

تحديد معاني الظاهرة الاتصالية بالبيانات الضخمة، ابستمولوجيا أم عدة منهجية

- انموذج التجنيس المنهجي المختلط في اللغة الضخمة -

Defining the meanings of communication phenomena with big data, epistemology or several methods

- The mixed systematic naturalization model in the big language-

رقاب محمد

جامعة الجلفة (الجزائر)، mohammed.reggab@univ-djelfa.dz

تاريخ النشر: 2025 / 03 / 31

تاريخ القبول: 2025 / 03 / 26

تاريخ الاستلام: 2025 / 01 / 13

ملخص:

تعتبر هذه المحاولة كنوع من المغامرة المعرفية لحقل بحثي يزداد تعقيدا لتعدد المتغيرات التي أصبحت تتموضع في حدوده الانطولوجية، والمقصود بهذا الحقل هو حقل الاتصال بكافة امتداداته، والتعقيد المقصود هو انتقال الظاهرة الاتصالية من الوضع الخطي الذي يتم فيه معالجة الواقع في الشكل الثابت الى الشكل المتشابك والذي لعبت فيه ظاهرة "الاتصال المفرط" دورا أساسيا في إعادة صياغته في ظل صعود مجتمع المنصة والذي تُفهم معانيه وبنيته من خلال أداء البيانات الضخمة، وعليه فإن جوهر ما سنتناوله يرتبط بتبرير البيانات الضخمة كجزء من نظرية المعرفة في حقل الاتصال من جهة، وكرافد ميتودولوجي ضمن النهج المختلطة من جهة أخرى، بحيث يتم اللجوء للغة الضخمة كشكل مناسب لإحداث التناهي المفترض. الكلمات المفتاحية: البيانات الضخمة، نظرية الاتصال، اللغة الضخمة، النهج المختلط، الخوارزميات.

Abstract:

This attempt can be considered a form of cognitive exploration into a research domain that is increasingly complex due to the multitude of variables now situated within its ontological boundaries. This domain in question is communication, including all of its extensions. The complexity in question involves the shift of the communicative phenomenon from a linear state, where reality is processed in a fixed form, to a more entangled state. In this new framework, the phenomenon of "hypercommunication" has played a fundamental role in its reconfiguration, especially with the rise of the platform society, where meanings and structures are interpreted through the analysis of big data. Therefore, the crux of our discussion relates to justifying the role of big data as an integral part of the theory of knowledge within the field of communication, and as a methodological resource within mixed-methods approaches, where big data serves as a suitable means for achieving the intended methodological integration.

Keywords: Big Data, Communication Theory, Big Language, Mixed Method; Algorithms.

قبل ضبط مستوى معالجة الغايات التي سنتناولها عبر هذا الطرح وجب الحديث عن البيانات الضخمة وعلاقتها بالظاهرة الاتصالية من خلال سؤال استفهامي يبدو الأكثر بروزاً لدى الباحثين في هذا المجال ومفاده: هل يمكن للبيانات الضخمة بناءً على الإمكانيات الحاسوبية التي تقدمها في تنظيم أشكال المجتمع والعلوم على حد سواء أن تولد مدخلات معرفية لإعادة مناقشة مفاهيم ونظريات الاتصال؟ قد يكون هذا السؤال البارز على ما يبدو ضرورياً للتفكير في مستقبل علوم الاتصال بعد حوالي قرن من الاجتهادات النظرية لفهم آثار الظاهرة الاتصالية على التكوينات النفسية والاجتماعية للأفراد في المجتمع الجماهيري، خصوصاً مع كثافة استخدام الوسائط، والذي تجسد في الاعتراف بظاهرة الاستيلاء التكنولوجي.

في ظل هذا الوضع كان من الطبيعي أن تزداد المناقشات كتعبير عن ضرورة تطويع الظاهرة الاتصالية انطولوجياً وتجنيسها منهجياً. ليصبح الحديث الأول يرتبط أساساً بتموضع نهج البيانات الضخمة ضمن نطاق العلم ومشروعية توظيف أسلوب البحث الكمي والكيفي في حدوده المنهجية، ومن باب توصيف المناقشات السائدة وجب الاقرار بأن Cukier و Myer-Schonberger Boyd و Crawford و Côté Gerbaudo و Pybus و Gil de Zúñiga هم بعض المؤلفين الذين برروا التفكير في قوة البيانات الضخمة والخوارزميات في تطابق الممارسات الاجتماعية والثقافية وتمدها. وفقاً لـ Kitchin هناك تسارع أكاديمي متزايد لتحليل كيفية تمثيل الخوارزميات في إنتاج المعرفة العلمية بشكل نقدي (Kitchin, 2017). والواقع أن من الصعب تحديد مجال لا يتأثر بصورة مباشرة أو غير مباشرة بسرعة وتنوع وحجم إنتاج البيانات وتحويلها إلى معلومات ذات قدرة انتشارية عالية. ابتداءً من العلم إلى الصحة العامة، ومن النظام المصرفي وصولاً إلى إنتاج الأخبار والتسويق (لعياضي، 2021). فكمية البيانات التي يتم إنتاجها تلقائياً أو تم صنعها من قبل المجتمع تنمو بطريقة غير متوقعة، مما يولد أنظمة سياسية جديدة لديها القدرة على تحدي الأعراف الاجتماعية الحالية. فنحن أمام ظاهرة تسجيل كل شيء في الواقع الاجتماعي (Cukier & Mayer, 2013)، فعلى سبيل المثال، تصبح شركة مثل جوجل قاعدة موثوقة معرفياً لمنع والسيطرة على وباء الإنفلونزا الذي كان لديه القدرة على الانتشار حول العالم (Helbing, 2015). فيما ينتج موقع فايسبوك _ من خلال أنشطة مستخدميهِ _ أكثر من 500 تيرا بايت من البيانات يومياً، يشمل ذلك المنشورات والتعليقات بالإضافة إلى الإعجابات (عبادنة، 2023).

وبالتالي، فإن الهدف من هاته الورقة البحثية يتدرج عبر مستويين الأول هو الإشارة بطريقة أولية إلى النقاط والمسارات الحاسمة في الخيارات المنهجية لباحث حقل الاتصال في الراهن الحالي، نظراً للإمكانيات الهائلة لجمع البيانات وتحليلها التي تمكنها منصات الإنترنت وأدواتها وتطبيقاتها. بحيث ينصب التركيز على تأثير البيانات الضخمة في نظريات المعرفة والأساليب التي تستند إليها الأبحاث التي تترجم واقع الاتصال، من خلال البيانات الواردة من الوسائط المتنوعة والتطبيقات الرقمية المتاحة للمجتمع اليوم، أما المستوى الثاني هو الافتراض بالمشروعية المنهجية للبيانات الضخمة في الظاهرة الاتصالية من المنطلق تحقيق التبرير وإحداث التجاسر لاستخدام النهج المختلط في اللغة الضخمة كشكل غالب على العملية التواصلية في ظل تحول المجتمع إلى "مجتمع المنصة".

أولاً: البيانات الضخمة وبنية الاتصال الرقمي ... حصر معاني الظاهرة بالعلم

يكشف النمو الهائل للبحوث والمقالات التي تتضمن قياس الحياة الاجتماعية من خلال البيئات الرقمية، مثل منصات التواصل الاجتماعي وأنظمة البيانات الحكومية المفتوحة، عن الحاجة الملحة للنقاش حول القضايا المعرفية والنظرية المنهجية التي ساعدت البيانات في صياغتها. ضمن هذا التوجه نجد Côté و

Pybus و Gerbaudo على سبيل المثال يضعون البيانات الضخمة على الطيف السياسي الذي يستهدف التلاعب بالبيانات الخوارزمية من قبل الوكلاء الاجتماعيين*. وانطلاقاً من السؤال الأكثر عمومية المتمحور حول علاقات القوة التي تؤثر على المعرفة الناتجة عن تحليل البيانات الرقمية، نجد أن الباحثين أعالوا النقاش لوجهة نظر نقدية مفادها استيلاء العلماء على البيانات الضخمة التي تفيد بفقدان خاصية الضمان المفترض في إطار الحياد والموضوعية الأخلاقية لجمع البيانات الهائل من الإنترنت (شاوش اخوان، 2022). إن ابرز تمثيلات هاته القوة المهيمنة على البيانات الضخمة تتجسد في قوة سوق وادي السيليكون ودوره الغالب في سلاسل الإنتاج الاقتصادية الجديدة والتفاوتات الهيكلية بين توليد البيانات من قبل المواطنين وسيطرة الشركات عليها، بالإضافة إلى التعطيم الخوارزمي الذي يقرر تكوينات الشبكة وتصورات المستخدم ليست سوى بعض العلامات على أننا دخلنا بالفعل في واجهة جديدة في علاقة المجتمع بالتكنولوجيا (Coté, 2016).

هناك حاجة ملحة في عصرنا الراهن لفهم كيف يمكن للخوارزميات والبيانات الضخمة أن تتيح عملية تمكين أشكال اجتماعية وثقافية جديدة وترسيخها ضمن فعل الاستخدام، يرافق هذا الاتجاه حتمية تطوير الإحساس النقدي بحدود وقدرات الخوارزميات والبيانات الضخمة والتي تنطوي في النهاية على فجوات دلالية وأحكام مسبقة (Uricchio, 2017). في الوقت نفسه يتم التأكيد باستمرار على آليات الهيمنة التي تمارسها ديناميكيات تحقيق الدخل الرأسمالي للبيانات الرقمية من قبل الوكلاء الاجتماعيين أنفسهم. أي أن استخدامات البيانات وتخصيصها من قبل المجالات الاقتصادية والسياسية اكتسبت أبعاداً ونفوذاً هائلين إذ أصبحت أيضاً بنى تشكل المعرفة. في الحركة الانعكاسية الحتمية التي تحكم الوعي الفردي والجماعي يقترح Milan و van der Velden مفهوم «نشاط البيانات»، وهو لمحة عن المعرفة التي تفلت من تجديد المستقبل، بناء على تطوير حاسم للعلوم والتكنولوجيا أسسه الباحثون أنفسهم. في هذا السياق يُنظر إلى زيادة توافر البيانات على أنها فرصة قوية وغير مسبقة لإحداث تغيير اجتماعي. يندرج هذا التوجه في الاقرار بدعم نشاط البيانات لظهور ثقافات معرفية جديدة" تتحدى القراءات السائدة للواقع وتشكل النهج الذي نتعامل به مع المعرفة والتحقق من صحتها" (Milan & van der Velden, 2016).

من وجهة نظر البحث في الاتصال السياسي في عصر وسائل التواصل الاجتماعي يقدم Gil de Zúñiga مجموعة من المحاولات الأكاديمية التي تقدم طرقاً فريدة للجمع بين الأساليب والتقنيات المتعلقة بالبيانات الضخمة والتي تأخذ في الاعتبار المكون الأخلاقي لاستخدام البيانات الشخصية في تشكيل عمليات البحث. ومع ذلك يقر الباحث بأن معرفة كيفية تفسير ضخامة البيانات التي توفرها آثارنا الرقمية لا تزال تمثل تحدياً هائلاً (Gil, 2015).

في سياقنا الحالي وقياساً بالتقييم العام لحيز العلم وطبيعة التماسك بينه وبين البيانات الضخمة سنلاحظ أن هناك مفاهيم على المحك مثل صحة البنيات العلمية، واختبارات الفرضيات والتعميم المرتبطة كرنولوجيا ببعض البنى النموذجية للعلوم الكلاسيكية. يهتم المؤلفون المذكورون سابقاً بقضايا مثل التأسيس لدحض عصر النظريات كنقطة انطلاق لإنتاج المعرفة نظراً لفقدانها للقوة التفسيرية للبعد الهيكلي للمجتمع بتحولاته الراهنة، في ظل صعود نماذج لجمع البيانات يحتمل أن تكون قادرة على جمع قواعد بيانات رقمية لا نهاية لها (Gil, 2015).

* الوكلاء الاجتماعيون هم وكلاء معرفيون حاسوبيون يتمتعون بقدرات معرفية إلى جانب القدرات التقليدية للعمل والتواصل وأشكال أخرى من التفاعل، مهمتهم تعزيز رسالة المستخدم من خلال مشاركة المحتوى الخاص به. غالباً لا يتعامل الوكلاء الاجتماعيون مع مستخدمي المنصات الاجتماعية، ولكنهم يعرضون محتوى المستخدمين مع بعضهم، غير أن الحديث بدأ يسود حول أدوار خفية للوكلاء فيما يتعلق بضمان بيانات المستخدمين.

إن عكس الإجراءات المنهجية التي تضع عمليات جمع وترابط البيانات المستفيضة على المجتمع باعتبارها تعبيرات معرفية محسنة يعتبر مؤشرا لثورة على المستوى الكليات العلمية، بالمعنى الذي ضمنه Kuhn لهذا المصطلح (Kuhn، 1998). يستكشف Kitchin إلى أي مدى تؤدي البيانات الضخمة إلى نظريات معرفية بديلة في العلوم الاجتماعية وذلك بناء على تقييم نقدي للتغيرات الحاصلة في ممارسات البحث. إذ يظهر تحدي الباحث واضحا للفكرة التي تظهر في هذا السياق والمبنية على طريقة عمل استقرائية بحتة، بحجة أنه بدلا من تعزيز النزعة التجريبية على البناء النظري (Kitchin، 2014)، فإن العلم القائم على البيانات هو نسخة معاد تشكيلها من المنهج العلمي التقليدي ذو النزعة الوصفية. ولهذا الغرض يجب جمع البيانات وتفسيرها في ضوء المسائل ذات الصلة بالمعارف القائمة. فالمجموعات المصنفة من «الصور الفوتوغرافية»، المتعلقة بظروف وسياقات محددة للواقع الرقمي، والتي توضح نوعا من الارتباط على الآثار الاجتماعية المرتبة على الإنترنت والتي يتصورها الباحث، قد لا تجيب بالضرورة على الأسئلة النظرية والعلمية. يؤكد الجانب المغربي للبحث القائم على البيانات الضخمة أيضًا على بناء الأساطير، وفقًا للمصطلح الذي استخدمه Boyd و Crawford، والذي يستند إلى "الاعتقاد السائد بأن مجموعات البيانات الكبيرة تقدم شكلاً أعلى من الذكاء والمعرفة التي يمكن أن تولد رؤى كان من المستحيل سابقا تجسيدها، مع هالة الحقيقة والموضوعية والدقة" (Crawford & Boyd، 2012). ويمثل هذا البعد الرمزي للبيانات تعزيزا لوهم الشمولية المعلوماتية والتمثيلية والجدارة بالثقة، بحيث يميل إلى أن يكون أساسا أنطولوجيا ويوجه ويشكل المعرفة العلمية. إذ أن هناك ضمان للتحقق من التحيز التجريبي للمعرفة الجديدة والقائم على وفرة الغوص الاستكشافي في تحسين مجموعات البيانات على الإنترنت. ومع ذلك في هذه الحركة يحذر Kitchin من أن "البيانات يتم فحصها من خلال عدسة معينة تؤثر على كيفية تفسيرها"، علاوة على ذلك فإن "الارتباطات بين المتغيرات داخل مجموعة البيانات يمكن أن تكون عشوائية بطبيعتها وليس لها ارتباط سببي" (Kitchin، 2014)

. إن تجنيس المغالطات البيئية الذي أشار إليه Kitchin باعتباره خطراً محتملاً على نظرية المعرفة العلمية التي تظهر داخل البيانات الضخمة ليس قضية معزولة للتفكير، فمن بين التحديات الأخرى التي لوحظت للممارسات العلمية الميدانية التقليدية الإفراط في فردية أساليب البحث وجمع البيانات. وهذا يؤدي في أقصى الحدود إلى عدم قابلية التحليلات المتعددة وبالتالي إعاقة الإجراءات الحاسمة التي تشكل المرتكز الذي يتم على أساسه اختبار النظريات وتحسينها.

من المفارقات أن الاستقلالية الممنوحة للباحثين فيما يتعلق بوضع أدواتهم الخاصة لتتبع البيانات وجمعها وتحليلها يمكن أن تصبح ما يستنفد منه العلم الحديث في إحدى مركباته الرئيسية والمتمثلة في إمكانية البحث عن الحقيقة من خلال مقارنة نتائج اختبارات التحقيق ودحض الفرضية. ووجه المفارقة يرتبط بالتصرف التجريبي على المنصات الرقمية بتأسيسها استباقيا كحقيقة من خلال تحويل أطر الواقع المتعددة إلى معرفة بلا منازع. غير أن هذا التجنيس البحثي تحيل نتائجه التي تم تحديدها بواسطة أدوات خوارزمية تجمع البيانات وتنظمها، إلى فقدان العدة التفسيرية والتي تعتبر السمة المميزة للتعميم والصياغة النظرية.

يلاحظ في راهن البحوث العلمية المرتبطة بالإنترنت والمهتمة باستكشاف البيانات المخزنة رقمياً، أن لديها نماذجها الخاصة وأدوات جمع البيانات الرقمية*. هناك وفرة من الخوارزميات التي تم إنشاؤها لاختيار بيانات الوسائط الاجتماعية ودمجها وليست جميعها مفتوحة المصدر. ينعكس الاهتمام المتزايد بالبحوث المتعلقة بالشبكات الاجتماعية عبر الإنترنت _ والتي اكتسبت مكانة بارزة في مجال دراسات الاتصالات على مدى العقد الماضي _ على مجموعة واسعة من الأدوات والبرامج المتاحة حصرياً لمعالجة هذا النوع من البيانات، كل منها بوظائفه من أبرزها مسار الترميز الخوارزمي والصيغة الجمالية*.

وعلى الرغم من أن هذا التنوع في النهج البحثية يمثل تجاوزاً للنماذج والتقنيات التي تحد من فهم الواقع الاجتماعي لكونها مبنية من منظور السيطرة على ذلك الواقع وثباته في أطره التفسيرية (وبالتالي تتعارض مع الحركة المتأصلة في الاتصال)، إلا أنه يجب الإقرار بأن البحث من خلال نهج البيانات يرتبط بشكل أساسي بما تتيحه الأدوات التقنية لحصر البيانات وتحليلها عن طريق الخوارزميات ضمن شكل غير متجانس وغير شفاف، والذي قد يكون مبرمجاً لأداء وظيفة الذكاء الاصطناعي بحيث لديه القدرة على توسيع نوايا الإرشاد لتصميمه. كما يشدد *Kitchin* على عمل الخوارزميات بطرق متعددة المتغيرات " يجب التعرف عليها على أنها وراثية، وأدائية، وشرطية أي أنه لم يتم إصلاحها أبداً في الطبيعة، ولكنها ناشئة إذ تم تصميم العديد من الخوارزميات لتكون تفاعلية وقابلة للتغيير للمدخلات " (*Kitchin, 2017*). وملخص هذا المنظور أن الأداء البحثي للخوارزميات يرتبط بعملية الكشف المستمر المستمد أساساً من الذكاء الاصطناعي التلقائي.

ومع ذلك، فإن الحكم العام على هاته التأثيرات في إطار علاقة العلم بالتلاعب بالبيانات الضخمة كثنائية مربكة لا زال محل تردد وتشكيك، ناهيك عن الأشكال المتعددة للاتصال المعاصر والبيانات الوصفية التي استخرجها منها من طرف باحثو الاتصال. نظراً لأن الترميز الخوارزمي الذي يعزز البيانات الضخمة يميل إلى الحجب من خلال جماليات تطبيقاتها وبرامجها وعدتها التقنية، فمن الصعب الكشف عن المقاصد الأصلية التي تؤدي إلى نشر البيانات ومشاركتها من قبل مختلف الوكلاء. تعتمد هذه الاتجاهات بالأساس على الترجمات العكسية للرموز المبرمجة سابقاً وهو مسلك شائك لم يغفله *Kitchin* ضمن مراجعته النقدية للخوارزميات. تعتمد برمجة المساحات الإلكترونية الرقمية على ترجمتين، بشكل أساسي: " الأول هو ترجمة مهمة أو مشكلة إلى صيغة منظمة مع مجموعة قواعد مناسبة (رموزائف). والثاني هو ترجمة هذا الرمز الكاذب إلى رمز مصدري عند تجميعه ليؤدي المهمة أو يحل المشكلة. يمكن أن تكون كلا الترجمتين صعبة، وتتطلبان تعريفاً دقيقاً لماهية المهمة/المشكلة (المنطق)، ثم تقسيمها إلى مجموعة دقيقة من التعليمات، مع مراعاة أي حالات طوارئ مثل كيفية أداء الخوارزمية في ظل ظروف مختلفة (التحكم) " (*Kitchin, 2017*).

على الرغم من أنها تبدو كعملية بناء تلقائية يمكن التحكم فيها بشكل فعال، إلا أن التحدي الذي تشكله الخوارزميات وفراط البيانات التي يقودونها هو قدرة المجتمع على التواصل بشكل فعال مع الترجمات التي يقوم بها المبرمجون. هذا الاتصال المحتمل بين معاني الخوارزميات والآثار التي تحدثها على الديناميكيات التنظيمية للمجتمع هو أحد أكثر الاحتياجات التي تم التقليل من شأنها في السياسة المعاصرة.

* من بين العديد من مراكز الأبحاث العاملة اليوم، هناك مختبر *Citizen Lab*، وهو مختبر في كلية مونك للشؤون العالمية بجامعة تورنتو؛ ومختبر وسائل التواصل الاجتماعي بجامعة رايرسون، أيضاً في تورنتو، ومركز الثقافة الرقمية في المكسيك؛ ومركز نورديك للإنترنت والمجتمع بجامعة سيجن بألمانيا؛ ومركز أبحاث الإنترنت في سنغافورة ومختبر أبحاث الوسائط التابع لمعهد ماساتشوستس للتكنولوجيا (MIT) في الولايات المتحدة.

* على سبيل المثال تشكل الأدوات التالية، من بين العديد من الأدوات الأخرى، قائمة البدائل لجمع ومعالجة وتحليل بيانات الإنترنت: *NodeXL*، و *iGraph*، و *Netlytic*، و *Uberlink*، و *Issue Crawler*، و *Gephi*، و *Cosmos*، و *Chorus*، و *VT Tracker*، و *Netvizz*، والأخير مخصص لجمع بيانات *Facebook*. الموارد والتطبيقات المسموح بها عديدة، مما يولد أكثر الخصائص التحليلية تنوعاً.

بالنظر إلى البانوراما المعروضة، يواجه مجال أبحاث الاتصال فئتين مهيمنتين على الأقل من القضايا المتعلقة بالتطورات الحالية والمستقبلية. تشير الأولى إلى التوافق الضروري بين الأنطولوجيا ومنهجية البحث، في هذا الصدد استخدم Hall هذا التوجه لمناقشة أسس البحث السياسي المقارن. واعتبر مسألة التفكير فيما يبدو أنه ظهور توجه معرفي في علوم الاتصال والتي على الرغم من تركيزها على التطور التقني لوسائل الإعلام لا تفكر بالضرورة في افتراضاتها الضمنية وآثارها على التنظيم النظري والمنهجي للبحوث (Hall, 2003).

أما الفئة الثانية التي ترتبط حتماً بالفئة الأولى فتربط بالخوف من التأزم النظري، مما يفضل معرفة بيانات الواقع في البحث العلمي على النحو المبين أعلاه. وفي هذا السياق من المناسب إعادة التفكير في الاتجاه نحو الاستخدام العشوائي للمفاهيم النظرية المستخلصة من سياقات إنتاجها الأصلية، والتي تصف مجموعات البيانات وتصممها بشكل متخم بالكلاسيكية المقيدة. والحجة أن هاته الصرامة أحد الضمانات لتطور المعرفة العلمية في حقل الاتصال.

غير أن تجارب الانغماس الاجتماعي في الإنترنت أظهرت أننا نواجه سياقاً جديداً، مع متطلبات وعمليات جديدة، تتخللها حتمية إنتاج البيانات ونشرها على جميع المستويات، في ظل ثورة ويب 3.0 الذي يتمثل كرونولوجيا كنوع من مراحل الربط بين البيانات، وهو بعد جديد للفضاء تحدث فيه العلاقات بطريقة متنوعة ومتفرقة ومتعددة (Amaral, 2012). يمكن القول أنه لا توجد نقاط مشتركة كثيرة بين عصر الاتصال الجماهيري وتلك التي تظهر على المسارات الرقمية. وبالنظر إلى هذا الاعتراف سيكون من الخطأ تفسير الواقع الأخير بناءً على الإرث النظري والمفاهيمي الذي تم تشكيله كشرح مشروع لنمط تواصل آخر للمجتمع.

حول المحتوى الذي يولده الوكلاء الاجتماعيون في الشبكات الرقمية، على سبيل المثال يقال أن " تصنيفهم المحتمل حتى مع التنوع الرسمي للمنصات يتجاوز التصنيفات الموحدة " (Amaral, 2012). إذا كانت البيانات الضخمة والبيانات الوصفية التي تسمح بتكوينها هي الدعم الجديد للروابط الاجتماعية، كما يفسره Amaral فمن الضروري التحرك بسرعة نحو فهم دلالاتها حول الحواس المتشعبة في تصورها وتطورها.

يشير هذا التوجس أيضاً إلى ما عبر عنه كل من Shaldbolt, Hall و Berners-Lee في إطار الحاجة المتزايدة لدمج البيانات الرقمية في قواسم مشتركة للمعنى. إذ يُنظر إلى الويب الدلالي الذي دعا إليه الباحثون بما يتجاوز ما يسمى بـ «الفولكسونوميات» (Shaldbolt, Hall, & Berners-lee, 2006)، وهي لغات يتشاركها عدد كبير من الأشخاص المهتمين على أنها تكرار لبيانات معينة من خلال المعلومات لترديدها في علم المعلومات، يتميز تنظيم تدفق المعلومات والبناء المشترك لدلالات الويب باللامركزية في الوصول، وسهولة نشر البيانات من خلال وسائل التواصل الاجتماعي. ومع ذلك فإن أيقونات هذه «الفولكسونوميات» متعددة. ومن علامات التصنيف إلى الميمات، يمكن أن تمثل الأشكال المختلفة للتمثيل الجمالي الديناميكيات العلائقية للبيانات على الإنترنت. ومع ذلك يبقى أن نتساءل ما هي أوجه التشابه والتوازي؟ وهي أنسب المقارنات وأكثرها مشروعية التي يجب إجراؤها بين هذه المجموعات الدلالية، والتي تتيح فهرستها إلى بناء الأنطولوجيات، كما حددها Shaldbolt, Hall و Berners-Lee، فالأنطولوجيات بهذا المعنى قادرة على إزالة أوجه الغموض بين المسارات التي لا حصر لها التي تنتقل من خلالها البيانات وتنتشر في الفضاء الرقمي. هذه هي الدلالات الأكثر ديمومة والتي تستند إليها التصنيفات مما يتيح فهم المجتمع في سياق الإنترنت (Shaldbolt, Hall, & Berners-lee, 2006).

من خلال هذه التعقيدات يبدو أن البيانات الضخمة ظاهرة واعدة لفهم واقع التواصل الاجتماعي المعاصر الذي يضع الاتصال وعملياته المنبثقة كحجر الزاوية ومحور نظري وتفسيري ذي صلة بباقي العلوم الإنسانية والاجتماعية. فظهور البيانات كنموذج منهجي مهيمن يكشف عن التقارب والمؤانسة بين العلوم كافة.

في هذا النطاق من البحث الذي يواجه الممارسات العلمية لا توجد طريقة لتجاهل الاتصال كعنصر تفسيري حيوي في إطار العلاقات الثنائية المتبادلة بين الأفراد والبيانات، ومع ذلك من الضروري استحداث تبريرات معرفية من قبل الباحثين في خياراتهم حول ما تم تحديده في الوسائط والأجهزة الرقمية .

1.1 من أنطولوجيات التواصل إلى تصاميم البحوث المنهجية

ما هو الواقع التواصل الذي يتحدث عنه الباحثون ؟ ما هي الأسئلة التي تقود التبريرات العلمية التي تستند إلى بيانات الإنترنت* ؟ ما هي الأنطولوجيا التي تدعم الخيارات المنهجية للباحث ؟ في مقاربة هاته الاستفهامات بالظاهرة الاتصالية. تبرز الحجة الأساسية هنا في الزاوية التي توجه منها الأعمال القائمة على البيانات الضخمة والتي تساعد في تكوين المسارات المنهجية والنظرية لتحقيق التبرير. على سبيل المثال يؤدي اتساع وجهة نظر الاتصال كظاهرة ثابتة، للتسليم بأن الواقع عبارة عن صور ثابتة تقع في خطط محددة مسبقاً من خلال المراقبة المباشرة والمحايدة مما ينعكس على التصاميم العامة للبحث والتي يُفضل فيها الآليات الخطية لتداول المعلومات. وعلى غرار ما لاحظته Hall في بحوث السياسات المقارنة الدولية، يمكننا أن نقول إن قيمة الأساليب المستخدمة في ميدان الاتصال أيضاً تتوقف على توافقها مع الطريقة التي يفهم بها الباحث الأسباب في العمليات التي تتحكم في نوع التحليلات المعتمدة (Hall, 2003). وبالتالي فإن التقاط البيانات الرقمية في لحظة زمنية معينة وفي سياق محدد من الأداء سيشكل العدة المناسبة لتحقيق التفسيرات الاتصالية المحتملة، بعبارة أخرى إذا كان منظور الواقع الخطي في الاتصال والتحكم في نتائجه هو ما يسود من الناحية الأنطولوجية في البحث، فلن يكون لدى البيانات الضخمة الكثير لتقدمه من حيث بناء المعرفة. وهذه الطريقة لتمثيل الواقع بصفته يتألف من كيانات ثابتة وينظم مسبقاً في أشكال خطية من الأسباب والعواقب التي تحد من القدرة التفسيرية لنهج البيانات الضخمة في ظل النظرة التي تتضمن نماذج خطية لتمثيل للواقع (Abbott, 1998). وبالتالي من هذا المنظور لن يكون من الممكن إحراز تقدم معرفي من حيث توسيع الإمكانيات التفسيرية للظاهرة الاتصالية.

في المقابل وعلى مسرح الأفضية الرقمية يبدو أن البيانات والمعلومات يمكن تحديدها على أنها وحدات ذات تعقيد اجتماعي تقطعها المتغيرات الأكثر تنوعاً وبالتالي فهي متعددة الخطوط. وبقدراً يتم تمثيلها من خلال التعبيرات النصية والتصويرية والسمعية وأيضاً من خلال هياكل العلاقة بين العقد، فإن البيانات التي تم جمعها في البيئات الرقمية تسمح لنا بإلقاء نظرة خاطفة - وحتى تمثيل - قوة التواصل البشري. لا يتعلق الأمر بالقول بأن ما يحدث على الإنترنت من حيث العمل الاجتماعي والتعبير يختلف اختلافاً جوهرياً عن العمليات التي تم تجربتها بالفعل في «المجتمع التناظري». ومع ذلك إذا أخذنا البعد الرقمي للوجود الاجتماعي كنقطة انطلاق لشرحه، يجب أن يكون هناك تعديل دقيق وكاف بين قوائم العدد التفسيرية التي تراكمت حتى الآن كمجال علمي، وهذا الواقع الانعكاسي والثابت هو نتاج الأفراد والمنظمات أساساً. لذلك فإن التبرير المستحدث هو استحالة الحفاظ على نفس المفاهيم حول الواقع التواصل الذي ساد في القرن العشرين .

إن نظرية المعرفة القائمة على براديفم التعقيد _ باعتباره فلسفة مورانية محايدة لمنطق التطور والتداخل العلمي _ تعتبر أكثر تماسكاً مع ديناميكيات وسائل التواصل الاجتماعي، بناءً على الحركات غير الخطية ومتعددة السببية والدائمة التي لوحظت من خلال التعبير عن التوترات والصلات المتنازع عليها. تتحد

* من المهم أن نؤكد أن البيانات المستمدة من منصات الشبكات الاجتماعية، أو وسائل الإعلام الاجتماعية (وفقاً للمفهوم الشائع في اللغة الإنجليزية)، ليست سوى نوع واحد من البيانات الضخمة. ولم يكن نموها كموضوع بحثي مهمة علماء الاتصال وحدهم، بل كان أيضاً علماء السياسة وعلماء الأنثروبولوجيا وعلماء النفس والجغرافيين وعلماء الكمبيوتر والأكاديميين من مختلف مجالات المعرفة.

شبكات العلاقات مثل Facebook و LinkedIn، ومنصات المعلومات مثل منصة X ومدونات الشبكات الاجتماعية مثل Medium، قد حلت محل المواقع التقليدية كمكان للعروض العامة والهوية من قبل الأفراد والمنظمات، مما عزز ظهور وبناء بيئة إعلامية جديدة ناتجة عن الانتشار المتزايد للتطبيقات كما حددتها Amaral في تحليلها لـ Web 3.0 والمحصلة أنها تجعل وسائل التواصل الاجتماعي أكثر وضوحًا (Amaral, 2012).

وبالتالي فإن تقييد وسائل التواصل الاجتماعي كموضوع لدراسات الاتصال يصبح التحدي الكبير المتمثل في صقل الأساليب والأدوات المتاحة من أجل ضمان إمكانية فهم ديناميات التعبير الأكثر تنوعًا، واعتمادات التكنولوجيات ومعانيها من قبل المرسلين والمستقبلين، وكذلك عدم احتمالات الاتصال التي تتخلل الواقع الرقمي (Luhmann, 2006). وبهذا المعنى فإن مبادرات كالتي كل من Procter, Vis و Voss تعتبر محاولات ملهمة، لأنها تقترح طرقًا قادرة على دمج تفسير وتمثيل الواقع الافتراضي بطريقة أكثر وضوحًا، والحركة المتأصلة في النقاشات والتقاء المعنى في عمليات التواصل الاجتماعي (Procter, Voss, 2013). وفي تقييم معرفي يمكن القول أن هؤلاء الباحثين نجحوا في تطوير أداة جمع منشورات X (التغريد وإعادة التغريد)، حيث يبدو الترميز الخوارزمي قادرًا على تتبع ديناميكيات الاضطرابات والشائعات في المنصة. علاوة على ذلك فإن تمثيل النتائج التحليلية في جدول زمني للرسائل المرسلة في X يعالج مرونة إجراءات تبادل المعلومات التي تشكل جزءًا من البيانات الضخمة.

وبهذه الطريقة يمكن لباحثي الاتصال أن يجدوا حيزًا مميزًا لتطبيق الأدوات التكنولوجية المبتكرة القادرة على تصور مسارات الرسائل المتداولة قد تصبح "أجهزة نقد الوسائط التفاعلية" التي حللها Braga وفقًا للنطاقات المنهجية والأدوات والتقنيات لجمع البيانات وتحليلها والتي تم تحسينها باستخدام البيانات الضخمة، مرئية في طبقات البناء الجماعي المختلفة (Braga, 2006).

وفقًا لـ Martino فإن الأخبار الإعلامية، التي تشكلت كبنى كاشفة للتواصل أو مواضيع ذات مصلحة عامة، تتقاطع بسرعة وأحيانًا ضمن اتجاهات غير متوقعة مع الحياة اليومية للإنترنت (Martino, 2009). إن فهم الآليات التي تظهر بها الأخبار في البيئات الرقمية، سواء كتعبير عن الإمكانيات الفردية والتنظيمية أو كأحداث عفوية متأصلة في ديناميكيات العلاقات الاجتماعية تصبح هدفًا بحثيًا مهمًا في مجال الاتصال، إذ يتمتع التعزيز التكنولوجي المستمر للتحليلات الرقمية للبيانات بالقدرة على إنشاء أدوات رسم الخرائط التي تغطي إطارات وسائط اتصال أكثر اكتمالًا، وتجعل مسارها تمتد عبر إرسال المعلومات مع اتجاهاتها اللاحقة من خلال وسائل التواصل الاجتماعي. بمعنى أنه ستمثل المسارات التي فتحتها وتتبعها الرسائل نفسها ضمن بنية رقمية تركز على زمان الاتصال. وبالتالي فإن الرسائل مع كل الفروق الدقيقة التفسيرية الناشئة بشكل طبيعي والمتغيرة ستكمل الإطار التحليلي المعقد للاتصال في وسائل التواصل الاجتماعي.

في هذا السياق يمكن لبحوث الاتصال أن تغلب أخيرًا على تحيزها الآداتي، بحيث يتجاوز دمجها الوصف والتحليل خلال الإرسال إلى نقاط التلاشي التي تفلت من محاولات تنظيم معانيها، كما حللها Fausto Neto إذ إن استخدام وسائل الإعلام - التي شكلت العقيدة في مجال الاتصال ولا تزال تهيمن على مجالات إنتاج المعرفة الموجهة من قبل أو من أجل مؤسسات السوق والدولة - لا يتوافق مع عمليات الاتصال التلقائية التي يمكن ملاحظتها بوضوح في جميع أنحاء الإنترنت (Fausto Neto, 2001).

2.1 شبكات وسائل التواصل الاجتماعي وتحدي نظريات المجتمع الجماهيري

من خلال ما تم تناوله سنستفيض بالحديث انطلاقًا من فكرة Kitchen التي تفيد بأنه لا يمكن للثورة النموذجية التي قدمتها البيانات الضخمة أن تبرر غياب التفكير النظري. فعلى العكس من ذلك يقال بأن ظهور مجتمع معتمد على البيانات يولد لحظة مواتية لتعميق وإعادة مناقشة صلابة المفاهيم وافتراساتها

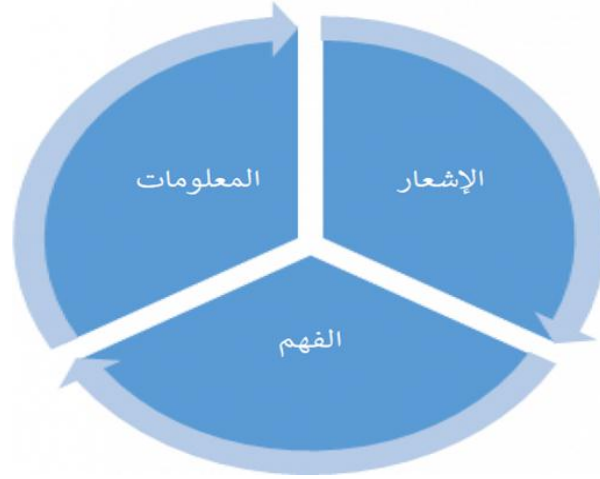
المعرفية، وبالتالي يجب أن تسود المآزق والأسئلة البحثية المستندة إلى النظريات كبوصلات توجه النظرة التجريبية، وألا يتم تجاهلها باسم " البيانات السيادية التي تتحدث عن نفسها " (Kitchin, 2014) . إن استمرار عملية إعادة النظر والنقاش حول الأساس النظري والمفاهيمي لمجال الاتصال هو ضمان لجدوى المعرفة العلمية، ومن المفترض أن هذا يتوقف على الدقة التفسيرية للبيانات والتعبير عن المشاعر، ووجهات النظر العامة ومحاولات الإقناع المتعمدة، فضلا عن الإجراءات الفردية والجماعية. في النهاية تتعلق وسائل التواصل الاجتماعي والمعلومات الموجودة لديها عن الواقع بمجموعة واسعة من التعبيرات الاجتماعية. من هذا المنطلق سيكون من المناسب والملح مناقشة كيف يمكن للأطر المفاهيمية والنظرية الكلاسيكية للاتصال أن تساهم في التفكير في قضايا البحث التي تنطوي على البيانات الضخمة، أو كيف يمكن _ في علاقة عكسية _ إعادة التفكير في نفس المقترحات النظرية من الواقع الاجتماعي المتشكل عبر الإنترنت. إن إحدى نقاط الانطلاق للتفكير وفقًا للحجج السابقة، هي الطبيعة المتقلبة للغاية وسريعة الزوال والطائرة لإنتاج البيانات التي تتخلل المساعي البحثية. تعد المنصات والوسائط العديدة التي يمكن للوكلاء الاجتماعيين من خلالها عرض أصواتهم وصورهم وإجراءاتهم المنسقة حجر الزاوية ضمن البيانات التي يتم التلاعب بها. وهي تتقاطع معها وتبنيها دوافع ومصالح متنوعة تظهر وتختفي كمعلومات قابلة للتحليل في تحركات غير منقطعة ومتراصة للتعبيرات الفردية والتنظيمية.

وكما تم عرضه سابقا تتنافس مجموعة متنوعة من وسائل التواصل الاجتماعي على جذب انتباه المستخدمين، في عالم مليء بإمكانيات التواصل أو ما يصطلح عليه " الاتصال المفرط أو الكثيف " ولكل منها مادية ورمزية. يوضح هذا السيناريو أن أنماط الحياة والمصالح ذات الأولوية في التواصل الاجتماعي والموقع المكاني والظروف الاقتصادية، من بين عوامل أخرى تشكل قائمة المتغيرات التي تؤثر على ديناميكيات المجتمع الرقمي. ويمكن أن تؤدي السياقات الثقافية والوطنية، على سبيل المثال إلى استخدام منبر واحد على حساب منبر آخر. في هذا الطرح يمكن الاستدلال بمنصة X كحالة مثيرة للاهتمام، فبينما في بلدان مثل الولايات المتحدة، يُنظر إلى المنصة على أنها مقياس للسياسة في المجال العام، فإن تأثيرها في البرازيل ليس مماثلاً. من ناحية أخرى فإن دراسة موضوع سياسة الأزدياء النادرة في البرازيل دون تضمين إشارة واحدة على الأقل للدور البارز لفيسبوك كمنصة للنقاش ستقصر في إعطاء الصورة الكاملة.

لذلك في هذا السياق متعدد الأوجه، يبدو أن العمل بطريقة محكمة مع مفاهيم المرسل والوسيط والرسالة والمستقبل، ومع تجاهل التحركات المتأصلة للاتصالات الهادفة في الوسائط الرقمية، يؤدي إلى نتائج عكسية فالتخصيص المفاهيمي الذي يوضح العضلات التي تواجهها النظرية يرتبط فكرة البث المباشر والتي يتم تطبيقها اليوم لفهم البيانات الناتجة عن شبكة إعلامية رقمية ومركزية وأحادية الاتجاه. يُعتقد أن مثل هذا المفهوم لا يعطي قيمة تذكر لفهم الاتصال عبر الإنترنت، لأن البيانات التي يعالجها عادةً لا يمكن أن تكشف بالضرورة عن استحواذ على المعلومات من قبل الجمهور المزعوم والمتمثل في شعار " ديمقراطية الاتصال "، ليرز مآزق انطولوجي لمفهوم «الجمهور» كما تمت الإشارة إليه في الأدبيات التقليدية.، والسؤال هل يمكننا حقًا تفسير تكوينات الأفضية الرقمية بناء على آليات مفاهيمية تعود إلى السياق التاريخي للاتصال في القرن العشرين مثل نماذج التأثير التي طرحتها نظريات مختلفة مثل الإبرة تحت الجلد، والتدفق المعلومات عبر مرحلتين والأجنحة ستينغ ودوام الصمت، وبالتالي تبني واستخدام المفاهيم الأولية للمجال لشرح الواقع الحالي ؟

من المهم أن نتذكر أنه لا يمكن استنتاج الاتصال من المعلومات بحد ذاتها، فالاتصال يتطلب وجود روابط مشتركة للمعاني يتم التوسط فيها بواسطة التقنيات أو الأفراد. لذلك لا يمكن اعتبار متابعي صفحة

على فيسبوك كجمهور متجانس مثل جمهور وسائل الإعلام الجماهيرية التقليدية، على الرغم من أن استخدام شركات مثل **Cambridge Analytica** لعلم النفس الكمي في تحليل البيانات الشخصية يكتسب أهمية كبيرة في عملية جذب انتباه الأفراد وإقناعهم. فالاتصال ليس عملية خطية أو ميكانيكية تركز على الإرسال، وإنما هو عملية ديناميكية ومتغيرة تعتمد على البيانات الضخمة لفهمها. وبالتالي يقال أن النموذج النظري السائد لإجراء أبحاث وسائل التواصل الاجتماعي في هذا السيناريو المعقد الذي يربط السياقات التقليدية للاتصال الجماهيري بالسياقات الرقمية لإنتاج الواقع هو الذي اقترحه **Luhmann** (Luhmann, 1995). من هذا المنظور لا يكتمل الاتصال إلا عندما يتم اختيار وفهم المعلومات الصادرة بشكل منهجي والقدرة على توليد معلومات جديدة، ضمن دورات اتصال جديدة.



الشكل 1: وحدة التواصل في نظرية نيكلاس لومان (1995)

إن فكرة الاتصال كعنصر أساسي في التنسيق الاجتماعي - وكوحدة مُجسّدة من خلال دورات مستمرة من إنتاج وإلغاء المعاني - هي الفكرة التي تُناسب أفضل تبادل للبيانات والمعلومات الفورية في وسائل التواصل الاجتماعي، بالإضافة إلى توافق مجالات الإرسال والاستقبال في بقينا عند المفاهيم التقليدية للمجال. يقدم **Sbardelotto** من خلال مسارات نظرية أخرى، فكرة مقارنة للمنظور اللوهمني عندما يدافع عن إمكانية تبادل الإنتاج والاستقبال في عمليات دوران الشبكة ، و "أن هذين القطبين موجودان فقط بشكل متبادل ، مما يشكل بعضهما البعض من خلال ديناميات الدوران" (Sbardelotto, 2017). لن يكون هناك دعم تجريبي في هذا السياق لإعادة تأكيد نماذج الاتصال النظرية التي يتم فيها عزل حالات إنتاج واستيعاب المعلومات عن بعضها البعض. " في التدفق الدائري لشبكات الاتصال على وجه التحديد لا يختفي قطبي الإنتاج والاستقبال، لكن لم يعد من الممكن تثبيتهما في موضوع اجتماعي محدد، مثل شركات الإعلام أو المؤسسات الاجتماعية التقليدية " (Sbardelotto, 2017). حقيقة الحياة الرقمية اليوم هي أنه من الممكن لوكيل أن يؤدي وظائف المرسل والمستقبل في نفس الوقت. بالإضافة إلى ذلك يمكن أن يكون هذا الوكيل زائفا يتم التحكم فيه بشكل مصطنع بواسطة الخوارزميات، أو يمثل منظمة أو مجموعة أيديولوجية إلخ. هناك العديد من جوانب من ينتج ويستهلك المعلومات.

دون إرهاق التفكير في كيفية توجيه نظريات ومفاهيم الاتصال للبحث في سياق التعقيد المتزايد الذي يكشفه الإنترنت، يجدر بنا التفكير النهائي في المعنى المفاهيمي للرسائل. هل نتحدث عن الرسائل بمعناها التقليدي عندما نختار أسئلة البحث التي تركز على الوسم، والمحادثات، والمشاركات، والإعجابات، والصور، وسجلات الوسائط المتعددة أو إعادة التغريدات في وسائل التواصل الاجتماعي ؟ تعيد هذه المسألة التأكيد

على الحجة المذكورة أعلاه حول مستويات علم البيانات الدلالية، بدون تحديد المعايير التي يتم بناء عليها المساحات الرقمية، يكون من الصعب فهم كيف يمكنها أن تساهم بشكل فعال في التواصل الاجتماعي. فرضية البحث هي أن اختيار منصة الاتصال التي يتم التعبير عن الرسالة فيها يربط أنواعاً محددة من المحتوى والجماليات، متوافقة أيضاً مع استخدامها الاجتماعي. لهذا تبعات لا تنحصر فقط على صياغة تصنيفات أوسع حول معاني الرسائل ومحتواها، جنباً إلى جنب مع ارتباطها بوسائل الإعلام (ستكون التطبيقات الآن مجرد أحد أنواع الوسائط)، دون اغفال الفصل بين هذا المتغير الذي "يحمل المعنى" والأبعاد الاجتماعية لدورانها (Shaldbolt, Hall, Berners-Lee, 2006).

ثانياً: النهج المختلط في اللغة التواصلية الضخمة... تجاوز اشكالية اللامعنى

للبدء في تفصيلات هذا الطرح المتشعب وجب تبرير اختيار اللغة الضخمة كنقطة التقاء في النهج الكمي والحاسوبي والنوعي، مصطلح اللغة تم اللجوء له عمدا بغض النظر عن النص أو الخطاب أو البلاغة، وذلك للنطاق الواسع الذي يفسحه في احتواء أشكال مختلفة من الاتصال المكتوب والمنطوق. وفي حين نعترف تماماً بأهمية هذه المجالات الفرعية، فإن النظري في مجال واحد فقط من شأنه أن يستبعد باحثي الاتصال الذين يستخدمون البيانات الضخمة لدراسة على سبيل المثال الاتصال بين الأشخاص والتنظيمات. بالإضافة إلى ذلك فإنه يخاطر بعزل ما يسمى بالدراسات الميدانية عن المحادثات المتخصصة الأكثر شيوعاً. وعلاوة على ذلك فإن التركيز على النص لن يكون مراعيًا للغة المنطوقة على الرغم من مجموعة متنامية من أبحاث العلوم الاجتماعية باستخدام البيانات الصوتية (Dagmar, Björn, 2021).

في حين أن تحليل اللغة من منطق البيانات الضخمة جديد نسبياً، فقد استخدم المهتمون بحقل الاتصال والإعلام منذ فترة طويلة كل من الأساليب النوعية والكمية لدراسة النصوص التي نشأت في وسائل الإعلام التقليدية ضمن الخطب السياسية والوثائق المكتوبة (بما في ذلك الإعلانات السياسية)، وأكثر من ذلك بكثير فقد ساعدت الأساليب الكمية لتحليل اللغة في تطوير نظريات مثل نظرية تحديد الأجندة والأطر، في حين توفر الأساليب النوعية القدرة على النظر في كيفية تضمين القوة في التواصل اللغوي والتعمق أكثر في تعقيدات الحياة الاجتماعية والاقتصادية والسياسية.

عند تنشيط بيانات اللغة الكبيرة، من الضروري أن يأخذ الباحثون في الاعتبار العلاقة بين النص ومعنى النص، خاصة إذا كانت الكلمة أو العبارة أو الجملة قادرة على استحضار معاني وتأويلات مختلفة لمتلقي الرسائل (James, 2004). وتختلف هذه المعاني الدلالية عبر المكان والزمان والسياق، وبالتالي فهي لا تتطلب خبرة في الموضوع فحسب بل تتطلب أيضاً القدرة البشرية على تمييز أنماط اللغة (على سبيل المثال، بين السخرية والمحاكاة الساخرة والتهكم، وما إلى ذلك). لذلك من الضروري أن يدرس الباحثون بعناية ما إذا كانت الأساليب المستخدمة في تحليل بيانات اللغة الضخمة قد تكون متحيزة دلالياً نتيجة لاختيار الكلمات الرئيسية. فعلى سبيل هذا الرافد، تتطلب الأساليب الحاسوبية التي تستخدم أساليب القاموس التقليدي (أي مجموعة من الكلمات الرئيسية) استخدام القواميس المؤصلة أحادياً*، والتي قد لا تتناسب مع الحالة أو السياق قيد البحث أو إنشاء قاموس مسبق، وكلاهما يعاني من توجهات على مستويات مختلفة. وقد أبرزت المناهج النقدية والثقافية في مجال الاتصال وغيره من التخصصات، أهمية النظر في تفسيرات وتأويلات متعددة للنص على أساس المعارف أو المواقف السياقية، وهو أمر غير ممكن باستخدام أساليب اختيار الكلمات الرئيسية، دون اغفال أن البعد التأويلي الذي نشأ وازدهر بعيداً عن الأساليب الكمية ضمن الدراسات

* عبارة عن مجموعة قواميس تستخدم المعنى الثابت غير القابل للتأويل والمقايضة الدلالية

الثقافية التقليدية والتي كانت دراسات تأسيسية لما آلت إليه التفرعات المشروعة في حقل الاعلام والاتصال وأصبحت نقطة التجاسر في البراديغم التأويلي.

1.2 تحليل بيانات اللغة الكبيرة وموانسة النهج التكراري مختلط الأساليب

نظراً للتحويل النقدي البارز في حقل الاتصال إلى جانب نقاط القوة والضعف لكل من الأساليب الكمية والنوعية، فقد أصبح لزاماً على المنشغلين في حقل الاتصال التأسيس لما يعرف بالتناهج في مجال البيانات الضخمة وأبرز سمة في التناهج هو الاعتماد على مناهج تكرارية مختلطة الأساليب. هذا الطرح لا يعني أنه لم تكن نقطة التقاء بين النهجين الكمي والكيفي من الناحية التاريخية فقد كانت موضع حدوث من خلال محاولة التعاون التي جرى الحديث عنها بين **Mills** و **Lazarsfeld** في السنوات الأولى من القرن العشرين (Michael, 2013). ومع ذلك لاحظ العلماء القبول المتزايد لأساليب المختلطة على الرغم من السرد السلبي حول الاختلاف الذي يلقي بظلاله غالباً على أوجه التشابه "بين المناهج الكمية والحاسوبية والنوعية، (Carpini, 2016). ونتيجة لهذا، فإننا نأخذ على محمل الجد تحذير **Ophir** و **Walter** و **Marchant** من ضرورة الاهتمام بشكل أفضل بنقاط القوة في النهج المختلط بدلاً من التركيز على نقاط الضعف الملموسة في كل نهج. ولتحقيق هذه الغاية فإنه من الضروري إبراز العلاقة بين الأساليب الكمية والحاسوبية والنوعية في تحليل بيانات اللغة الضخمة.

أولاً، هناك تكامل بين المناهج الحاسوبية والكمية النوعية من منظور انطولوجي ومعرفي. فمن الناحية الانطولوجية، تُظهر لنا المناهج الحاسوبية ما يوجد داخل مجموعة معينة من خلال أدوات مثل كشط الويب، ونمذجة الموضوعات، وتحليل الشبكات، في حين توفر المناهج النوعية مثل تحليل النصوص والإثنوغرافيا منظوراً بنائياً اجتماعياً للوجود والواقع والعالم الاجتماعي. أما من الناحية المعرفية توفر الأساليب الحاسوبية نهجاً "موضوعياً وتجريبياً" أكثر لتحليل النصوص، في حين توفر الأساليب النوعية تحليلاً ذاتياً وتفسيرياً أكثر للأيدولوجيات والثقافة التي تتموضع ضمن حجب النصوص، ففي لحظة أخذها معا كعدة منهجية مختلطة ستوفر صورة أكثر ثراءً وذات مغزى للأشخاص والمحتوى والمعنى الاجتماعي والثقافي مقارنة باستخدام طريقة واحدة فقط من هذه الطرق بمعزل عن غيرها.

ثانياً، يمكن استخدام الأساليب الحاسوبية جنباً إلى جنب مع الأساليب النوعية لأن "الباحثين النوعيين يسعون في الأساس إلى الحفاظ على الشكل والمحتوى وتجربة الفعل الاجتماعي وتحليله، بدلاً من إخضاعه للمخرجات الاحصائية أو غيرها من التحولات الشكلية. فالحديث الفعلي والإيماءات وما إلى ذلك هي المواد الخام للتحليل". ففي هذا التقليد البحثي لا يهتم الباحثون بتكرارات الأشياء بقدر ما يهتمون بالمعنى. وبالتالي توفر الأساليب الحاسوبية أدوات لجمع وتحليل وتفسير وتصور هذه المواد الخام التي يمكن أن تساعد الباحثين على فهم تحليلي لمجموعات البيانات الكبيرة والمعقدة والسريعة الحركة في كثير من المواقف المنهجية

في حقل الاتصال يمكن أن تساعد هذه الأساليب المختلطة الباحثين الذين يعملون مع المنصات عبر الإنترنت والذين يريدون فحص الممارسات الاتصالية التي ينشرها مستخدمو وسائل التواصل الاجتماعي (مثل منصة X وفيسبوك ويوتيوب) أو منتديات الويب (مثل ريديت) ضمن مجموعات كبيرة تحتوي على عشرات الآلاف أو حتى الملايين من المنشورات. يمكن استخدامها أيضاً لتحديد الأطر المشتركة التي تستخدمها إحدى وسائل الإعلام أو أكثر عبر عديد التقارير أو لتحليل نصوص البرامج التلفزيونية الطويلة الأمد بحثاً عن مواضيع مشتركة ومحو تمثيلي. في حين أن هناك العديد من الطرق لمزج الأساليب البحثية، فباحث الاتصال من منطلق مرونته القبلية وفهمه لطبيعة الظواهر الاتصالية الناشئة ستكون لديه العديد من الفرص للجمع بين الأساليب النوعية والكمية. على سبيل المثال يمكن للباحث الذي لديه مجموعة بيانات الاخبارية في

منشورات منصة **subreddit** إجراء تحليل استقرائي استكشافي باستخدام نهج حاسوبي مثل نمذجة الموضوع واستكمالها بتحليل نصي متعمق للعناصر الأكثر تمثيلاً إحصائياً في مجموعة البيانات كما تم تحديدها بواسطة نموذج الموضوع.

ومن الأمثلة على ذلك مجال الإنسانيات الرقمية المتاخم لحقل الاتصال حيث يتم نشر مناهج الأساليب المختلطة على نطاق واسع، إذ تشير مجموعة من الأبحاث ذات الصلة إلى قدرة الأساليب الحاسوبية على معالجة المخاوف النوعية المفهومة منهجياً وتقنياً. غالباً ما يقع مجال دراسات الاتصال على الخط الفاصل بين العلوم الإنسانية والعلوم الاجتماعية، ومع ذلك عندما يتم تجيله من منظور حاسوبي نقدي يحافظ على الالتزامات المعيارية، يمكن لدراسات الاتصال أن تشارك في نقد تحويلي باستخدام الأساليب والأدوات والممارسات الناشئة (Lothian, Phliips, 2015)، بما في ذلك مناهج الأساليب المختلطة. لقد اهتم باحثو العلوم الإنسانية الرقمية الذين يستخدمون الأساليب الحاسوبية منذ فترة طويلة بالقضايا التي تهم العلماء النقيدين والنسويين واستعانوا بهذه النظريات للتحليل الاستقرائي وبناء النظريات. لقد تم تطبيق هذا النهج التكراري المختلط الأساليب في مجال الاتصالات بشكل غير مستديم وبوضعية متخمة بالتوجس، على الرغم من أن جوهره يتمتع بالعديد من المزايا المهمة لتحليل بيانات اللغة الكبيرة الهامة في مجالات الاتصالات ودراسات الإعلام.

2.2 مواضيع الأساليب المختلطة في بيانات اللغة الضخمة

كما تم تناوله آنفاً يتيح حقل الاتصال _ كونه يتوسط الدراسات الإنسانية والاجتماعية والرقمية _ للباحثين ميزة تمكنهم من الاستفادة من الأساليب التقليدية والحاسوبية كاستخدام أساليب مثل تحليل الخطاب وتحليل المحتوى (Klaus, 2019)، لدراسة مجموعة واسعة من الاتصالات البشرية، بما في ذلك التفاعلات الشخصية (Joseph, 2017)، أو كانت خطاباً (Marco, 2008)، أخبار (Fatemeh, 2021)، عملاً إبداعياً (ennifer, and others, 2010) بالإضافة إلى الإعلانات. وفي الوقت نفسه قد يكون من الصعب استخدام هذه الأساليب التقليدية، نظراً للكُم الهائل من بيانات اللغة التي تنتجها المجتمعات الآن، ومن الناحية العملية ليس من الممكن تصنيف هذه الكمية من البيانات يدوياً لصعوبة محاكاتها بشرياً (راحو، 2022).

ولحسن الحظ، فقد أدرك المنشغلون في حقل الاتصال تاريخياً الحاجة إلى توسيع الأساليب التي يستخدمونها لدراسة التواصل اللغوي (Klaus Krippendor). في العصر الرقمي يشمل هذا الاستخدام أدوات البرمجة والكمبيوتر، وينطبق هذا على الباحثين الكميّين والحاسوبيين الذين يستخدمون Python كما ينطبق على الباحثين النوعيين، الذين يعتمد العديد منهم على برامج مثل MAXQDA و Nivivo (R and Python) لنسخ وترميز الميزات الخطابية المختلفة يدوياً. كما ساعد توافر هذه الأدوات على تيسير دمج مناهج التحليل اللغوي الكمي والنوعي. على سبيل المثال، يمكن نقل الرموز المشتقة نوعياً والمخزنة في أداة التحليل النصي النوعي "QDA Miner" إلى "برنامجها الشقيق"، وهو البرنامج الكمي Wordstat. وقد أدت شعبية هذه الأدوات أيضاً إلى إثارة الاهتمام بالبرمجيات المجانية والمفتوحة المصدر، مما يجعل تحليل اللغة بمساعدة الكمبيوتر أكثر سهولة في الوصول إليه (Lindlof & Taylor, 2018).

تتمثل إحدى مزايا هذه الأساليب الحاسوبية في القدرة على توسيع نطاق تحليلات المحتوى التقليدية باستخدام البيانات المصنفة وبواسطة العنصر البشري لتهيئة خوارزميات التصنيف (Damian, Jeroen, 2018). هناك العديد من التطبيقات التي تتألف مع كيفية اكتشاف أولوية سمة لغوية معينة. ومن التطبيقات الشائعة لهذا التحليل العاطفي: استخدام المصنفات الخوارزمية لتحديد الرسائل باعتبارها ذات

قيمة إيجابية أو سلبية. كما أن الأساليب غير الخاضعة للإشراف مثل نمذجة الموضوع تجعل من الأسهل استكشاف البيانات من خلال، من بين أمور أخرى تسليط الضوء على المعلومات الأكثر تمثيلاً إحصائياً في مجموعة بيانات معينة للتحليل (Daniel and others, 2018).

ومن الجدير بالذكر أن الأساليب الحاسوبية اختزالية بطبيعتها، ففي التعامل مع النص واللغة كمصدرين للبيانات المراد استخراجها من أجل تحديد سمات لغوية محددة باعتبارها الأساليب الحاسوبية مثل التعلم الآلي الخاضع للإشراف لا تهدف إلى فهم اللغة وتطويع معانيها، بل تهدف إلى تحديد "إشارات" محددة ومتكررة مضمنة في التواصل اللغوي. ولهذا السبب يشكل اختزال اللغة خطوة أساسية في المعالجة المسبقة لهذه التحليلات (Justin, Brandon , 2013). وهذا يمثل مأزقاً بالنسبة لظواهر اللغة المعقدة مثل اكتشاف السخرية والمفارقة (Christian and others, 2022)، ولكنها قد تساعد في تحديد سمات اللغة ذات المعنى التصريحي. بعبارة أخرى كل نماذج اللغة سيئة المفترضة لا تبطل فكرة أنها بعضها مفيد.

إن إحدى الطرق التي يمكن للباحثين من خلالها زيادة فائدة هذه الأساليب هي إشراك العنصر البشري في العملية التحليلية في كافة أطوارها. وفي علوم الحاسوب (HITL) تسمى هذه الأساليب بـ "الإنسان كجزء من الحلقة" بحيث تظهر تقييمات تفيد بأن البشر ضروريون للتحقق من صحة أي نتائج حاسوبية. وعلى نحو مماثل فإن أبحاث الاتصال الحاسوبية التي تدرس بيانات اللغة الضخمة هي أبحاث سايبروتيكية*. تتألف من عمل حاسوبي وبشري. وتماشياً مع طرح D'Ignazio و Klein: "فإنه من الإيجابي جعل العمل البشري مرئياً وشفافاً طوال العملية" (d'Ignazio, Klein).

لإبراز قيمة كل من مناهج تحليل اللغة التقليدية والحاسوبية، سيكون الخط المنهجي عبارة عن مزيج مترافق ضمن النهج الممكنة التي تم ذكرها سابقاً، مع أجزاء استقرائية واستنتاجية متجاسرة تشابكا أو ترادفاً. وفي هذا الصدد تم استحداث خمس خطوات للتناهج المختلط كالتالي:

- تركز الخطوة الأولى على استكشاف بيانات النص، فغالبا ما تكون هذه الخطوة وصفية واستقرائية، وهي عملية أساسية للباحث لفهم ما هو موجود في بيانات اللغة تقليدياً، قام المنشغلون في الظاهرة الاتصالية بتنفيذ هذه الخطوة باستخدام مناهج نوعية مثل النظرية الأساسية والتميز النوعي وتحليل الخطاب. ومع ذلك سلطت الدراسات الحديثة الضوء أيضاً على قيمة استخدام الأساليب الحسابية الاستكشافية مثل التعلم الآلي غير الخاضع للإشراف وتحليل الشبكات للتركيز على النص الأكثر تمثيلاً إحصائياً في مجموعة معينة. وتشترط هذه الخطوة بالضرورة أيضاً، كما يقترح D'Ignazio و Klein إشراك خبراء في الموضوع من البشر لوضع المشروع البحثي في السياق الأوسع الذي يتم فيه تنفيذه (d'Ignazio, Klein).
- بمجرد استكشاف بيانات النص، فإن الخطوة الثانية بالنسبة للباحثين هي تحديد الأنماط ذات المغزى التي تثير اهتمامهم. يجب أن يكون هذا النهج مدفوعاً بسؤال بحثي وهو الخطوة الأولى لترجمة التعريف المفاهيمي وتنفيذه. خلال هذه المرحلة، سيستفيد الباحثون كثيراً من كتابة المذكرات، وهي العملية التي غالباً ما ترتبط بتميز النظرية الأساسية النوعية (Melanie and others, 2008). ومع ذلك فإن كتابة المذكرات تفيد أيضاً الباحثين الذين يستخدمون الأساليب الكمية والحاسوبية، حيث تعد المذكرات سجلاً للأساليب المنهجية المتبعة وانطباعات الباحثين في مراحل مختلفة من التحليل. كما يمكن أن تعمل كتابة المذكرات على تعزيز قدرتنا على إعادة التفكير في الثنائيات التي قد يتم دمجها في جمع البيانات وتصنيف الأساليب كمية، فضلاً عن تبني التعددية في ظل التصارع

* أساليب يتم فيها دمج المهارات البشرية والقدرة التفسيرية للعقل الإنساني مع القدرة غير المحددة للأدوات الحاسوبية

المفترض مع المواقف الخاصة والظهور المحتمل لنقاط بيانات متناقضة ظاهرياً أو "حقائق متعددة" (d'Ignazio, Klein).

- الخطوة الثالثة هي إنشاء كتاب تعليمات برمجية للترميز البشري اليدوي. تتوافق هذه الخطوة بشكل وثيق مع تحليل المحتوى التقليدي، من خلال تحديد العلامات ذات المعنى للترميز والحصول على موثوقية المبرمجين، ثم ترميز المحتوى يدوياً (Kimberly, 2017). ومع ذلك فإن الفارق الرئيسي بين هذه الخطوة وتحليل المحتوى التقليدي هو كيفية استخدام هذه العلامات اليدوية، ففي تحليل المحتوى التقليدي تكون هذه العلامات هي البيانات، بينما في النهج المختلط تكون هذه العلامات عبارة عن عينة يتألف معها الحاسوب بعد ذلك.

- الخطوة الرابعة هي استخدام العلامات البشرية التي تم إنتاجها في تحليل المحتوى لبناء مصنف نصي للتعليم الآلي الخاضع للإشراف. يعد نهج النمذجة الحاسوبية ذو فاعلية عالية للتوسع في تحليل المحتوى الذي تم إجراؤه في الخطوة السابقة وغالباً ما يكون ضرورياً لتحليل بيانات اللغة الكبيرة على نطاق واسع، حيث من غير المرجح أن يتم إجراء تحليل المحتوى على مجموعة بيانات لغة كبيرة بالكامل. على نحو متزايد أصبح التعلم الآلي الخاضع للإشراف حلاً شائعاً لدراسة لغة وسائل التواصل الاجتماعي بسبب كمية الرسائل المنتجة (Ward, Toni, 2016). ومع ذلك، قام الباحثون أيضاً بتطبيق التعلم الآلي الخاضع للإشراف على بيانات الأخبار (Björn, Daan Odijk, 2014). ورغم أن هذا النمذجة الحاسوبية يشكل غالباً النقطة المحورية في الورقة الأكاديمية، غير أنها لا تعتبر إجراء بحثياً ختامياً.

- الخطوة الأساسية الخامسة هي تقييم جودة المصنف، والمحدد أن يتم ذلك باستخدام مزيج من الأساليب الكمية والنوعية. على سبيل المثال قارن Claire وزملاؤه بين تصنيف بيانات النصوص بواسطة الباحثين البشريين وتصنيفات التعلم الآلي، ووجدوا أن تصنيفات التعلم الآلي كانت أقل فائدة وأقل موثوقية في تحديد الأشكال المعقدة للغة (Claire and others, 2018).

إن الأساليب النوعية مفيدة في هذه الخطوات، ولكنها قد تكون مفيدة بشكل خاص للباحثين في مجال الاتصالات الحاسوبية، حيث يتم التحقق من صحة نتائجهم وإعادة وضع نتائجهم في سياق نظام التواصل. وكما لاحظنا، فإن الأساليب الحاسوبية اختزالية بطبيعتها فهي ممتازة في تحديد سمات لغوية معينة، ولكنها أقل فائدة لفهم كيفية وضع هذه السمات داخل الموقف أو الظاهرة الاتصالية الأوسع. إن العودة إلى التحليلات النوعية في النهاية يمكن أن تعزز فهم الباحث لتلك السمات اللغوية.

إحدى الطرق للقيام بذلك هي تطبيق نهج Fairclough لتحليل الخطاب (Fairclough, 1985). ويمكن القيام بذلك من خلال التعامل مع المصنف باعتباره تحليلاً على مستوى النص الجزئي، ولكن مع وضع هذه النتائج في سياق الاعتبارات على المستوى المتوسط والعوامل المجتمعية (المستوى الكلي). وفي حين أن هذه العناصر ذات صلة طوال عملية البحث، فإن هذه الخطوة الأخيرة لها أيضاً فائدة توفير مساحة للباحثين لفحص السلطة، وتحديد الطرق الممكنة لتمكين المعرفة في شقها المجتمعي (d'Ignazio, Klein).

II. خاتمة

في ختام هاته المحاولة البحثية وجب الحديث على أن الاستفاضة في النقاش حول القوة الحاسوبية للبيانات الضخمة وموقعها من الموقف العلمي الراهن، يتبلور من الجدل الفلسفي الذي أثاره الباحث Dusek

(Dusek, 2006) ، والذي بموجبه تشكل جميع وسائل الاتصال التكنولوجية التجارب الاجتماعية إلى حد ما إذ يقترح ضرورة إعادة التفكير جذريا في النظرة الكلاسيكية لهيمنة القوة البشرية في عمليات التواصل من منطلق التشكيك في وجهة النظر الموضوعية للتكنولوجيا، وعليه فتطويع البيانات الضخمة كعدة منهجية لن يمر إلا بتبرير انطولوجي لحضورها في بنية الاتصال الرقمي.

الإحالات والمراجع:

المقالات

1. راحو محمد عبد الوهاب، تحليل البيانات الضخمة من شبكات التواصل الاجتماعي كعامل لبناء معرفة تسويقية، مجلة المدير، المجلد 09، جامعة المدرسة العليا للتسيير والاقتصاد الرقمي، الجزائر، 2022، نص 280.
2. شاوش اخوان جهيدة، تحليل البيانات الضخمة ومستقبل الدراسات الاجتماعية، المجلد 22، العدد 02، مجلة العلوم الانسانية، جامعة بسكرة، 2022، ص 537.
3. عيادنة محمد لمين، العلاقات العامة 2.0 في زمن البيانات الضخمة نحو اتصال استراتيجي للعلامة التجارية تحكمه الخوارزميات، مجلد 27، عدد 55، جامعة الأمير عبد القادر قسنطينة، الجزائر، 2023، ص 782.
4. لعياضي نصر الدين، هل ستقضي المناهج الحاسوبية على نظريات علوم الاعلام والاتصال، مجلة رقمنة للدراسات الاعلامية والاتصالية، العدد 2، جامعة الجزائر 3، 2021، ص 21.

المراجع الاجنبية

5. Abbott, A, Transcending general linear reality. *Sociological Theory*, 6, 1998, pp 169-186.
6. Amaral, I, Participação em rede: do utilizador ao "consumer 2.0" e ao "prosumer". *Comunicação e Sociedade*, 2012, pp 22-131-142.
7. Björn Burscher, Daan Odijk, Rens Vliegthart, Maarten De Rijke, and Claes H. De Vreese, "Teaching the Computer to Code Frames in News: Comparing Two Supervised Machine Learning Approaches to Frame Analysis," *Communication Methods and Measures* 8, no. 3, 2014, pp 190-206.
8. Boyd, D. & Crawford, K, Critical questions for Big Data. *Information, Communication & Society*, 15(5), 2012, pp 662-679.
9. Braga, J. L, A sociedade enfrenta sua mídia. Dispositivos sociais de crítica midiática. São Paulo: Paulus, 2006.
10. Carpinì and Delli, Breaking Boundaries; Catherine d'Ignazio, and Lauren F. Klein, "Feminist Data Visualization," Workshop on Visualization for the Digital Humanities (VIS4DH) (Baltimore. IEEE, 2016.
11. Christian Baden, Christian Pipal, Martijn Schoonvelde, and Mariken AC G. van der Velden, "Three Gaps in Computational Text Analysis Methods for Social Sciences: A Research Agenda," *Communication Methods and Measures* 16, no. 1, 2022, pp 1-18
12. Claire Lauer, Eva Brumberger, and Aaron Beveridge, "Hand Collecting and Coding Versus Data-Driven Methods in Technical and Professional Communication Research," *IEEE Transactions on Professional Communication* 61, no. 4, 2018, pp 389-408.
13. Coté, M., Gerbaudo, P. & Pybus, J, Introduction: politics of big data. *Digital Culture & Society*, 2(2), 2016, pp 5-15.
14. Cukier, K. & Mayer-Schönberger, V, Big Data: a revolution that will transform how we live, work and think. Nova Iorque: Houghton Mifflin Harcourt, 2013.
15. Dagmar M. Schuller, and Björn W. Schuller, "A Review on Five Recent and Near-Future Developments in Computational Processing of Emotion in the Human Voice," *Emotion Review* 13, no. 1, 2021, pp 45-50.
16. Damian Trilling, and Jeroen GF Jonkman, "Scaling up Content Analysis," *Communication Methods and Measures* 12, no. 2-3, 2018, pp 47-158.
17. Daniel Maier, Annie Waldherr, Peter Miltner, Gregor Wiedemann, Andreas Niekler, Alexa Keinert, Barbara Pfetsch et al, "Applying LDA Topic Modeling in Communication Research: Toward A Valid and Reliable Methodology," *Communication Methods and Measures* 12, no. 2-3, 2018, pp 93-118.
18. Dusek, V, Philosophy of technology: an introduction. Malden, MA: Blackwell Publishing, 2006, p 205.
19. Jennifer A. Manganello, Vani R. Henderson, Amy Jordan, Nicole Trentacoste, Suzanne Martin, Michael Hennessy, and Martin Fishbein, "Adolescent Judgment of Sexual Content on Television: Implications for Future Content Analysis Research," *Journal of Sex Research* 47, no. 4, 2010, pp 73-364.
20. Fatemeh Torabi Asr, Mohammad Mazraeh, Alexandre Lopes, Vasundhara Gautam, Junette Gonzales, Prashanth Rao, and Maite Taboada, "The Gender Gap Tracker: Using Natural Language Processing to Measure Gender Bias in Media," *PloS One* 16, no. 1, 2011.
21. Fausto Neto, A, Comunicação das organizações: da vigilância aos pontos de fuga. In I. Oliveira & A. Soares (Eds.), *Interfaces e tendências da comunicação no contexto das organizações*, 2011, pp 43-67.

22. Gil de Zúñiga, H, Citizenship, social media, and big data: current and future research in the social sciences. *Social Science Computer Review*, 2015, pp 1-2-7.
23. Hall, P. A, Aligning ontology and methodology in comparative research. In J. Mahoney & D. Rueschemeyer (Eds.), Cambridge: Cambridge University Press, *Comparative historical analysis in the Social Sciences*, 2003, pp 373-404.
24. Helbing, D, The automation of society is next. How to survive the digital revolution. Charleston: CreateSpace Independent Publishing Platform; 2015.
25. James Paul Gee, *An Introduction to Discourse Analysis: Theory and Method*, Routledge, 2004
26. Joseph N. Cappella, "Vectors into the Future of Mass and Interpersonal Communication Research: Big Data, Social Media, and Computational Social Science," *Human Communication Research* 43, no. 4, 2017, pp 58-545.
27. Justin Grimmer, and Brandon M. Stewart, "Text As Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," *Political Analysis* 21, no. 3, 2013, pp 97-267.
28. Kimberly A. Neuendorf, *The Content Analysis Guidebook*, Sage, 2017.
29. Kitchin, R, Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, April-June, 2014, pp 7-12.
30. Kitchin, R, Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 2017, pp 14-17-21-29.
31. Klaus Krippendorff, "The Changing Landscape of Content Analysis: Reflections on Social Construction of Reality and Beyond," *Communication & Society* 47, 1, 2019, p1
32. Kuhn, T, *A estrutura das revoluções científicas*. São Paulo: Perspectiva, 1998.
33. Lindlof & Taylor, *Qualitative Communication Methods*, Psychology, Thousand Oaks, CA: Sage, Vol.9 No.13, December 6, 2018, p 41.
34. Lothian and Phlips, *Transformative Critique*; Roopika Risam, "Beyond the Margins: Intersectionality and the Digital Humanities," *Digital Humanities Quarterly* 9, no. 2, 2015.
35. Luhmann, N, *Social systems*. Stanford, CA: Stanford University Press, 1995.
36. Luhmann, N, *A improbabilidade da comunicação*. Lisboa: Vega, 2006.
37. Marco Guerini, Carlo Strapparava, and Oliviero Stock, (2008), "Corps: A Corpus of Tagged Political Speeches For Persuasive Communication Processing," *Journal of Information Technology & Politics* 5, no. 1, 2008, pp 19-32.
38. Martino, L. C, *Atualidade mediática: o conceito e suas dimensões*. Comunicação apresentada no XVIII Compós - Encontro da Associação Brasileira de Programas de Pós-graduação em Comunicação, Belo Horizonte, 2009.
39. Melanie Birks, Ysanne Chapman, and Karen Francis, "Memoing in Qualitative Research: Probing Data and Processes," *Journal of Research in Nursing* 13, no. 1, 2008, pp 98-75.
40. Michael X. Delli Carpini, "Breaking Boundaries: Can We Bridge the Quantitative Versus Qualitative Divide Through the Study of Entertainment and Politics?," *International Journal of Communication* 7, 2013, p21.
41. Milan, S. & van der Velden, L, The alternative epistemologies of data activism. *Digital Culture & Society*, 2(2), 2016, pp 57-64.
42. Norman L. Fairclough, "Critical and Descriptive Goals in Discourse Analysis," *Journal of Pragmatics* 9, no. 6, 1985, pp 63-739.
43. Procter, R., Vis, F. & Voss, A, Reading the riots on Twitter: methodological innovation for the analysis of big data. *International Journal of Social Research Methodology*, 16(3), 2103, pp 197-214.
44. R and Python are free, open-source programming languages commonly used in computational methods like machine learning. MaxQDA and Nvivo are paid data-analysis software packages commonly used by qualitative and mixed-method researchers ,
45. Sbardelotto, M, *Circulação em rede: a comutabilidade dos polos de produção e recepção no fluxo comunicacional digital*. Comunicação apresentada no XXVI Compós - Encontro da Associação Brasileira de Programas de Pós-graduação em Comunicação, São Paulo, 2017, pp 4-19.
46. Shaldbolt, N., Hall, W. & Berners-Lee, T, The semantic web revisited. *IEEE Intelligent Systems*, 2006, pp 96-101.
47. Uricchio, W, Data, culture and the ambivalence of algorithms. In M. T. Schäfer & K. van Es. (Eds.), *The datafied society, studying culture through data*, Amsterdam University Press, 2017, pp 125-137.

48. Ward van Zoonen, and G. L. A. Toni, (2016), "Social Media Research: The Application of Supervised Machine Learning in Organizational Communication Research," Computers in Human Behavior 63, pp 41-132.