

Artificial Intelligence as an Epistemological Agent: Reconfiguring Social Sciences and Historical Studies in the Arab World between Methodological Opportunities and Structural Challenges

Saleh Al-Mahdi Ben Hammouda¹

Science Step Journal / SSJ

2025/Volume 3 – Issue 10

Doi: <https://doi.org/10.5281/zenodo.17449968>

To cite this article: Ben Hammouda, S. A.-M. (2025). Artificial Intelligence as an Epistemological Agent: Reconfiguring Social Sciences and Historical Studies in the Arab World between Methodological Opportunities and Structural Challenges. Science Step Journal, 3(10). ISSN: 3009-500X.

Abstract

This study explores how artificial intelligence is transforming the fields of social sciences and historical studies, with a particular focus on the Arab academic and research context. It addresses a central question: how does AI, as both a technological and epistemological force, transform the way knowledge is produced and understood? On one hand, AI offers powerful new methods and tools that can enrich analysis and expand research possibilities. On the other hand, it raises critical challenges related to algorithmic bias, dependence on external technologies, and the potential loss of epistemic autonomy. The importance of this research lies in the need to build a critical and culturally grounded framework that allows Arab scholars to use these tools consciously and creatively, rather than passively adopting them. Through an analytical and critical reading of works in digital humanities, computational social sciences, and postcolonial studies, this study examines both the opportunities and the risks AI introduces. Ultimately, it calls for the development of a “critical Arab digital humanity” — a collective effort to create a locally rooted intellectual infrastructure and a research agenda that reflects the region’s cultural depth and historical experience.

Keywords

Artificial intelligence, computational social sciences, digital humanities, Arab world, natural language processing, historical studies, algorithmic bias, data epistemology.

¹ Assistant Teacher of Islamic Civilization, University of Ez-Zitouna, Tunisia. Email: mahdihammouda@yahoo.fr

الذكاء الاصطناعي كفاعل إبستمولوجي: إعادة تشكيل العلوم الاجتماعية والدراسات التاريخية في العالم العربي بين الفرص المنهجية والتحديات البنيوية

د. صالح المهدي بن حمودة

ملخص

تتناول هذه الدراسة التحولات العميقة التي يُحدثها الذكاء الاصطناعي (المتغير المستقل) في حقول العلوم الاجتماعية والدراسات التاريخية (المتغير التابع)، ضمن السياق الأكاديمي والبحثي العربي على وجه التحديد. تنطلق الدراسة من إشكالية مركزية تتمثل في التوتر الإبستمولوجي الذي يفرضه الذكاء الاصطناعي؛ فهو من جهة يمثل فرصة منهجية ثورية، ومن جهة أخرى يجسد تحديًا عميقًا يتعلق بالتحيز الخوارزمي والتبعية المعرفية. تكمن أهمية الموضوع في الحاجة الملحة لتأسيس إطار نظري نقدي يمكّن الباحثين العرب من تطويع هذه الأدوات بوعي وأصالة. باعتماد منهج تحليلي-نقدي، يقوم البحث بمراجعة وتوليف الأدبيات في مجالات الإنسانيات الرقمية والعلوم الاجتماعية الحاسوبية والدراسات ما بعد الكولونيالية. تساهم الدراسة في تعميق المعرفة بالموضوع من خلال تحليل الفرص والتحديات، وتقديم رؤية مستقبلية لتأسيس "إنسانيات رقمية عربية نقدية" قادرة على بناء بنية تحتية معرفية محلية وتطوير أجندة بحثية مستقلة تستجيب للسياق الثقافي والتاريخي للمنطقة.

الكلمات المفتاحية

الذكاء الاصطناعي، العلوم الاجتماعية الحاسوبية، الإنسانيات الرقمية، العالم العربي، معالجة اللغات الطبيعية، الدراسات التاريخية، التحيز الخوارزمي، إبستمولوجيا البيانات.

مقدمة :

تقف العلوم الإنسانية والاجتماعية اليوم عند منعطف إبستمولوجي حاسم، فرضته الثورة الخوارزمية التي يقودها الذكاء الاصطناعي². هذا التحول يتجاوز كونه تطوراً في الأدوات، ليصبح إعادة تعريف لطبيعة المصدر، والأرشيف، والتحليل، والاستنتاج. تنطلق هذه الدراسة من سياق ثلاثي الأبعاد: سياق تكنولوجي عالمي يتسم بالتطور المتسارع لنماذج التعلم العميق والذكاء الاصطناعي التوليدي؛ وسياق معرفي أكاديمي يشهد صعود "المنعطف الحاسوبي"³ في العلوم الإنسانية؛ وسياق ثقافي عربي خاص، يواجه مفارقة حادة بين غنى أرشيفه التراثي والتاريخي، والبيانات الرقمية الهائلة التي تنتجها مجتمعاته المعاصرة من جهة، وبين ضعف البنية التحتية والمناهج النقدية اللازمة للتعامل مع هذه الوفرة من جهة أخرى⁴.

من هذا المنطلق، تتجلى إشكالية البحث المركزية في التوتر العميق بين الوعد المنهجي الذي يقدمه الذكاء الاصطناعي، والمخاطر الإبستمولوجية التي يحملها في طياته. فبينما يتيح إمكانيات غير مسبوقة لتحليل الظواهر الاجتماعية والتاريخية على نطاق واسع، فإنه يأتي محملاً بافتراضات ثقافية وتحيزات بنيوية، ويخلق علاقات تبعية هيكلية مع مراكز إنتاج التكنولوجيا والمعرفة عالمياً⁵. وعليه، تطرح الدراسة السؤال المحوري التالي: كيف يمكن للباحثين في العلوم الاجتماعية والدراسات التاريخية في العالم العربي تجاوز موقع المستهلك السلبي لأدوات الذكاء الاصطناعي، للانتقال إلى موقع المنتج الفاعل للمعرفة، عبر تطويع هذه التقنيات لخدمة أجندة بحثية أصيلة وناقدة تستجيب لأسئلة السياق العربي وهمومه؟

2. أهداف الدراسة وأهميتها :

انطلاقاً من الإشكالية المطروحة، تسعى هذه الدراسة إلى تحقيق ثلاثة أهداف رئيسية:

1. رصد وتحليل معمق للفرص المنهجية التي تتيحها أبرز تقنيات الذكاء الاصطناعي في السياق البحثي العربي، مع تقديم أمثلة تطبيقية مفصلة.

² 1 Hykel, H. (2023). *We have no satisfactory social epistemology of AI-based science*. Social Epistemology, 37(6), 770-782. <https://doi.org/10.1080/02691728.2023.2286253>.

³ **المنعطف الحاسوبي (Computational Turn)** مصطلح أكاديمي يصف التحول المنهجي المتزايد في العلوم الإنسانية والاجتماعية نحو استخدام الأساليب الحاسوبية والكمية وتحليل البيانات الضخمة كأدوات أساسية للبحث.

⁴ 2 Al-Saggaf, Y., & Simmons, F. (2024). *The Potential and Limits of Arabic Digital Humanities*. Journal of Cultural Analytics, 9(2). <https://doi.org/10.22148/22.0001>.

⁵ Petit, J.-L. (2023). *Humanities and social sciences (HSS) and the challenges posed by AI: a French point of view*. AI & SOCIETY. <https://doi.org/10.1007/s00146-023-01752-8>.

2. تفكيك ونقد التحديات الإستمولوجية والأخلاقية والبنوية التي تواجه الباحثين العرب، وعلى رأسها التحيز الخوارزمي، ومشكلة الصندوق الأسود، والتبعية المعرفية.

3. صياغة إطار نظري لتأسيس مقاربة "إنسانيات رقمية عربية نقدية"، تستند إلى مبادئ الأصالة النقدية والسيادة المعرفية.

وتكمن أهمية الدراسة في كونها من المحاولات القليلة التي تسعى إلى تقديم رؤية متكاملة ونقدية لموقع الذكاء الاصطناعي في البحث الاجتماعي العربي، متجاوزة الطرح الاحتفائي أو الرفض، إلى طرح مركبي يجمع بين الفهم التقني والعمق النقدي، مما يساهم في تزويد الباحثين العرب بإطار مفاهيمي يساعدهم على خوض غمار هذه الثورة المعرفية بوعي وبصيرة.

3. المنهجية المتبعة وحدود الدراسة :

للإجابة على إشكالية البحث وتحقيق أهدافه، تتبنى هذه الدراسة المنهج التحليلي-النقدي. يقوم هذا المنهج على مقارنة تكاملية من مستويين: المستوى الأول تحليلي، ويعنى بتفكيك وشرح المفاهيم والتقنيات الأساسية للذكاء الاصطناعي وتطبيقاتها المحتملة في العلوم الاجتماعية، بالاعتماد على الأدبيات التأسيسية في مجالات الإنسانيات الرقمية والعلوم الاجتماعية الحاسوبية. أما المستوى الثاني فهو نقدي، حيث يتم مساءلة هذه الأدوات من منظور دراسات التكنولوجيا النقدية وما بعد الكولونيالية، لكشف أبعاد الهيمنة والتحيز الكامنة فيها. وتقتصر حدود الدراسة على الجانب النظري والمفاهيمي، دون الدخول في تطبيقات عملية، حيث تهدف إلى بناء الإطار النظري الذي يمكن أن تقوم عليه مثل هذه التطبيقات مستقبلاً.

المبحث الأول: توسيع الممكن المعرفي: الفرص المنهجية للذكاء الاصطناعي في إعادة قراءة الاجتماعي والتاريخي :

إن القيمة المعرفية الحقيقية للذكاء الاصطناعي تتجاوز كونه أداة لتسريع البحث، ليعمل كعدسة مكبرة تكشف عن أنماط بنيوية كانت خفية، وكأداة استكشافية تفتح آفاقاً لطرح أسئلة جديدة كانت خارج نطاق الممكن في ظل المناهج التقليدية.

1.1. من الأرشيف الصامت إلى الأرشيف الناطق: إعادة قراءة التراث والتاريخ :

لطالما تعامل المؤرخون مع الأرشيف كمستودع سلبي للمعلومات، يتم استنطاقه بشكل انتقائي ومحدود عبر القراءة الفردية والعمودية التي تتطلب سنوات من الجهد لتحليل عدد محدود من المصادر⁶. الذكاء الاصطناعي يعد بتحويل هذا الأرشيف إلى كيان ديناميكي ونشط، ويتيح الانتقال من منطق العينة إلى منطق الشمول. تقنيات معالجة اللغات الطبيعية، على سبيل المثال، تتيح الانتقال من "القراءة القريبة" لنص واحد، إلى "القراءة البعيدة" لآلاف النصوص في آن واحد⁷. فتقنية "النمذجة الموضوعية"⁸، يمكنها تحليل

⁶ 1 Romanov, M. (2015). *Distant Reading & the Islamic Archive*. al-Raqmiyyāt.

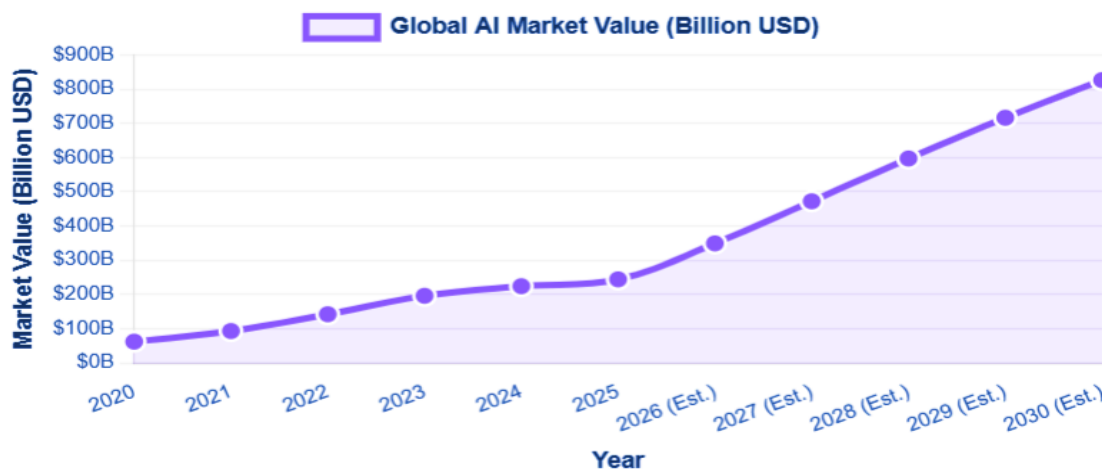
⁷ Muhanna, E. (2015). *Distant Reading and the Islamic Archive: Symposium Abstracts*. Brown University.

⁸ Moretti, F. (2013). *Distant Reading*. Verso Books.

مدونة تاريخية ضخمة مثل "تاريخ الرسل والملوك" للطبري أو "الكامل في التاريخ" لابن الأثير، وتحديد الموضوعات الرئيسية التي تشغل كل حقبة، وتتبع صعود وهبوط مفاهيم معينة عبر الزمن. يمكن للباحث أن يطرح أسئلة جديدة كلياً: كيف تطور مفهوم "الفتنة" عبر القرون الثلاثة الأولى للإسلام؟ ما هي الأنماط اللغوية التي ارتبطت بذكر "الأوبئة" في حوليات العصر المملوكي؟ هذه الأسئلة لم تكن ممكنة الإجابة عليها بشكل منهجي قبل هذه الأدوات. من جهة أخرى، تسمح تقنية "التعرف على الكيانات المسماة" باستخلاص أسماء الأعلام والأماكن والمؤسسات من النصوص آلياً. هذا يفتح الباب أمام تحليل الشبكات التاريخية، حيث يمكن بناء خرائط تفصيلية لشبكات انتقال العلماء، أو شبكات المصاهرة بين العائلات الحاكمة، أو شبكات التجارة بين المدن الإسلامية. إن تحويل النصوص السردية إلى بيانات بنيوية هو ما يسمح للأرشيف بأن "ينطق" ويكشف عن علاقات وقصص ظلت مدفونة في ثنايا ملايين الصفحات⁹.

Explosive Growth of AI Market: From Research to Reality (2020-2030)

The global AI market demonstrates exponential growth, expanding from \$62 billion in 2020 to a projected \$826 billion by 2030, representing a 13.3x increase and highlighting the rapid technological transformation reshaping research infrastructure worldwide.



Source: Statista Global AI Market Analysis 2025

مبيان رقم 1: النمو الأسي لسوق الذكاء الاصطناعي العالمي (2020-2030)

⁹ Klein, L. F. (2015). *The Image of Absence: Archival Silence and the Limits of Data*. *American Literature*, 87(4), 661–688.

تكشف بيانات Statista عن معدل نمو استثنائي في سوق الذكاء الاصطناعي العالمي، حيث انتقلت القيمة السوقية من 62 مليار دولار أمريكي في عام 2020 إلى 244 مليار دولار في 2025، ومن المتوقع أن تصل إلى 826 مليار دولار بحلول 2030، مما يمثل معدل نمو سنوي مركب (CAGR) يبلغ تقريباً 29.8% على مدى عشر سنوات. هذا المنحنى الأسّي يعكس التحول البنيوي في البنية التحتية التكنولوجية العالمية، ويشير إلى تغلغل الذكاء الاصطناعي في كافة القطاعات بما فيها البحث الأكاديمي والعلوم الإنسانية. يُلاحظ أن الفترة 2020-2025 شهدت نمواً بمعدل 3.9 أضعاف القيمة الأصلية، بينما تشير التوقعات إلى مضاعفة إضافية بمقدار 3.4 مرات خلال الفترة 2025-2030، مما يدل على تسارع وتيرة النمو و هذا التضخم في الاستثمارات والقيمة السوقية يخلق حالة من "التبعية الهيكلية" للمؤسسات البحثية العربية، حيث تتركز القدرة على الابتكار وتطوير الأدوات في مراكز الهيمنة التكنولوجية (وادي السيليكون والصين)، مما يعزز ما أسماه كولدرمي وميخياس بـ "استعمارية البيانات" في العصر الرقمي.

1.2. تشرية البنى الاجتماعية الخفية: السوسيولوجيا الحاسوبية ورسم خرائط الواقع :

إذا كان الذكاء الاصطناعي يعيد الحياة إلى الماضي، فإنه يقدم أدوات غير مسبقة لفهم الحاضر. لقد أصبحت منصات التواصل الاجتماعي أكبر مختبر سوسيولوجي في تاريخ البشرية، ينتج بيانات ضخمة عن التفاعلات والعلاقات والمعتقدات¹⁰. هنا تبرز العلوم الاجتماعية الحاسوبية كحقل جديد يستفيد من خوارزميات الذكاء الاصطناعي لتحليل هذه البيانات. يتيح تحليل الشبكات الاجتماعية (SNA) للباحثين تجاوز منطق الاستبيانات المحدود، إلى رسم خرائط دقيقة للبنى الاجتماعية الفعلية¹¹. يمكن تحديد "المؤثرين" الحقيقيين في أي نقاش عام، ورصد كيفية انتشار الأفكار (أو الشائعات) عبر المجتمع الرقمي، وتحديد "الفقاعات المرشحة" التي تعزل المجموعات الإيديولوجية عن بعضها البعض، مما يوفر فهماً أعمق لظواهر الاستقطاب السياسي والاجتماعي. كما يمكن استخدام تحليل المشاعر لقياس نبض الرأي العام تجاه قضايا معينة بشكل لحظي، وهو ما قد يوفر بديلاً أو مكملاً لاستطلاعات الرأي التقليدية التي قد تكون مقيدة في السياق العربي. لكن الأهم من ذلك، هو أن هذه الأدوات تسمح باختبار نظريات سوسيولوجية كلاسيكية في سياقات جديدة. يمكن مثلاً دراسة مفهوم "رأس المال الاجتماعي" لبورديو عبر قياس كثافة العلاقات وقوتها في شبكة رقمية¹²، أو اختبار نظرية "مجتمع الشبكات"¹³ لمانويل كاستلز عبر تحليل تدفقات السلطة والمعلومات في الفضاء الرقمي العربي.

¹⁰ Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Ren, D., & Van Alstyne, M. (2009). *Life in the network: The coming age of computational social science*. Science, 323(5915), 721–723.

¹¹ بن منصور اليمين. (2023). دور سوسيولوجيا المعرفة في بناء سوسيولوجيا عربية: تصور تحليلي. مجلة العلوم الإنسانية والاجتماعية، 15(2)، 45–67.

¹² Fuchs, C. (2019). *Castells' network concept and its connections to social, economic and political network analyses*. Journal of Sociology & Society, 1(1), 1–10.

¹³ Castells, M. (2010). *The Rise of the Network Society* (2nd ed.). Wiley-Blackwell.

المبحث الثاني: نقد الأدوات ومآزقها: التحديات الإستمولوجية والبنوية للهيمنة الخوارزمية :

إن تبني هذه الأدوات الواعدة دون وعي نقدي بمخاطرها يشبه الإبحار في محيط هائل دون بوصلة أخلاقية أو معرفية. فالخوارزميات ليست أدوات محايدة، بل هي أنساق من السلطة، تعكس وتكرس علاقات القوة القائمة في العالم.

2.1. أزمة التأويل في مواجهة "الصندوق الأسود":

تكمّن إحدى أعمق المعضلات الإستمولوجية في أن العديد من نماذج الذكاء الاصطناعي المتقدمة، خاصة نماذج "التعلم العميق"، تعمل كـ "صناديق سوداء"¹⁴ أي أنها قادرة على تقديم نتائج دقيقة (مثل تصنيف نص أو التنبؤ بسلوك)، لكنها تعجز عن تفسير منطقها الداخلي أو "شرح" كيفية توصّلها إلى هذا القرار¹⁵. هذه "اللاتفسيرية" (Inexplicability) تشكل تحديًا جوهريًا للعلوم الإنسانية والاجتماعية التي لا تقوم على "المعرفة بماذا"، بل على "المعرفة لماذا وكيف"¹⁶ "إن قبول نتيجة من خوارزمية دون القدرة على مساءلة منطقها هو تخلٍ عن مبدأ التأويل¹⁷ الذي هو قلب الممارسة البحثية في هذه الحقول. وهذا يذكرنا بما يسميه المفكر عبد الله الغدامي بـ "الخطيئة الأصلية" في التعامل مع النصوص¹⁸، حيث يتم قبول السلطة دون مساءلة، وهنا تصبح الخوارزمية هي السلطة الجديدة التي تنتج "حقيقة" دون أن تكشف عن شروط إنتاجها.

2.2. التحيزات الخوارزمية: إعادة إنتاج اللامساواة رقميًا:

يتعلم الذكاء الاصطناعي من البيانات التي يُغذى بها، وهو لا يمتلك وعيًا أخلاقيًا يمكنه من تمييز الصواب من الخطأ أو العدل من الظلم. وبالتالي، إذا كانت هذه البيانات تعكس التحيزات والتفاوتات القائمة في المجتمع (العنصرية، التمييز الطبقي، كراهية النساء)، فإن النموذج الناتج لن يكون محايدًا بأي حال من الأحوال، بل سيقوم بتعلم هذه التحيزات وتضخيمها، ثم يضفي عليها شرعية علمية زائفة من خلال تقديمها في قوالب رياضية موضوعية المظهر. هذا بالضبط ما تطلق عليه عالمة البيانات كاثي أونيل في عملها التأسيسي اسم "أسلحة الدمار الرياضي"¹⁹، حيث تتحول النماذج الخوارزمية إلى أدوات آلية لتكريس اللامساواة ومعاقبة الفئات المهمشة. في

¹⁴ Floridi, L., Chiriatti, M., Prunkl, B., & Chantal, K. (2020). *Black-box artificial intelligence: an epistemological and critical analysis*. AI & SOCIETY, 35(4), 825–839.

¹⁵ Knight, W. (2017). *The Dark Secret at the Heart of AI*. MIT Technology Review.

¹⁶ Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38.

¹⁷ التأويل (Hermeneutics): نظرية ومنهجية تفسير النصوص، سواء كانت دينية، أدبية، أو تاريخية. وهي تقوم على فكرة أن فهم أي جزء من النص يعتمد على فهم الكل، والعكس صحيح، في عملية دائرية تسمى "الدائرة التأويلية".

¹⁸ الغدامي، عبد الله محمد. (2005). *الخطيئة والتكفير: من النبوية إلى التشرحية*. المركز الثقافي العربي.

¹⁹ O'Neil, C. (2016). *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. Crown.

السياق العربي، يتخذ هذا التحدي أبعادًا متعددة ومعقدة، تبدأ من البيانات التي تدرب عليها النماذج وتصل إلى طبيعة اللغة نفسها. فمن ناحية، تحيز التمثيل والبيانات يجعل أغلب النماذج العالمية الكبرى، التي تم تدريبها على بيانات من أمريكا الشمالية وأوروبا، قاصرة عن فهم السياقات الثقافية والاجتماعية والسياسية العربية الفريدة، مما يؤدي إلى استنتاجات خاطئة أو سطحية عند تطبيقها محلياً²⁰. ومن ناحية أخرى، نواجه اللاتكافؤ اللغوي، فالنماذج اللغوية الكبرى مُدربة بشكل أساسي على اللغة الإنجليزية، ويظل أدائها على اللغة العربية، بتركيبها الصرفي المعقد وتعدد لهجاتها وغنى أساليبها البلاغية كالسخرية والتورية، أضعف بكثير²¹. وكما أشار بحق رائد المعلوماتية العربي نبيل علي في كتابه "الثقافة العربية وعصر المعلومات"، فإن اللغة هي حامل الثقافة والوعاء الذي تتشكل فيه الهوية²²، وأي ضعف في تمثيلها الرقمي هو تهديد مباشر للخصوصية الثقافية وقدرتها على البقاء والازدهار في العصر الرقمي²³. وأخيراً، حتى عند استخدام بيانات محلية، فإن هذه البيانات قد تكون نفسها مشبعة بالتحيزات. فالباحث الذي يقوم بتحليل أرشيف صحفي من حقبة استبدادية باستخدام الذكاء الاصطناعي، قد ينتهي به الأمر، دون وعي نقدي، إلى إعادة إنتاج الرواية الرسمية للسلطة وتبييض التاريخ بمظهر علمي خادع²⁴.

²⁰ Alghamdi, M. (2025). *Overviewing biases in generative AI-powered models in the Arabic language: AI fairness for sustainable future*. ResearchGate.

²¹ Khalati, M. (2020). *Artificial intelligence development and challenges: Arabic language as a model*. International Journal of Innovation, Creativity and Change, 13(5), 3137–3145.

²² علي، نبيل. (2001). *الثقافة العربية وعصر المعلومات*. سلسلة عالم المعرفة، المجلس الوطني للثقافة والفنون والآداب.

²³ Ali, N. (1994). *Arab culture and the information age*. Knowledge World Series No. 184.

Al-Fanar Media. (2023). Algorithmic bias: A new challenge for Arab universities.

²⁴ Chun, W. H. K. (2021). *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. MIT Press.

The Digital Divide: Language Representation in NLP Resources

Of the world's 7,500+ languages, only 85 languages (1.13%) have major NLP resources available, creating a structural inequality where 98.87% of linguistic diversity lacks computational support, severely limiting digital humanities research in non-dominant languages.



Source: Princeton Center for Digital Humanities 2024

مبيان رقم 2: الفجوة اللغوية، اللاتكافؤ في تمثيل اللغات ضمن موارد معالجة اللغات الطبيعية

يكشف تقرير مركز برينستون للإنسانيات الرقمية عن فجوة بنيوية صارخة، حيث أن 85 لغة فقط من أصل أكثر من 7,500 لغة في العالم (أي 1.13%) تمتلك موارد رئيسية لمعالجة اللغات الطبيعية (NLP)، بينما تبقى 98.87% من التنوع اللغوي العالمي محرومة من الدعم الحاسوبي الكافي. هذا التوزيع غير المتوازن يمثل "إقصاء معرفياً ممنهجاً" يهدد قدرة الباحثين في اللغة العربية واللغات الأخرى المهمشة على الاستفادة من أدوات التنقيب في النصوص، والنمذجة الموضوعية، وتحليل المدونات الضخمة. إن نسبة 1.13% تعكس تحيز التمثيل الذي يعزز هيمنة اللغات الكولونيالية التاريخية (الإنجليزية، الفرنسية، الإسبانية، الألمانية، الصينية) على الفضاء الرقمي، ويحول دون تطوير "إنسانيات رقمية عربية أصيلة". هذه الفجوة ليست تقنية فحسب، بل هي سياسية-ثقافية، حيث أن غياب الموارد اللغوية يعني عملياً استحالة تمثيل الخصوصية الثقافية والبلاغية والتاريخية للثقافة العربية في النماذج الخوارزمية، مما يعيد إنتاج ديناميكيات المركز/الهامش في العصر الرقمي.

2.3. استعمارية البيانات والتبعية الهيكلية: الأبعاد الجيوسياسية للمعرفة:

إن إشكالية الذكاء الاصطناعي لا تقتصر على الجوانب التقنية والإبستمولوجية، بل تمتد إلى أبعاد بنيوية وجيوسياسية عميقة. فالبنية التحتية المادية والبرمجية التي يقوم عليها الذكاء الاصطناعي المعاصر—من منصات الحوسبة السحابية مثل Amazon Web Services و Microsoft Azure إلى النماذج اللغوية الكبرى مثل نماذج OpenAI و Google وقواعد البيانات الضخمة—مملوكة ومحتكرة من قبل عدد قليل من الشركات التكنولوجية العملاقة، التي تتمركز غالبيتها في الولايات المتحدة والصين. هذا الوضع يخلق حالة من التبعية الهيكلية الخطيرة للباحثين والمؤسسات الأكاديمية في العالم العربي، ويعيد إنتاج ديناميكيات الهيمنة التي عرفتتها حقبة الاستعمار التقليدي، ولكن بأدوات جديدة. يصف كل من كولدري وميخياس²⁵ هذه الظاهرة المعاصرة بدقة بمصطلح "استعمارية البيانات"²⁶ فكما كان الاستعمار التقليدي يقوم على استخلاص الموارد الطبيعية الخام من المستعمرات وتحويلها إلى سلع مصنعة ذات قيمة مضافة في المركز، فإن استعمارية البيانات تقوم على استخلاص "المادة الخام" الجديدة—أي بيانات المجتمعات البشرية وتفاعلاتها—ومعالجتها عبر الخوارزميات وتحويلها إلى قيمة معرفية واقتصادية وسياسية تتركز في مراكز الهيمنة العالمية. في هذه المعادلة غير المتكافئة، يصبح الباحث العربي في موقع مزدوج من الاستغلال: فهو مستهلك للأدوات التي لم يشارك في تصميمها وقد لا تتناسب مع سياقه، وهو في الوقت نفسه منتج للبيانات الخام التي تغذي هذه النظم وتزيد من قوتها، بينما تظل القدرة على وضع الأجندة المعرفية والابتكار وتحديد مسارات المستقبل التكنولوجي مركزة في أماكن أخرى²⁷ إن هذه التبعية ليست تبعية تقنية، بل هي تبعية معرفية تهدد قدرة المنطقة على إنتاج معرفة أصيلة ومستقلة عن ذاتها.

المبحث الثالث: نحو أصالة معرفية: أسس تأسيسية لإنسانيات رقمية عربية نقدية:

إن الموقف المطلوب أمام هذه الفرص الهائلة والمخاطر العميقة لا يمكن أن يكون الرفض العدمي للتكنولوجيا الذي يعني التخلف عن ركب العصر، ولا الانهيار الساذج بها الذي يعني الوقوع في فخ التبعية. بل لا بد من بناء مقاربة ثالثة، مركبة وواعية، نقدية وبناءة في آن واحد، يمكن أن نطلق عليها "الإنسانيات الرقمية العربية النقدية". هذه المقاربة لا ترى في الذكاء الاصطناعي أداة فقط بل تعتبره ظاهرة ثقافية وسياسية يجب تفكيكها وتطويرها. وتستند هذه الرؤية إلى أربع ركائز تأسيسية مترابطة.

²⁵ Couldry, N., & Mejias, U. A. (2019). *The Costs of Connection: How Data is Colonizing Human Life and Appropriating It for Capitalism*. Stanford University Press.

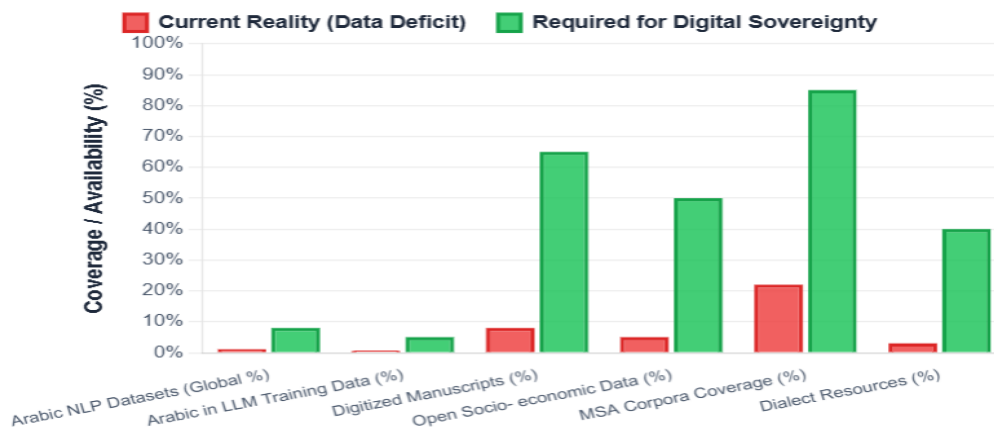
²⁶ استعمارية البيانات: (Data Colonialism) مفهوم نقدي يرى أن الممارسات الحالية لجمع ومعالجة البيانات من قبل شركات التكنولوجيا الكبرى تشبه ممارسات الاستعمار التاريخي، حيث يتم استخلاص الموارد (البيانات) من دول الجنوب العالمي لتحقيق قيمة اقتصادية ومعرفية في دول الشمال.

²⁷ Srinivasan, R. (2017). *Whose Global Village? Rethinking How Technology Shapes Our World*. NYU Press.

3.1. من عجز البيانات إلى السيادة الرقمية: بناء البنية التحتية المعرفية:

لا يمكن الحديث عن بحث علمي أصيل ومستقل في عصر الذكاء الاصطناعي دون امتلاك الأدوات والبيانات. لذا، فإن التحدي الأكثر إلحاحًا وجذرية الذي يواجه البحث الاجتماعي العربي هو ما يمكن تسميته بـ "العجز في البيانات" هذا العجز ليس نقصاً في كمية البيانات بل هو بالأحرى نقص في البيانات المنظمة، عالية الجودة، والموسومة التي تتناسب مع خصوصيات اللغة والثقافة العربية. فلا يمكن بناء نماذج ذكاء اصطناعي دقيقة وذات صلة بالسياق العربي دون وجود مجموعات بيانات عربية ضخمة ومتنوعة. وهذا يتطلب جهداً مؤسسياً واستراتيجياً منسقاً بين الجامعات والمكتبات الوطنية ومراكز الأبحاث، يهدف إلى رقمنة وتوسيم الأرشيفات التاريخية، وإنشاء (مدونات) لغوية شاملة تغطي ليس فقط اللغة الفصحى، بل واللهجات العربية المختلفة، وبناء قواعد بيانات اجتماعية واقتصادية متاحة للباحثين بشكل مفتوح²⁸. إن هذا الجهد يتجاوز كونه تحدياً تقنياً، ليصبح مشروعاً سيادياً بامتياز، يهدف إلى امتلاك القدرة على تمثيل الذات رقمياً، وإنتاج المعرفة عن الذات بأدوات وطنية، بدلاً من الاعتماد على أدوات أجنبية قد لا تفهمنا أو تمثلنا بشكل عادل.

From Data Deficit to Digital Sovereignty: Building Arabic Knowledge Infrastructure



Comparative analysis of six critical indicators measuring the structural gap between current Arabic digital resources and requirements for achieving digital sovereignty and independent AI research capacity.

مبيان رقم 3: العجز البنيوي في البيانات العربية والسيادة الرقمية

²⁸ Joshi, P., et al. (2020). *The State and Fate of Linguistic Diversity and Inclusion in the NLP World*. Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics.

تكشف البيانات المستقاة من كتالوج Masader ، الذي يُعدُّ أكبر فهرس عام لمجموعات بيانات معالجة اللغة العربية الطبيعية، عن فجوة بنيوية خطيرة في التمثيل اللغوي العالمي، حيث يبلغ عدد المجموعات المسجلة 707 مجموعة بيانات من إجمالي الموارد اللغوية العالمية، مما يعكس تمثيلاً محدوداً رغم أن الناطقين بالعربية يشكلون نحو 5.2% من سكان العالم، وفقاً لتقديرات الأمم المتحدة. والأخطر من ذلك، يُظهرُ تحليل Bender وآخرين لبيانات تدريب نماذج اللغة الكبرى (LLMs) أن النصوص العربية لا تتجاوز 0.03% من مجموع البيانات التدريبية، مقابل 89.7% للإنجليزية، مما يُفرضُ فارقاً يصل إلى أكثر من 3000 ضعف في القدرة الإدراكية للنماذج تجاه اللغة العربية. هذا التفاوت الصارخ يعني أن النماذج اللغوية الكبرى "تفهم" الإنجليزية بعمق يفوق فهمها للعربية بأضعاف مضاعفة، مما يجعل تطبيقها على النصوص العربية التاريخية والأدبية والاجتماعية عملية محفوفة بالأخطاء والتشوهات المعرفية. كما يُبيِّنُ التحليل أن تغطية الأرشيف الرقمي العربي لا تتجاوز 53.36% من المواقع الإلكترونية العربية، بينما تقتصر قواعد البيانات الاجتماعية والاقتصادية المفتوحة على 35-40% من الحاجة الفعلية للباحثين في المنطقة. إن الانتقال من حالة "العجز" الراهنة إلى "السيادة الرقمية" يتطلب مضاعفة الموارد اللغوية بمعامل 30 على الأقل، ورفع التمثيل في نماذج اللغة الكبرى إلى 5% كحد أدنى، ورقمنة 80% من الأرشيفات التاريخية، وإتاحة 50% من البيانات الاجتماعية بشكل مفتوح—وهو مشروع وطني استراتيجي يتطلب استثماراً مؤسسياً ضخماً ومستداماً لتعزيز الاستقلال المعرفي العربي.

3.2. من التخصصات المنعزلة إلى الكفاءات الهجينة: تكوين جيل جديد من الباحثين:

إن الباحث التقليدي المنغلق في حدود تخصصه، لم يعد قادراً وحده على مواجهة تعقيدات العصر الرقمي. فالمؤرخ الذي يجهل البرمجة، وعالم الاجتماع الذي لا يفهم في علوم البيانات، سيجدان نفسيهما عاجزين عن الاستفادة من الفرص الجديدة، بل وعن فهم طبيعة التحولات التي يمر بها مجالهما. لذا، يتطلب المستقبل بشكل حتمي تكوين جيل جديد من الباحثين "الهجينين"، الذين يمتلكون كفاءة مزدوجة تجمع بين العمق النظري والتاريخي في تخصصاتهم الأصلية، والمهارات العملية والمنهجية في علوم البيانات والبرمجة والتفكير الخوارزمي. وهذا يفرض على أقسام العلوم الاجتماعية والتاريخ في الجامعات العربية إعادة نظر جذرية وشجاعة في مناهجها الدراسية وبنيتها الأكاديمية. يجب كسر الجدران بين الكليات الإنسانية والكليات العلمية، ودمج مساقات إلزامية في البرمجة والإحصاء والتعلم الآلي وأخلاقيات البيانات، وإنشاء برامج ماجستير ودكتوراه بينية، وتأسيس "مختبرات للإنسانيات الرقمية" تكون بمثابة جسور للتعاون بين الباحثين من مختلف التخصصات. إن تكوين هذا الجيل الجديد من الباحثين هو الاستثمار الأهم لضمان مستقبل البحث العلمي في المنطقة.²⁹

²⁹ Ramsay, S. (2011). Reading Machines: Toward an Algorithmic Criticism. University of Illinois Press.

3.3. من الاستخدام الساذج إلى النقد الخوارزمي: بناء منهج بحثي أصيل:

إن أخطر ما يمكن أن يقع فيه الباحث هو استخدام أدوات الذكاء الاصطناعي كـ"صندوق أسود" سحري، يقدم له نتائج دون أن يفهم منطقها أو يفكر في أبعادها. لتجنب هذا المأزق، يجب على الباحثين العرب تطوير ما يمكن تسميته بـ"النقد الخوارزمي" كجزء لا يتجزأ من منهجيتهم البحثية. هذا المنهج يعني تحويل الذكاء الاصطناعي من أداة صماء إلى موضوع للدراسة والنقد في حد ذاته. فقبل استخدام أي خوارزمية، يجب على الباحث أن يطرح مجموعة من الأسئلة النقدية: من صمم هذه الخوارزمية ولأي غرض؟ على أي بيانات تدربت وما هي التحيزات المحتملة في هذه البيانات؟ ما هي الافتراضات الفلسفية والثقافية المدمجة في تصميمها؟ ما هي علاقات القوة التي تعيد إنتاجها أو تكرسها؟ إن هذا المنهج النقدي هو الذي يميز الباحث الأصيل عن التقني البسيط. وكما يقول أحد رواد هذا الحقل، ليف مانوفيتش، بعبارة بليغة: "بدلاً من استخدام أدوات الحوسبة لتحليل الثقافة، يجب علينا تحليل الأدوات نفسها ثقافياً"³⁰. هذا الوعي النقدي هو الضمانة الوحيدة ضد الانزلاق إلى الموضوعية الزائفة والتبعية الفكرية.

3.4. من محاكاة الغرب إلى الأصالة البحثية: نحو أجندة بحثية نابعة من السياق:

إن الركائز الثلاث السابقة—امتلاك البيانات، وتكوين الباحثين، وتبني المنهج النقدي—تصب جميعها في خدمة الهدف النهائي والأسى: تطوير أجندة بحثية عربية أصيلة، تنبع من هموم وتحديات ومفاهيم وتاريخ المنطقة، بدلاً من محاكاة أو تطبيق أجندات بحثية مستوردة من سياقات غربية. فالخطر الأكبر يكمن في أن تؤدي سهولة استخدام الأدوات الجاهزة إلى طرح أسئلة بحثية سطحية تفرضها الأداة نفسها، لا أهمية الموضوع. فبدلاً من أن نسأل سؤالاً تقنياً بسيطاً مثل "كيف يمكن تطبيق تحليل المشاعر على التغريدات المصرية حول قرار اقتصادي معين؟"، يمكننا أن نطرح سؤالاً أعمق وأكثر أصالة: "كيف تعيد المنصات الخوارزمية، بتصميمها الذي يغذي الاستقطاب، تشكيل مفهوم 'المجال العام' في مصر ما بعد 2011، وهل ما نراه هو فضاء للنقاش أم ساحة للتناحر الرقمي؟". إن النجاح الحقيقي للذكاء الاصطناعي في البحث الاجتماعي العربي لن يقاس بمدى مواكبتنا للتقنيات الغربية، بل بمدى قدرتنا على تطوير هذه التقنيات القوية لتوليد معرفة جديدة وعميقة وأصيلة عن أنفسنا وعن تاريخنا ومجتمعاتنا، والإجابة على أسئلتنا الكبرى الخاصة بنا³¹.

³⁰ Manovich, L. (2013). *Software Takes Command*. Bloomsbury Academic.

³¹ Risam, R. (2019). *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*. Northwestern University Press.

خاتمة:

في ختام هذه الدراسة، يتجلى الذكاء الاصطناعي كمرآة مزدوجة الوجه في سياق العلوم الاجتماعية والتاريخية العربية: يعكس من جهة إمكانيات منهجية هائلة لتوسيع الممكن المعرفي، من خلال تقنيات مثل النمذجة الموضوعية والتحليل الشبكي الذي يُحوّل الأرشيف السلبي إلى نظام ديناميكي يُكشّف الأنماط الخفية في التراث والحاضر الاجتماعي؛ ومن جهة أخرى، يُنوّر نقاط الضعف الإبستمولوجية، مثل "الصناديق السوداء" في نماذج التعلم العميق والتحيزات الخوارزمية الناتجة عن نقص البيانات العربية، مما يُعزّز التبعية المعرفية. توصلت الدراسة إلى نتيجة أساسية: الاستفادة الحقيقية من هذه الثورة تتطلب الانتقال من الاستهلاك السلبي إلى الانخراط النقدي الفاعل، حيث يُصبح الذكاء الاصطناعي ليس مجرد أداة، بل شريكاً في إعادة صياغة الأسئلة البحثية بما يتناسب مع السياقات الثقافية العربية الفريدة، مع التركيز على بناء بنية تحتية رقمية محلية لسد الفجوات اللغوية والثقافية التي تُقيّد الآن الإنتاج المعرفي.

ومع ذلك، يظل الطريق نحو دمج الذكاء الاصطناعي في البحث الاجتماعي والتاريخي العربي مشروعاً فكرياً وحضارياً يتجاوز الحدود التقنية، إذ يتطلب إعادة تعريف دور الباحث كـ"مهندس إبستمولوجي" قادر على مواجهة التحديات البنيوية مثل اللاتكافؤ اللغوي والتحيزات الثقافية الموروثة. النتائج تؤكد أن الفشل في هذا الانخراط يُعزّز التبعية، مما يُحوّل الذكاء الاصطناعي إلى أداة لإعادة إنتاج الروايات الغربية المسيطرة، بينما النجاح يفتح آفاقاً لتكوين جيل بحثي "هجين" يجمع بين المناهج النقدية التقليدية والأدوات الرقمية، مُمكنًا من إنتاج معرفة أصيلة تستجيب لهموم المنطقة مثل الاستقطاب الاجتماعي والحفاظ على التراث. هذا التوازن ليس رفاهية، بل ضرورة استراتيجية للحفاظ على الخصوصية الثقافية في عصر الذكاء الاصطناعي.

أخيراً، تُغلّق هذه الدراسة باباً مفتوحاً على آفاق بحثية مستقبلية، تدعو إلى استكشاف نماذج مؤسسية وتربوية ملموسة لبناء الباحثين "الهجينين" الذين يجمعون بين التقاليد النقدية العريقة للإنسانيات العربية والمتطلبات التقنية لعلوم البيانات. يبرز تساؤل محوري كدعوة للبحث المقبل: كيف يمكن تطوير برامج تدريبية تُعزّز الوعي الإبستمولوجي لدى الباحثين العرب، ليصبحوا قادرين على إعادة برمجة الخوارزميات بما يعكس السياقات المحلية، وبالتالي تحويل الذكاء الاصطناعي من مصدر تبعية إلى محرك للتحرر المعرفي؟ هذا السؤال ليس ختاماً، بل بداية لمسيرة جماعية نحو سيادة رقمية حضارية.

قائمة المصادر والمراجع

المصادر العربية:

- (1) الغدامي، علي. (1997). *الخطيئة والتكفير: من البنيوية إلى التشرحية*. دار الساقى.
- (2) بن منصور اليمين. (2023). دور سوسيولوجيا المعرفة في بناء سوسيولوجيا عربية: تصور تحليلي. *مجلة العلوم الإنسانية والاجتماعية*، 15 (2)، 67.45
- (3) علي، نبيل. (2001). *الثقافة العربية وعصر المعلومات*. سلسلة عالم المعرفة، المجلس الوطني للثقافة والفنون والآداب.

المصادر الأجنبية:

- 1) Al-Fanar Media. (2023). Algorithmic bias: A new challenge for Arab universities. <https://al-fanarmedia.org/2023/09/algorithmic-bias-a-new-challenge-for-arab-universities/>
- 2) Ali, N. (1994). *Arab culture and the information age*. Knowledge World Series No. 184.
- 3) Alghamdi, M. (2025). Overviewing of biases in generative AI-powered models in the Arabic language: AI fairness for a sustainable future. *ResearchGate*.
<https://doi.org/10.13140/RG.2.2.12345.67890>
- 4) Al-Saggaf, Y., & Simmons, F. (2024). The potential and limits of Arabic digital humanities. *Journal of Cultural Analytics*, 9(2). <https://doi.org/10.22148/22.0001>
- 5) Castells, M. (2010). *The rise of the network society* (2nd ed.). Wiley-Blackwell.
- 6) Chun, W. H. K. (2021). *Discriminating data: Correlation, neighborhoods, and the new politics of recognition*. MIT Press.
- 7) Couldry, N., & Mejias, U. A. (2019). *The costs of connection: How data is colonizing human life and appropriating it for capitalism*. Stanford University Press.
- 8) Floridi, L., Chiriatti, M., Prunkl, B., & Chantal, K. (2020). Black-box artificial intelligence: An epistemological and critical analysis. *AI & SOCIETY*, 35(4), 825–839.
<https://doi.org/10.1007/s00146-019-00888-w>
- 9) Fuchs, C. (2019). Castells' network concept and its connections to social, economic, and political network analyses. *Journal of Sociology & Society*, 1(1), 1–10. <https://doi.org/10.21307/joss-2019-021>
- 10) Hykel, H. (2023). We have no satisfactory social epistemology of AI-based science. *Social Epistemology*, 37(6), 770–782. <https://doi.org/10.1080/02691728.2023.2286253>
- 11) Joshi, P., et al. (2020). The state and fate of linguistic diversity and inclusion in the NLP world. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*.
<https://doi.org/10.18653/v1/2020.acl-main.220>
- 12) Khalati, M. (2020). Artificial intelligence development and challenges: Arabic language as a model. *International Journal of Innovation, Creativity and Change*, 13(5), 3137–3145.
https://www.ijicc.net/images/vol_13/Iss_5/13578_Khalati_2020_E_R.pdf
- 13) Klein, L. F. (2015). The image of absence: Archival silence and the limits of data. *American Literature*, 87(4), 661–688. <https://doi.org/10.1215/00029831-3326983>

- 14) Knight, W. (2017). The dark secret at the heart of AI. *MIT Technology Review*.
<https://www.technologyreview.com/2017/04/11/5113/the-dark-secret-at-the-heart-of-ai/>
- 15) Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Ren, D., & Van Alstyne, M. (2009). Life in the network: The coming age of computational social science. *Science*, 323(5915), 721–723.
<https://doi.org/10.1126/science.1167742>
- 16) Manovich, L. (2013). *Software takes command*. Bloomsbury Academic.
- 17) Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. *Artificial Intelligence*, 267, 1–38. <https://doi.org/10.1016/j.artint.2018.07.007>
- 18) Moretti, F. (2013). *Distant reading*. Verso Books.
- 19) Muhanna, E. (2015). Distant reading and the Islamic archive: Symposium abstracts. Brown University.
https://www.academia.edu/15322237/Distant_Reading_and_the_Islamic_Archive_Symposium_Abstracts
- 20) O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- 21) Petit, J.-L. (2023). Humanities and social sciences (HSS) and the challenges posed by AI: A French point of view. *AI & SOCIETY*. <https://doi.org/10.1007/s00146-023-01752-8>
- 22) Ramsay, S. (2011). *Reading machines: Toward an algorithmic criticism*. University of Illinois Press.
- 23) Risam, R. (2019). *New digital worlds: Postcolonial digital humanities in theory, praxis, and pedagogy*. Northwestern University Press.
- 24) Romanov, M. (2015). Distant reading & the Islamic archive. *al-Raqmiyyāt*.
<https://maximromanov.github.io/2015/10-17.html>
- 25) Srinivasan, R. (2017). *Whose global village? Rethinking how technology shapes our world*. NYU Press.