. ففهم التداعيات اللسانية الاجتماعية والجنائية للذكاء الاصطناعي على اللغة العربية لم يعد ترفاً فكرياً، بل ضرورة استراتيجية لمعالجة قضايا حرجة تتعلق بالعدالة اللغوية الرقمية، والحفاظ على التنوع الثقافي واللغوي، ومكافحة التضليل، وتعزيز الأمن المجتمعي. وتهدف هذه الدراسة إلى تقديم إطار تحليلي مبتكر يدمج منظورات اللسانيات الاجتماعية وعلم اللغة الجنائي لفهم هذه التحديات، وتقييم الفرص المتاحة، واستخلاص دروس من تجارب دولية، واقتراح سياسات عملية لتعزيز الأمن اللغوي العربي في عصر الذكاء الاصطناعي.

**أسئلة البحث:**

تسعى هذه الدراسة للإجابة على الأسئلة المحورية التالية:

1. ما أبرز التهديدات اللسانية الاجتماعية والجنائية التي يفرضها الذكاء الاصطناعي على الأمن اللغوي العربي، وكيف تتجلى هذه التهديدات في أنماط الاستخدام اللغوي، وموثوقية الخطاب الرقمي، وديناميكيات الهوية؟
2. ما الفرص العملية التي يمكن لتقنيات الذكاء الاصطناعي أن توفرها لدعم حيوية اللغة العربية، وتعزيز وظائفها الاجتماعية، وتطوير أدوات لسانية جنائية تخدم السياق العربي؟

3. كيف تعاملت تجارب دولية مختارة (مثل المملكة العربية السعودية والإمارات العربية المتحدة العربية المتحدة، والاتحاد الأوروبي، ومنطقة شمال إفريقيا) مع تحديات حوكمة التقاطع بين الذكاء الاصطناعي واللغة الوطنية من منظور لساني اجتماعي، وما الدروس المستفادة التي يمكن تكييفها للسياق العربي؟
4. ما إطار العمل المقترح لسياسات وطنية عربية استباقية ومتكاملة، تكون قادرة على توجيه التفاعل مع الذكاء الاصطناعي بما يخدم تعزيز الأمن اللغوي العربي بأبعاده الاجتماعية والجنائية والثقافية؟

**الإطار النظري ومراجعة الأدبيات**

يتطلب بناء فهم لساني اجتماعي وجنائي نقدي لتأثير الذكاء الاصطناعي (AI) على الأمن اللغوي العربي تأسيسًا معرفيًا راسخًا يتجاوز الانطباعات العامة أو التحليلات التقنية المنعزلة أو التفاؤل التكنولوجي غير المدروس. فلا يكفي رصد التحديات الهندسية أو تعداد الفرص المحتملة، بل لا بد من الغوص في الأبعاد اللسانية الاجتماعية والجنائية لهذه الظاهرة المعقدة والمتسارعة، وفهم كيف تتشابك اللغة والتكنولوجيا والهوية والسلطة والعدالة في تشكيل مستقبل الخطاب والممارسات اللغوية في المجتمعات العربية. يضطلع هذا القسم بمهمة حاسمة في تحقيق هذا التأسيس من خلال غايتين متكاملتين: أولاً، مراجعة تحليلية نقدية للأدبيات الراهنة في الحقول المعرفية المتقاطعة (اللسانيات الحاسوبية العربية، اللسانيات الاجتماعية الرقمية، علم اللغة الجنائي الرقمي، السياسات اللغوية، الدراسات النقدية للذكاء الاصطناعي، دراسات الهوية والأمن اللغوي العربي)، بهدف تحديد خطوط الصدع المعرفي، وكشف مناطق الإغفال المنهجي والنظري، وتحديد الفجوة المعرفية التي تسعى هذه الدراسة لمعالجتها. وثانياً، بناء وتقديم الإطار النظري اللساني

الاجتماعي والجنائي التكاملي الذي سيشكل العدسة التحليلية للدراسة، مع تبرير اختيار النظريات وشرح مفاهيمها المحورية وكيفية توظيفها في تحليل الديناميكيات المعقدة.

**مراجعة نقدية للأدبيات: تحديد ملامح الحقل البحثي وفجواته من منظور لساني اجتماعي وجنائي**

تتكشف إشكالية البحث عند تقاطع تيارات بحثية متعددة، تتميز بالديناميكية ولكن أيضاً بالتشتت، مما يستدعي قراءة نقدية موجهة بخلفيتنا اللسانية الاجتماعية والجنائية.

**في محور الذكاء الاصطناعي ومعالجة اللغة العربية (NLP)**

شهد هذا الحقل تطورات تقنية لافتة (Habash, 2010; Zaghouani, 2021; El-Haj et al., 2022)، بما في ذلك مبادرات كنماذج "جيس" و"فالكون" (Alhafni et al., 2023; Antoun et al., 2022). ورغم أهمية هذه الجهود، يظل النقد الجوهري من منظورنا هو تركيزها الغالب على الأداء التقني مع إهمال نسبي للأبعاد اللسانية الاجتماعية العميقة. فالأسئلة المتعلقة بتأثير هذه النماذج على المعايير اللغوية الاجتماعية، وتفاقم الصراع بين الفصحى والعاميات، والتمثيل اللهجي، وديناميكيات القوة اللغوية (من يمتلك التكنولوجيا ويوجهها؟)، وكيفية تفاعل هذه التقنيات مع السياقات الاجتماعية المتنوعة وإعادة تشكيلها، لا تزال هامشية. هذا يترك فجوة في فهم تأثير الذكاء الاصطناعي على الأمن اللغوي كبناء لساني اجتماعي يتشكل بفعل عوامل تتجاوز الدقة الخوارزمية (Brown et al., 2020). علاوة على ذلك، فإن البعد الجنائي المتعلق بإمكانية استخدام هذه النماذج في توليد نصوص عربية يصعب تمييزها عن تلك البشرية، وما يترتب على ذلك من تحديات في تحديد المؤلف (Authorship Attribution) ومكافحة التضليل، يكاد يكون غائباً تماماً عن هذا الخطاب التقني.

**في الدراسات النقدية للذكاء الاصطناعي**

يتناول تيار بحثي نقدي متنامٍ، خاصة في السياقات الغربية، التأثيرات الاجتماعية والأخلاقية للذكاء الاصطناعي، كاشفاً عن آليات التحيز والتمييز (Noble, 2018; Benjamin, 2019)، وديناميكيات استعمار البيانات (Couldry & Mejias, 2019). ورغم أهميتها، تبرز فجوة عند تطبيقها على السياق العربي، وتحديداً من منظور لساني اجتماعي وجنائي. فالدراسات التي تحلل كيف يتجلى التحيز الخوارزمي في المحتوى العربي المولّد آلياً وتأثيره على تمثيل الهويات اللغوية، أو كيف يؤثر استخلاص البيانات اللغوية العربية على السيادة الرقمية (Ali, 2016)، لا تزال محدودة. والأهم من ذلك، أن التحليل النقدي لتأثير الذكاء الاصطناعي على "سلامة وموثوقية الخطاب الرقمي العربي" من منظور جنائي لغوي، مثل دراسة انتشار الخطاب المضلل المولّد آلياً أو انتحال الهوية اللغوية، لا يزال مجالاً بكراً. وبينما نجد أعمالاً عربية قيمة تتناول قضايا الهوية والأمن اللغوي بشكل عام (المسدي، 2014؛ العزير، 2023)، فإن التحليل اللساني الاجتماعي والجنائي النقدي لتأثير الذكاء الاصطناعي تحديداً يظل نادراً.

**في حقل السياسات والتخطيط اللغوي (LPP) في العصر الرقمي**

شهد هذا الحقل محاولات للتكيف مع التحولات الرقمية (Spolsky, 2004, 2012; Pimienta, 2009; Gazzola et al., 2021). لكن النقد الموجه لهذا الحقل، خاصة في السياق العربي، يتمثل في بطء استجابته للتحديات النوعية التي يطرحها الذكاء الاصطناعي (Habash, 2010). فالسياسات اللغوية التقليدية قد تبدو قاصرة أمام التأثير البنيوي لنماذج اللغة الكبيرة على الممارسات اللغوية، أو أمام التحديات القانونية والأخلاقية الناشئة عن المحتوى المولّد آلياً. علاوة على ذلك، تفتقر النقاشات حول السياسات اللغوية الرقمية العربية غالباً إلى تحليل لساني اجتماعي نقدي لآليات التنفيذ وتحديات التطبيق، أو

إلى دمج منظور علم اللغة الجنائي في صياغة تشريعات تتعلق بمصداقية المحتوى الرقمي (القايدي، 2023).

**في مجال اللسانيات الاجتماعية الرقمية والهوية**

توجد أدبيات متنامية تدرس استخدامات الإنترنت والذكاء الاصطناعي وتأثيرها على الهوية والتفاعل اللغوي (Wheeler, 2012; Al-Saggaf, 2016). ورغم أهميتها، فإن دمج دور الذكاء الاصطناعي كفاعل مؤثر في هذه الديناميكيات لا يزال محدوداً. فالأسئلة حول كيف تعيد الخوارزميات تشكيل تصورات الهوية اللغوية الرقمية العربية، وكيف يتفاعل المستخدمون مع المحتوى المولّد آلياً وتأثير ذلك على ممارساتهم التواصلية، وما إذا كانت هذه التقنيات تعزز أشكالاً معينة من "الهابيتوس اللغوي الرقمي" على حساب أخرى، لا تزال تتطلب استكشافاً أعمق من منظور لساني اجتماعي نقدي.

إن هذه المراجعة النقدية، الموجهة بخلفيتنا اللسانية الاجتماعية والجنائية، تكشف عن حاجة ملحة لبحث يربط بين الأبعاد التقنية واللسانية الاجتماعية والجنائية والسياساتية؛ ويتجاوز الوصف إلى تحليل علاقات القوة والتحيزات والتأثيرات على سلامة الخطاب؛ ويركز على خصوصيات الواقع اللغوي العربي؛ ويسعى لربط الفهم النظري باقتراح حلول عملية.

**الإطار النظري اللساني الاجتماعي والجنائي التكاملي**

إن فهم الطبيعة المتعددة الأوجه لتأثير الذكاء الاصطناعي على الأمن اللغوي العربي، بأبعاده اللسانية الاجتماعية والجنائية، يتطلب تبني إطار نظري تكاملي. ينطلق هذا البحث من قناعة بأن التكنولوجيا ليست محايدة، بل هي مُشكِّلة ومُتشكِّلة اجتماعياً (Winner, 1980; Jasanoff, 2016)، وأن اللغة ليست مجرد أداة، بل هي ممارسة اجتماعية مشبعة بالمعنى ومرتبطة بالهوية والسلطة وسلامة الخطاب

(Fairclough, 2013; Bourdieu, 1991)، وهي علاقات يمكن للتكنولوجيا الخوارزمية أن تعيد تشكيلها وتعكس تحيزاتها (Noble, 2018) أو تستغلها لأغراض تضليلية. وعليه، ينسج هذا الإطار خيوطًا من منظورات لسانية اجتماعية وجنائية متعددة:

**اللسانيات الاجتماعية النقدية وتحليل الخطاب النقدي (CDA)**

يوفر هذا المنظور أدوات لتحليل كيف يعكس الخطاب حول الذكاء الاصطناعي وتطبيقاته اللغوية، وكيف يساهم في بناء أو تكريس علاقات القوة والهيمنة اللغوية (Fairclough, 2013; Van Dijk, 2001). سنستخدمه لفحص كيف تُؤطَّر "مشكلة" الأمن اللغوي، ومن هم الفاعلون الشرعيون، وما هي الحلول المقترحة (أو المستبعدة) في الخطابات الرسمية والتقنية والإعلامية. كما سيساعد في تحليل كيف يمكن للخوارزميات أن تفرض "معايير لغوية" مهيمنة بشكل ضمني، وتهمش أشكالاً لغوية أخرى (مثل اللهجات)، مما يعكس ديناميكيات الهيمنة والاستعمار الرقمي اللغوي (Phillipson, 2003; Couldry & Mejias, 2019).

**نظريات الهوية اللغوية والتنوع اللغوي في العصر الرقمي**

لفهم التأثير على الهوية الفردية والجماعية، ندمج رؤى من نظريات الهوية اللغوية (Giles & Johnson, 1987; Le Page & Tabouret-Keller, 1985) وتأثير التكنولوجيا عليها (Jones & Hafner, 2012).

سنحلل كيف يؤثر الذكاء الاصطناعي على وظيفة اللغة كعلامة للانتماء الجماعي، وكيف يتفاوض الأفراد العرب حول هوياتهم اللغوية المتعددة (فصحى، عاميات، لغات أخرى) في تفاعلهم مع هذه التقنيات.

كما سنهتم بمسألة الحفاظ على التنوع اللغوي العربي (بما في ذلك اللهجات) في مواجهة ضغوط التجانس التي قد تفرضها نماذج اللغة الكبيرة (UNESCO, 2005).

**مبادئ علم اللغة الجنائي وتحليل المؤلف (Forensic Linguistics & Authorship Analysis)**

يوفر هذا الحقل الأدوات النظرية والمنهجية لتحليل النصوص من منظور قانوني وجنائي (Coulthard & Johnson, 2007; Olsson, 2008). سنستلهم من هذا الحقل لتحليل التحديات المتعلقة بموثوقية النصوص العربية المولدة بواسطة الذكاء الاصطناعي. ويشمل ذلك:

- **تحديد المؤلف (Authorship Attribution/Stylometry):** كيف يمكن التمييز بين النصوص التي كتبها بشر وتلك التي ولدتها الآلة؟ وما هي التحديات التي تواجه تقنيات تحديد المؤلف في سياق النصوص العربية المولدة آلياً؟ (Koppel, Schler, & Argamon, 2009).
- **تحليل الخطاب المضلل (Deceptive Language Analysis):** ما هي السمات اللغوية التي قد تميز الخطاب العربي المضلل المولّد آلياً؟ وكيف يمكن تطوير أدوات لكشفه؟ (Newman et al., 2003).
- **الأبعاد القانونية والأخلاقية للمحتوى المولّد آلياً:** ما هي الآثار المترتبة على انتشار محتوى عربي مولّد آلياً قد يكون مسيئاً، أو منتحلاً، أو محمياً بحقوق ملكية فكرية؟

**نظريات التخطيط والسياسة اللغوية (Language Policy and Planning - LPP):**

يوفر هذا المجال إطاراً لفهم وتقييم الجهود المبذولة للتأثير على وضع اللغة واستخدامها وتطويرها (Spolsky, 2004; Cooper, 1989). سنستخدم هذا الإطار لتحليل السياسات اللغوية الرقمية العربية القائمة أو المقترحة، وتقييم مدى استجابتها للتحديات اللسانية الاجتماعية والجنائية التي يطرحها الذكاء الاصطناعي، مع التركيز على أبعاد تخطيط المكانة (Status Planning)، وتخطيط المتن (Corpus Planning)، وتخطيط الاكتساب (Acquisition Planning) في البيئة الرقمية.

إن التكامل بين هذه المنظورات اللسانية الاجتماعية والجنائية، مع التأصيل المستمر في السياق العربي، هو ما يمنح هذا البحث إطاره المميز. فهو يسمح بربط التحليل على مستويات متعددة (النص، الخطاب، الممارسة الاجتماعية، السياسة)، ويساعد في توجيه عملية التحليل وتفسير النتائج، والانتقال من الفهم النظري إلى اقتراح حلول وسياسات تستجيب للواقع المعقد للأمن اللغوي العربي في عصر الذكاء الاصطناعي.

**تعريف المفاهيم الإجرائية**

لضمان الوضوح والدقة في التحليل، وتأسيس أرضية مشتركة للفهم ضمن الإطار اللساني الاجتماعي والجنائي المعتمد، تُعرَّف المفاهيم المحورية التالية إجرائياً في سياق هذه الدراسة تحديداً، مع الاستنارة بالنقاشات الأكاديمية ذات الصلة:

**الأمن اللغوي العربي (Arab Linguistic Security):** يُقصد به في هذا البحث الحالة التي تمتلك فيها اللغة العربية، بتنوعاتها، والمجتمعات الناطقة بها، القدرة الشاملة والمستدامة على الحفاظ على حيوية وظائفها الاجتماعية، وسلامة بنيتها واستخداماتها، وموثوقية الخطاب المنتج بها في البيئة الرقمية المتغيرة. ويشمل ذلك القدرة على استخدامها كوسيط فعال للتعبير عن الهوية الفردية والجماعية، والمشاركة

المواطنة، ودعم التنوع الثقافي والمعرفي، وحماية الفضاء التواصلي من التضليل والتحيز الخوارزمي، وذلك في مواجهة التحديات اللسانية الاجتماعية والجنائية التي يطرحها عصر الذكاء الاصطناعي. يتقاطع هذا الفهم مع مفاهيم الحيوية اللغوية (Language Vitality)، والعدالة اللغوية الرقمية (Pimienta, 2009)، وسلامة الخطاب (Discourse Integrity).

**الذكاء الاصطناعي اللغوي (Linguistic Artificial Intelligence):** يُشار به هنا إلى مجموعة التقنيات والنماذج الخوارزمية، وبشكل خاص تقنيات معالجة اللغات الطبيعية (NLP) ونماذج اللغة الكبيرة (LLMs)، المصممة لمعالجة اللغة العربية بفصحاها ولهجاتها (فهماً وتحليلاً وتوليداً). تركز الدراسة ليس فقط على القدرات التقنية لهذه الأدوات، بل بشكل حاسم على تفاعلها كنظم "سوسيوتقنية" (Sociotechnical Systems) مع المستخدمين والممارسات اللغوية الاجتماعية، وتأثيرها على المعايير اللغوية، وتمثيل الهويات، وإمكانية استخدامها في إنتاج نصوص يصعب تحديد مصدرها أو قد تكون مضللة، مع إدراك أنها ليست أدوات محايدة بل تحمل إمكانات للتحيز والتأثير على ديناميكيات القوة اللغوية (Bender et al., 2021; Crawford, 2021).

**السياسات اللغوية الرقمية (العربية) (Arab Digital Language Policies):** تُعرّف بأنها مجمل الإجراءات والتدابير والاستراتيجيات، الرسمية وغير الرسمية، التي تهدف إلى التأثير بشكل مقصود على وضع اللغة العربية (Status Planning)، وبنيتها وتطوير مواردها الرقمية (Corpus Planning)، وأنماط استخدامها وحيويتها الاجتماعية (Acquisition Planning) في البيئة الرقمية. يشمل ذلك بشكل خاص السياسات المتعلقة بدعم المحتوى العربي الرقمي، وتطوير تقنيات اللغة العربية، وتنظيم تفاعلها مع تقنيات الذكاء الاصطناعي بما يستجيب للتحديات اللسانية الاجتماعية والجنائية (مثل

التحيز، والتضليل، وحماية التنوع اللهجي)، وضمان العدالة اللغوية (building on Spolsky, 2004; Gazzola et al., 2021).

**الهوية اللغوية الرقمية العربية (Arab Digital Linguistic Identity):** تُفهم على أنها الكيفية التي يعبر بها الأفراد الناطقون بالعربية عن انتمائهم اللغوي والثقافي ويتفاوضون حوله من خلال ممارساتهم اللغوية عبر المنصات الرقمية، بما في ذلك اختياراتهم بين الفصحى والعاميات والمزج اللغوي، وأساليبهم التعبيرية. تُعتبر هذه الهوية ديناميكية ومتأثرة بالبنى التحتية الرقمية والخوارزميات، والمحتوى المولّد آلياً، مما يؤثر على كيفية بناء "الذات اللغوية" وتقديمها، وقدرتها على مقاومة التنميط أو التهميش في الفضاء الرقمي (cf. Jones & Hafner, 2012; Deumert, 2014).

**الاستعمار الرقمي اللغوي (Linguistic Digital Colonialism):** يُستخدم هذا المصطلح لوصف الآليات البنيوية والاقتصادية والثقافية في النظام الرقمي العالمي التي تؤدي إلى هيمنة لغات ونماذج تكنولوجية (غالباً مرتبطة بمراكز القوة الغربية) على الفضاء الرقمي العربي. يتجلى ذلك في استخلاص البيانات اللغوية العربية، وفرض معايير تكنولوجية تعكس تحيزات ثقافية ولغوية (Algorithmic Bias)، وتهميش المحتوى واللهجات المحلية، مما يقوض السيادة اللغوية والمعرفية ويهدد التنوع اللغوي العربي، ويعمق أشكال التبعية الرقمية (Couldry & Mejias, 2019; Phillipson, 2003).

**تحليل المؤلف الآلي (Automated Authorship Analysis):** يشير إلى تطبيق تقنيات الذكاء الاصطناعي واللسانيات الحاسوبية لتحديد أو التحقق من هوية مؤلف نص ما بناءً على سماته الأسلوبية واللغوية. في سياق هذه الدراسة، يُعنى هذا المفهوم بالتحديات والفرص المتعلقة بتمييز النصوص

العربية التي كتبها بشر عن تلك التي ولدتها نماذج اللغة الكبيرة، وتداعيات ذلك على قضايا الانتحال، والمسؤولية عن المحتوى، وموثوقية المعلومات (Koppel, Schler, & Argamon, 2009).

**الخطاب الرقمي المضلل (Digital Misinformation/Disinformation):** يُقصد به المعلومات الخاطئة أو الكاذبة التي يتم نشرها عبر المنصات الرقمية، سواء عن قصد (Disinformation) أو عن غير قصد (Misinformation). في إطار هذه الدراسة، يتم التركيز على دور الذكاء الاصطناعي اللغوي في توليد ونشر مثل هذا الخطاب باللغة العربية، وتأثيره على الرأي العام، والثقة المجتمعية، والأمن اللغوي بمفهومه الشامل (Wardle & Derakhshan, 2017).

**منهجية البحث:**

إن الإجابة العلمية الرصينة على أسئلة البحث المعقدة التي تطرحها هذه الدراسة حول تفاعل الذكاء الاصطناعي والأمن اللغوي العربي من منظور لساني اجتماعي وجنائي، واستخلاص استنتاجات ذات مصداقية وقيمة نظرية وتطبيقية، يتوقف بشكل حاسم على مدى دقة وملاءمة المنهجية المتبعة. يهدف هذا القسم إلى تقديم عرض تفصيلي ومنهجي للخيارات الإجرائية التي تم اعتمادها، وتبرير هذه الخيارات في ضوء طبيعة الإشكالية البحثية (كما حُددت في المقدمة ومراجعة الأدبيات)، وأهداف الدراسة وأسئلتها، والإطار النظري اللساني الاجتماعي والجنائي التكاملي الذي تم تبنيه. سيتم توضيح تصميم البحث العام، واستراتيجيات جمع البيانات وتحليلها، ومبررات اختيار الحالات المقارنة، مع مناقشة الاعتبارات الأخلاقية ومعايير ضمان الجودة والمحددات الكامنة.

**تصميم البحث والمنهج المتبع: نحو فهم نوعي نقدي من منظور لساني اجتماعي وجنائي**

تعتمد هذه الدراسة بشكل أساسي على **تصميم بحثي نوعي (Qualitative Research Design)**، وهو الخيار الأمثل لاستكشاف وفهم الظواهر اللسانية الاجتماعية المعقدة بعمق، مثل التفاعلات الدقيقة بين الذكاء الاصطناعي، والممارسات اللغوية، والهوية، وديناميكيات القوة، وسلامة الخطاب في السياق العربي (Creswell & Poth, 2018; Silverman, 2020). فالهدف هنا ليس قياس العلاقات السببية إحصائياً، بل الغوص في كيفية تشكّل تأثير الذكاء الاصطناعي على الأمن اللغوي العربي من خلال تحليل الخطابات المتداولة، وتفسير المعاني التي ينسبها الفاعلون لتجاربهم، وفهم السياقات المؤسسية والثقافية التي تؤطر هذه التفاعلات، وتقييم التحديات الناشئة من منظور جنائي لغوي. يتيح المنهج النوعي، بأدواته المرنة، التقاط هذه التعقيدات وتفسيرها بشكل ثري وسياقي.

ويندرج هذا التصميم النوعي ضمن **منظور تفسيري-نقدي (Interpretive-Critical Perspective)** موجه باللسانيات الاجتماعية وعلم اللغة الجنائي. فهو **تفسيري** بمعنى أنه يسعى لفهم وتأويل المعاني الاجتماعية والتفاعلات الرمزية التي يبنيها الأفراد والجماعات حول الذكاء الاصطناعي وتأثيره على لغتهم وهويتهم وممارساتهم الخطابية (Schwandt, 2000). وهو **نقدي** لأنه لا يكتفي بالوصف أو التفسير، بل يتجاوزهما إلى تحليل علاقات القوة غير المتكافئة، والتحيزات الخوارزمية والاجتماعية، والأيديولوجيات المهيمنة (مثل الاستعمار الرقمي اللغوي)، وتحديات موثوقية الخطاب (مثل التضليل المولّد آلياً) التي تشكل الظاهرة وتؤثر في توزيع منافعها ومخاطرها، مسترشداً في ذلك بالأطر النظرية اللسانية الاجتماعية والجنائية التي تم تبنيها (Fairclough, 2013; Coulthard & Johnson, 2007; Couldry & Mejias, 2019). يهدف هذا المنظور، في نهاية المطاف، إلى المساهمة في رفع الوعي وإمكانية التغيير نحو سياسات لغوية رقمية أكثر عدالة وأمناً.

**استراتيجيات جمع البيانات: الاعتماد المنهجي على المصادر الثانوية وتحليلها**

نظراً لطبيعة البحث التي تركز على تحليل الاتجاهات الكبرى والسياسات الوطنية والمقارنات الدولية المتعلقة بتقاطع الذكاء الاصطناعي واللغة العربية من منظور لساني اجتماعي وجنائي، وبسبب القيود العملية أحياناً، تعتمد الدراسة بشكل منهجي على جمع وتحليل **البيانات الثانوية (Secondary Data)** المتاحة من مصادر متعددة. يتيح هذا النهج، عند استخدامه بشكل نقدي ومنهجي، بناء فهم شامل للظاهرة، وتتبع تطور الخطابات والسياسات، وهو مناسب بشكل خاص للإجابة على أسئلة البحث المطروحة (Bryman, 2016). تشمل استراتيجيات جمع البيانات المعتمدة ما يلي:

**المراجعة النقدية الممنهجة للأدبيات:** تُعدّ هذه المراجعة، التي بُدء بها في القسم السابق وسيتم تعميقها، أداة رئيسية ليس فقط لتحديد الفجوة المعرفية ولكن أيضاً كـ"بيانات" في حد ذاتها. سيتم التركيز على الدراسات الأكاديمية المحكمة (عربية وأجنبية) في مجالات اللسانيات الاجتماعية الرقمية، وعلم اللغة الجنائي الرقمي، والسياسات اللغوية، ومعالجة اللغة العربية، والدراسات النقدية للذكاء الاصطناعي.

**تحليل الوثائق متعددة المصادر (Document Analysis):** يمثل هذا المصدر الأساسي للبيانات حول السياسات والخطابات والممارسات المتعلقة بتقاطع الذكاء الاصطناعي والأمن اللغوي العربي. تشمل أنواع الوثائق:

**الوثائق الرسمية والسياساتية:** استراتيجيات الذكاء الاصطناعي الوطنية (في المملكة العربية السعودية والإمارات العربية المتحدة، والاتحاد الأوروبي، ودول شمال إفريقيا المعنية)، السياسات اللغوية والثقافية، التشريعات المتعلقة بالبيانات ومكافحة التضليل.

**وثائق المنظمات الدولية والإقليمية:** تقارير اليونسكو، الإيسيسكو، الألكسو، وغيرها، المتعلقة بأخلاقيات الذكاء الاصطناعي، والتنوع اللغوي، ومحو الأمية الرقمية.

**وثائق وتقارير القطاع الخاص والصناعة:** تقارير الشركات المطورة لنماذج اللغة الكبيرة العربية، لفهم توجهاتها وخطاباتها (مع قراءة نقدية).

**الخطاب العام والإعلامي:** مقالات تحليلية ونقاشات حول تأثير الذكاء الاصطناعي على اللغة العربية، وحالات التضليل، وتحديات تحديد المؤلف.

**منهج دراسة الحالة المقارنة (Comparative Case Study):** لإضفاء عمق على التحليل، تتبنى الدراسة منهج دراسة الحالة المقارنة لسياسات ومقاربات: **المملكة العربية السعودية والإمارات العربية المتحدة العربية المتحدة** (كممثل للخليج العربي)، **التوجهات الأوروبية** (كمجموعة دول متقدمة في التنظيم)، و**منطقة شمال إفريقيا** (لتمثيل تنوع إضافي في السياق العربي). يبرر هذا الاختيار التباين في النماذج والخبرات، مما يتيح فهم مقاربات مختلفة لحوكمة تقاطع الذكاء الاصطناعي واللغة الوطنية من منظور لساني اجتماعي وجنائي، واستخلاص دروس مستفادة (Yin, 2018; Bartlett & Vavrus, 2017). ستركز المقارنة على محاور مثل: الأهداف السياساتية المتعلقة باللغة، آليات التعامل مع التحيز اللغوي، استراتيجيات مكافحة التضليل المولّد آلياً، ودعم التنوع اللغوي.

**استراتيجيات تحليل البيانات النوعية**

سيتم تحليل البيانات النوعية المجمعة من خلال عملية منهجية وتكرارية، وتوظيف استراتيجيات تحليل متعددة ومتكاملة لضمان ثراء وعمق النتائج، بما يخدم الإجابة على أسئلة البحث:

**التحليل الموضوعي الموجه نظرياً (Theory-Driven Thematic Analysis):** هذه هي الاستراتيجية المحورية لتحديد الأنماط والموضوعات الرئيسية المتعلقة بالتهديدات والتحديات اللسانية الاجتماعية والجنائية، والفرص، والسياسات. سيتم توجيه عملية الترميز (Coding) وتوليد الموضوعات بالإطار النظري اللساني الاجتماعي والجنائي للدراسة (مفاهيم الهيمنة الرقمية اللغوية، الهوية اللغوية الرقمية، تحليل المؤلف، الخطاب المضلل، السياسات اللغوية) (Braun & Clarke, 2006, 2019).

**التحليل النقدي للخطاب (Critical Discourse Analysis - CDA):** سيتم توظيف أدوات التحليل النقدي للخطاب بشكل موجه لتحليل وثائق السياسات والخطابات الرسمية والإعلامية. سيركز التحليل على كيفية تعريف وتأطير "مشكلة" الأمن اللغوي العربي في عصر الذكاء الاصطناعي، وكيف يعكس الخطاب علاقات القوة، وما هي الافتراضات الأيديولوجية المضمرة المتعلقة باللغة والتكنولوجيا (Fairclough, 2013; Van Dijk, 2001).

**التحليل المقارن المنهجي عبر الحالات (Systematic Cross-Case Analysis):** لتحقيق أهداف الدراسة المقارنة، سيتم تصميم مصفوفة تحليل بناءً على المحاور المحددة للمقارنة (الأهداف، الآليات، التحديات اللسانية الاجتماعية والجنائية المعالجة). سيتم بعد ذلك إجراء مقارنة منهجية لتحديد الأنماط وأوجه الاختلاف، والعلاقات المحتملة بين السياق والسياسات المتبعة (Miles, Huberman, & Saldaña, 2014). الهدف هو استخلاص دروس مستفادة قابلة للتطبيق بحذر في السياق العربي.

سيتم التأكيد طوال عملية التحليل على الحوار التفاعلي بين البيانات والنظرية (Abductive Reasoning). فالإطار النظري اللساني الاجتماعي والجنائي سيوفر العدسات لتوجيه التحليل، بينما ستساهم النتائج التجريبية في اختبار وتخصيص هذه النظريات.

**النتائج:**

يُقدم هذا القسم عرضاً تحليلياً للنتائج المستخلصة من تطبيق المنهجية الموصوفة، والتي استهدفت فهم التفاعل المعقد بين الذكاء الاصطناعي والأمن اللغوي العربي من منظور لساني اجتماعي وجنائي. يركز هذا العرض على إبراز الأنماط الرئيسية المستمدة من مراجعة الأدبيات، وتحليل الوثائق، والمقارنة بين تجارب المملكة العربية السعودية والإمارات العربية المتحدة العربية المتحدة، والتوجهات الأوروبية، ومنطقة شمال إفريقيا، وذلك للإجابة بشكل مباشر على أسئلة البحث المطروحة.

**الأمن اللغوي العربي في مرمى الذكاء الاصطناعي: تشريح لساني اجتماعي وجنائي للتهديدات والفرص**

يكشف التحليل المعمق للأدبيات المتاحة والوثائق ذات الصلة والخطابات السائدة عن أن اللغة العربية، بأمنها اللغوي متعدد الأبعاد، تقف اليوم عند منعطف حرج يتشكل بفعل التغلغل المتسارع لتقنيات الذكاء الاصطناعي. هذا التفاعل، بعيدًا عن كونه مجرد تحديث تقني، يمثل ظاهرة "سوسيوتقنية" (Sociotechnical) معقدة تعيد تشكيل علاقات القوة اللغوية، والممارسات الخطابية، والمعايير الاجتماعية للغة، والهويات اللغوية، وتطرح تحديات جنائية لغوية جديدة. ويتضح أن فهم هذه الظاهرة يتطلب تفكيكًا نقديًا للتهديدات الكامنة والفرص الممكنة من منظور لساني اجتماعي وجنائي يربط بين البنى الكلية والممارسات الجزئية، ويأخذ في الاعتبار التحديات الداخلية والخارجية.

أحد أبرز التهديدات اللسانية الاجتماعية يتعلق بتعميق **الهيمنة الرقمية اللغوية وتكريس التبعية المعرفية**، وهو ما يجد تفسيره في إطار منظور الاستعمار الرقمي اللغوي (Couldry & Mejias, 2019; Phillipson, 2003). تشير الدلائل إلى هيمنة اللغات ذات الموارد الرقمية الضخمة، وخاصة

الإنجليزية، في تطوير وتدريب نماذج الذكاء الاصطناعي (Zaghouani, 2021). هذا "الفقر التمثيلي" للغة العربية لا يؤدي فقط إلى قصور تقني (Alhafni et al., 2023)، بل الأخطر أنه يؤدي إلى **تحيزات خوارزمية** معرفية وثقافية مضمنة، قد تنعكس في توليد صور نمطية أو تقديم سرديات منحازة (Noble, 2018; Benjamin, 2019). علاوة على ذلك، فإن عملية "استخلاص" البيانات اللغوية العربية دون شفافية أو حوكمة تشاركية تمثل تحديًا للسيادة اللغوية الرقمية (Ali, 2016).

على المستوى الجنائي اللغوي، تبرز تهديدات جدية تتعلق **بموثوقية الخطاب الرقمي العربي وسلامته**. فقدرة نماذج اللغة الكبيرة على توليد نصوص عربية "متقنة ظاهريًا" (Brown et al., 2020) تفتح الباب أمام:

1. **صعوبة تحديد المؤلف (Authorship Attribution):** مما يعقد مسائل المسؤولية والانتحال.
2. **انتشار الخطاب المضلل (Disinformation/Misinformation):** حيث يمكن استخدام هذه التقنيات لإنتاج أخبار زائفة أو دعاية مغرضة باللغة العربية على نطاق واسع، مما يقوض الثقة ويؤثر على الرأي العام (Wardle & Derakhshan, 2017).
3. **انتحال الهوية اللغوية (Linguistic Impersonation):** إمكانية استخدام النصوص المولدة آليًا لانتحال شخصيات أو إنشاء حسابات وهمية ذات مصداقية لغوية ظاهرية لأغراض خبيثة.

ويتشابك هذا مع تحولات لسانية اجتماعية عميقة تتعلق بإعادة **تشكيل المعايير والممارسات اللغوية اليومية**. فالانتشار المتزايد لأدوات الذكاء الاصطناعي التي تتوسط تفاعلاتنا اللغوية يؤثر على كيفية استخدام اللغة، وقد يؤدي إلى "تطبيع" لغة آلية مبسطة أو تراكيب مستوردة، مما يثير قلقًا حول الحفاظ على

الثراء البلاغي والتنوع الأسلوبي للعربية. كما أن تهميش اللهجات العربية في معظم النماذج الحالية يعزز من الشعور بدونيتها الرقمية ويمثل شكلًا من "العنف الرمزي" اللغوي (Bourdieu, 1991).

يترتب على هذه التحولات تأثيرات بالغة الأهمية على **بنية الهوية اللغوية والثقافية العربية** (المسدي، 2014). فالتعرض المستمر لمحتوى مولّد أو مصفى بواسطة خوارزميات قد تحمل تحيزات أو تعكس رؤى عالمية مهيمنة يضع الهوية في حالة "تفاوض" وتوتر، ويثير تساؤلات حول "الأصالة" اللغوية والثقافية (Jones & Hafner, 2012).

ولكن، وسط هذه التهديدات، يكشف التحليل أيضًا عن **فرص كامنة** يمكن استثمارها من منظور لساني اجتماعي وجنائي:

- **في مجال اللسانيات الاجتماعية التطبيقية:** في مجال اللسانيات الاجتماعية التطبيقية، يفتح الذكاء الاصطناعي آفاقًا واعدة يمكن استثمارها بشكل استراتيجي لخدمة اللغة العربية وتعزيز حيويتها الرقمية. إذ يمكن توظيف تقنياته المتقدمة في تطوير أدوات تعليمية تفاعلية ومبتكرة للغة العربية، تستجيب للاحتياجات المتنوعة للمتعلمين وتسهم في تجسير الفجوة بين الفصحى واللهجات، وهو ما تتطرق إليه دراسات حديثة حول دور التكنولوجيا في تعزيز الكفاءة اللغوية (انظر El Zahraa, 2025; AlAfnan, 2025). علاوة على ذلك، يساهم الذكاء الاصطناعي في تيسير الوصول إلى المعرفة العالمية عبر تطوير أنظمة ترجمة آلية محسنة من وإلى العربية، مع الأخذ في الاعتبار التحديات المتعلقة بالدقة والملاءمة الثقافية، خاصة عند التعامل مع مصطلحات متخصصة أو محتوى ذي حمولة أيديولوجية (Al-Jarf, 2025). كما أن دعم رقمنة التراث اللغوي العربي الهائل، بما في ذلك المخطوطات النادرة والنصوص التاريخية، يصبح أكثر فعالية وكفاءة باستخدام

أدوات الذكاء الاصطناعي في التعرف الضوئي على الحروف وفهرسة المحتوى وتحليله. ومن الجوانب ذات الأهمية البالغة، قدرة الذكاء الاصطناعي على تطوير أدوات متخصصة لفهم ومعالجة اللهجات العربية المختلفة، وهو مجال يشهد اهتمامًا متزايدًا كما يتضح في جهود بناء مجموعات بيانات لهجية وتطوير نماذج قادرة على التعامل مع هذا التنوع الثري (Shang et al., 2024; Faisal et al., 2024; Tibi & Messaoud, 2025). إن مثل هذه الأدوات لا تقتصر فائدتها على الجانب التقني، بل تمتد لتعزيز التنوع اللغوي كقيمة ثقافية واجتماعية، ودعم الحيوية اللغوية للهجات في الفضاء الرقمي، وتمكين الناطقين بها من التعبير عن هوياتهم المتعددة، والمساهمة في بناء مجتمعات معرفية أكثر شمولاً وإنصافًا.

- **في مجال علم اللغة الجنائي:**

أما في مجال علم اللغة الجنائي، فإن توجيه أبحاث الذكاء الاصطناعي بشكل استراتيجي يفتح الباب أمام تطوير أدوات تحليلية متقدمة وحاسمة لمواجهة التحديات المتزايدة لسلامة الخطاب الرقمي العربي. فمع تزايد قدرة نماذج اللغة الكبيرة على توليد نصوص يصعب تمييزها عن تلك التي ينتجها البشر، تصبح الحاجة ملحة لتطوير تقنيات قادرة على كشف النصوص المولدة آليًا، وهو ما تسعى إليه مبادرات عالمية بدأت تمتد لتشمل اللغة العربية ( PAN, 2024 ) حول التحقق من مؤلفي النصوص المولدة بالذكاء الاصطناعي). كما أن تطبيقات الذكاء الاصطناعي يمكن أن تقدم مساعدة قيمة في مهمة تحديد المؤلف (Authorship Attribution)، وهي قضية محورية في سياق المسؤولية القانونية والأخلاقية عن المحتوى المنشور. بالإضافة إلى ذلك، يبرز دور الذكاء الاصطناعي في رصد وتتبع الخطاب المضلل والمعلومات المغلوطة باللغة العربية، والتي تشكل

تهديدًا متناميًا للاستقرار الاجتماعي والثقة العامة، لا سيما في ظل تقنيات المقاومة الخوارزمية التي يلجأ إليها ناشرو هذا المحتوى (Hosny & Nasef, 2025) . إن بناء نماذج قادرة على تحليل الأنماط اللغوية الدالة على الخداع أو التضليل، وتحديد مصادر الحملات الممنهجة، يمكن أن يسهم بشكل كبير في تعزيز سلامة الفضاء الرقمي العربي وحمايته من التلاعب والاستغلال.

**سياسات الذكاء الاصطناعي واللغة الوطنية: تحليل مقارن لتجارب المملكة العربية السعودية والإمارات العربية المتحدة والتوجهات الأوروبية ومنطقة شمال إفريقيا**

إن فهم التحديات والفرص المعقدة التي تواجه الأمن اللغوي العربي في عصر الذكاء الاصطناعي لا يمكن أن يكتمل بمعزل عن السياق العالمي المتسارع وديناميكيات القوة والهيمنة اللغوية الرقمية التي تشكله. فاللغة العربية، بتاريخها وثقلها الحضاري، تجد نفسها اليوم في تفاعل مباشر مع تقنيات غالبًا ما تُطور وتُدرب ضمن سياقات لغوية وثقافية مغايرة، مما يثير تساؤلات جوهرية حول التمثيل والتحيز (Mihalcea et al., 2025; Demidova et al., 2024). وفي هذا الإطار، يكتسب التحليل المقارن للاستجابات السياساتية المتباينة أهمية قصوى، ليس بهدف الاستنساخ الحرفي للتجارب، بل لاستخلاص دروس مستفادة وفهم البدائل المتاحة والعوامل الحاسمة التي تؤثر في فعالية السياسات المعنية بتقاطع الذكاء الاصطناعي واللغة الوطنية. إن دراسة نماذج متنوعة، مثل تلك المتبعة في المملكة العربية السعودية والإمارات العربية المتحدة كممثلين لمنطقة الخليج الطموحة، والتوجهات الأوروبية الرائدة في التنظيم الأخلاقي والسعي للتنافسية، بالإضافة إلى منطقة شمال إفريقيا التي تعكس فسيفساء من المقاربات والتحديات ضمن السياق العربي الأوسع، توفر رؤى بالغة الأهمية لتوجيه المساعي العربية نحو حماية وتعزيز لغتها

في الفضاء الرقمي (انظر Trigui et al., 2024 حول استكشاف حوكمة الذكاء الاصطناعي في منطقة الشرق الأوسط وشمال إفريقيا).

يُظهر التحليل المقارن، كما يلخصه جدول 1 مدعومًا بنتائج الدراسات الحديثة، تباينًا واضحًا في مستويات النضج الاستراتيجي والموارد المخصصة. ففي منطقة الخليج، تتجلى الريادة في إطلاق استراتيجيات وطنية طموحة للذكاء الاصطناعي، مصحوبة بمبادرات ملموسة لتطوير نماذج لغوية عربية متقدمة مثل "جيس" و"فالكون"، وإنشاء منصات توليدية عربية متعددة الوسائط كـ "Fanar" (Fanar Team et al., 2025)، وتوفير مجموعات بيانات لغوية ضخمة (Aloui et al., 2024; Alshahrani, 2024)، مما يعكس إدراكًا متزايدًا لأهمية اللغة كجزء من الهوية الوطنية الرقمية وأداة للقوة الناعمة (AlAfnan, 2025). وفي المقابل، تركز التوجهات الأوروبية بشكل كبير على بناء أطر تنظيمية وتشريعية قوية، مثل قانون الذكاء الاصطناعي الأوروبي (انظر Stelling et al., 2025)، بهدف ضمان تطوير "ذكاء اصطناعي موثوق وأخلاقي" وحماية التنوع اللغوي والثقافي في مواجهة هيمنة النماذج العالمية، مع دعم الابتكار وبناء نماذج لغوية سيادية (Bondarenko et al., 2025). أما منطقة شمال إفريقيا، فتشهد جهودًا واعدة ولكنها غالبًا ما تكون متفرقة ومحدودة الموارد، مثل تطوير نماذج لهجية كـ "Atlas-Chat" للهجة المغربية (Shang et al., 2024)، مع استمرار التحديات المتعلقة بالفجوة الرقمية واللغوية والحاجة إلى تنسيق إقليمي أكبر وسياسات لغوية رقمية أكثر شمولاً (Trigui et al., 2024; Khalfan et al., 2024).

على الرغم من هذا التباين، تواجه جميع هذه المناطق تحديات مشتركة بدرجات متفاوتة فيما يتعلق بمعالجة الأبعاد اللسانية الاجتماعية والجنائية للذكاء الاصطناعي. فمكافحة التحيز الخوارزمي في النماذج

اللغوية (Demidova et al., 2024; Nacar et al., 2025)، وضمان تمثيل عادل للهجات والتنوعات اللغوية (Faisal et al., 2024)، وحماية الفضاء الرقمي من التضليل المعلوماتي المولّد آلياً (Hosny & Nasef, 2025) ومن التحديات المتعلقة بتحديد المؤلف (PAN 2024)، تشكل هاجساً متزايداً. وتؤكد الدراسات الحديثة على أهمية تطوير أدوات ومقاربات خاصة بالسياق العربي لمواجهة هذه التحديات، مثل بناء معايير تقييم ثقافية للنماذج اللغوية العربية (Nacar et al., 2025) وتطوير تقنيات لكشف التلاعب اللغوي. إن الدروس المستفادة من هذه المقارنات تشير إلى أن فعالية أي سياسة لا تكمن فقط في التقدم التقني، بل في قدرتها على بناء حوكمة رشيدة (Trigui et al., 2024)، وتعزيز السيادة اللغوية الرقمية (Bondarenko et al., 2025)، وتأهيل الكفاءات البشرية، وتوفير البنى التحتية اللازمة لضمان أمن لغوي مستدام وشامل يخدم المجتمعات الناطقة بالعربية ويحمي هويتها في العصر الرقمي.

**جدول 1:** مقارنة بين سياسات الذكاء الاصطناعي واللغة الوطنية في المملكة العربية السعودية والإمارات العربية المتحدة (الخليج العربي)، والتوجهات الأوروبية، ومنطقة شمال إفريقيا

| المحور المقارن | المملكة العربية السعودية والإمارات العربية المتحدة | التوجهات الأوروبية | منطقة شمال إفريقيا |
|---|---|---|---|
| **الأهداف الرئيسة للذكاء الاصطناعي** | ريادة إقليمية وعالمية، اقتصاد قائم على المعرفة، تحسين جودة الحياة، تعزيز الهوية الوطنية الرقمية **(انظر Trigui et al., 2024 حول استراتيجيات AI العربية).** | بناء "ذكاء اصطناعي موثوق وأخلاقي" **(انظر Stelling et al., 2025 حول ممارسات قانون AI الأوروبي)**، تعزيز التنافسية، الحفاظ على القيم الديمقراطية وحقوق الإنسان، دعم الابتكار. | متفاوتة بين الدول: تحديث اقتصادي، تطوير خدمات، مع اهتمام متزايد (وإن كان متباينًا) بالبعد الرقمي والسيادة التكنولوجية **(Trigui et al., 2024).** |

| | | | |
|---|---|---|---|
| | | مهمة كجزء من التنوع الثقافي | |
| **مكانة اللغة الوطنية في سياسات AI** | متزايدة الأهمية كجزء من الهوية الوطنية واستراتيجية القوة الناعمة، دعم العربية في البيئة الرقمية (مثل نماذج "جيس"، "فالكون"، ومنصات مثل **"Fanar" كمنصة توليدية متعددة الوسائط عربية (Fanar Team et al., 2025)**). | الأوروبي، حماية اللغات الوطنية من هيمنة الإنجليزية، تعزيز تعدد اللغات في البيئة الرقمية (مثل مبادرات LT-Innovate، ومناقشة نماذج سيادية **(Bondarenko et al., 2025)**). | تتباين الأهمية المعلنة؛ جهود لدعم العربية واللغات الوطنية الأخرى (كالأمازيغية)، مع مبادرات لتطوير نماذج لهجية مثل **"Atlas-Chat" للهجة المغربية (Shang et al., 2024)**، ولكن غالبًا ما تكون الموارد والتنسيق محدودين. |
| **أبرز السياسات/الآليات (ذات الصلة لغويًا)** | استراتيجيات وطنية طموحة (استراتيجية المملكة العربية السعودية والإمارات العربية المتحدة للذكاء الاصطناعي)، مبادرات لتطوير نماذج لغوية عربية معيارية ولهجية **(AlAfnan, 2025; Nacar et al., 2025)**، دعم المحتوى العربي الرقمي، تعريب الخدمات، وتوفير مجموعات بيانات ضخمة **(Aloui et al., 2024; Alshahrani, 2024)**. | أطر تنظيمية وتشريعية (مثل قانون الذكاء الاصطناعي الأوروبي، قانون الخدمات الرقمية DSA، انظر **Stelling et al., 2025)**، دعم البحث والتطوير (مع تركيز على الجوانب الأخلاقية)، دعم الصناعات الثقافية واللغوية، مبادرات لإنشاء بنى تحتية لغوية أوروبية. | مبادرات وطنية متفرقة، محاولات لوضع استراتيجيات للذكاء الاصطناعي **(Trigui et al., 2024)**، بعض المشاريع لدعم المحتوى المحلي والبحث في معالجة اللغات واللهجات (مثل **Atlas-Chat (Shang et al., 2024)**)، ولكن غالبًا ما تفتقر إلى التمويل الكافي والتنسيق الإقليمي. |
| **معالجة التحديات اللسانية الاجتماعية والجنائية** | وعي متزايد بأهمية البيانات العربية والتحيز **حول التحيز في LLMs تجاه العربية (Demidova et al., 2024)**، جهود لتطوير نماذج محلية شاملة ثقافياً **(Nacar et al., 2025)**. الاهتمام بقضايا التضليل الرقمي والمقاومة الخوارزمية **(Hosny & Nasef, 2025)** يبرز في سياسات أوسع للأمن السيبراني. | تركيز قوي على مكافحة التحيز الخوارزمي، وضمان الشفافية، وحماية الخصوصية. قوانين صارمة لمكافحة التضليل الرقمي وخطاب الكراهية. نقاش حول تنظيم المنصات لحماية التنوع اللغوي العميق، | وعي ناشئ بالتحديات **(Trigui et al., 2024)**، لكن غالبًا ما تكون الأطر القانونية والمؤسسية غير مكتملة أو قيد التطوير لمواجهة التحديات اللسانية الجنائية المتطورة للذكاء الاصطناعي (مثل التزييف العميق، تحديد المؤلف، تحديات |

| | | | |
|---|---|---|---|
| | | **Mihalcea et al.,) 2025 حول AI للجميع)**. | الاعتدال اللغوي **(Shahid et al., 2025)**. |
| **التحديات الرئيسية (من منظور لغوي)** | بناء بيانات عربية ضخمة ومتنوعة وتمثيلية **(Aloui et al., 2024; Alshahrani, 2024)** تشمل اللهجات والمتون الضخمة **حول Faisal et al., 2024) Dialectbench)**، توفير كفاءات متخصصة، تنسيق المبادرات، إدارة التنوع اللهجي بفعالية، الموازنة بين الانفتاح والسيادة اللغوية الرقمية **(Bondarenko et al., 2025)**. | بطء وتيرة التنظيم مقارنة بسرعة التكنولوجيا، تحديات في بناء بدائل أوروبية قوية للمنصات العالمية، تطبيق التشريعات بفعالية عبر الدول الأعضاء **(Stelling et al., 2025)**. | فجوة رقمية ولغوية، نقص الاستثمار في البحث والتطوير اللغوي، هجرة الأدمغة، ضعف البنية التحتية الرقمية في بعض المناطق، تحديات التعدد اللغوي والتخطيط اللغوي **(Khalfan et al., 2024 حول تحديات الإعلام الرقمي)**. |
| **دروس/قابلية للتطبيق عربيًا (عام)** | أهمية القيادة السياسية، الاستثمار في البحث والتطوير **(انظر جهود بناء النماذج والبيانات)**، الشراكات، الحاجة الملحة لمعالجة تحديات البيانات والكفاءات والتنوع اللهجي. | أهمية الإطار المؤسسي والتشريعي، التركيز على الأخلاقيات والجودة، التعاون الإقليمي، ضرورة حماية المستخدمين والتنوع الثقافي. (مع تحدي القدرة التنافسية العالمية). | ضرورة تعزيز الاستثمار في البنية التحتية والموارد البشرية، أهمية التعاون الإقليمي **(Trigui et al., 2024)**، الحاجة إلى سياسات لغوية رقمية واضحة تأخذ بعين الاعتبار السياقات المحلية المتنوعة واللهجية **(Shang et al., 2024)**. |

*ملاحظة:* تم إعداد هذا الجدول بناءً على تحليل الوثائق والأدبيات المتاحة حتى تاريخ إعداد هذه الدراسة.

يكشف التحليل المقارن المعمق للنماذج السياساتية المتبعة في التعامل مع تقاطع الذكاء الاصطناعي واللغة الوطنية، والذي يستند إلى فحص الوثائق الرسمية والخطابات المصاحبة، عن غياب وصفة سحرية قابلة للتطبيق عالميًا. فكل مقاربة هي نتاج سياقها الاجتماعي والاقتصادي والسياسي الفريد، وتحمل بالضرورة مفاضلاتها وتحدياتها الخاصة. ومع ذلك، يتضح أن القاسم المشترك بين التجارب الأكثر وعيًا وتقدمًا، سواء في منطقة الخليج العربي أو ضمن التوجهات الأوروبية، يتمثل في الاعتراف الاستراتيجي الصريح بالأهمية الحيوية لتقاطع الذكاء الاصطناعي واللغة. يظهر هذا الاعتراف جليًا في وجود رؤية

سياسية واضحة، كما تشير إليه استراتيجيات الذكاء الاصطناعي الطموحة في دول مثل المملكة العربية السعودية والإمارات العربية المتحدة (انظر Trigui et al., 2024 حول استراتيجيات AI العربية)، وكذلك في الجهود الأوروبية لبناء "ذكاء اصطناعي موثوق وأخلاقي" (انظر Stelling et al., 2025 حول ممارسات قانون AI الأوروبي). ويترافق هذا الاعتراف باستعداد ملموس لتخصيص موارد كبيرة، سواء لتطوير نماذج لغوية عربية كبرى ومنصات توليدية متعددة الوسائط مثل ""Fanar (Fanar Team et al., 2025 )، أو لإنشاء مجموعات بيانات ضخمة ومتنوعة كقاعدة أساسية لهذه النماذج ( Aloui et al., 2024; Alshahrani, 2024).

وفي المقابل، فإن غياب هذه العناصر من الاعتراف الاستراتيجي، والرؤية السياسية المتكاملة، والاستثمار الكافي، كما قد يُستشف من التحديات التي تواجهها بعض السياقات في منطقة شمال إفريقيا على الرغم من مبادرات واعدة مثل "Atlas-Chat" ( Shang et al., 2024)، يترك حتمًا المجال مفتوحًا أمام التأثيرات غير الموجهة للتكنولوجيا العالمية. هذا الوضع لا يهدد فقط بتعميق فجوة التبعية المعرفية والتكنولوجية، بل يفاقم من مخاطر تهميش اللغة العربية، ويقوض الجهود الرامية لمعالجة التحيزات الكامنة في النماذج العالمية ( Demidova et al., 2024; Mihalcea et al., 2025)، ويزيد من صعوبة مواجهة التحديات اللسانية الجنائية المتعلقة بانتشار الخطاب المضلل والمقاومة الخوارزمية (Hosny & Nasef, 2025; Shahid et al., 2025). إن فهم هذه الديناميكيات يؤكد على ضرورة بناء نماذج سيادية للغات ( Bondarenko et al., 2025) تأخذ بعين الاعتبار الخصوصيات الثقافية واللغوية، وتضمن أمنًا لغويًا شاملاً في مواجهة الهيمنة الرقمية.

**المناقشة**

يتجاوز هذا القسم مجرد عرض النتائج لينخرط في مناقشة تحليلية نقدية لها في ضوء الإطار النظري اللساني الاجتماعي والجنائي المعتمد وأسئلة البحث المطروحة. يهدف إلى تفسير الأنماط الملاحظة، وربط النتائج بالأدبيات ذات الصلة، واستخلاص استنتاجات أعمق حول مستقبل الأمن اللغوي العربي في عصر الذكاء الاصطناعي.

**تأكيد التهديدات اللسانية الاجتماعية والجنائية وتعميق فهمها**

تؤكد النتائج المستخلصة بقوة أن الأمن اللغوي العربي يواجه تهديدات بنيوية متعددة الأوجه تتجاوز القصور التقني البحت. فظاهرة الهيمنة الرقمية اللغوية، كما توثقها دراسات نقدية (Couldry & Mejias, 2019)، ليست مجرد مسألة كمية تتعلق بقلة البيانات العربية المتاحة لتدريب نماذج الذكاء الاصطناعي، بل هي انعكاس عميق لعلاقات قوة عالمية وتاريخية ((Langmia, 2016 حول الهيمنة الثقافية واللغوية؛ (Sarieddine, 2023) حول الهيمنة الصينية واللغة. هذا "الفقر التمثيلي" للغة العربية، وخاصة تنوعاتها اللهجية، في النماذج اللغوية الكبيرة لا يؤدي فقط إلى أداء دون المستوى المطلوب، بل، وهو الأخطر من منظور لساني اجتماعي، يكرس تحيزات خوارزمية قد تهمش هويات لغوية وثقافية معينة داخل الفضاء العربي نفسه أو تقدم صورة نمطية عن العرب واللغة العربية للعالم (Noble, 2018؛ انظر أيضاً Demidova et al., 2024 حول التحيز في نماذج اللغة الكبيرة متعددة اللغات تجاه الأسماء والمفاهيم العربية؛ Shahid et al., 2025 حول التحيزات الاستعمارية في الاعتدال الآلي للغات منخفضة الموارد). ويتجلى هذا في كيفية تأثير هذه النماذج على المعايير اللغوية الاجتماعية، حيث قد يؤدي الاستخدام المتزايد لأدوات الذكاء الاصطناعي إلى تبني أنماط لغوية مبسطة أو مستوردة، أو إلى تعزيز هيمنة أشكال لغوية معينة (مثل الفصحى المعيارية المبسطة) على حساب التنوع اللهجي الغني، وهو ما يمثل

شكلاً من أشكال "العنف الرمزي" (Bourdieu, 1991) الذي يقلل من قيمة رأس المال اللغوي المرتبط باللهجات، ويقوض الجهود الرامية لتمثيل هذا التنوع في الفضاء الرقمي (Mihalcea et al., 2025).

أما على الصعيد الجنائي اللغوي، فإن النتائج تسلط الضوء بجلاء على خطورة التحديات الناشئة عن قدرة الذكاء الاصطناعي المتطورة على توليد نصوص عربية يصعب، وأحيانًا يستحيل، تمييزها عن النصوص البشرية الأصيلة. هذا التقدم التقني يفتح الباب على مصراعيه أمام تفاقم مشكلة الخطاب المضلل والمعلومات المغلوطة (Wardle & Derakhshan, 2017)، حيث يمكن إنتاج محتوى زائف بكميات هائلة وبتكلفة زهيدة، مما يشكل تهديدًا مباشرًا لسلامة الفضاء المعلوماتي العربي والثقة المجتمعية، وهو ما تؤكده دراسات حديثة حول تحديات الاعتدال اللغوي والمقاومة الخوارزمية في السياق العربي (Hosny & Nasef, 2025). كما أن صعوبة تحديد المؤلف الحقيقي للنصوص المولدة آليًا تطرح إشكاليات قانونية وأخلاقية معقدة تتعلق بالمسؤولية عن المحتوى الضار، أو المنتهك لحقوق الملكية الفكرية، أو المستخدم في أنشطة احتيالية (Coulthard & Johnson, 2007)، وهي تحديات تتناولها الأبحاث المتخصصة في التحقق من مؤلفي النصوص المولدة بالذكاء الاصطناعي (انظر PAN 2024). إن هذه التحديات الجنائية اللغوية لا تهدد فقط الأفراد والمؤسسات، بل تقوض أسس الحوار العام البنّاء، وتزيد من استقطاب المجتمعات، وتستدعي تطوير أدوات وتقنيات مضادة قادرة على كشف هذا النوع من التزييف اللغوي.

وتتفاقم هذه التهديدات اللسانية الاجتماعية والجنائية بفعل ديناميكيات أوسع تتعلق بالسيادة اللغوية الرقمية والحاجة الملحة لبناء قدرات عربية مستقلة في مجال الذكاء الاصطناعي. فالاعتماد المفرط على نماذج وتقنيات مطورة في سياقات غربية لا يعرض اللغة العربية لخطر التحيز والتهميش فحسب، بل يعمق أيضاً التبعية التكنولوجية ويقوض الجهود الرامية لتحقيق "سيادة لغوية رقمية" حقيقية (Bondarenko et

al., 2025). إن بناء مجموعات بيانات عربية ضخمة ومتنوعة وتمثيلية (Aloui et al., 2024; Alshahrani, 2024)، وتطوير نماذج لغوية عربية تأخذ في الاعتبار الخصوصيات الثقافية واللهجية (Nacar et al., 2025; Shang et al., 2024)، وتأهيل الكفاءات الوطنية القادرة على الابتكار والتوجيه الأخلاقي والقانوني لهذه التقنيات (Trigui et al., 2024)، لم تعد ترفاً فكرياً بل ضرورة استراتيجية. فبدون هذه الجهود، ستبقى اللغة العربية ومستخدموها عرضة لتأثيرات غير مرغوب فيها، وستظل الجهود المبذولة لمكافحة التضليل أو تعزيز الهوية الرقمية قاصرة ما لم تُبنَ على أسس تكنولوجية ومعرفية ذاتية تضمن الفهم العميق للسياق اللغوي والثقافي العربي (AlAfnan, 2025; El Zahraa, 2025).

**الفرص الكامنة: نحو توجيه استراتيجي للذكاء الاصطناعي اللغوي**

رغم هذه التهديدات المتعددة الأوجه، تشير النتائج أيضًا إلى أن الذكاء الاصطناعي يحمل في طياته فرصًا واعدة يمكن، بل يجب، توجيهها استراتيجيًا من منظور لساني اجتماعي وجنائي لخدمة اللغة العربية ومجتمعاتها. ففي مجال اللسانيات الاجتماعية التطبيقية، تبرز إمكانات هذه التقنيات في المساهمة الفعالة في تطوير أدوات تعليمية مبتكرة وتفاعلية للغة العربية، تستجيب للاحتياجات المتغيرة للمتعلمين وتعزز من كفاءتهم اللغوية (انظر El Zahraa, 2025; AlAfnan, 2025). كما يمكن للذكاء الاصطناعي أن يدعم بشكل كبير جهود رقمنة التراث اللغوي العربي الهائل، وتيسير الوصول إليه، وتوفير تقنيات مساعدة للمتحدثين باللهجات المختلفة أو لذوي الاحتياجات الخاصة، مما يعزز من شمولية الفضاء الرقمي. وفي هذا السياق، يكتسب تطوير نماذج لغوية عربية تأخذ في الاعتبار التنوع اللهجي الغني أهمية بالغة، إذ يمكن أن يساهم بشكل مباشر في تعزيز "العدالة اللغوية الرقمية" وتقليل التهميش الذي قد تتعرض له بعض التنويعات

اللغوية (انظر Shang et al., 2024 حول تطوير نماذج للهجات؛ Faisal et al., 2024 حول معايير تقييم اللهجات؛ Tibi & Messaoud, 2025 لدراسات تصنيف اللهجات).

أما من منظور علم اللغة الجنائي، فإن تطوير أدوات ذكاء اصطناعي متخصصة في تحليل النصوص العربية يمكن أن يمثل خط دفاع حيوي ومهم ضد موجات التضليل والتزييف المتصاعدة. فبناء نماذج قادرة على كشف السمات اللغوية الدقيقة والمميزة للنصوص المولدة آليًا، وتحديد أنماط الخداع اللغوي، والمساعدة في تحليل المؤلف بشكل أكثر كفاءة ودقة، يمكن أن يوفر أدوات لا تقدر بثمن للباحثين والمحققين والصحفيين وعامة الناس للتحقق من مصداقية المحتوى الرقمي (انظر PAN 2024 حول التحقق من مؤلفي النصوص المولدة بالذكاء الاصطناعي). إن القدرة على تمييز المحتوى الأصيل من المحتوى المصطنع أو المضلل، خاصة في سياق الحملات المنظمة التي قد تستغل تقنيات المقاومة الخوارزمية (Hosny & Nasef, 2025)، تعد ركيزة أساسية للحفاظ على الثقة وسلامة الخطاب في المجتمعات العربية.

**السياسات المقارنة: دروس مستفادة وضرورة التكييف السياقي**

يكشف التحليل المقارن لتجارب المملكة العربية السعودية والإمارات العربية المتحدة والتوجهات الأوروبية ومنطقة شمال إفريقيا عن تباين واضح في مستويات الوعي والنضج المؤسسي والقدرات التقنية في التعامل مع تقاطع الذكاء الاصطناعي واللغة. فبينما تُظهر **المملكة العربية السعودية والإمارات العربية المتحدة** طموحًا رياديًا في تبني الذكاء الاصطناعي وتطوير نماذج لغوية عربية، فإن التحدي يكمن في تعميق المعالجة اللسانية الاجتماعية والجنائية بشكل منهجي ضمن هذه الاستراتيجيات، وضمان استدامة هذه الجهود. ومن جانبها، تقدم **التوجهات الأوروبية** نموذجًا للتركيز على الأطر التنظيمية والأخلاقية ومكافحة

التحيز والتضليل، ولكنها تواجه تحديات في سرعة مواكبة التطور التكنولوجي والتنافسية العالمية. أما **منطقة شمال إفريقيا**، فتعكس التحديات الكبيرة المتعلقة بالفجوة الرقمية واللغوية، ونقص الموارد، والحاجة إلى تعزيز التنسيق ووضع سياسات لغوية رقمية شاملة تأخذ بعين الاعتبار التنوع اللغوي والثقافي الفريد لهذه المنطقة.

الدرس الأساسي المستفاد هو أنه لا توجد حلول جاهزة يمكن استيرادها، وأن فعالية أي سياسة تعتمد على مدى تكييفها مع السياق اللساني الاجتماعي والثقافي والمؤسسي والقانوني المحلي. ومع ذلك، فإن القواسم المشتركة بين المقاربات الواعدة تشمل: الاعتراف الاستراتيجي بأهمية الأمن اللغوي، وتخصيص الاستثمارات اللازمة، وبناء القدرات البشرية، وتعزيز التعاون بين مختلف الأطراف المعنية (الحكومات، المؤسسات الأكاديمية، القطاع الخاص، المجتمع المدني)، ووضع أطر تنظيمية مرنة وقادرة على مواكبة التطورات.

**نحو إطار متكامل للأمن اللغوي العربي في عصر الذكاء الاصطناعي**

إن الإجابة الفعالة على التحديات اللسانية الاجتماعية والجنائية التي يطرحها الذكاء الاصطناعي تتطلب تجاوز الجهود المتفرقة وتبني **مقاربة استراتيجية متكاملة للأمن اللغوي العربي**. يجب أن ترتكز هذه المقاربة على فهم عميق لديناميكيات القوة والهيمنة في الفضاء الرقمي، وأن تدمج بين التدابير التقنية والسياساتية والقانونية والتعليمية والتوعوية. كما يجب أن تعزز هذه المقاربة **"السيادة اللغوية الرقمية"** للعربية، ليس بمعنى الانغلاق، بل بمعنى تمكين المجتمعات الناطقة بالعربية من التحكم في مصير لغتها في البيئة الرقمية، والمشاركة بفعالية في تطوير واستخدام تقنيات الذكاء الاصطناعي بما يخدم مصالحها وقيمها، ويحميها من المخاطر الناشئة.

**محددات الدراسة**

على الرغم من الجهود المبذولة لضمان الدقة والشمولية، تظل هناك بعض المحددات الكامنة في هذه الدراسة والتي ينبغي أخذها في الاعتبار عند تفسير النتائج وتعميمها:

- **الاعتماد على البيانات الثانوية:** استندت الدراسة بشكل أساسي إلى تحليل الوثائق والأدبيات المتاحة. ورغم السعي لتثليث المصادر، فإن بعض هذه البيانات قد تحمل تحيزات كامنة أو لا تعكس الواقع الميداني بشكل كامل، خاصة فيما يتعلق بالممارسات اللغانية الفعلية للأفراد وتفاعلاتهم مع تقنيات الذكاء الاصطناعي.
- **حداثة الظاهرة وسرعة تطورها:** يشهد مجال الذكاء الاصطناعي اللغوي تطورات متسارعة للغاية. وبالتالي، فإن التحليلات والنتائج المقدمة تمثل "لقطة" في لحظة زمنية محددة، وقد تتغير الديناميكيات والسياسات بسرعة، مما يستدعي تحديثًا ومتابعة مستمرة.
- **تحديات المقارنة بين سياقات متنوعة:** رغم الحرص على اختيار حالات مقارنة (المملكة العربية السعودية والإمارات العربية المتحدة، التوجهات الأوروبية، شمال إفريقيا) تعكس تنوعًا، فإن المقارنة بين سياقات سياسية واقتصادية وثقافية ولغوية شديدة الاختلاف تحمل خطر التبسيط. تم التركيز على المقارنة التحليلية للمقاربات والآليات بدلاً من النتائج النهائية، مع التأكيد على أهمية تكييف أي دروس مستفادة مع الخصوصيات المحلية.

- **الوصول المحدود لبعض البيانات التفصيلية:** قد يكون الوصول إلى بيانات تفصيلية حول بعض المبادرات أو السياسات اللغوية الرقمية في بعض السياقات المقارنة، خاصة تلك غير المنشورة باللغات التي يتقنها الباحث، محدودًا، مما قد يؤثر على عمق بعض التحليلات المقارنة.
- **التركيز على المستوى الكلي والمتوسط:** ركزت الدراسة بشكل كبير على تحليل السياسات والخطابات والاتجاهات العامة. ورغم أهمية ذلك، فإنها لم تتعمق بشكل كافٍ في دراسة التأثيرات على المستوى الميكرو-لساني (مثل التغيرات الدقيقة في التراكيب اللغوية أو الأساليب الفردية) أو إجراء تحليلات جنائية لغوية تطبيقية على مجموعات بيانات محددة، وهو ما يمثل مجالات واعدة لبحوث مستقبلية.

إن إدراك هذه المحددات لا يقلل من أهمية النتائج التي توصلت إليها الدراسة، ولكنه يدعو إلى التعامل معها بحذر وإلى ضرورة إجراء المزيد من البحوث الميدانية والتطبيقية لتعميق الفهم.

**الخاتمة والتوصيات**

ختاماً لهذه الرحلة البحثية التي سعت لاستكشاف التفاعلات المعقدة بين الذكاء الاصطناعي (AI) والأمن اللغوي العربي من منظور لساني اجتماعي وجنائي نقدي، يمكن القول بأننا نقف أمام لحظة تاريخية فارقة تتطلب وعياً عميقاً واستجابة استراتيجية مدروسة. لقد كشفت الدراسة، من خلال مراجعتها النقدية للأدبيات، وبناء إطار نظري لساني اجتماعي وجنائي تكاملي، وتحليل الوثائق والمقارنة بين التجارب الدولية، أن تأثير الذكاء الاصطناعي على اللغة العربية يتجاوز كونه مجرد تحدٍ تقني، ليشكل ظاهرة "لغوية تقنية اجتماعية" جذرية تعيد تشكيل المشهد اللغوي والثقافي والمعرفي، وتطرح تحديات جدية على سلامة الخطاب الرقمي العربي.

وقد خلص التحليل إلى أن الأمن اللغوي العربي، بمفهومه الشامل الذي يدمج الأبعاد اللسانية الاجتماعية والجنائية، يواجه تهديدات بنيوية حقيقية ناجمة عن ديناميكيات الهيمنة الرقمية العالمية، والتحيزات الخوارزمية المضمنة في التقنيات، والتأثير على المعايير والممارسات اللغوية، والتحديات التي تطرحها على الهوية اللغوية، وموثوقية الخطاب الرقمي. تتفاقم هذه التهديدات بفعل تحديات داخلية تتعلق بضعف الاستثمار الموجه والتنسيق الاستراتيجي. ولكن، وفي المقابل، أبرزت الدراسة أيضاً أن هذه التكنولوجيا نفسها تحمل في طياتها فرصاً غير مسبوقة لتعزيز حيوية اللغة العربية، وتوسيع نطاق استخدامها، وتعزيز حضورها الرقمي، وتطوير أدوات لسانية جنائية لمكافحة التضليل، شريطة أن يتم التعامل معها بإرادة واعية ورؤية استراتيجية تتجاوز مجرد رد الفعل أو الاستهلاك السلبي.

وتُؤكّدُ الدراسةُ، استناداً إلى منظوراتٍ نظريةٍ لسانية اجتماعية وجنائية مُعتمدةٍ كنقد الخطاب، وتحليل الهيمنة اللغوية، ومبادئ علم اللغة الجنائي، أنّ هذه التأثيراتِ ليستْ قدراً محتوماً؛ إنّما هي نتاجٌ لعمليات بناءٍ اجتماعي وتفاعلاتِ قوى وتطورات تكنولوجية يُمكنُ، بل ويتوجّبُ، التدخّلُ لتوجيهِ مساراتِها. وفي السياقِ ذاتِه، أظهرَ التحليلُ المقارنُ لتجاربِ المملكة العربية السعودية والإمارات العربية المتحدة والتوجهات الأوروبية ومنطقة شمال إفريقيا، على الرّغمِ من تباينِها، أنّ الوعيَ الاستراتيجيَّ والتدخّلَ السياساتيَّ المدروسَ يُشكّلانِ ركيزةً أساسيةً لمجابهةِ التحدّياتِ اللغويةِ وتعزيز الأمن اللغوي في العصرِ الرقميِّ.

بناءً على هذه الاستخلاصات، تقدم هذه الدراسة مساهمات على المستويين النظري والتطبيقي. نظرياً، تساهم في إثراء النقاش الأكاديمي من خلال تقديم إطار لساني اجتماعي وجنائي تكاملي ونقدي لتحليل تقاطع الذكاء الاصطناعي واللغة في سياق عربي. وتطبيقياً، توفر لصناع القرار والسياسات

والمؤسسات المعنية في العالم العربي تحليلاً معمقاً للواقع الراهن ورؤى مستمدة من تجارب دولية، وتخلص إلى مجموعة من التوصيات الاستراتيجية والعملية التي تهدف إلى بناء مقاربة عربية شاملة وفعالة لتعزيز الأمن اللغوي بأبعاده المتعددة في عصر الذكاء الاصطناعي.

**التوصيات الاستراتيجية والعملية:**

1. **بناء رؤية استراتيجية وطنية وإقليمية واضحة للأمن اللغوي الرقمي:**

- **التوصية:** تطوير واعتماد استراتيجيات وطنية (تُنسق إقليمياً) تعترف صراحة بالأمن اللغوي العربي (بأبعاده اللسانية الاجتماعية والجنائية) كأصل استراتيجي، وتدمجه ضمن خطط الذكاء الاصطناعي والتحول الرقمي والأمن السيبراني.
- **آليات التنفيذ المقترحة:** تشكيل لجان وطنية عليا متعددة التخصصات (تضم خبراء لغة، لسانيين اجتماعيين، خبراء في علم اللغة الجنائي، تقنيين، قانونيين، صناع سياسات) لوضع الخطط ومؤشرات الأداء، مع دعم سياسي رفيع المستوى وتخصيص ميزانيات واضحة.

2. **الاستثمار المكثف والمستدام في الموارد اللغوية الرقمية العربية الموثوقة والمتنوعة:**

- **التوصية:** إطلاق مبادرات كبرى لبناء وتوسيم وإتاحة مجموعات بيانات (Corpora) عربية ضخمة وعالية الجودة وتمثيلية لمختلف اللهجات والمجالات، مع التركيز على بناء بيانات تدريبية تساعد في كشف التحيز والتضليل، ودعم تطوير أدوات لغوية حاسوبية أساسية مفتوحة المصدر.

- **آليات التنفيذ المقترحة:** إنشاء "مرصد عربي للبيانات اللغوية والذكاء الاصطناعي" بتمويل مشترك، وإطلاق برامج لدعم تطوير أدوات لمعالجة اللهجات وتحليل المشاعر وكشف التزييف العميق.

3. **بناء القدرات البشرية المتخصصة في اللسانيات الحاسوبية وعلم اللغة الجنائي الرقمي:**

- **التوصية:** تأهيل جيل جديد من الكفاءات العربية في مجالات تقاطع اللغة والذكاء الاصطناعي، بما في ذلك اللسانيات الاجتماعية الرقمية، وعلم اللغة الجنائي التطبيقي، وزيادة تمويل البحث العلمي الموجه لتطوير نماذج وتقنيات تخدم اللغة العربية وتلبي احتياجات الأمن اللغوي.
- **آليات التنفيذ المقترحة:** تطوير برامج دراسات عليا متخصصة، وورش عمل تدريبية للمهنيين (في الإعلام والقانون والأمن)، وإطلاق حاضنات أعمال للشركات الناشئة في التقنيات اللغوية العربية التي تركز على الموثوقية والأمن.

4. **تطوير حوكمة أخلاقية وقانونية رشيدة لتطبيقات الذكاء الاصطناعي اللغوية:**

- **التوصية:** الإسراع في وضع أطر تنظيمية وتشريعية تضمن الشفافية والمساءلة والعدالة في تطبيقات الذكاء الاصطناعي اللغوية، وتحمي المستخدمين من التحيز والتضليل، وتعالج قضايا الملكية الفكرية، وتحدد المسؤوليات القانونية عن المحتوى المولّد آلياً.
- **آليات التنفيذ المقترحة:** تشكيل لجان خبراء وطنية لوضع مبادئ توجيهية أخلاقية وقانونية للذكاء الاصطناعي اللغوي العربي، مسترشدة بالتوصيات الدولية مع تكييفها للسياق العربي، وتعزيز التعاون مع المنصات الرقمية لمكافحة المحتوى الضار.

5. **تعزيز المحتوى العربي الرقمي عالي الجودة ودعم التنوع اللغوي وحماية سلامة الخطاب:**

- **التوصية:** التركيز على النوعية والأصالة والقيمة المضافة في المحتوى الرقمي العربي، وتشجيع تطوير واستخدام تطبيقات ذكاء اصطناعي تدعم اللهجات العربية المختلفة، وتطوير أدوات للتحقق من صحة المعلومات ومكافحة التضليل.
- **آليات التنفيذ المقترحة:** إطلاق جوائز ومبادرات لدعم المحتوى العربي المتميز والموثوق، وتمويل مشاريع تستخدم الذكاء الاصطناعي لتوثيق اللهجات أو لإنتاج محتوى تعليمي بها، ودعم مبادرات التربية الإعلامية والمعلوماتية.

6. **رفع الوعي المجتمعي وتعزيز التفكير النقدي والمشاركة في حوكمة الذكاء الاصطناعي اللغوي:**

- **التوصية:** تضمين مهارات التفكير النقدي والوعي الإعلامي والرقمي في المناهج التعليمية، وفتح نقاش مجتمعي واسع حول تأثيرات الذكاء الاصطناعي على اللغة والهوية وسلامة الخطاب، وإشراك مختلف الفئات في صياغة الرؤى ومراقبة السياسات.
- **آليات التنفيذ المقترحة:** تطوير وحدات دراسية حول "الذكاء الاصطناعي والمجتمع واللغة" للمدارس والجامعات، وتنظيم حملات توعية عامة حول مخاطر التضليل وكيفية التعرف على المحتوى المولّد آلياً.

7. **تفعيل التعاون الإقليمي العربي كضرورة استراتيجية لمواجهة التحديات المشتركة:**

- **التوصية:** إنشاء منصات ومبادرات عربية مشتركة لتبادل الخبرات وتنسيق الجهود وتوحيد الموارد في المجالات الحيوية (بناء البيانات، تطوير التقنيات، وضع المعايير، مكافحة التضليل العابر للحدود)، وتبني موقف تفاوضي موحد في المحافل الدولية.
- **آليات التنفيذ المقترحة:** عقد مؤتمر سنوي دوري حول "الذكاء الاصطناعي والأمن اللغوي العربي: التحديات اللسانية الاجتماعية والجنائية"، وإنشاء شبكة بحثية عربية مشتركة، وتشكيل فريق عمل عربي للتفاوض في المنظمات الدولية المعنية بحوكمة الذكاء الاصطناعي والإنترنت.

**آفاق بحثية مستقبلية:**

تفتح هذه الدراسة الباب واسعاً أمام العديد من المسارات البحثية المستقبلية، منها:

1. دراسات لغوية اجتماعية ورقمية معمقة حول تفاعلات المستخدمين العرب مع المحتوى المولّد آلياً وتأثير ذلك على ممارساتهم اللغوية.
2. تحليلات جنائية لغوية تطبيقية على مجموعات بيانات عربية لتطوير نماذج لكشف التزييف وتحديد المؤلف.
3. دراسة مقارنة تفصيلية لفعالية التشريعات والمبادرات المتخذة داخل المنطقة العربية لمكافحة التضليل الرقمي.
4. بحث معمق في الأبعاد الاقتصادية لتطوير تقنيات لغوية عربية آمنة وموثوقة.
5. تطوير مؤشرات ومقاييس علمية لتقييم ورصد حالة "الأمن اللغوي الرقمي العربي" بأبعاده اللسانية الاجتماعية والجنائية.

نؤكد مجدداً أن ضمان مستقبل آمن ومزدهر للغة العربية في عصر الذكاء الاصطناعي المتسارع ليس مجرد خيار تكنولوجي، بل هو ضرورة استراتيجية ومسؤولية حضارية مشتركة. يتطلب ذلك مقاربة استباقية تستند إلى وعي نقدي بديناميكيات القوة والتأثيرات اللسانية الاجتماعية والجنائية للتكنولوجيا، ورؤية استراتيجية تعلي من شأن اللغة كركيزة للهوية والأمن. يأمل هذا البحث، بكل تواضع، أن يكون قد أضاء بعض جوانب هذا التحدي المعقد والمتشعب، وأن تمثل تحليلاته وتوصياته مساهمة معرفية متواضعة ولكنها محفزة، ومنطلقاً لجهود مستقبلية أكثر عمقاً وفعالية وتأثيراً، بما يخدم في النهاية الهدف الأسمى المتمثل في: **"بأمن لغتنا، نصون هويتنا ومستقبلنا."**

## قائمة المراجع

### المراجع العربية:


التركي، عبد الله. (2001). *الأمن الفكري وعناية المملكة العربية السعودية به* [محاضرة]. مدينة تدريب الأمن العام، مكة المكرمة.

العزير، أكرم محمد محمد. (2023). *الأمن اللغوي للغة العربية*. (ط.1). صنعاء.

عطية، محمد، وراغب، أحمد وآخرون. (2019). *العربية والذكاء الاصطناعي*. مركز الملك عبد الله بن عبد العزيز الدولي.

القايدي، آمنة مصبح علي. (2023). *اللغة العربية والذكاء الاصطناعي (رؤى وآمال)*. ورقة بحثية مقدمة إلى مؤتمر اللغة العربية التاسع، المجلس الدولي للغة العربية، دبي، 6-8 نوفمبر 2023.

المسدي، عبد السلام. (2014). *الهوية العربية والأمن اللغوي: دراسة وتوثيق*. المركز العربي للأبحاث ودراسة السياسات.

زاهد، زهير غازي. (2000). *العربية والأمن اللغوي*. مؤسسة الورّاق للنشر والتوزيع.


### المراجع الأجنبية:


Abdul-Mageed, M., & Diab, M. T. (2012). SAMAR: A system for subjectivity and sentiment analysis of Arabic social media. *Computational Linguistics, 38*(1), 20–36. https://doi.org/10.1162/COLI_a_00118

AlAfnan, M. A. (2025). Artificial intelligence and language: Bridging Arabic and English with technology. *Journal of Ecohumanism*. Retrieved from https://ecohumanism.co.uk/

Al-Saggaf, Y. (2016). The role of social media in the political landscape in the Arab world. In *Social media in the Arab world* (pp. 1–14). Springer VS. https://doi.org/10.1007/978-3-658-11911-1_1

Alhafni, B., Abdul-Mageed, M., El-Kishky, A., Elmadany, A., Nagoudi, E. M., & Glass, J. (2023). Towards understanding the capabilities and limitations of large language models for

Arabic. *Findings of the Association for Computational Linguistics: EMNLP 2023*, 8721–8744. https://aclanthology.org/2023.findings-emnlp.584

Ali, M. (2016). A brief introduction to decolonial computing. *XRDS: Crossroads, The ACM Magazine for Students, 22*(4), 16–21. https://doi.org/10.1145/2930886

Aloui, M., Chouikhi, H., Chaabane, G., Kchaou, H., Abdul-Mageed, M., Al-Mohannadi, N., Al-Shorbaji, N., Belinkov, Y., Bouamor, H., Habash, N., Obeid, O., Sagot, B., Samih, Y., & Zaghouani, W. (2024). *101 billion arabic words dataset*. arXiv preprint arXiv:2402.15108. https://arxiv.org/abs/2402.15108

Alshahrani, S. F. A. (2024). *Towards representative pre-training corpora for Arabic natural language processing*. ProQuest Dissertations Publishing.

Antoun, W., Baly, F., Hajj, H., Habash, N., & El-Haj, W. (2022). AraBench: Benchmarking dialectal Arabic-English machine translation. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 4854–4866. https://aclanthology.org/2022.emnlp-main.327

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *Journal of Economic Perspectives, 29*(3), 3–30. https://doi.org/10.1257/jep.29.3.3

Ayele, A. A., Babakov, N., Bevendorff, J., Casals, X. B., Potthast, M., Rosso, P., Specht, G., Stamatatos, E., & Stein, B. (2024). *Overview of PAN 2024: Multi-author writing style analysis, multilingual text detoxification, oppositional thinking analysis, and generative AI authorship verification*. Working Notes of CLEF 2024 - Conference and Labs of the Evaluation Forum.

Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim Code*. Polity Press.

Bijker, W. E., Hughes, T. P., & Pinch, T. J. (Eds.). (1987). *The social construction of technological systems: New directions in the sociology and history of technology*. MIT Press.

Bondarenko, M., Lushnei, S., Paniv, Y., SkKuznetsov, V., SkKuznetsova, I., & Yermakov, D. (2025). *Sovereign large language models: Advantages, strategy and regulations*. arXiv preprint arXiv:2501.05057. https://arxiv.org/abs/2501.05057

Bourdieu, P. (1991). *Language and symbolic power* (G. Raymond & M. Adamson, Trans.; J. B. Thompson, Ed.). Harvard University Press.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901. https://proceedings.neurips.cc/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Bucher, T. (2018). *If... then: Algorithmic power and politics*. Oxford University Press. https://doi.org/10.1093/oso/9780190657823.001.0001

Castells, M. (2004). *The power of identity* (2nd ed.). Blackwell Publishing.

Couldry, N., & Hepp, A. (2017). *The mediated construction of reality*. Polity Press.

Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.

Crawford, K. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Cunliffe, D., & Hijazi, M. (2018). Language policies for the digital world. In F. Grin, L. Cardinal, & K. Rédai (Eds.), *The Routledge handbook of language policy and planning* (pp. 499–514). Routledge. https://doi.org/10.4324/9781315752818

Demidova, A., Atwany, H., Rabih, N., Yaghoobzadeh, Y., & Fraser, A. (2024). John vs. Ahmed: Debate-induced bias in multilingual LLMs. *Proceedings of the 2024 Conference of the*

*North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 3345-3358. https://aclanthology.org/2024.naacl-main.186

El Zahraa, F. (2025). Leveraging artificial intelligence and digital technologies to enhance sociolinguistic competence and Arabic language skills. *Proceeding of the International Conference on Arabic Language Processing (ICALP)*. (Retrieved from prosiding.aripafi.or.id - specific DOI will depend on final publication).

El-Haj, W., Rayson, P., Aboelezz, M., El-Beltagy, S., & Gabr, M. (2022). Arabic language processing: State of the art, resources, and tools. *ACM Transactions on Asian and Low-Resource Language Information Processing, 21*(4), Article 83. https://doi.org/10.1145/3487895

Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kaziunas, E. (2018). Conceptualizing algorithmic auditing: How can we know if algorithms are working the way we want? *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 1–13. https://doi.org/10.1145/3173574.3174188

European Commission. (2021). *Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)* (COM(2021) 206 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52021PC0206

Fairclough, N. (2013). *Critical discourse analysis: The critical study of language* (2nd ed.). Routledge. https://doi.org/10.4324/9781315834368

Faisal, F., Ahia, O., Srivastava, A., Ahuja, K., Kann, K., & Palmer, A. (2024). *Dialectbench: A nlp benchmark for dialects, varieties, and closely-related languages*. arXiv preprint arXiv:2402.04309. https://arxiv.org/abs/2402.04309

Fanar Platform Team. (2025). *Fanar: An Arabic-centric multimodal generative AI platform*. arXiv preprint arXiv:2501.03043. https://arxiv.org/abs/2501.03043 (Note: Listing as corporate author as per previous discussion structure, though arXiv papers usually list individual authors).

Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972-1977* (C. Gordon, Ed.). Pantheon Books.

French Ministry of Culture. (2020). *Plan d'action pour la langue française et le plurilinguisme* [Action plan for the French language and multilingualism]. Ministère de la Culture.

Gazzola, M., Wickström, B.-A., & Spolsky, B. (Eds.). (2021). *The Routledge handbook of language policy and planning*. Routledge. https://doi.org/10.4324/9781003091706

Gibbs, G. R. (2007). *Analyzing qualitative data*. SAGE Publications.

Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford University Press.

Habash, N. (2010). *Introduction to Arabic natural language processing*. Morgan & Claypool Publishers. (A newer edition has DOI: https://doi.org/10.1007/978-3-031-02198-0)

Hosny, R., & Nasef, M. A. (2025). Lexical algorithmic resistance: Tactics of deceiving Arabic content moderation algorithms on Facebook. *Big Data & Society*, *12*(2), 20539517251318277.

Jasanoff, S. (2016). *The ethics of invention: Technology and the human future*. W. W. Norton & Company.

Jones, R. H., & Hafner, C. A. (2012). *Understanding digital literacies: A practical introduction*. Routledge. https://doi.org/10.4324/9780203095330

Khalfan, W., Lugo-Ocando, J., & Al-Najjar, A. (2024). Coding and capability-building in nonprofit digital-native news organizations in the Middle East and North Africa (MENA) and Latin America (LATAM). *Digital Journalism*.

Kincheloe, J. L., & McLaren, P. (2005). Rethinking critical theory and qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The SAGE handbook of qualitative research* (3rd ed., pp. 303–342). SAGE Publications.

Langmia, K. (2016). *Globalization and cyberculture: An Afrocentric perspective*. Palgrave Macmillan.

Lee, K.-F. (2018). *AI superpowers: China, Silicon Valley, and the new world order*. Houghton Mifflin Harcourt.

Mihalcea, R., Ignat, O., Bai, L., Borah, A., Brahman, F., Chen, P., ... & Zhao, J. (2025). Why AI is WEIRD and shouldn't be this way: Towards AI for everyone, with everyone, by

everyone. *Proceedings of the AAAI Conference on Artificial Intelligence*. (Full publication details and DOI will be available upon formal publication).

Mills, M., van de Bunt, G. G., & de Bruijn, J. (2006). Comparative research: Persistent problems and promising solutions. *International Sociology, 21*(5), 619–631. https://doi.org/10.1177/0268580906067833

Mohammed bin Rashid Al Maktoum Knowledge Foundation. (2021). *Arab Reading Index Report*.

Nacar, O., Sibaee, S. T., Ahmed, S., Elgohary, A., & Habash, N. (2025). Towards inclusive Arabic LLMs: A culturally aligned benchmark in Arabic large language model evaluation. *Proceedings of the LREC-COLING 2024 Workshop on Resources and Evaluation for Call Center AI (RECALL AI)*.

Noble, S. U. (2018). *Algorithms of oppression: How search engines reinforce racism*. NYU Press.

O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown Publishing Group.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*. Penguin Books.

Phillipson, R. (2003). *Linguistic imperialism continued.*

Routledge. https://doi.org/10.4324/9780203427666

Pimienta, D. (2009). Linguistic diversity in cyberspace. In A. C. d. Ó, M. P. d. A. Pereira, & M. I. T. Morado (Eds.), *Diversidad lingüística, políticas y tecnologías* [Linguistic diversity, policies and technologies]. Axac.

Pinch, T. J., & Bijker, W. E. (1984). The social construction of facts and artefacts: Or how the sociology of science and the sociology of technology might benefit each other. *Social Studies of Science, 14*(3), 399–441. https://doi.org/10.1177/030631284014003004

Prieur, C., & Savage, M. (2013). Emerging forms of cultural capital. *European Societies, 15*(2), 246–267. https://doi.org/10.1080/14616696.2012.748930

Sarieddine, T. (2023). Shades of red: Assessing China's hegemony in the Belt and Road Initiative. *Journal of World-Systems Research, 29*(2), 430–460. https://doi.org/10.5195/jwsr.2023.1193

Schwab, K. (2017). *The fourth industrial revolution*. Crown Business.

Shahid, F., Elswah, M., & Vashistha, A. (2025). *Think outside the data: Colonial biases and systemic issues in automated moderation pipelines for low-resource languages*. arXiv preprint arXiv:2501.13836. https://arxiv.org/abs/2501.13836

Shang, G., Abdine, H., Khoubrane, Y., Mohamed, A., Mourchid, Y., Al Kwatli, S., & Rust, S. (2024). *Atlas-Chat: Adapting large language models for low-resource Moroccan Arabic dialect*. arXiv preprint arXiv:2402.08583. https://arxiv.org/abs/2402.08583

Skutnabb-Kangas, T. (2000). *Linguistic genocide in education—or worldwide diversity and human rights?* Lawrence Erlbaum Associates.

Spolsky, B. (2004). *Language policy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511616161

Spolsky, B. (Ed.). (2012). *The Cambridge handbook of language policy*. Cambridge University Press. https://doi.org/10.1017/CBO9780511979026

State Council of China. (2017). *New generation artificial intelligence development plan*. The State Council of the People's Republic of China. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

Stelling, L., Yang, M., Gipiškis, R., Staufer, L., Wagner, B., & Ződi, Z. (2025). *Existing industry practice for the EU AI Act's general-purpose AI code of practice safety and security measures*. arXiv preprint arXiv:2502.02475. https://arxiv.org/abs/2502.02475

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Brooks/Cole.

Tibi, N., & Messaoud, M. A. B. (2025). Arabic dialect classification using an adaptive deep learning model. *Bulletin of Electrical Engineering and Informatics*, *14*(2), 1108-1116.

Trigui, H., Guerfali, F., Harigua-Souiai, E., Qasrawi, R., Hlaoua, A., & Elloumi, S. (2024). Exploring AI governance in the Middle East and North Africa (MENA) region: Gaps, efforts, and initiatives. *Data & Policy, 6*, e11. https://doi.org/10.1017/dap.2024.5

UAE National Program for Artificial Intelligence. (n.d.). *UAE Strategy for Artificial Intelligence 2031*.

UNESCO. (2005). *Convention on the Protection and Promotion of the Diversity of Cultural Expressions*. https://en.unesco.org/creativity/convention

UNESCO. (2021). *Recommendation on the Ethics of Artificial Intelligence* (SHS/BIO/PI/2021/1). https://unesdoc.unesco.org/ark:/48223/pf0000381137

United Nations Economic and Social Commission for Western Asia (ESCWA). (2022). *National AI Strategies in the Arab Region: A Comparative Analysis*.

van Dijck, J., Poell, T., & de Waal, M. (2018). *The platform society: Public values in a connective world*. Oxford University Press. https://doi.org/10.1093/oso/9780190889760.001.0001

Wardle, C., & Derakhshan, H. (2017). *Information Disorder: Toward an interdisciplinary framework for research and policymaking* (Report DGI(2017)09). Council of Europe.

Warschauer, M. (2003). *Technology and social inclusion: Rethinking the digital divide*. MIT Press.

Wheeler, D. L. (2012). Social media and the Arab Spring. *Brown Journal of World Affairs, 19*(1), 285–300.

Winner, L. (1980). Do artifacts have politics? *Daedalus, 109*(1), 121–136. http://www.jstor.org/stable/20024652

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

Zaghouani, W. (2021). A review of Arabic corpus linguistics. *Journal of King Saud University - Computer and Information Sciences, 33*(6), 681–694. https://doi.org/10.1016/j.jksuci.2019.09.001

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.