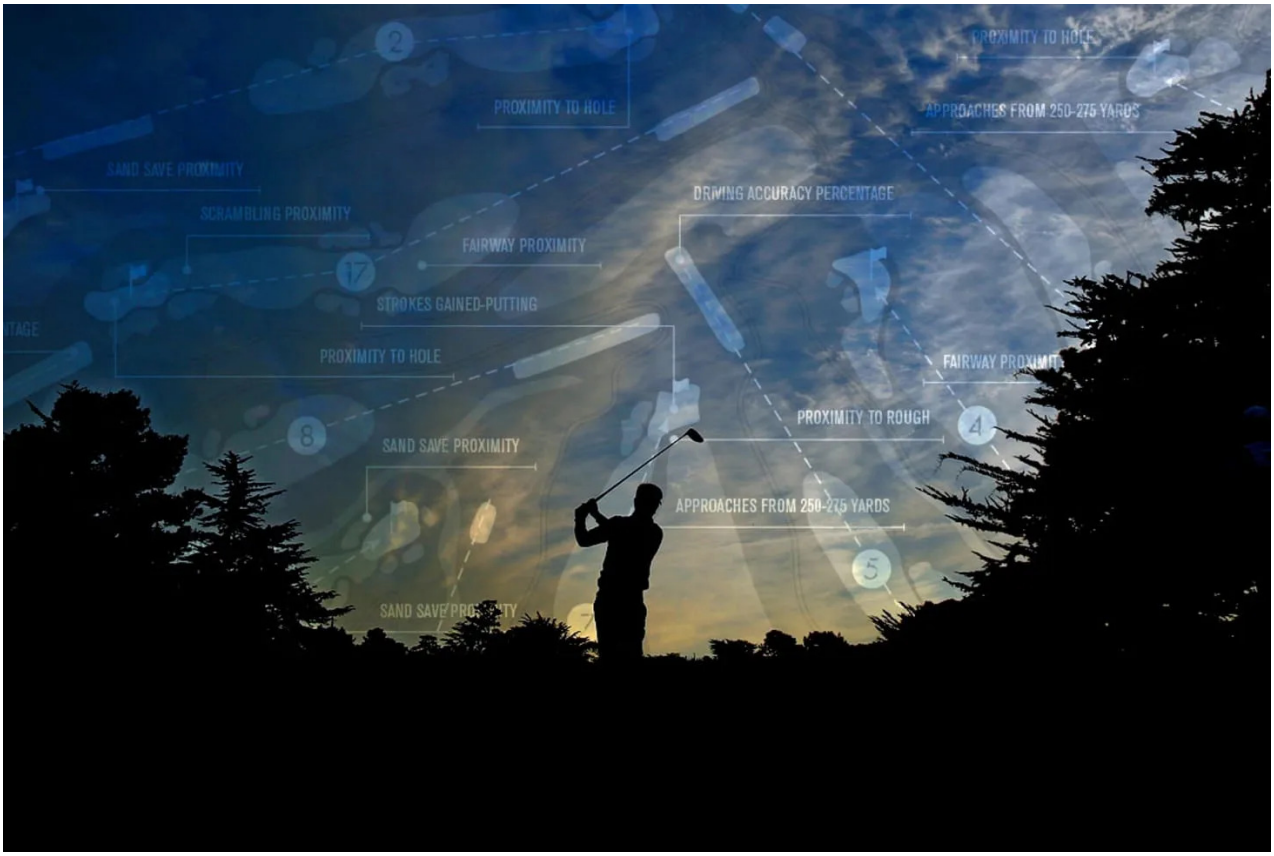
الصحافة في عصر البيانات الضخمة



عندما نتحدث عن البيانات الضخمة، فنحن نتحدث عما يدخل ضمن وحدة القياس إكسابايت (exabyte) وأعلى التي، تيرابايب مليون أو جيجابايت مليار منها الوحدة تعادل التي، ما نتحدث عنه يفوق قدرة ليس فقط البشر على التحليل والمعالجة، ولكن كذلك معظم الحواسيب والنظم التقليدية، فوحدها الأنظمة المتخصصة بمعالجة البيانات الضخمة هي ما يمكن أن تحلل هذه الكمية المهولة من المعطيات والأرقام وتحولها لصورتها الأخيرة التي تمكننا من الاستفادة منها.

وعندما يتعلق الأمر بالصحافة، فالإثارة ذاتها كما تبدو في بقية المجالات التي وجدت الـData Big طريقها إليها، مثل الصحة والتعليم والأحوال الجوية والأنشطة التجارية والقطاع الحكومي وغير ذلك.

تكمن حساسية الموقف عندما نتحدث عن علاقة البيانات الضخمة بالصحافة، بخطورة محتملة عند تجميع هذا الكم الهائل من البيانات القادر عبر استخدامه إعلاميًا، على التحكم بالرأي العام، كما كشفته فضيحة "كامبريدج أناليتيكا"، وهي الشركة الوسيطة التي أتاحت لها شركة فيسبوك إمكانية الوصول لبيانات مهولة لملايين المستخدمين، ما مكنها عقب فرزها وتحليلها من معرفة توجهاتهم السياسية والفكرية ومحاولة التأثير عليها وتوجيهها في الانتخابات الأمريكية 2016.

لكن إذا تجاوزنا هذه الفضيحة، وفي ظل غياب الدراسات الموسعة حتى الآن بشأن الأدوار المحتملة للبيانات الضخمة في الصناعة الصحفية، يمكن الحديث عن بعض الأدوار التي بدأت بعض المؤسسات الصحفية بممارستها بالفعل، وعلى رأس تلك الأدوار:

تخصيص المحتوى وتحسين تجربة المستخدم

تراقب الخوارزميات سلوك المستخدم عبر المنصة طوال الوقت وتسجل كل حركة لإصبعه أو للمؤشر،

إنها كمية مهولة أخرى من المعلومات التي تجمعها وسائل الإعلام دون شعورنا. تتعلم الخوارزميات من المستخدم أي الأخبار والقصص تستهويه ويهتم لها، وأيها لا، أيها يتفاعل معها وأيها يلقي عليها نظرة خاطفة، وتلك التي لا يعبأ بها، أي أساليب العناوين تجذبه، أي التصنيفات يتابع، هل يقرأ الخبر كاملًا أم يكتفي بملخص؟ هل يشاهد الفيديو كاملًا؟ وهل يقلب ألبوم الصور؟ هل يضغط على الروابط؟ ماذا بشأن الإعلانات؟

الآن بعد أن تقوم الخوارزميات بجمع وفرز وتحليل تلك البيانات، باتت جاهزة لتقدم للمستخدم ما تتوقع أنه يريده تحديدًا، تبعًا لاختياراته السابقة واهتماماته طيلة فترة التعلم.

خلق صحافة روبوتية

"صحافة لا تمسها الأيدي"، تضطلع بها الخوارزميات، بدءًا من عملية جمع البيانات مرورًا بفرزها وتحليلها، وانتهاءً بتقديمها على صورة قصص يسردها روبوت مدرب بعد أن أحال إلى العطالة عددًا من المحررين البشريين البلداء الذين بالكاد ينتج الواحد منهم بضع قصص في اليوم، فيما ينتج هذا الآلي المئات منها بلا تأفف!

لا يكتفي الروبوت بسردد القصص التي بناها مما يجمع من بيانات، بل يستطيع التفاعل مع القارئ، أو لنقل مع ملايين القراء بالتزامن، ويستطيع بفضل الذكاء الصناعي ليس فقط تلبية حاجهم للمعرفة والمعلومات والأخبار، ولكن أيضًا يفعل ذلك بطريقة تفاعلية مؤثرة.

زيادة الإيرادات

لا يوجد ذكاء صناعي دون بيانات ضخمة، وتوفير هذه الأخيرة عن مستخدمي المنصة الإعلامية وقرائها وعاداتهم الانقرائية وسلوكهم كمستخدمين من لحظة فتح المنصة (تطبيق أو موقع إلكتروني) إلى لحظة خروجهم، وتفضيلاتهم وما أثار اهتمامهم وما نفروا منه، ثم الإحاطة بكل ذلك وتحليله وتقديم خلاصة عنه، يساعد صناع القرار على اتخاذ الخطوة التالية نحو صناعة ولاء الجمهور وزيادة الوقت الذي يمضيه على المنصة، من خلال تحسين تجربة المستخدم وتطوير الواجهة إلى النحو الذي يستهوي القراء ويدفعهم للبقاء أكثر والقراءة أكثر، وبالطبع النقر على الإعلانات واتخاذ إجراءات أخرى كالاشتراك أو شراء منتج ما.

زيادة الإيرادات المدعومة من الذكاء الصناعي يمكن أن تنعكس على استقلالية الصحافة، فوسيلة إعلام تمول ذاتها وتحافظ على ولاء قرائها وتزيد نسبة مشتركيها أو مرتاديها، قادرة على التماسك أمام عواصف الضغط أكثر من تلك التي تعتمد على تمويل الحكومات أو رجال الأعمال النافذين.

إذن، فنحن لا نتحدث عما بات يُعرف بصحافة البيانات التي تركز على إعادة تقديم الأرقام والإحصاءات بصريًا عبر أشكال وخرائط وإنفوجرافيك وجداول وغيرها، وبالتالي البيانات بحجمها الصغير التي يمكن للصحفيين الأفراد أو الأطقم الصغيرة التعامل معها بسهولة، بل نتحدث عن عالم مختلف، يسود فيه الروبوت والخوارزميات والذكاء الصناعي، ويشاهد فيها الأفراد مكتوفي الأيدي تقريبًا!

على أن صحافة البيانات هذه ستشهد طفرة مذهلة مع دخولها عالم البيانات الضخمة.

نذهب الآن إلى تجارب عملية قامت بها أربع وسائل إعلام كبرى، استغلت الكمية المهولة من البيانات التي لديها عن عملائها ومستخدميها، ثم أعادت تقييم تجربتها بناء على ذلك، فأحيانًا قدمت منتجات جديدة، وأحيانًا أخرى ألغت بعض المزايا، وفي بعض الأحيان عدلت في بعض برامجها.

بي بي سي: تحليل وجوه المشاهدين

لدى "بي بي سي"، مئات العروض والبرامج التليفزيونية، وترى أن استطلاعات الرأي ليست أفضل طريقة

لتعرف حقيقة رضى الجمهور عنها، وما إن كان برنامج تليفزيوني ينال استحسان المشاهدين أم لا على وجه الدقة، وبهذا الخصوص يقول دايفيد بويلي نائب الرئيس التنفيذي للبحث في شركة بي بي سي العالمية:

"أخبرنا البعض أنهم لم يكونوا متحمسين لبعض البرامج، رغم كونهم متحمسين على المستوى العاطفي والداخلي"، وأضاف "لا يريد أحد الاعتراف بإعجابه بالعروض التليفزيونية، ولكن الأطفال يتفاعلون مع تلك العروض".

من هنا، ابتكرت بي بي سي أداة لتحليل انفعالات الجمهور تعتمد على البيانات الضخمة، وتكمن مهمتها بالتعرف على تعبيرات الوجه للمشاهدين في أثناء العروض التليفزيونية.

حلّلت بي بي سي، وفقًا لكتاب Practice in Data Big لـ Marr Bernard، نحو 50 عرضًا تليفزيونيًا في 4 دول مختلفة، باستخدام كاميرات مصممة لتسجيل أوجه المشاهدين وتفسير عواطفهم، وبمساعدة شركة CrowdEmotion، وهي شركة متخصصة بقياس وملاحظة التغيرات التي تطرأ على تعبيرات الوجه، وتمكنت من تسجيل تعبيرات الوجه للمشاهدين في أثناء مشاهدتهم للعروض لتحديد الأجزاء التي تثير فضولهم وتلمس اهتماماتهم.

تتعلم الخوارزميات وتتعرف على وجه كل مشاهد للتمييز بين انفعال الوجه في الحالة الطبيعية وفي حالة التفاعل العاطفي مع العروض التليفزيونية، ثم تقوم بقياس تفاعل واحد مثل السعادة أو الخوف، وتقوم أيضًا بالعمل على تطوير قياسات أكثر تعقيدًا مثل "التفاعل العاطفي الكلي" أو "السرحان العاطفي"، وبمقدور CrowdEmotion تحليل نحو 2.5 مليون وجه في الدقيقة عبر السحابة.

بمقارنة بيانات التفاعل الوجهي الضمني بآراء العملاء الصريحة، تمكنت بي بي سي من تحديد أي العروض كانت ناجحة، وأيها تحتاج إلى إعادة تأهيل، ففي تجربة واحدة، عدد من الأسر في أستراليا تم تسجيلهم للحصول على ردود فعل المشاهدين عندما شاهدوا إعلان الموسم الأول من برنامج تشيرلوك.

اكتشف الباحثون من هذه التجربة في أستراليا أن المشاهدين الذين استمروا في إعطاء العرض تقييم عالٍ أظهروا ردة فعل أكبر في المشاهد على الشاشة التي يمكن وصفها بأنها مفاجئة أو حزينة، أكثر من كونها مضحكة، هذا دفع منتج البرنامج في بي بي سي ليجعله أكثر ظلمة وإثارة وأقل كوميدية.

فايننشال تايمز: اشتراكات الموقع تفوق اشتراكات الصحيفة

الجريدة التي يزيد عمرها على قرن وربع، نجحت بفضل البيانات في نمو وزيادة الاشتراكات عبر الإنترنت من خلال فهم العميل واهتماماته، ومن ثم خدمته بشكل أفضل، وباستخدام البيانات الضخمة التي تم جمعها، تمكنت مجلة فايننشال تايمز من خلق طابعها الخاص في القراءات والمحتويات التي تصل إلى العملاء والقراء، وباتت الاشتراكات في موقع com.FT تفوق الاشتراكات في مجلة FT، طبقًا لمدير قسم البيانات في فايننشال تايمز توم بيت.

تمكنت فايننشال تايمز من خلال البيانات الضخمة من تنظيم حملات إعلانية مخصصة بشكل أفضل، بالإضافة إلى تطوير منتجات جديدة بالاعتماد على تلك البيانات التي توضح اهتمامات القراء.

ساعدها جمع البيانات وتحليلها في:

1. فهم ما يفضله القارئ من محتويات.
2. تحسين العلاقة بين العملاء والجريدة.
3. تخصيص المحتويات للمستخدمين.

4. تقديم ما يتوافق مع اهتمامات العملاء.

وفي حين واجهت الجريدة العريقة مشكلة بوجود حلقة مفقودة بين وقت النشر ووقت القراءة، أمكن للذكاء الصناعي وتحليل البيانات الضخمة ردم هذه الهوة وسد تلك الثغرة، عبر معرفة ما يلائم القراء سواء لناحية وقت النشر أم الاهتمامات والمحتوى.

هاف بوست: البيانات هي الملك

تعتبر مجلة الهافنغتون بوست أحد أكثر من يطوّع البيانات الضخمة لصالح نموه وتوسعه، وهي تؤمن بأن "البيانات كل شيء"، على ما قال مسؤول بارز فيها لتقرير رويترز، استخدمت HuffPost البيانات الضخمة لتحسين المحتوى وتوثيق التعليقات وتحقيق الفعالية المطلوبة من الإعلانات وتنظيم عملية نشر ووضع الإعلانات وإضفاء الطابع الشخصي.

قد لا يعد سبقًا أن تستخدم هاف بوست، البيانات الضخمة من أجل التعرف على اهتمامات متصفحيها عبر مراقبة سلوكياتها داخل مواقعها وتطبيقاتها، بغية تخصيص المحتويات المناسبة للجمهور وتقديم ما يجذبه، الواقع أن هذا الاستخدام للـData Big سيغدو سمة لمعظم وسائل الإعلام الكبيرة، وربما المتوسطة أيضًا بغضون سنوات قليلة، لكن الهاف لم تقف هنا، إذ بالإضافة لذلك عملت على الاستفادة من البيانات الضخمة حتى في الجانب التحريري، فضلًا عن ضبط عملية التفاعل المسيء، ومن ذلك:

العناوين: اختبرت هاف بوست التأثير الإيجابي للعناوين الرئيسية في جذب الكثير للقراءة، عبر تجربة لاستهداف الوالدين عبر مقال "10 طرق تجعل العيش مع الطفل الصغير كالعيش في السجن"، وعبر اختبار بطريقة B/A حاولوا معرفة أي أنواع الوالدين يكونون نشطاء على الإنترنت في أيام الأسبوع عدا الأحد والسبت، لذا، للوصول إلى هذا الجمهور المستهدف، وبعد تقييم البيانات الضخمة التي سجلتها الخوارزميات من اختبار B/A، نشروا المقال في الأوقات التي يكون الوالدان نشطاءً بها بحيث يتم جذب الكثير من الزيارات، وكانت النتيجة أن المقالة قرأها 24000 زائر في 7 ساعات ونحو 41% من المشاهدات أتت عن طريق مواقع التواصل الاجتماعي.

التعليقات: حللت هاف بوست مئات ملايين التعليقات التي تردها سنويًا على مقالاتها، وتبيّن لها أن التعليقات الجيدة تأتي من أشخاص معروفي الهوية، فيما تأتي التعليقات السبية وعديمة الفائدة من الأشخاص مجهولي الهوية، ساعدها معرفة ذلك في اتخاذ قرار بإلزام المعلقين بتسجيل عضوية وإعلان هويتهم، ومنعت التعليقات من أشخاص مجهولي الهوية.

وسرعة بذكاء الغولف أخبار :PGA TOUR

تقوم منصة Quill الذكية التي طورتها شركة Science Narrative بكتابة مقالات صحفية رياضية مباشرة من إحصاءات المباريات! حيث تلخص الجولة بشكل مخصص لكل لاعب في الملعب ولكل جولة على حدة، Quill وتحول أداء اللاعب وما يتوافر من انطباعات عنه وتقديمها للجمهور والمشجعين بحيث يمكن الوصول إليها من خلال محركات البحث، وهذا الخبر قامت Quill بإنشائه وتحريريه لصالح منصة Tour PAG المعنية بأخبار رياضة الغولف، التي أعلنت أنها تستخدم Quill في تقديم أفضل تغطية للمشجعين.

إذن كيف يتم الأمر؟

في معظم الأسابيع، يكون هناك 150 لاعب جولف تقريبًا يتنافسون في دوري TOUR PGA. تبدأ الجولة مبكرًا في تمام الثامنة صباحًا وتنتهي في السادسة مساءً بشكل يومي، وبسبب ضيق وقت الجولات وكثرة عدد اللاعبين، لم يكن لدى القائمين على الدوري القدرة على تحليل أداء جميع اللاعبين سابقًا، بل كانوا يحللون أداء اللاعبين المتصدرين فقط، لذا، لم يكن هناك الإمكانية لتحقيق أمنيات

جميع المشجعين وتحليل أداء لاعبيهم المفضلين.

قال سكوت جيترمان، نائب رئيس العمليات الالكترونية لدى TOUR: "تعمل منصة القصة الأتوماتيكية على تلبية احتياجات الجماهير وتحقيق رغبتهم في تحليل أداء لاعبيهم المفضلين وتقديم ما يمكن قراءته للجمهور، رغم أنه من الممكن ألا يصل هؤلاء اللاعبين إلى الصدارة، وبمساعدة منصة Quill، تمكنا من عمل تحليل وتقرير عن كل لاعب في الملعب بعد نهاية كل جولة".

يعمل فريق Science Narrative مع فريق تطوير المحتوى والمنتجات لدى TOUR لفهم ما يجعل لعبة الجولف مثيرة وفهم ما يهتم به المشجعون في هذه اللعبة، لذا، تم تطوير منصة Quill لتحليل البيانات المستفيضة التي يتم الحصول عليها بشأن اللاعب في كل جولة لإنشاء قصة وتقرير عن أدائه.

ومن خلال القدرة الفائقة للمنصة على تحليل كميات ضخمة من البيانات، أصبح لديها القدرة على صناعة محتويات تثير اهتمام المشجعين لكل لاعب ولهذه اللعبة بشكل عام، بالإضافة لذلك، يتم تطوير وإعداد التقرير ومحتواه بشكل مثير عن اللاعب قبل بداية الجولة التالية، دون الحاجة إلى الانتظار حتى نهاية اليوم.

شبكة PGATour تعاونت أيضًا مع شركة Sports WSC المتخصصة بتطوير منصة يمكنها إنشاء الفيديوهات الذكية التي تعتمد على الذكاء الصناعي من أجل تحليل فيديوهات TOUR التي يتم بثها بالإضافة إلى البيانات التي يتم جمعها بواسطة نظام ShotLink (مملوك لشركة التكنولوجيا CDW وهو نظام ثوري لجمع ونشر الدرجات والبيانات الإحصائية عن كل لقطة يقوم بها كل لاعب في الوقت الفعلي) من أجل إنشاء وتطوير محتويات مناسبة لجميع المنصات الرقمية الخاصة بـTOUR PGA، فخلال دقائق، تقوم منصة إنشاء الفيديوهات بإنشاء مجموعة مختلفة من الفيديوهات التي تشتمل على فيديوهات مدتها 5 دقائق عن اللاعب، وهذه المهمة كانت تأخذ ساعات ليتم إعدادها ونشرها على المنصات الإلكترونية في السابق.

لتوضيح أهمية المشروع، قال سكوت جيترمان نائب رئيس العمليات الرقمية في TOUR PGA: "يساعد بث ملفات الفيديو المباشرة الوقتية باستخدام التكنولوجيا المتقدمة التي تقدمها WSC في الحفاظ على المشجعين والجماهير وجذب تفاعلهم، تقديم الأفضل للجماهير عبر الإنترنت، والمساعدة في وصول العالم مستوى على أكثر جماهير إلى TOUR".

لا يزال أمام قطاع الإعلام، الحيوي، الكثير ليتعلمه ويستفيده من الـ Data Big لا سيما مع تطوّر المجال الأخير وأدواته (جمعًا وتحليلًا وفرزًا وتقديمًا) والتي تصبح يومًا بعد يوم أكثر سهولة في استخدامها من قبل الفرق الصغيرة، وربما الأفراد في وقت قريب، ما يعني آفاقًا جديدة كليًا لتطوير أشكال مبتكرة للعمل الصحفي، ليس لناحية الوصول والتأثير وحسب، بل حتى على جوهر العمل الصحفي ذاته برمته.